Springer Monographs in Mathematics

Arkady L. Onishchik · Rolf Sulanke

Projective and Cayley-Klein Geometries

With 69 Figures

Springer

Arkady L. Onishchik
Faculty of Mathematics
Yaroslavl State University
Sovetskaya 14
150000 Yaroslavl, Russia
e-mail: aonishch@aha.ru

Rolf Sulanke
Institut für Mathematik
Humboldt-Universität zu Berlin
Rudower Chaussee 25
10099 Berlin, Germany
e-mail: sulanke@mathematik.hu-berlin.de

Library of Congress Control Number: 2006930592

Mathematics Subject Classification (2000): 51-01, 51N15, 51N30, 51M10, 51A50, 57S25

ISSN 1439-7382

ISBN-10 3-540-35644-4 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-35644-8 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media
springer.com

Typesetting by the authors using a Springer TeX macro package
Production: LE-TeX Jelonek, Schmidt & Vöckler GbR, Leipzig
Cover design: Erich Kirchner, Heidelberg

Printed on acid-free paper 44/3100YL - 5 4 3 2 1 0

Preface

Projective geometry, and the Cayley-Klein geometries embedded into it, are rather ancient topics of geometry, which originated in the 19th century with the work of V. Poncelet, J. Gergonne, Ch. v. Staudt, A.-F. Möbius, A. Cayley, F. Klein, S. Lie, N. I. Lobatschewski, and many others. Although this field is one of the foundations of algebraic geometry and has many applications to differential geometry, it has been widely neglected in the teaching at German universities — and not only there. In the more recent mathematical literature these classical aspects of geometry are also scarcely taken into account. In the present book, the synthetic projective geometry and some of its recent applications, e.g. of the finite geometries, are mentioned only in passing, i.e., they form the content of some remarks. Instead, we intend to present a systematic introduction of projective geometry as based on the notion of vector space, which is the central topic of the first chapter. In the second chapter the most important classical geometries are systematically developed following the principles founded by A. Cayley and F. Klein, which rely on distinguishing an absolute and then studying the resulting invariants of geometric objects. These methods, determined by linear algebra and the theory of transformation groups, are just what is needed in algebraic as well as differential geometry. Furthermore, they may rightly be considered as an integrating factor for the development of analysis, where we mainly have in mind the harmonic or geometric analysis as based on the theory of Lie groups. Even though, wherever it does not require any extra effort, we allow for vector spaces over an arbitrary skew field, we nevertheless mainly deal with geometries over the real, the complex, or the quaternionic numbers; we also discuss the consequences of extensions as well as restrictions of the scalar domain to the geometries in question. Apart from the real projective geometry we also treat some of the complex and quaternionic geometries in detail, which are rarely presented on an elementary level. The elementary conformal or Möbius geometry is extensively discussed, and even some aspects of the symplectic projective geometry are studied. The concluding Section 2.9 contains a brief introduction into the theory of Lie transformation groups with an outline of the classification of transitive Lie group actions on the n-dimensional spheres and projective spaces. The *octonions* and the *octonion geometries*, which correspond to the exceptional groups rather than to the classical series of Lie groups remain outside of the scope of this book; for this topic we refer to H. Freudenthal [39], M. M. Postnikov [89], and H. Salzmann, D. Betten, Th. Grundhöfer, H. Hähl, R. Löwen, M. Stroppel [97].

With this systematic presentation we hope to provide a useful tool for all who intend to acquire the necessary background material all by themselves or in special seminars. The systematic organization, the detailed proofs, interspersed exercises, and an extensive index starting with a list of symbols, are meant to facilitate choosing the topic of respective interest as well as its study. The appendix contains the definitions of several notions from algebra and geometry used in the text, which are not treated in greater detail there. The book is intended to be a manual and work book, and was by no means written as the text on which a lecture should be based.

A flaw of the book which the reader will soon notice is the missing or, at best, loose connection of the material with its history. The presentation by no means follows the course of the historic development; that would have expanded the size of the book considerably and also claimed too much time and working capacity from the authors as well as the readers. Instead, we decided to take the systematics of the subject from the current point of view as the guiding principle, and to start from the structures also used in other parts of mathematics, e.g. groups or vector spaces, which in fact developed in connection with geometry. We consider labelling propositions, principles, etc., by the names of mathematicians rather as the usual way of denotation than stating any priority or authorship of the mathematician in question for the result. Interpreting these propositions as historically correct claims of authorship would, implicitly, assert to have studied the complete literature, even the correspondence or the the oral traditions up to this moment, in order to exclude the existence of any predecessor, a venture the authors do not want to pride themselves on.

We do not intend to include a complete bibliography of this field. Today it is not difficult to gather arbitrarily extensive lists of references from the available data banks. We only refer to the titles used to work on this subject, where chance and personal choice play an essential part. Frequently, we relied on the well-known basic textbooks by E. Artin [4] and R. Baer [5]. Of course, we have to thank our teachers for the suggestions and orientations we received since the time of our studies and over the long years of our cooperation, which, perhaps sometimes unknowingly, influenced the presentation. Particularly to be mentioned here are the lectures, seminars, and writings by W. Blaschke, E. B. Dynkin, L. A. Kaloujnine, A. G. Kurosch, P. K. Raschewski, and H. Reichardt.

We also want to quote some textbooks and monographs dealing with our subject from a different point of view or with differing emphasis. First, there is the report of results by J. Dieudonné [36] containing more general and more detailed results than this book as well as a comprehensive bibliography. The very interesting and brief textbook by A. Beutelspacher and U. Rosenbaum [12] contains the analytical as well as the *synthetic foundations of projective geometry,* and, in addition, recent applications to *coding theory* and *cryptography.* The *finite geometry* is also discussed there, a subject to which the complete two-volume treatise [10], [11] by A. Beutelspacher is devoted, see also L. M. Batten, A. Beutelspacher [6]. The book [7] by W. Benz presents interesting propositions characterizing geometric transformations under weak assumptions, e.g. the Theorems of Beckmann and Quarles for Euclidean isometries, and the theorem of A. D. Alexandrow concerning Lorentz transformations. In that book the geometries of Lie and Laguerre are also treated, which are not included here. Interesting concrete material discussed from the point of view of Klein's Erlanger Programm can be found in the book [8] by W. Benz. There, apart from the classical geometries, the reader may also find descriptions of the Einstein cylinder universe and the de-Sitter space. The rich summary [69] by Linus Kramer describes the recent development of projective geometry; the structure of the projective group, the relation between incidence properties and algebraic properties of the scalar domain, generalizations of the fundamental theorem, Tits buildings, Moufang planes, polar spaces, and quadratic forms are discussed there, and non-commutative scalar domains receive particular emphasis.

As prerequisites for the understanding of this book the reader is assumed to be familiar with the fundamental notions of linear algebra as well as the affine and Euclidean geometries based on it as they are usually presented in the first year of any course in mathematics or physics. This also includes the knowledge of the affine classification of quadrics and its Euclidean refinement by means of the principal axes transformation. We will frequently use these notions and results without special reference. Moreover, some experience of the metric Euclidean, non-Euclidean, and spherical elementary geometries are useful for understanding this book; as a brief and clearly written, very rich presentation we mention the recently published textbook [1] by I. Agricola and Th. Friedrich. Obviously, for the authors and the readers, who are familiar with or have at their disposal both volumes of the series "Algebra und Geometrie" [82, 83], it is advisable as well as convenient to make use of the relations to the topics discussed there. For this reason we frequently refer to these sources, e.g., Proposition 1 in § 4.2 of [82] is quoted as Proposition I.4.2.1, and, correspondingly, § 6.2 from [83] as § II.6.2. Instead of these books the reader may of course draw on other textbooks which as a rule contain similar results. Familiarity with differentiable manifolds or Lie groups, however, is not required to understand this text. The authors tried to include an extensive index. Nevertheless, the entries usually do not occur twice; e.g., "orthogonal group" is only listed as "group, orthogonal". References to the book itself have the usual format; i.e., Proposition 5 stands for the corresponding proposition in the same section, Proposition 6.3 refers to Proposition 3 in Section 6 of the same chapter, and Corollary 1.3.4 cites Corollary 4 in Section 3 of

Chapter 1. The same holds for formulas: (2.5.19) is formula (19) in Section 5 of Chapter 2.

The nearly 70 illustrations accompanying the text mostly were produced using the software *Mathematica* by St. Wolfram [114]. This program provides numerous possibilities for numerical as well as symbolic calculations and contains a varied graphic apparatus for the visualization of plane and spatial geometric objects, which is naturally tied to Euclidean geometry. On the homepage

http://www-irm.mathematik.hu-berlin.de/~sulanke

some notebooks written using this software can be found. They were developed writing this book and extend the possibilities of Mathematica. They include notebooks for pseudo-Euclidean geometry, by means of which relativity theory, Möbius geometry, Lie's sphere geometry, and, via their Killing forms, semi-simple Lie algebra may be studied. In great detail the three-dimensional Euclidean and the Möbius geometry of k-spheres, $k = 0, 1, 2$, is treated. Some of the formulas obtained by means of symbolic calculations using these notebooks, expressing Möbius invariants in terms of Euclidean invariants similarly to the Coxeter distance of hyperspheres, are included here. On the homepage one also finds a very fast orthogonalization algorithm going back to Erhard Schmidt as well as a procedure to orthogonalize sequences of vectors in pseudo-Euclidean spaces. This opens up possibilities which, because of the scope and complexity of the calculations involved, are impossible to reach by the traditional methods, i.e. "by hand". The so-called "artificial intelligence" reaches its limits as soon as it leaves the finite algorithmic ground and turns towards the abstract conceptual thinking abounding in modern mathematics. Already implementing the naive set theory in Mathematica meets serious difficulties, as the example of a system of Mathematica notebooks and packages designed for this purpose shows, which can also be found on the homepage.

The authors gratefully acknowledge the support of their work provided by the Humboldt-Universität zu Berlin. In particular, we thank Heike Pahlisch, Hubert Gollek, Jürgen Gehne, and Hans-Joachim Spitzer for their assistance and technical help. We express our sincere gratitude to Bernd Wegner from the Technische Universität Berlin for his interest in our work as well as his endeavor to publish it. We thank the Springer-Verlag for including the English translation of the book into its publishing program and the permission to publish the German manuscript on the homepage above for private use. Last but not least we thank Andreas Nestke for his careful translation of the German manuscript and the always friendly and fruitful cooperation with us during his work. We think and hope that this book may cause some of our colleagues and many students to considerations and investigations of their own.

Berlin, June 2006

A. L. Onishchik
R. Sulanke

Contents

List of Figures

List of Tables

1

Projective Geometry

The presentation of projective geometry in this chapter will be based on linear algebra. We will start with some remarks concerning the central projection intended to be a motivation, which will then lead to the definition of a projective geometry as the lattice of subspaces of a vector space. The first fundamental result to be proved is the Main Theorem of projective geometry relating its geometric properties closely to the algebraic ones: If the dimension is at least two, i.e. for the plane and all higher-dimensional projective spaces, then geometric isomorphy implies algebraic isomorphy. There exists a collineation between two spaces, i.e. a bijective map transforming lines into lines, if and only if, first, both their dimensions coincide, and, second, their scalar fields are isomorphic. The one-dimensional case, the projective line, requires special considerations, since the collineation condition cannot work there; we will introduce the cross ratio of four collinear points and prove Staudt's Main Theorem: Invariance of the harmonic position alone implies invariance of the cross ratio. The duality of vector spaces has a projective counterpart, the duality of projective spaces. This duality enables to study correlations, i.e. bijective maps from point spaces to spaces of hyperplanes preserving incidence. Symmetric auto-correlations will be classified by reducing the problem to the classification of skew-symmetric and Hermitean biforms. The results of both classifications will be interpreted geometrically: as linear line complexes in the skew-symmetric case, and as hyperquadrics in the Hermitean case. Over the scalar domains of the real numbers, the complex numbers, and the quaternions we will obtain a complete classification. Finally, we will also investigate the geometric implications brought about by extending or restricting the scalar fields. In particular, we will describe the Hopf fibrations, which play a prominent part in topology.

1.1 Projective Spaces

1.1.1 Definitions and First Properties

There are two, in their course essentially different approaches to projective geometry, the *synthetic* and the *analytic* one. The latter should better be called the linear-algebraic approach, since linear algebra rather than analysis forms its basis and provides the methods. Nevertheless, the name "analytic geometry" became firmly established in the course of history. In synthetic geometry sets of geometric elements of various kinds, e.g. points, lines, planes, are given, and the existing relations between them are fixed axiomatically; these axioms correspond to familiar geometric concepts. Instead, here we prefer an approach based on linear algebra, which is more abstract but at the same time much more general and, in principle, even simpler. There is no need to build a special theory whose axioms have to be changed frequently, but everything relies on the notion of vector space. This structure is fundamental for analysis as well as many applications of mathematics, hence its knowledge is compulsory for every mathematician. Moreover, any undergraduate course would include this topic. Synthetic projective geometry is already systematically treated in the comprehensive two-volume treatise [106] by O. Veblen and J. W. Young that appeared in 1910 and 1918. The two-dimensional case, i.e. plane projective geometry, plays a prominent part there, cf. G. Pickert [87]. Later in this section we will present a brief discussion of its simple axiomatics. In this context, interesting connections between this synthetic axiomatics and algebraic properties of the scalar domain will arise, that can be constructed on the basis of these axioms. Conversely, these relations will become evident by looking for the analytic projective geometries satisfying certain synthetic axioms. A vivid introduction into the plane projective geometry can be found in H. S. M. Coxeter [34].

In the sequel we will always suppose that a vector space $\boldsymbol{V}$ over a skew field K is given[1]. The set of one-dimensional subspaces of $\boldsymbol{V}$ will be denoted by $\boldsymbol{P}(\boldsymbol{V})$; we will call it the *projective space associated with the vector space* $\boldsymbol{V}$. It is common to define the projective space as an extension of the affine point space by means of a so-called *hyperplane at infinity.* The one-dimensional subspaces of the vector space associated with the affine geometry bijectively correspond to the families of parallel lines in the affine space, which are determined by attaching the vectors of the subspace to the points in the affine point space. Expressed differently, each of these families is the orbit of a one-dimensional vector subspace in point space. Intuitively, one thinks of all the

[1] The only skew field essential for the following is the skew field of quaternions $\mathbf{H}$. To distinguish between a left and a right vector space is only necessary in the case of a non-commutative field K. For now it will suffice to take the fields of real or of complex numbers as K, hence the whole discussion will remain within the usual framework of vector spaces.

lines within such a family as passing through a common single point at infinity. The set of all these points at infinity corresponding to all the different families of parallels constitute the hyperplane at infinity to be attached. To make this concept more precise is elementary, but lengthy and laborious. It is much simpler to start from the definition above and then embed the affine geometry into the projective one by distinguishing an arbitrary hyperplane, which, consequently, is to be called the *hyperplane at infinity,* cf. Section 5 of this chapter. The projective plane will serve as an example to illustrate these observations. In Example 2 below we will see that the projective extension of point space is necessary and expedient to describe central projections:

Example 1. The Real Projective Plane. Let $\boldsymbol{V} = \mathbf{R}^3$ be the three-dimensional vector space over the field $\mathbf{R}$ of real numbers, and denote by $\boldsymbol{A}^3$ the associated affine space; we will frequently consider this to be the physical space. By $\boldsymbol{e} \in \boldsymbol{A}^3$ we denote the point with coordinates $(0,0,1)$, i.e. the unit point on the z-axis. Attaching the one-dimensional subspaces of $\boldsymbol{V}$ to the point $\boldsymbol{e}$ defines a bijection from $\boldsymbol{P}(\boldsymbol{V})$ onto the set of lines passing through $\boldsymbol{e}$, cf. Figure 1.1. Attaching in a similar way the two-dimensional subspaces of $\boldsymbol{V}$ to $\boldsymbol{e}$ defines a bijection between the set of these subspaces and the set of planes through $\boldsymbol{e}$.

Each line through $\boldsymbol{e}$, which is not parallel to the x,y-plane $\boldsymbol{H}$ $(z=0)$, meets this plane in a unique point. Obviously, the lines through $\boldsymbol{e}$ parallel to $\boldsymbol{H}$ form a one-parameter family, whereas to describe the others we need to specify two parameters, e.g. the coordinates of their intersection point with the x,y-plane $\boldsymbol{H}$. All lines through $\boldsymbol{e}$ lying in a plane $\boldsymbol{B}$ through $\boldsymbol{e}$, i.e. whose vectors belong to the same two-dimensional subspace $\boldsymbol{W} \subset \boldsymbol{V}$, intersect $\boldsymbol{H}$ in one line, the line of intersection $\boldsymbol{B} \cap \boldsymbol{H}$. The only exception to this is the vector space $\boldsymbol{U}$ of $\boldsymbol{H}$ itself, which, by attaching it to $\boldsymbol{e}$, determines the plane $\boldsymbol{B}_0 = \boldsymbol{e} + \boldsymbol{U}$ parallel to $\boldsymbol{H}$. Thus it appears natural to extend the affine plane by a “line at infinity”, whose points just correspond to the lines through $\boldsymbol{e}$ parallel to $\boldsymbol{H}$. We will denote this extension by $\hat{\boldsymbol{H}}$. In this picture, the line at infinity appears as the intersection of the similarly extended plane $\hat{\boldsymbol{B}}_0$ with $\hat{\boldsymbol{H}}$. It is common to introduce the *real projective plane* this way: Start from the affine plane and attach a line at infinity to it, whose points correspond to the non-oriented directions of the plane, i.e. the equivalence classes of parallel lines in the affine plane. So all parallel lines in such a class appear to meet in a unique point of the line at infinity. Since there is a bijective correspondence

$$p : \boldsymbol{x} \in \boldsymbol{P}(\boldsymbol{V}) \mapsto p(\boldsymbol{x}) \in \hat{\boldsymbol{H}}$$

between the elements of $\boldsymbol{P}(\boldsymbol{V})$ and the points of the extended plane $\hat{\boldsymbol{H}}$, we arrive at a clear picture of $\boldsymbol{P}(\boldsymbol{V})$[2]. Note that, in the projective sense, the line at infinity in the picture above is equivalent to any of the other lines in the plane. □

[2] The topological shape of the real projective plane will be described in Examples 2.5.1 and 2.5.4 in the following chapter.

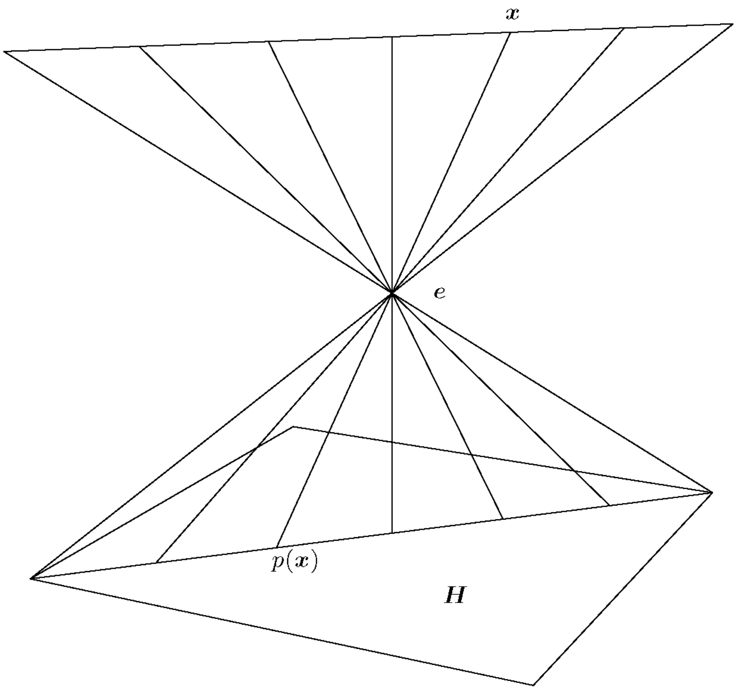


Fig. 1.1. Line bundle and the projective plane.

The bijection p described in the previous example induces a similar bijection between the two-dimensional subspaces $\boldsymbol{W} \subset \boldsymbol{V}$ of the vector space $\boldsymbol{V}$ and the set of lines $\boldsymbol{h} \subset \hat{\boldsymbol{H}}$ in the projective plane $\hat{\boldsymbol{H}}$:

$$\boldsymbol{W} \subset \boldsymbol{V} \mapsto \boldsymbol{B} = \boldsymbol{e} + \boldsymbol{W} \mapsto \boldsymbol{h} = \boldsymbol{B} \cap \hat{\boldsymbol{H}}.$$

Here, the subspace $\boldsymbol{U}$ of $\boldsymbol{H}$ corresponds to the plane parallel to $\boldsymbol{H}$ through $\boldsymbol{e}$ and the line at infinity of $\hat{\boldsymbol{H}}$. These considerations can easily be generalized to the n-dimensional case; they suggest to interpret the *lattice of subspaces of the vector space* $\boldsymbol{V}$ (cf. § I.4.2) in the following projective-geometric way extending the affine point of view that, basically, amounts to renaming the structures provided by the lattice of subspaces:

Definition 1. Let $\boldsymbol{V}$ be a right vector space over the skew field K. By $\mathfrak{P}(\boldsymbol{V})$ we will denote the lattice of subspaces of $\boldsymbol{V}$ and call it the *projective geometry of* $\boldsymbol{V}$: its elements are called *projective subspaces.* The *projective dimension* Dim of an element $\boldsymbol{W} \in \mathfrak{P}(\boldsymbol{V})$ is equal to the dimension of the vector space diminished by one:

$$\operatorname{Dim} \boldsymbol{W} := \dim \boldsymbol{W} - 1.$$

Since all right (and left, respectively) vector spaces $\boldsymbol{V}$ of equal finite dimension $n+1$ over the same skew field K are isomorphic, we will usually just write

$$\mathfrak{P}_K^n := \mathfrak{P}(\boldsymbol{V}^{n+1})$$

This is the *n-dimensional projective geometry over K*. The zero-dimensional elements in $\mathfrak{P}(\boldsymbol{V})$ are the *points* of the *projective space*

$$\boldsymbol{P}(\boldsymbol{V}) := \{\boldsymbol{a} \in \mathfrak{P}(\boldsymbol{V}) |\ \mathrm{Dim}\, \boldsymbol{a} = 0\}.$$

For the n-dimensional projective space we will also use the notation

$$\boldsymbol{P}_K^n := \boldsymbol{P}(\boldsymbol{V}^{n+1}) \qquad (n < \infty).$$

If the skew field K is clear from the context, we will omit the index K. In general, $\boldsymbol{W}$ is called a *projective k-plane* if $\mathrm{Dim}\, \boldsymbol{W} = k$; the set of k-planes of an n-dimensional projective geometry will be denoted by $\boldsymbol{P}_{n,k}$ and referred to as the *Graßmann manifold.*[3] As usual, the elements of $\boldsymbol{P}_{n,1}$ are called *lines*, those of $\boldsymbol{P}_{n,2}$ are the *planes*, and for $k = n - 1$ we have the *hyperplanes*; the word "projective" will be omitted if there is no danger of confusion. Two elements $\boldsymbol{U}, \boldsymbol{W} \in \mathfrak{P}(\boldsymbol{V})$ are said to be *incident*, written as

$$\boldsymbol{U} \,\iota\, \boldsymbol{W}, \text{negation: } \boldsymbol{U} \,\bar{\iota}\, \boldsymbol{H},$$

if $\boldsymbol{U} \subset \boldsymbol{W}$ or $\boldsymbol{W} \subset \boldsymbol{U}$ holds. Keeping the notation we transfer the *inclusion relation*[4] $\subset$ of subspaces of $\boldsymbol{V}$ to $\mathfrak{P}(\boldsymbol{V})$. The vector space $\boldsymbol{V}$ itself is the largest element with respect to the order $\subset$, and the trivial subspace is the smallest, $\boldsymbol{o} := \{\mathfrak{o}\} \in \mathfrak{P}(\boldsymbol{V})$; we decided to called it the *nopoint* in projective space. The *section* of two projective subspaces $\boldsymbol{U}, \boldsymbol{W} \in \mathfrak{P}(\boldsymbol{V})$ is understood to be their (set-theoretic) intersection:

$$\boldsymbol{U} \wedge \boldsymbol{W} := \boldsymbol{U} \cap \boldsymbol{W}.$$

Their *join* is defined to be the smallest subspace containing both, i.e. the linear span of their union (cf. § I.4.2):

$$\boldsymbol{U} \vee \boldsymbol{W} := \mathfrak{L}(\boldsymbol{U} \cup \boldsymbol{W}) = \boldsymbol{U} + \boldsymbol{W}.$$

□

It is straightforward to extend the notions *section* and *join* to arbitrary families of subspaces $(\boldsymbol{H}_\iota)_{\iota \in I}$:

---

[3] In correspondence with the dimension convention, the Graßmann manifold has a differing notation in vector algebra, $G_{n+1,k+1} (= \boldsymbol{P}_{n,k})$.

[4] We do not distinguish between $\subset$ and $\subseteq$, i.e., $A \subset B$ is the same as $A \subseteq B$, equality is not excluded.

$$\bigwedge_{\iota \in I} \boldsymbol{H}_\iota := \bigcap_{\iota \in I} \boldsymbol{H}_\iota,$$
$$\bigvee_{\iota \in I} \boldsymbol{H}_\iota := \sum_{\iota \in I} \boldsymbol{H}_\iota.$$

The familiar rules of linear algebra (cf. § I.4.2) remain valid; this is summarized by stating that projective geometry is a "complete lattice" with respect to the inclusion $\subset$.

The nopoint $\boldsymbol{o}$ has only a formal, hardly any geometric meaning. Two projective subspaces are called *skew* if their intersection is the nopoint. By Definition 1 we have $\operatorname{Dim} \boldsymbol{o} = -1$. This definition together with the dimension theorem of linear algebra (Proposition I.4.6.3) immediately implies the analogous result for projective subspaces:

Proposition 1. *Consider $\boldsymbol{U}, \boldsymbol{W} \in \mathfrak{P}(\boldsymbol{V})$. Then the following dimension formula holds:*

$$\operatorname{Dim}(\boldsymbol{U} \wedge \boldsymbol{W}) + \operatorname{Dim}(\boldsymbol{U} \vee \boldsymbol{W}) = \operatorname{Dim} \boldsymbol{U} + \operatorname{Dim} \boldsymbol{W}.$$

□

Exercise 1. Use the dimension formula (Proposition 1) to discuss all possible position relations of projective subspaces $\boldsymbol{F}^k$, $\boldsymbol{H}^m \subset \boldsymbol{P}^n$ for $0 \leq k \leq m < n \leq 4$. Find an example for each of the resulting cases. The two subspaces are skew if and only if their join is not contained in any $(k+m)$-dimensional subspace. Compare the results with the corresponding ones in affine space (s. § I.4.6).

The next example clearly shows the advantage of the historically younger concept of projective geometry: From this point of view, the properties of central projection are derived by direct reference to the dimension formula. From the affine point of view, i.e. looking at the projective plane as the extended affine plane, it is necessary to distinguish several cases. We will nevertheless discuss them, as they also provide a useful bridge between affine and projective geometry.

Example 2. Central Projection. We start by considering the three-dimensional projective space $\boldsymbol{P}^3$. Let $\boldsymbol{H}$ and $\boldsymbol{B}$ be different projective planes in $\boldsymbol{P}^3$, and let $\boldsymbol{e}$ be a point not incident with either plane. The *central projection* from $\boldsymbol{B}$ onto $\boldsymbol{H}$ with *projection center* $\boldsymbol{e}$ is the map:

$$q : \boldsymbol{x} \in \boldsymbol{B} \longmapsto \hat{\boldsymbol{x}} = q(\boldsymbol{x}) := (\boldsymbol{x} \vee \boldsymbol{e}) \wedge \boldsymbol{H} \in \boldsymbol{H}.$$

In Section 1.3 we will give a more general definition of the central projection, which plays a very important part in projective geometry. Interpreting the definition projectively, it is easy to see that $q : \boldsymbol{B} \to \boldsymbol{H}$ is a bijection mapping lines to lines; this is a direct consequence of the dimension formula.

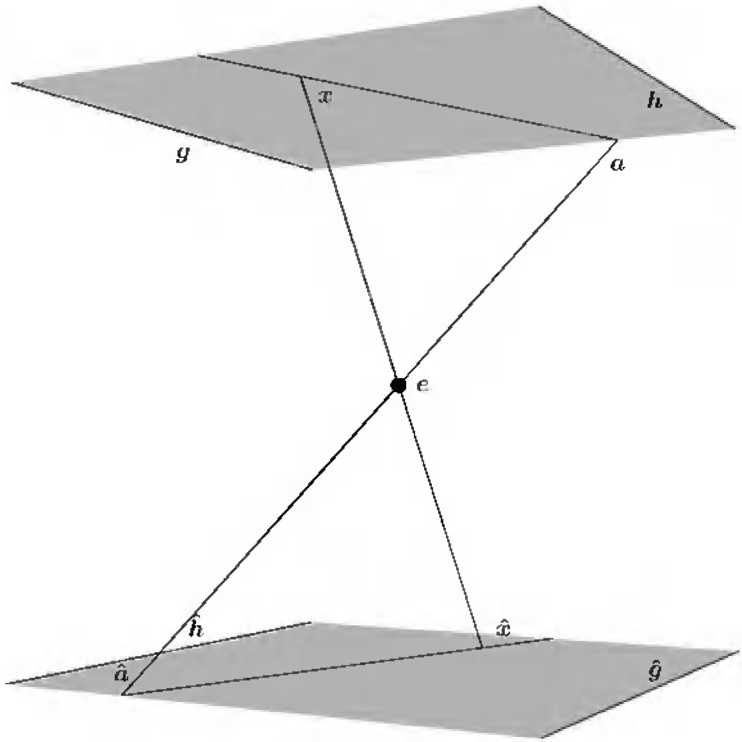


Fig. 1.2. Central projection of parallel planes.

From the affine point of view the situation turns out to be more varied. First we have to distinguish two cases.

Case 1: The planes $\boldsymbol{B}$ and $\boldsymbol{H}$ are parallel. In Figure 1.2, let the upper plane $\boldsymbol{B} = \boldsymbol{g} \vee \boldsymbol{h}$ be the pre-image, and take the lower one as the image $\boldsymbol{H} = \hat{\boldsymbol{g}} \vee \hat{\boldsymbol{h}}$. For the images of the lines $\boldsymbol{g}, \boldsymbol{h}$ and points $\boldsymbol{a}, \boldsymbol{x}$ of $\boldsymbol{B}$ above we then have:

$$q(\boldsymbol{g}) = \hat{\boldsymbol{g}},\ q(\boldsymbol{h}) = \hat{\boldsymbol{h}},\ q(\boldsymbol{a} \vee \boldsymbol{x}) = q(\boldsymbol{a}) \vee q(\boldsymbol{x}) = \hat{\boldsymbol{a}} \vee \hat{\boldsymbol{x}}.$$

In this case, q is actually an affine map between the affine, not just the extended planes; it thus maps parallel lines to parallel lines. The projectively extended planes $\boldsymbol{B}, \boldsymbol{H}$ intersect in their common line at infinity, $\boldsymbol{h}_\infty = \boldsymbol{B} \wedge \boldsymbol{H}$, containing the fixed points of the central projection q.

Case 2: The pre-image plane $\boldsymbol{B} = \boldsymbol{c} \vee \boldsymbol{d} \vee \boldsymbol{a}$ and the image plane $\boldsymbol{H} = \boldsymbol{c} \vee \boldsymbol{d} \vee \hat{\boldsymbol{a}}$ are not parallel; they intersect in a line $\boldsymbol{c} \vee \boldsymbol{d} = \boldsymbol{B} \wedge \boldsymbol{H}$ formed by the fixed points of q, cf. Fig. 1.3. Let $\boldsymbol{e} \vee \boldsymbol{u} \vee \boldsymbol{v}$ be the plane parallel to $\boldsymbol{H}$ through the projection center $\boldsymbol{e}$; it intersects $\boldsymbol{B}$ in the line $\boldsymbol{u} \vee \boldsymbol{v}$. For $\boldsymbol{w} \in \boldsymbol{u} \vee \boldsymbol{v}$ the line $\boldsymbol{w} \vee \boldsymbol{e}$ is parallel to $\boldsymbol{H}$; hence it meets the line at infinity $\boldsymbol{h}_\infty \subset \boldsymbol{H}$ of the image plane in the sense of extension as described in the previous example. Moreover,

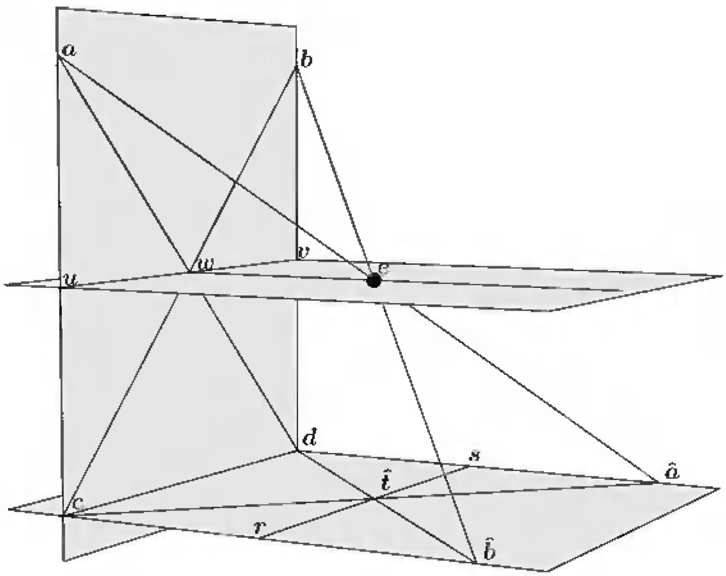


Fig. 1.3. Central projection of non-parallel planes.

$$q(\boldsymbol{u} \vee \boldsymbol{v}) = \boldsymbol{h}_\infty.$$

All lines in the pre-image plane $\boldsymbol{B}$ incident with $\boldsymbol{w}$, e.g. the lines $\boldsymbol{a} \vee \boldsymbol{d}$ and $\boldsymbol{b} \vee \boldsymbol{c}$, are mapped to parallel lines:

$$q(\boldsymbol{a} \vee \boldsymbol{d}) = \hat{\boldsymbol{a}} \vee \boldsymbol{d},\ q(\boldsymbol{b} \vee \boldsymbol{c}) = \hat{\boldsymbol{b}} \vee \boldsymbol{c}.$$

Thus the pencil of lines through $\boldsymbol{w}$ corresponds to the pencil of parallel lines in the common, non-oriented direction of the image plane $\boldsymbol{H}$ determined by their images. On the other hand, the images of parallel lines from $\boldsymbol{B}$ intersect:

$$q(\boldsymbol{a} \vee \boldsymbol{c}) \wedge q(\boldsymbol{b} \vee \boldsymbol{d}) = (\hat{\boldsymbol{a}} \vee \boldsymbol{c}) \wedge (\hat{\boldsymbol{b}} \vee \boldsymbol{d}) = \hat{\boldsymbol{t}}.$$

Denote by $\boldsymbol{b}_\infty$ the line at infinity of the pre-image plane $\boldsymbol{B}$ and by $\boldsymbol{t}$ the point of intersection on it determined by the parallel lines $\boldsymbol{a} \vee \boldsymbol{c}$ and $\boldsymbol{b} \vee \boldsymbol{d}$. Then we have $q(\boldsymbol{t}) = \hat{\boldsymbol{t}}$. Hence the line of intersection $\boldsymbol{r} \vee \boldsymbol{s}$ of the image plane $\boldsymbol{H}$ with the plane parallel to the pre-image plane $\boldsymbol{B}$ through $\boldsymbol{e}$ is the image of the line at infinity of $\boldsymbol{B}$: $q(\boldsymbol{b}_\infty) = \boldsymbol{r} \vee \boldsymbol{s}$.

These considerations imply that q cannot be an affine map, since it does not preserve parallelism. The situation just described is known in photography as the occurrence of plunging lines in pictures of high buildings taken with inclined camera. From the point of view of geometry, central projection is the ideal model for photographing with a pinhole camera; unfortunately, the luminosity decreases with increasing precision of the reproduction. Central projections and their generalizations are a natural point of departure for projective geometry as they map lines into lines and preserve incidence. □

Exercise 2. Introduce a cartesian coordinate system adapted to the configuration shown in Fig. 1.3 and prove the statements in the previous example by computations based on the methods of affine geometry.

For formal reasons it is often advisable to adjoin the nopoint to projective space. Therefore we define the disjoint union

$$\boldsymbol{P}^n_o := \boldsymbol{P}^n \cup \{\boldsymbol{o}\}. \tag{1}$$

As a first application consider the *canonical map*,

$$\pi : \mathfrak{x} \in \boldsymbol{V}^{n+1} \mapsto \boldsymbol{x} := [\mathfrak{x}] \in \boldsymbol{P}^n_o, \tag{2}$$

where $[\mathfrak{x}]$ denotes the linear hull $[\mathfrak{x}] = \mathfrak{L}(\{\mathfrak{x}\})$ in the vector space $\boldsymbol{V}$. The map π allows to identify the elements $\boldsymbol{H} \in \mathfrak{P}(\boldsymbol{V})$ with the point sets $\pi(\boldsymbol{H}) \subset \boldsymbol{P}^n_o$; this identification preserves the order. If not explicitly stated otherwise, we will always assume that the nopoint $\boldsymbol{o}$ is included into each projective subspace of $\boldsymbol{P}^n_o$. Since the section of projective subspaces is also a projective subspace, setting

$$M \in \mathcal{P}(\boldsymbol{P}^n_o) \longmapsto \bigvee(M) := \bigwedge_{M \subset \boldsymbol{H} \in \mathfrak{P}(\boldsymbol{V})} \boldsymbol{H} \in \mathfrak{P}(\boldsymbol{V}) \tag{3}$$

defines the *projective hull* of the subset $M \subset \boldsymbol{P}^n_o$. Using the identification above we obtain

$$\bigvee(M) = \pi(\mathfrak{L}(\pi^{-1}(M))) = \bigcap_{M \subset \boldsymbol{H} \in \mathfrak{P}(\boldsymbol{V})} \boldsymbol{H} \in \mathcal{P}(\boldsymbol{P}^n_o).$$

$\bigvee(M)$ will also be called *the projective subspace spanned by* M.

Exercise 3. Prove that (3) defines a *hull operator*, i.e. that the following properties hold:

a) $M \subset \bigvee(M)$;

b) $M \subset L$ implies $\bigvee(M) \subset \bigvee(L)$;

c) $\bigvee(\bigvee(M)) = \bigvee(M)$.

Prove, in addition, that:

d) $\bigvee(\emptyset) = \boldsymbol{o} \in \bigvee(M)$ for all $M \subset \boldsymbol{P}^n_o$.

e) The equality $\bigvee(M) = M$ holds if and only if M is a projective subspace.

Exercise 4. Prove that a subset $M \subset \boldsymbol{P}^n_o$ is a projective subspace if and only if, firstly, $\boldsymbol{o} \in M$ and, secondly, with every $\boldsymbol{x}, \boldsymbol{y} \in M, \boldsymbol{x} \neq \boldsymbol{y}$, the connecting line $\boldsymbol{x} \vee \boldsymbol{y}$ also belongs to M.

Exercise 5. Let $\boldsymbol{H}^{n-1} \subset \boldsymbol{P}^n_o$ be a projective hyperplane. Prove that

$$\bigvee(\boldsymbol{P}^n_o \setminus \boldsymbol{H}^{n-1}) = \boldsymbol{P}^n_o.$$

Now we will transfer the notion of linear independence to projective geometry.

Definition 2. Let $(\boldsymbol{x}_\iota)_{\iota \in I}$ be a family (or set) of points $\boldsymbol{x}_\iota \in \boldsymbol{P}^n_o$. $(\boldsymbol{x}_\iota)_{\iota \in I}$ is said to be *in general position* if for each of its subsequences (or subsets) with $k+1$ points, $0 \le k \le n$,

$$\text{Dim } \boldsymbol{x}_{\iota_0} \vee \boldsymbol{x}_{\iota_1} \vee \ldots \vee \boldsymbol{x}_{\iota_k} = k.$$

A sequence (or set) $(\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_r)$ of points is called *projectively independent* if

$$\text{Dim } \boldsymbol{x}_0 \vee \boldsymbol{x}_1 \vee \ldots \vee \boldsymbol{x}_r = r.$$

□

Exercise 6. Let $\boldsymbol{x}_i = [\mathfrak{x}_i] \in \boldsymbol{P}^n_o$. Prove: a) The sequence $(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_k)$, $k \le n$, is in general position if and only if the vectors $(\mathfrak{x}_0, \ldots, \mathfrak{x}_k)$ are linearly independent. – b) An arbitrary sequence $(\boldsymbol{x}_i)_{i \in I}$ is in general position if and only if every subsequence of at most $n+1$ points is projectively independent. – c) Each subsequence of a sequence in general position is itself in general position. (Similar statements to a), b), c) hold for sets.)

Corollary 2. *In n-dimensional projective space there are $n+1$ projectively independent points; every sequence or set with $r > n+1$ points is projectively dependent. There is at least one set M containing $n+2$ points in general position in $\boldsymbol{P}^n_o$.*

P r o o f. First, note that the points $(\mathfrak{a}_i)_{i=0,\ldots,n}$ of $\boldsymbol{V}^{n+1}$ corresponding to any basis $\boldsymbol{a}_i = [\mathfrak{a}_i]$ are projectively independent. Then add the *unit point* $\boldsymbol{e} = [\mathfrak{a}_0 + \mathfrak{a}_1 + \ldots + \mathfrak{a}_n]$ to the set $\{\boldsymbol{a}_0, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ of these so-called *base points*. Together, we obtain $n+2$ points in general position. □

Example 3. **Projective Lines.** Consider the projective line $\boldsymbol{P}^1$ over the skew field K. Let $(\mathfrak{a}_0, \mathfrak{a}_1)$ be a basis of the associated vector space $\boldsymbol{V}^2$. The *projective scale* corresponding to $(\mathfrak{a}_0, \mathfrak{a}_1)$ is defined as follows[5]:

$$\xi : \boldsymbol{x} = [\mathfrak{a}_0 x^0 + \mathfrak{a}_1 x^1] \in \boldsymbol{P}^1 \mapsto \xi(\boldsymbol{x}) := \begin{cases} x^1(x^0)^{-1} \in K & \text{if } x^0 \neq 0, \\ \infty & \text{if } x^0 = 0. \end{cases} \tag{4}$$

Obviously, this definition is independent of the chosen representatives; to both, $\mathfrak{x}$ and $\mathfrak{x}\lambda$, $\lambda \neq 0$,, the same value is assigned. Setting $\hat{K} = K \cup \{\infty\}$

[5] As is common in tensor calculus, the upper indices label the coordinates of a vector, here 0, 1; they are not to be confused with powers.

the map $\xi : \boldsymbol{P}^1 \to \hat{K}$ becomes a bijection establishing a close and natural relation between the properties of K and those of the projective geometry. The point $\boldsymbol{a}_0 = [\mathfrak{a}_0]$ is called the *zero point*, $\boldsymbol{a}_1 = [\mathfrak{a}_1]$ the *point at infinity*, and $\boldsymbol{e} = [\mathfrak{a}_0 + \mathfrak{a}_1]$ the *unit point* of the projective scale; the following relations are obvious:

$$\xi(\boldsymbol{a}_0) = 0,\ \xi(\boldsymbol{a}_1) = \infty,\ \xi(\boldsymbol{e}) = 1.$$

In the cases $K = \mathbf{R}, \mathbf{C}$ or $\mathbf{H}$, a metric as well as a topology are defined on the scalar domain. Then there is a standard topology on $\hat{K}$, that of the so-called *Alexandroff compactification*, cf. W. Rinow [94], § 28: This is nothing but the usual topology of K to which the complements in $\hat{K}$ of all compact sets in K are added as the neighborhoods of ∞. Considering them as homeomorphisms the projective scales transfer the topologies defined on $\hat{K}$ onto the corresponding projective lines. The space obtained this way is completely different from the affine and the Euclidean one: Topologically, the real projective line $\boldsymbol{P}^1_{\mathbf{R}}$ is a circle S^1, the complex one, $\boldsymbol{P}^1_{\mathbf{C}}$, is the *Riemann sphere* S^2, and the *quaternionic projective line* $\boldsymbol{P}^1_{\mathbf{H}}$ is homeomorphic to the four-dimensional sphere S^4. Later we will see that the above topology does not depend on the chosen projective scale, i.e. the basis $(\mathfrak{a}_0, \mathfrak{a}_1)$, cf. Exercise 2.4. □

A first example for the *duality* in projective geometry, which will be treated in detail in Section 1.6 below, is the following proposition:

Proposition 3. *Let $\mathfrak{P}^n$ be a projective geometry over K. Then:*

a) For each couple of points $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{P}^n$, $\boldsymbol{x} \neq \boldsymbol{y}$, there is a unique line $\boldsymbol{h} \in \boldsymbol{P}_{n,1}$ incident with both, the connecting line $\boldsymbol{h} = \boldsymbol{x} \vee \boldsymbol{y}$.

b) For each couple of hyperplanes $\boldsymbol{X}, \boldsymbol{Y} \in \boldsymbol{P}_{n,n-1}$, $\boldsymbol{X} \neq \boldsymbol{Y}$, there is a unique $(n-2)$-plane $\boldsymbol{H} \in \boldsymbol{P}_{n,n-2}$ incident with both, the $(n-2)$-plane of intersection $\boldsymbol{H} = \boldsymbol{X} \wedge \boldsymbol{Y}$.

Proof. a) $\boldsymbol{x} \neq \boldsymbol{y}$ implies $\boldsymbol{x} \wedge \boldsymbol{y} = \boldsymbol{o}$. Since $\operatorname{Dim} \boldsymbol{o} = -1$, the dimension formula yields $\operatorname{Dim} \boldsymbol{x} \vee \boldsymbol{y} = 1$. If $\boldsymbol{h}_1$ is any line incident with $\boldsymbol{x}$ and $\boldsymbol{y}$, then we have $\boldsymbol{h} = \boldsymbol{x} \vee \boldsymbol{y} \subset \boldsymbol{h}_1$, and hence $\operatorname{Dim} \boldsymbol{h} = \operatorname{Dim} \boldsymbol{h}_1 = 1$ implies $\boldsymbol{h} = \boldsymbol{h}_1$.
b) From $\boldsymbol{X} \neq \boldsymbol{Y}$ we conclude $\boldsymbol{X} \vee \boldsymbol{Y} = \boldsymbol{P}^n$. Since $\operatorname{Dim} \boldsymbol{P}^n = n$, the dimension formula yields $\operatorname{Dim} \boldsymbol{X} \wedge \boldsymbol{Y} = n-2$. If $\boldsymbol{H}_1$ is any $(n-2)$-plane incident with $\boldsymbol{X}$ and $\boldsymbol{Y}$, then we have $\boldsymbol{H} = \boldsymbol{X} \wedge \boldsymbol{Y} \supset \boldsymbol{H}_1$, and hence $\operatorname{Dim} \boldsymbol{H} = \operatorname{Dim} \boldsymbol{H}_1 = n-2$, i.e. $\boldsymbol{H} = \boldsymbol{H}_1$. □

1.1.2 Plane Projective Geometry

Example 4. Plane Incidence Geometry. For $n = 2$ Proposition 3 and Corollary 2 comprise the following statements:

P.1. *Any two different points are incident with exactly one line.*
P.2. *Any two different lines are incident with exactly one point.*
P.3. *There are four points no three of which lie on the same line.*

These three statements may serve as axioms for a synthetic projective *plane incidence geometry* Π. Such a geometry has two basic sets: points $x \in P$ and lines $X \in G$, together with a single relation, the incidence ι. Points and lines may be incident, $x \, \iota \, X$, or they may not be incident, $x \, \bar{\iota} \, Y$. The properties of this relations are described by the axioms P.1–3. However, these axioms do not suffice to guarantee the existence of a skew field K whose plane projective geometry is isomorphic to the one determined by the axioms. The existence of such a scalar domain can be achieved by additional axioms. In this context, interesting interrelations between more general scalar domains and synthetic-geometric properties of the projective planes arise, cf. G. Pickert [87]. For more recent developments concerning this subject we refer to the survey article [69] by L. Kramer, where in particular the case of non-commutative scalar domains is stressed. A very vivid presentation of the real plane projective geometry based on a synthetic system of axioms goes back to H. S. M. Coxeter [34]. The multifaceted textbook [12] by A. Beutelspacher and U. Rosenbaum treats the synthetic as well as the analytic foundations of projective geometry. The books [106] by O. Veblen and J. W. Young contain a synthetic presentation of higher-dimensional projective geometry. Note however, that the axioms necessary to describe incidence in three-dimensional space restrict the scalar field decisively stronger than in the case of planar incidence geometry. □

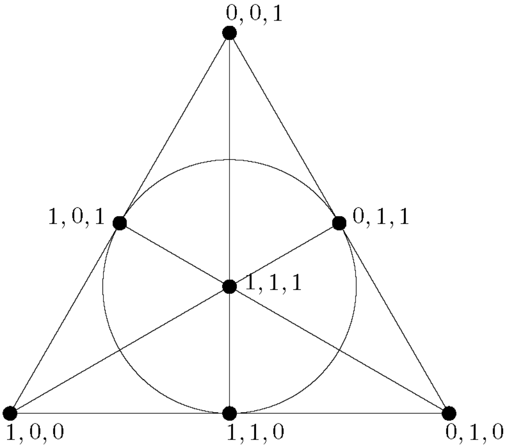


Fig. 1.4. The projective plane over $\mathbf{Z}_2$.

Exercise 7. Prove that in every plane projective incidence geometry Π satisfying the axioms P.1–3 there are four lines no three of which meet in a point. Hint: First prove that on the lines $A \in G$ connecting the points in a configuration as in axiom P.3 there are at least three different points.

Exercise 8. Consider a projective geometry $\mathfrak{P}_K^n$ over a field K. Prove: a) If there exists a line $\boldsymbol{h} \in \mathfrak{P}_K^n$ containing exactly $p+1$ points, p any prime, then $K = \mathbf{Z}_p$, the field of cosets mod p. – b) The projective geometry $\mathfrak{P}_{\mathbf{Z}_2}^2$ consists of exactly seven points and seven lines with the incidence relation described by Fig. 1.4. The circle together with the six segments represent the seven lines. This geometry is also known as the *Fano plane* . – c) In $\mathfrak{P}_{\mathbf{Z}_2}^n$ there are exactly $2^{n+1} - 1$ points.

Definition 3. A set of points $M \in \boldsymbol{P}_o^n$ is called *collinear*, if there is a line $\boldsymbol{h} \in \boldsymbol{P}_{n,1}$ such that $M \subset \boldsymbol{h}$. A set $\mathfrak{M} \subset \mathfrak{P}^n$ is called *concentric*, if there is a point $\boldsymbol{z} \in \boldsymbol{P}^n$ such that $\boldsymbol{z}\iota\boldsymbol{A}$ for all $\boldsymbol{A} \in \mathfrak{M}$. □

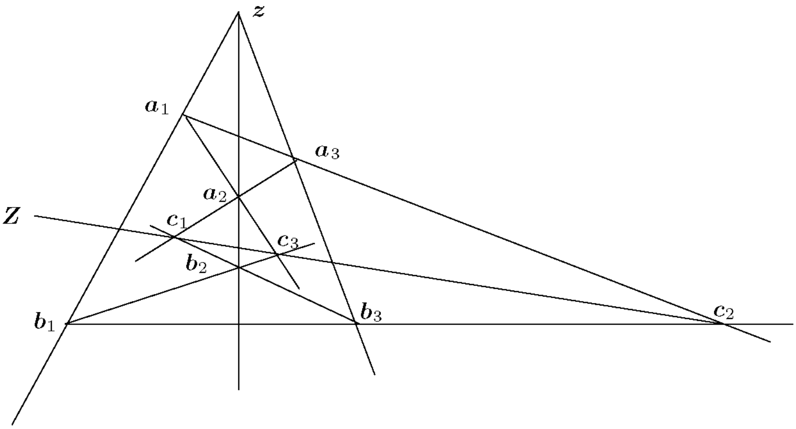


Fig. 1.5. Desargues' Theorem.

The proposition stated in the next exercise goes back to *Desargues* and is particularly interesting for the plane projective geometry, since it does not follow from the Axioms P.1–3 of plane incidence geometry; incidence planes satisfying it are called *Desarguesian planes*. It can be shown that Desargues' Property is equivalent to the associativity of the scalar domain, cf. G. Pickert [87]. In the incidence geometry of three-dimensional space Desargues' Theorem already follows from the other incidence axioms, cf. L. Heffter [48], Section A.5. Hence there are no projective spaces of dimension $n > 2$ with a non-associative scalar domain. The projective plane over the *octonions* is perhaps the most important example for a plane not having Desargues' Property, cf. H. Freudenthal [39] or H. Salzmann et al. [97]. In the geometries based on skew fields as they are discussed here the following proposition holds in all dimensions $n \geq 2$.

Exercise 9 (Desargues' Theorem). Let $\mathfrak{P}_K^n$ be a projective geometry over a skew field K, $n \geq 2$. Consider six different points $\boldsymbol{a}_i, \boldsymbol{b}_i \in \boldsymbol{P}^n$, $i = 1, 2, 3$, whose connecting lines $\boldsymbol{h}_i := \boldsymbol{a}_i \vee \boldsymbol{b}_i$, $i = 1, 2, 3$, are concentric. Prove that "corresponding sides" intersect in points, i.e.

$$\begin{aligned} \boldsymbol{c}_1 &= (\boldsymbol{a}_2 \vee \boldsymbol{a}_3) \wedge (\boldsymbol{b}_2 \vee \boldsymbol{b}_3), \\ \boldsymbol{c}_2 &= (\boldsymbol{a}_1 \vee \boldsymbol{a}_3) \wedge (\boldsymbol{b}_1 \vee \boldsymbol{b}_3), \\ \boldsymbol{c}_3 &= (\boldsymbol{a}_1 \vee \boldsymbol{a}_2) \wedge (\boldsymbol{b}_1 \vee \boldsymbol{b}_2), \end{aligned}$$

and that, moreover, these points are collinear, cf. Fig. 1.5.

For the proof of *Pappos'* Theorem, which does not follow from the incidence axioms P.1–3 as well, the commutativity of the scalar field is essential (cf. Exercise 2.10):

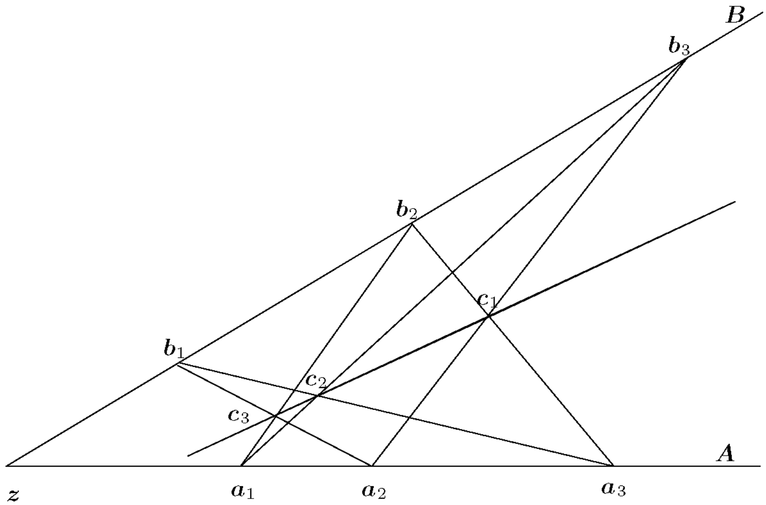


Fig. 1.6. Pappos' Theorem.

Exercise 10 (Pappos' Theorem). Let $\boldsymbol{P}_K^2$ be a projective plane over a field K. Let $\boldsymbol{A}$, $\boldsymbol{B}$ be two lines intersecting in the point $\boldsymbol{z} = \boldsymbol{A} \wedge \boldsymbol{B}$. Consider three points, different from one another and from $\boldsymbol{z}$, on each of the lines $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{a}_i \in \boldsymbol{A}$, $\boldsymbol{b}_i \in \boldsymbol{B}$, $i = 1, 2, 3$,, cf. Fig. 1.6. Prove that under these conditions the points

$$\begin{aligned} \boldsymbol{c}_1 &= (\boldsymbol{a}_2 \vee \boldsymbol{b}_3) \wedge (\boldsymbol{a}_3 \vee \boldsymbol{b}_2), \\ \boldsymbol{c}_2 &= (\boldsymbol{a}_1 \vee \boldsymbol{b}_3) \wedge (\boldsymbol{a}_3 \vee \boldsymbol{b}_1), \\ \boldsymbol{c}_3 &= (\boldsymbol{a}_1 \vee \boldsymbol{b}_2) \wedge (\boldsymbol{a}_2 \vee \boldsymbol{b}_1) \end{aligned}$$

are collinear.

The *finite geometries* some examples of which we have been mentioning here are studied systematically in the monographs [10], [11] by A. Beutelspacher. For more recent developments we in addition refer to L. M. Batten,

A. Beutelspacher [6], K. Metsch [76], and J. W. P. Hirschfeld [54]. The latter also contains an extensive bibliography. Finite geometries are applied to numerous kinds of problems and subjects, among others in the field of combinatorics.

1.2 Homogeneous Coordinates

Let K be an arbitrary skew field. Then $K\boldsymbol{P}_o^n$ (and $K\boldsymbol{P}^n$, respectively) denotes the projective space associated with the right vector space K^{n+1} of $(n+1)$-tuples: $K\boldsymbol{P}^n := \boldsymbol{P}(K^{n+1})$. The elements of $K\boldsymbol{P}_o^n$ are called *homogeneous* $(n+1)$*-tuples*; they are the orbits of the multiplicative group K^* of K for the right action

$$((x^i), \lambda) \in K^{n+1} \times K^* \longmapsto (x^i \lambda) \in K^{n+1}, \tag{1}$$

where in the case of $K\boldsymbol{P}^n$ the zero tuple is excluded. Just as the vector space K^m of m-tuples from K serves as the standard coordinate space for all the m-dimensional vector spaces over K, the space $K\boldsymbol{P}_o^n$ will now be used as the coordinate space for any n-dimensional projective space over K. Using homogeneous $(n+1)$-tuples, i.e. admitting a superfluous degree of freedom in the form of an arbitrary factor λ, allows to work with just a single coordinate system for the whole projective space. Moreover, it turns out to be possible to describe projective coordinate transformations in a way similar to the familiar representation of the transformations of vector coordinates, i.e. by means of matrix calculus.

1.2.1 Definition. Simplices

Now we consider an arbitrary $(n+1)$-dimensional right vector space $\boldsymbol{V}$ over K and the associated projective space $\boldsymbol{P}^n = \boldsymbol{P}(\boldsymbol{V})$. Let $(\mathfrak{a}_i)$ be a basis of $\boldsymbol{V}$, and let $(\mathfrak{c}_i)$ be the standard basis of K^{n+1}, $i = 0, \ldots, n$. Linearly extending the relations $\varphi(\mathfrak{a}_i) = \mathfrak{c}_i$, $i = 0, \ldots, n$, defines a linear isomorphism $\varphi : K^{n+1} \to \boldsymbol{V}$ of vector spaces, which in turn determines vector coordinates on $\boldsymbol{V}$. Denoting by π both the respective canonical maps the linearity of φ immediately implies

$$\tilde{\varphi}(\pi(\mathfrak{x})) = \pi(\varphi(\mathfrak{x})) \qquad (\mathfrak{x} \in \boldsymbol{V}), \tag{2}$$

i.e., the map $\tilde{\varphi} : \boldsymbol{P}_o^n \to K\boldsymbol{P}_o^n$ does not depend on the chosen representative $\mathfrak{x} \in \pi^{-1}(\boldsymbol{x})$. $\tilde{\varphi}$ is called the *homogeneous coordinate system determined by* $(\mathfrak{a}_i)$ on the projective space $\boldsymbol{P}^n$; correspondingly, the x^i defined by $\tilde{\varphi}(\boldsymbol{x}) = (x^i)K^*$ are called the *homogeneous coordinates* of $\boldsymbol{x}$ with respect to $\tilde{\varphi}$. They are determined only up to a common factor $\lambda \in K^*$. The points $\boldsymbol{a}_i := \pi(\mathfrak{a}_i)$ are called the *base points*, and $\boldsymbol{e} := \pi(\sum_{i=0}^n \mathfrak{a}_i)$ is the *unit point* of the homogeneous coordinate system. The j-th base point has homogeneous coordinates $(\delta_j^i)K^*$, where δ_j^i is the *Kronecker symbol*:

$$\delta^i_j = \begin{cases} 0 & \text{for } i \neq j, \\ 1 & \text{for } i = j, \end{cases} \quad i, j = 0, \dots, n.$$

The coordinates of the unit point are all equal $(1, \dots, 1)K^*$. Obviously, the sequence $(\boldsymbol{a}_0, \dots, \boldsymbol{a}_n; \boldsymbol{e})$ formed by the base points and the unit point of a homogeneous coordinate system is in general position, cf. Exercise 1.6. Hence any family of $k + 1$ base points of a coordinate system spans a *coordinate k-plane*:

$$\boldsymbol{H}_{i_0 \dots i_k} = \boldsymbol{a}_{i_0} \vee \dots \vee \boldsymbol{a}_{i_k}.$$

In general, two projective subspaces $\boldsymbol{A}$, $\boldsymbol{B}$ are called *complementary*, if $\boldsymbol{A} \wedge \boldsymbol{B} = \boldsymbol{o}$ and $\boldsymbol{A} \vee \boldsymbol{B} = \boldsymbol{P}^n$. For every coordinate k-plane $\boldsymbol{H}_{i_0 \dots i_k}$ there is exactly one complementary coordinate plane, namely $\boldsymbol{H}_{j_1 \dots j_l}$, where $\{j_1, \dots, j_l\}$ is the complement of $\{i_0, \dots, i_k\}$ in the set $\{0, \dots, n\}$; its dimension is $n - k - 1$.

Exercise 1. Let $\boldsymbol{A}, \boldsymbol{B} \subset \boldsymbol{P}^n$ be complementary subspaces. Prove: a) For every point $\boldsymbol{z} \in \boldsymbol{P}^n \setminus (\boldsymbol{A} \cup \boldsymbol{B})$ there is exactly one line $\boldsymbol{H}$ such that $\boldsymbol{z} \in \boldsymbol{H}$, $\boldsymbol{H} \cap \boldsymbol{A} \neq \boldsymbol{o}$ and $\boldsymbol{H} \cap \boldsymbol{B} \neq \boldsymbol{o}$. – b) Setting

$$p : \boldsymbol{x} \in \boldsymbol{P}^n_o \longmapsto (\boldsymbol{x} \vee \boldsymbol{B}) \wedge \boldsymbol{A} \in \boldsymbol{A}$$

defines a surjective map $p : \boldsymbol{P}^n_o \to \boldsymbol{A}$; the equality $p(\boldsymbol{x}) = \boldsymbol{o}$ holds if and only if $\boldsymbol{x} \in \boldsymbol{B}$. – c) The map p satisfies $p^2 = p$.

The following example illustrates the situation:

Example 1. Simplices. A projectively independent sequence of $k+1$ points $(\boldsymbol{b}_0, \dots, \boldsymbol{b}_k)$ is called a *k-dimensional simplex*, *k-simplex* for short. The points themselves are the *vertices*, and the subsequences form the *faces* of the simplex. Frequently we will also use the term face of the simplex to denote the whole projective subspace spanned by the face of the simplex as well. One-dimensional faces will also be called *edges*. For each vertex $\boldsymbol{b}_j$ of a k-simplex the $(k-1)$-face not containing $\boldsymbol{b}_j$ is called its *opposite face* $\boldsymbol{B}_j$. The base points of a homogeneous coordinate system for $\boldsymbol{P}^n$ form the *coordinate simplex*. Denoting by $\boldsymbol{B}_j$ the opposite face of $\boldsymbol{a}_j$ the points in the hyperplane spanned by $\boldsymbol{B}_j$ are characterized by the equation $x^j = 0$. The *j-th projection* p_j of a point $\boldsymbol{x}$ onto the face $\boldsymbol{B}_j$ (from the point $\boldsymbol{a}_j$) is defined by

$$p_j : \boldsymbol{x} \in \boldsymbol{P}^n \longmapsto \boldsymbol{x}_j := \boldsymbol{B}_j \wedge (\boldsymbol{a}_j \vee \boldsymbol{x}) \in \boldsymbol{B}_j; \tag{3}$$

thus $p_j(\boldsymbol{a}_j) = \boldsymbol{o}$. The unit point in the face $\boldsymbol{B}_j$ is the image of the unit point under the corresponding projection, $\boldsymbol{e}_j := p_j(\boldsymbol{e})$. Fig. 1.7 illustrates the situation for the plane. Take the face $\boldsymbol{B}_0$ as the line at infinity, cf. Example 1.1. Its complement, the affine plane $\boldsymbol{A}^2 := \boldsymbol{P}^2 \setminus \boldsymbol{B}_0$, is characterized by $x^0 \neq 0$. The homogeneous coordinates there can be normalized so that $x^0 = 1$ holds. The coordinates x^1, x^2 of the triples normalized in this way then become the cartesian coordinates of the point $\boldsymbol{x} \in \boldsymbol{A}^2$. They can also be computed as the scale values of the projections of $\boldsymbol{x}$ onto the axes of the scales chosen as described in Example 1.3. The point $\boldsymbol{a}_0$ will be called the *origin* of the cartesian coordinate system. Note that the projections from $\boldsymbol{a}_j$ onto the axes $\boldsymbol{B}_j$, $j = 1, 2$, appear as parallel projections in the affine picture. The quadrangle $(\boldsymbol{a}_0, \boldsymbol{e}_1, \boldsymbol{e}, \boldsymbol{e}_2)$ thus becomes the unit parallelogram. In the n-dimensional case the situation is quite similar. □

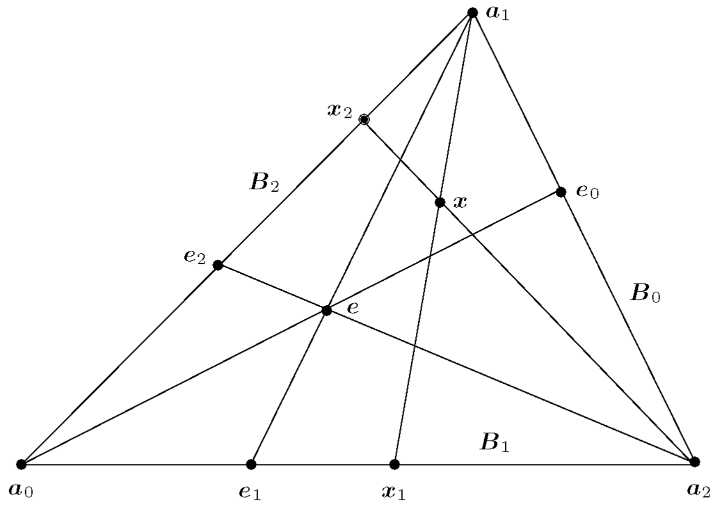


Fig. 1.7. Coordinate triangle in $\boldsymbol{P}^2$.

The following proposition justifies the definition of homogeneous coordinates:

Proposition 1. *The map $\tilde{\varphi} : \boldsymbol{P}^n_o \to K\boldsymbol{P}^n_o$ uniquely defined by (2) is a bijection. If $(\hat{\mathfrak{a}}_i)$ is a basis of $\boldsymbol{V}$ determining the homogeneous coordinate system $\tilde{\psi}$, then for $n \geq 1$ the maps $\tilde{\varphi}$ and $\tilde{\psi}$ coincide if and only if there exists an element μ in the center $Z(K^*)$ of the group K^* such that $\hat{\mathfrak{a}}_i = \mathfrak{a}_i \mu$, $i = 0, \ldots, n$.*

P r o o f. The bijectivity of $\tilde{\varphi}$ is a direct consequence of the bijectivity of φ. Now suppose $\tilde{\varphi} = \tilde{\psi}$. Then for the base points we have $\tilde{\varphi}^{-1}([\mathfrak{c}_i]) = \tilde{\psi}^{-1}([\mathfrak{c}_i]) = \boldsymbol{a}_i$, hence $\hat{\mathfrak{a}}_i = \mathfrak{a}_i \lambda_i$ for certain $\lambda_i \in K^*$. For the unit point we similarly obtain

$$\boldsymbol{e} = [\mathfrak{a}_0 + \ldots + \mathfrak{a}_n] = [\hat{\mathfrak{a}}_0 + \ldots + \hat{\mathfrak{a}}_n],$$

thus

$$\sum_{i=0}^{n} \hat{\mathfrak{a}}_i = \sum_{i=0}^{n} \mathfrak{a}_i \lambda_i = (\sum_{i=0}^{n} \mathfrak{a}_i)\mu.$$

This implies $\lambda_i = \mu$ for $i = 0, \ldots, n$. All we have to show now is $\mu \in Z(K^*)$. Starting[1] from $\mathfrak{x} = \mathfrak{a}_i x^i = \hat{\mathfrak{a}}_i \hat{x}^i = (\mathfrak{a}_i \mu)\hat{x}^i$ for all $\boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{P}^n$ we obtain

$$\tilde{\varphi}(\boldsymbol{x}) = (x^i)K^* = \tilde{\psi}(\boldsymbol{x}) = (\hat{x}^i)K^* \text{and } (x^i) = (\mu\hat{x}^i).$$

Hence $(\mu\hat{x}^i)K^* = (\hat{x}^i)K^*$ holds for all $(\hat{x}^i) \in K^{n+1}$. Consequently, there has to be a map $\tau : K^{n+1} \setminus \{(0, \ldots, 0)\} \to K^*$ with the property

$$(\mu x^i) = (x^i)\tau(x^i). \tag{4}$$

For linearly independent $(a^i), (b^i)$ this leads to the following:

$$\begin{aligned} (\mu(a^i + b^i)) &= (a^i + b^i)\tau(a^i + b^i) = (a^i)\tau(a^i + b^i) + (b^i)\tau(a^i + b^i), \\ &= (\mu a^i) + (\mu b^i) = (a^i)\tau(a^i) + (b^i)\tau(b^i); \end{aligned}$$

since $n \geq 1$ we thus obtain the equality $\tau(a^i) = \tau(\mathfrak{c}_0)$ for all $(a^i) \notin [\mathfrak{c}_0]$. Again as $n \geq 1$ we also have $(a^i) = \mathfrak{c}_0 s$ for $(a^i) \in [\mathfrak{c}_0]$, $s \neq 0$, and, similarly, $\tau(\mathfrak{c}_0 s) = \tau(\mathfrak{c}_1) = \tau(\mathfrak{c}_0)$. Hence $\tau(x^i) = \tau \in K^*$ is constant, and for $(x^i) = \mathfrak{c}_0 t$, $t \in K^*$, (4) immediately implies: 1.) $\mu = \tau$ — just take $t = 1$, and 2.) $\mu t = t\mu$ for all $t \in K^*$. But this just means $\mu \in Z(K^*)$. The converse is trivial. □

Remark: In the case of a one-dimensional vector space ($n = 0$), in general, maps of the form

$$\tau(x^0) = (x^0)^{-1}\mu x^0$$

are not constant solutions of (4).

Proposition 2. *The sequence $(\boldsymbol{a}_0, \ldots, \boldsymbol{a}_n; \boldsymbol{e})$ formed by the base points and the unit point of a homogeneous coordinate system is in general position. Conversely, for every sequence $(\boldsymbol{a}_0, \ldots, \boldsymbol{a}_n; \boldsymbol{e})$ of $n + 2$ points in general position there is a system of homogeneous coordinates $\tilde{\varphi}$ such that $\boldsymbol{a}_i$ is its i-th base point and $\boldsymbol{e}$ its unit point. Two coordinate systems $(x^i)K^*$, $(\hat{x}^i)K^*$ with equal sequences $(\boldsymbol{a}_0, \ldots, \boldsymbol{a}_n; \boldsymbol{e})$ of base and unit points are related by a left dilation: $(x^i)K^* = (\mu\hat{x}^i)K^*$; consequently, for $\mu \in Z(K^*)$ or in the case of a field they coincide.*

Proof. The definitions for the base points and the unit point of a homogeneous coordinate system immediately imply that any $n+1$ vectors representing these points are linearly independent. Hence, according to Exercise 1.1.6, they are in general position. Conversely, let $(\boldsymbol{a}_0, \ldots, \boldsymbol{a}_n; \boldsymbol{e})$ be a sequence of points in general position. Then any vectors $\mathfrak{a}_i$ representing the $\boldsymbol{a}_i$ are linearly independent. Choose a sequence of representatives $(\mathfrak{a}_i, i = 0, \ldots, n)$, $\boldsymbol{a}_i = [\mathfrak{a}_i]$.

[1] For simplicity we will use the *sum convention of tensor calculus* from now on: If an index occurs twice in an expression, as a lower as well as an upper index, then this has to be interpreted as a sum over this index for its complete domain of definition (in the present formula from 0 to n). To indicate that for a particular formula the sum convention is not in action we will use the notation $\not\Sigma$.

These vectors then form a basis of $\boldsymbol{V}$. Let $\mathfrak{e}$ be a vector with $\boldsymbol{e} = [\mathfrak{e}]$ and write it as

$$\mathfrak{e} = \sum_{i=0}^{n} \mathfrak{a}_i \eta_i .$$

Since the sequence $(\boldsymbol{a}_0, \dots, \boldsymbol{a}_n; \boldsymbol{e})$ is in general position, $\eta_i \neq 0$ for all i. Hence, setting $\mathfrak{b}_i := \mathfrak{a}_i \eta_i$ yields a sequence of vectors $(\mathfrak{b}_0, \dots, \mathfrak{b}_n; \mathfrak{e})$ also representing the chosen sequence of points. The definition of the $\mathfrak{b}_i$ implies that they form a basis, and, moreover, that $\boldsymbol{e}$ is the unit point of the corresponding homogeneous coordinate system. An arbitrary basis with base points $\boldsymbol{a}_i$ necessarily has the form $(\hat{\mathfrak{b}}_i)$ with $\hat{\mathfrak{b}}_i = \mathfrak{b}_i \lambda_i$. If $\boldsymbol{e}$ is to be the unit point for this basis as well, then $\boldsymbol{e} = [\sum \mathfrak{b}_i] = [\sum \hat{\mathfrak{b}}_i] = [\sum \mathfrak{b}_i \lambda_i]$ has to hold. Hence, $\lambda_i = \mu \in K^*$ for $i = 0, \dots, n$. This implies $\mu \hat{x}^i = x^i$ for the corresponding coordinates. The final statement follows from Proposition 1. □

The close relation between sequences of $n+2$ points in general position and homogeneous coordinate systems stated in Proposition 2 suggests to call such a sequence $(\boldsymbol{a}_0, \dots, \boldsymbol{a}_n; \boldsymbol{e})$ a *projective frame.* Thus, every homogeneous coordinate system uniquely determines a projective frame. Obviously, the following holds.

Corollary 3. *If K is a field, then for each projective frame there is a unique homogeneous coordinate system corresponding to it (and vice versa).* □

In the case of a skew field this relation is not bijective, in general. Let us just look at the projective line: The bases $(\mathfrak{a}_0, \mathfrak{a}_1)$, $(\mathfrak{a}_0 \mu, \mathfrak{a}_1 \mu)$ determine one and the same projective frame, whereas the corresponding homogeneous coordinates are related by an *inner automorphism*: The equalities $\mathfrak{x} = \hat{\mathfrak{a}}_i \hat{x}^i = \mathfrak{a}_i \mu \hat{x}^i = \mathfrak{a}_i x^i$ imply

$$(x^i)K^* = (\mu \hat{x}^i)K^* = (\sigma_\mu(\hat{x}^i))K^*,$$

where $\sigma_\mu : t \in K \mapsto \mu t \mu^{-1} \in K$ denotes the inner automorphism of K associated with $\mu \in K^*$. Additionally, for later use we point out the following.

Corollary 4. *Let $\boldsymbol{P}^1$ be a projective line. Then for each projective scale of $\boldsymbol{P}^1$ there is a uniquely determined projective frame $(\boldsymbol{a}_0, \boldsymbol{a}_1; \boldsymbol{e})$* (cf. Example 1.3). *The set of all scale values for $\boldsymbol{x}$, $\{\xi(\boldsymbol{x})\}$, forms a conjugacy class in K. Here ξ runs through all projective scales with the same projective frame $(\boldsymbol{a}_0, \boldsymbol{a}_1; \boldsymbol{e})$.* □

To prove this, note that, by its very definition (1.1.4), the projective scale is the ratio of both homogeneous coordinates.

1.2.2 Coordinate Transformations. The Projective Linear Group

Now we turn to the transformations of homogeneous coordinates.

Denote by $\boldsymbol{GL}(n+1, K)$ the linear group of order[2] $n+1$ over the skew field K, cf. § II.7.4. As is well-known, this group acts as a matrix group from the left on the vector space K^{n+1}:

$$((a_i^j), (x^i)) \in \boldsymbol{GL}(n+1, K) \times K^{n+1} \longmapsto (a_i^j)(x^i) \in K^{n+1},\ i, j = 0, \ldots, n.$$

Consider the homogeneous coordinate systems $\tilde{\varphi}, \tilde{\psi}$ determined by the bases $(\mathfrak{a}_i)$ and $(\hat{\mathfrak{a}}_j)$, respectively. For the corresponding coordinates (x^i), $(\hat{x}^j)$ of any point $\boldsymbol{x} \in \boldsymbol{P}^n$ we then have the following relations:

$$\boldsymbol{x} = [\mathfrak{a}_i x^i] = [\hat{\mathfrak{a}}_j \hat{x}^j] = [\hat{\mathfrak{a}}_j a_i^j x^i].$$

Here (a_i^j) denotes the transformation matrix for the vector coordinates. Hence,

$$\tilde{\psi}(\boldsymbol{x}) = (\hat{x}^j)K^* = (a_i^j)(x^i)K^* = (a_i^j)\tilde{\varphi}(\boldsymbol{x}). \tag{5}$$

The situation is thus the same as in the case of a transformation of vector coordinates: The group $\boldsymbol{GL}(n+1, K)$ acts from the left on the set $\tilde{\Phi}$ of homogeneous coordinate systems, cf. § I.5.7.

Lemma 5. *The action of* $\boldsymbol{GL}(n+1, K)$ *on* $\tilde{\Phi}$ *described above is transitive. A matrix* $(a_i^j) \in \boldsymbol{GL}(n+1, K)$ *acts as the identity on* $\tilde{\Phi}$ *if and only if*

$$(a_i^j) = (\delta_i^j)\mu,\ \mu \in Z(K^*). \tag{6}$$

P r o o f. The transitivity immediately follows from the transitivity of the action of $\boldsymbol{GL}(n+1, K)$ on the set of all bases. By Proposition 1 we have $(a_i^j)\tilde{\varphi} = \tilde{\varphi}$ if and only if there exists a $\mu \in Z(K^*)$ with the property (6). This obviously holds for all $\tilde{\varphi} \in \tilde{\Phi}$, if it is true for any single one. □

Definition 1. Let K be a skew field, and denote by $Z(n+1)$ the subgroup

$$Z(n+1) := \{(\delta_i^j)\mu | \mu \in Z(K^*), i, j = 0, \ldots, n\}. \tag{7}$$

The quotient group

$$\boldsymbol{PGL}(n+1, K) := \boldsymbol{GL}(n+1, K)/Z(n+1),$$

is called the *projective linear group of order* $n+1$ over K. Note also the following exercise.

Exercise 2. By Lemma 5 the center $Z(n+1)$ is the kernel of the action under consideration (cf. Definition I.1.4.2) and hence a normal subgroup. The action of $\boldsymbol{PGL}(n+1)$ on $\tilde{\Phi}$ defined via representatives is thus simply transitive, cf. § I.1.4. Prove that $Z(n+1)$ is the center of the group $\boldsymbol{GL}(n+1, K)$ (cf. Example I.1.4.7).

[2] This notion is not to be confused with that of the order of a finite group.

1.2.3 Inhomogeneous Projective Coordinates

In this section we will consider the inhomogeneous coordinates associated with a homogeneous coordinate system

$$\tilde{\varphi}(x) = (x^i(\boldsymbol{x}))K^*$$

and calculate their transformation rule. The i-th *coordinate hyperplane* $\boldsymbol{B}_i$ is defined as the coordinate hyperplane complementary to the base point $\boldsymbol{a}_i$. Let $U_i = \boldsymbol{P}^n \setminus \boldsymbol{B}_i$ be its complement in $\boldsymbol{P}^n$, cf. Example 1:

$$U_i = \{\boldsymbol{x} \in \boldsymbol{P}^n | \, x^i(\boldsymbol{x}) \neq 0\}.$$

Obviously, the i-th *chart* φ_i on $\boldsymbol{U}_i$, is correctly defined:

$$\boldsymbol{x} \in U_i \mapsto \varphi_i(\boldsymbol{x}) := (x^j(\boldsymbol{x})(x^i(\boldsymbol{x}))^{-1} | j \neq i, j = 0, \ldots, n) \in K^n. \tag{8}$$

For brevity we set

$$\xi_i^j(\boldsymbol{x}) := x^j(\boldsymbol{x})(x^i(\boldsymbol{x}))^{-1} \text{ for } \boldsymbol{x} \in U_i$$

and prove the following.

Lemma 6. *a) Each homogeneous coordinate system $\tilde{\varphi}$ on the projective space $\boldsymbol{P}^n$ determines $n+1$ charts, $\varphi_i : U_i \to K^n$, defined by* (8), *that are bijective maps.*
b) The domains of definition of these charts cover the projective space, i.e. $\bigcup_{i=0}^n U_i = \boldsymbol{P}^n$.
c) Let, as above, $\tilde{\psi}$ denote the homogeneous coordinate system associated with the basis $(\hat{\boldsymbol{a}}_i)$, let (a_i^j) be the matrix of the coordinate transformation, and let $\psi_i = (\hat{\xi}_i^j)$ be the charts $\psi_i : V_i \to K^n$ determined by $\tilde{\psi}$. Then the coordinate transformations

$$\psi_b \circ \varphi_a^{-1} : \varphi_a(U_a \cap V_b) \longrightarrow \psi_b(U_a \cap V_b)$$

defined on the intersections of coordinate neighborhoods, $U_a \cap V_b$, are fractional linear transformations; moreover,

$$\hat{\xi}_b^k(\boldsymbol{x}) = a_j^k \xi_a^j(\boldsymbol{x})(a_i^b \xi_a^i(\boldsymbol{x}))^{-1}. \tag{9}$$

P r o o f. Note that in this formula, according to the sum convention, the sum in the numerator runs over j and that in the denominator over i. Moreover, $\xi_i^i = 1$, $\not{\Sigma}$. For arbitrary $(y^1, \ldots, y^n) \in K^n$ the point $\boldsymbol{x}$ with homogeneous coordinates

$$x^0(\boldsymbol{x}) = y^1, \, x^1(\boldsymbol{x}) = y^2, \, \ldots, x^{i-1}(\boldsymbol{x}) = y^i,$$
$$x^i(\boldsymbol{x}) = 1, \, x^{i+1}(\boldsymbol{x}) = y^{i+1}, \, \ldots, \, x^n(\boldsymbol{x}) = y^n$$

is the only point in U_i mapped to $(y^1, \ldots, y^n)$. This proves a). Since for every point $\boldsymbol{x} \in \boldsymbol{P}^n$ at least one among its homogeneous coordinates does not vanish, b) holds. The remaining statement c) immediately follows from the transformation rule (5) and the definition of inhomogeneous coordinates. $\square$

Example 2. The Projective Line. For $n = 1$ there are two inhomogeneous coordinates essentially coinciding with the corresponding projective scales. For example, the value $\xi(\boldsymbol{a}_1)$ is not yet defined by $\xi := \xi_0^1$. Setting, as seems obvious, $\xi(\boldsymbol{a}_1) := \infty$ yields the projective scale defined in Example 1.3. In this case, formula (9) describes the transformation of projective scales on $\boldsymbol{P}^1$, $\theta_A : \hat{K} \to \hat{K}$, entailed by a change of bases in $\boldsymbol{V}^2$:

$$\theta_A : \xi \mapsto \hat{\xi} := (a\xi + b)(c\xi + d)^{-1} \text{ with } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \boldsymbol{GL}(2, K). \tag{10}$$

To complete the picture, we extend θ_A in the following way:

$$\theta_A(-c^{-1}d) = \infty,\ \theta_A(\infty) = \begin{cases} ac^{-1}, & \text{if } c \neq 0, \\ \infty, & \text{if } c = 0. \end{cases} \tag{11}$$

□

Exercise 3. In the notation of Lemma 6, let $\tilde{\varphi}$, $\tilde{\psi}$ be homogeneous coordinate systems on $\boldsymbol{P}^n$ satisfying the additional condition $U_0 = V_0$. Prove that the coordinate transformation $\psi_0 \circ \varphi_0^{-1}$ is an affine transformation in K^n, cf. I, (5.7.36).

Exercise 4. Prove that the map of $\hat{K}$ to itself defined by (10), (11) is a bijection and satisfies the equality $(\theta_A)^{-1} = \theta_{A^{-1}}$. (Note that, in general, K is a skew field and hence determinants cannot be used.) Denote by $\mathcal{S}(\hat{K})$ the group of all bijective maps on $\hat{K}$. Prove that the map

$$F : A \in \boldsymbol{GL}(2, K) \mapsto \theta_A \in \mathcal{S}(\hat{K})$$

is a homomorphism whose image is isomorphic to $\boldsymbol{PGL}(2, K)$. Show that for $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$ every θ_A is a homeomorphism with respect to the topology on $\hat{K}$ described in Example 1.3.

Exercise 5. A projective plane $\boldsymbol{P}^2$ is said to have *Pappos' Property* if Pappos' Theorem holds in it, cf. Exercise 1.10, i.e., if the points $\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3$ in a configuration as depicted in Figure 1.6 are always collinear. Prove that $\boldsymbol{P}^2$ has the Pappos' Property if and only if the skew field on which it is based is commutative, i.e., it is a field. — Hint. Choose homogeneous coordinates in a way that $(\boldsymbol{z}, \boldsymbol{a}_3, \boldsymbol{b}_3; \boldsymbol{c}_2)$ are the base points and the unit point of the coordinate system, respectively, and express then the condition in these coordinates.

1.2.4 The Projective Linear Group over a Field

Example 3. Let K be a field. Then we have $Z(K^*) = K^* \cong Z(n+1)$. Consider the canonical homomorphism resulting from Definition 1,

$$p : (a^i_j) \in \boldsymbol{GL}(n+1, K) \longmapsto (a^i_j)K^* \in \boldsymbol{PGL}(n+1, K).$$

The matrices from $\boldsymbol{GL}(n+1, K)$ contained in the same coset of the factor group $\boldsymbol{PGL}(n+1, K) = \boldsymbol{GL}(n+1, K)/Z(n+1)$ only differ in a factor $\mu \in K^*$.

Now choose a "fixed" projective frame R in the projective space $\boldsymbol{P}^n$ over K and, in addition, a variable "moving" projective frame $\hat{R}$; let (x^i) and $(\hat{x^i})$ be the homogeneous coordinates associated with R and $\hat{R}$, respectively, defined by the bases $(\mathfrak{a}_i)$ and $(\hat{\mathfrak{a}}_i)$ of $\boldsymbol{V}$. According to Proposition 1, these bases are determined by the homogeneous coordinates only up to a common factor $\mu \in K^*$. Hence the cosets of $\boldsymbol{GL}(n+1, K)$ by $Z(n+1)$, i.e. the elements of $\boldsymbol{PGL}(n+1, K)$, correspond bijectively to the projective frames. Thus, the set of these frames is a geometric realization of the group $\boldsymbol{PGL}(n+1, K)$, just like the set of linear frames, $(\hat{\mathfrak{a}}_i)$, is a geometric realization of the group $\boldsymbol{GL}(n+1, K)$.

Now we will try to find a distinguished representative in each coset of $\boldsymbol{GL}(n+1, K)$ by $Z(n+1)$ in a natural way. To this aim, we consider the determinant as a function of μ. Looking at the well-known formula $\det(A\mu) = \det(A)\mu^{n+1}$, the obvious condition to impose on a distinguished representative should be

$$\det(A\mu) = 1 \text{ for } A = (a_i^j) \in \boldsymbol{GL}(n+1, K) \text{ and } \mu \in K^* \tag{12}$$

Hence the intended choice amounts to deciding the solvability of

$$\mu^{n+1} = (\det(A))^{-1},$$

and, apart from n, this again essentially depends on the algebraic properties of the field K.

Let now $K = \mathbf{R}$. If n is even, for given A the equation (12) has exactly one real solution μ. Hence, each coset contains a unique representative with determinant one, and we arrive at the isomorphism

$$\boldsymbol{PGL}(n+1, \mathbf{R}) \cong \boldsymbol{SL}(n+1, \mathbf{R}) \ (n \text{ even}). \tag{13}$$

Hence, in this case, the projective linear group is isomorphic to the *special linear group*,

$$\boldsymbol{SL}(n+1, \mathbf{R}) := \{A \in \boldsymbol{GL}(n+1, \mathbf{R}) \mid \det(A) = 1\}. \tag{14}$$

The situation is different in the odd-dimensional case: Now the sign of $\det(A)$ is constant on every coset, and this allows to introduce an *orientation*: Two projective frames, or the homogeneous coordinate systems associated with them, respectively, are called *equally oriented* if the coordinate transformation A transferring one into the other has positive determinant, i.e. $\operatorname{sign}(\det(A)) = 1$. Apparently, there are two classes of equally oriented homogeneous coordinate systems (cf. the following exercise); a real projective space of odd dimension is called *oriented* if one of these classes is distinguished. This orientation, introduced within the framework of linear algebra, is in complete accordance with the topological notion of orientation, cf. Appendix 4.2. It can be shown that a real projective space, considered as a differentiable manifold, is orientable if and only if its dimension is odd [103], Exercise 1.4.2. □

Exercise 6. a) Prove that the relation "equally oriented", introduced in the preceding example, is an equivalence relation, and that there are precisely two equivalence classes for odd n. – b) Use the definition

$$|\boldsymbol{SL}|(N,\mathbf{R}) := \{A \in \boldsymbol{GL}(N,\mathbf{R}) \mid \quad |\det(A)| = 1\} \tag{15}$$

to prove the isomorphy

$$\boldsymbol{PGL}(n+1,\mathbf{R}) \cong |\boldsymbol{SL}|(n+1,\mathbf{R})/\{\pm I\} \qquad (n \text{ odd}), \tag{16}$$

where I denotes the unit matrix.

Exercise 7. In this exercise we consider the action of the group $\boldsymbol{GL}(\boldsymbol{V}^{n+1})$ of linear automorphisms of the right vector space $\boldsymbol{V}^{n+1}$ over the skew field K. Prove: a) Setting

$$(g, \boldsymbol{x}) \in \boldsymbol{GL}(\boldsymbol{V}^{n+1}) \times \boldsymbol{P}^n \longmapsto g\boldsymbol{x} := [g(\mathfrak{x})] \in \boldsymbol{P}^n \text{ for } \boldsymbol{x} = [\mathfrak{x}] \tag{17}$$

defines a transitive action; its kernel is isomorphic to the center $Z(K^*)$ of K^*. – b) If K is a field, then the restriction of this action to the *special linear group*

$$\boldsymbol{SL}(\boldsymbol{V}^{n+1}) := \{g \in \boldsymbol{GL}(\boldsymbol{V}^{n+1}) | N(g) = 1\} \tag{18}$$

is transitive; here $N(g)$ denotes the *norm* of the linear automorphism g, which is equal to the determinant of the matrix for g with respect to an arbitrary basis of $\boldsymbol{V}^{n+1}$. – c) Let (ξ^i) be the inhomogeneous coordinates of $\boldsymbol{x} \in \boldsymbol{P}^n$, and let (η^j), $i, j = 1, \ldots, n$, be the inhomogeneous coordinates of the image point $\boldsymbol{y} = g\boldsymbol{x}$ under a linear automorphism $g \in \boldsymbol{GL}(\boldsymbol{V}^{n+1})$. Prove that these coordinates are related by a *fractional linear transformation* of the following form:

$$\eta^j = (a^j + \Sigma_1^n a_i^j \xi^i)(b + \Sigma_1^n b_i \xi^i)^{-1}. \tag{19}$$

Here the coefficients a^j, a_i^j, b, b_i are the entries of the matrix for g with respect to the basis for $\boldsymbol{V}^{n+1}$ determining both the homogeneous and the corresponding inhomogeneous coordinates, which again depends on the charts chosen so that $\boldsymbol{x} \in U_k, \boldsymbol{y} \in U_l$.

1.3 Collineations

Collineations are the isomorphisms of projective geometry; sometimes they are even introduced that way, cf. Exercise 5 below. Here we will start with a much weaker definition: a map from one projective point space to another is collinear if it commutes with the operation of joining points. If, in addition, it is bijective, then we will call it a collineation: A bijective map transferring lines into lines. This property even implies that a collineation maps projective k-planes onto projective k-planes. Since the projective k-planes are embedded into the subsets lattice of the projective point spaces $\boldsymbol{P}_o^n$ via the canonical map π, each collineation generates an isomorphism of the lattices called projective geometries, and vice versa. The essential result here is the Main Theorem of

Projective Geometry describing the collineations algebraically in dimensions $n \geq 2$. One consequence of this is that geometric isomorphy implies algebraic isomorphy: Just the existence of a collineation $f : \boldsymbol{P}(\boldsymbol{V}) \mapsto \boldsymbol{Q}(\boldsymbol{W})$, where $\boldsymbol{V}, \boldsymbol{W}$ are finite-dimensional vector spaces of dimensions larger than two, suffices to conclude that the scalar domains of the vector spaces are isomorphic, and, moreover, that the spaces have equal dimensions. Note that the isomorphy of the scalar domains for the vector spaces is not a precondition, but a conclusion of the Main Theorem. In Subsection 3 this will lead to complete descriptions of the group of auto-collineations for the projective point space $\boldsymbol{P}_o^n$ and of the group of automorphisms of the projective geometry $\mathfrak{P}(\boldsymbol{V})$, $2 < \dim \boldsymbol{V} < \infty$.

1.3.1 Collinear Maps

In this section we consider maps $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^m$ preserving the geometric operation of join, i.e. satisfying $f(\boldsymbol{A} \vee \boldsymbol{B}) = f(\boldsymbol{A}) \vee f(\boldsymbol{B})$ for all projective subspaces $\boldsymbol{A}, \boldsymbol{B} \subset \boldsymbol{P}_o^n$. It suffices to require this property for the join of points (cf. Corollary 5). We define:

Definition 1. Let $\boldsymbol{P}_o^n, \boldsymbol{Q}_o^m$ be projective spaces over the respective (possibly different) skew fields K_1 and K_2. A map $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^m$ is called *collinear* if it satisfies the following two conditions:

$$f(\boldsymbol{o}) = \boldsymbol{o}, \tag{1}$$

$$f(\boldsymbol{x} \vee \boldsymbol{y}) = f(\boldsymbol{x}) \vee f(\boldsymbol{y}) \text{ for all } \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{P}_o^n. \tag{2}$$

A bijective collinear map is called a *collineation.* □

Any bijective map $f : \boldsymbol{P}^n \to \boldsymbol{Q}^m$ transferring lines into lines can always be considered as a collineation by setting $f(\boldsymbol{o}) := \boldsymbol{o}$. Conversely, because of (2), each collineation maps lines into lines. If the dimension of the image is less than two, $\operatorname{Dim} f(\boldsymbol{P}_o^n) < 2$, then, obviously, the definition has not much substance; e.g., every bijective map between projective lines satisfying (1) is a collineation. For this reason, the one-dimensional case will be treated separately in the following section.

The collinear maps with $\operatorname{Dim} f(\boldsymbol{P}_o^n) \geq 2$ can be represented as the superposition of a collineation with a (general) central projection to be introduced in the following example.

Example 1. General central projections. Let $\boldsymbol{Q}_o^m \subset \boldsymbol{P}_o^n$, $0 \leq m < n$, be a projective subspace, and let $\boldsymbol{B} \subset \boldsymbol{P}_o^n$ a complementary subspace to $\boldsymbol{Q}_o^m$. The *central projection* p from $\boldsymbol{P}_o^n$ onto $\boldsymbol{Q}_o^m$ with *center* $\boldsymbol{B}$ is defined to be the map

$$p(\boldsymbol{x}) := (\boldsymbol{x} \vee \boldsymbol{B}) \wedge \boldsymbol{Q}_o^m, \qquad \boldsymbol{x} \in \boldsymbol{P}_o^n. \tag{3}$$

In case $\boldsymbol{x} \notin \boldsymbol{B}$ Proposition 1.1 immediately implies

$$\mathrm{Dim}(\boldsymbol{x} \vee \boldsymbol{B}) = \mathrm{Dim}\,\boldsymbol{x} + \mathrm{Dim}\,\boldsymbol{B} + 1 = \mathrm{Dim}\,\boldsymbol{B} + 1 = n - m$$

as well as $\mathrm{Dim}((\boldsymbol{x} \vee \boldsymbol{B}) \wedge \boldsymbol{Q}_o^m) = 0$, so that (3) indeed uniquely determines a point of $\boldsymbol{Q}_o^m$. Obviously,

$$p^{-1}(\boldsymbol{o}) = \boldsymbol{B}. \tag{4}$$

Let $\boldsymbol{W} := \pi^{-1}(\boldsymbol{Q}_o^m)$ and $\boldsymbol{U} := \pi^{-1}(\boldsymbol{B})$ be the corresponding vector subspaces. Since $\boldsymbol{Q}_o^m$ and $\boldsymbol{B}$ are complementary, we have

$$\boldsymbol{V}^{n+1} = \boldsymbol{W} \oplus \boldsymbol{U},\ \dim \boldsymbol{W} = m + 1,\ \dim \boldsymbol{U} = n - m. \tag{5}$$

Denoting by $pr : \boldsymbol{V} \to \boldsymbol{W}$ the projection of vectors determined by the decomposition (5) it can easily be checked that the following equation holds for all $\mathfrak{x} \in \boldsymbol{V}$:

$$p([\mathfrak{x}]) = [pr(\mathfrak{x})]. \tag{6}$$

This or Definition (3) itself can be used for a straightforward proof that p is a collinear map satisfying $p \circ p = p$. If the center $\boldsymbol{B}$ is a point, we obtain the central projection discussed in Example 1.2. For a coordinate m-plane $\boldsymbol{Q}_o^m$, e.g. $\boldsymbol{Q}_o^m = \boldsymbol{a}_0 \vee \ldots \vee \boldsymbol{a}_m$, and its complementary coordinate $(n-m-1)$-plane $\boldsymbol{B} = \boldsymbol{a}_{m+1} \vee \ldots \vee \boldsymbol{a}_n$ the projection pr occurring in (6) is just the corresponding vector coordinate projection:

$$pr : \mathfrak{x} = \mathfrak{a}_i x^i \in \boldsymbol{V}^{n+1} \mapsto pr(\mathfrak{x}) = \mathfrak{a}_\alpha x^\alpha \in \boldsymbol{W}^{m+1}.$$

Here $\boldsymbol{W}^{m+1}$ denotes the vector subspace spanned by the $m+1$ basis vectors $\mathfrak{a}_\alpha$, $\alpha = 0, \ldots, m$, $i = 0, \ldots, n$, and the sum convention is in action. Restricting the coordinate functions $(x^\alpha(x))$, $\alpha = 0, \ldots, m$, to the coordinate m-plane $\boldsymbol{Q}^m$ yields a homogeneous coordinate system for this projective subspace with the base points $\boldsymbol{a}_\alpha$ whose unit point $\boldsymbol{e}_Q$ is the image under the projection of the unit point $\boldsymbol{e}$ for the coordinate system in $\boldsymbol{P}_o^n$ we started from:

$$\boldsymbol{e}_Q = p(\boldsymbol{e}) = (\boldsymbol{e} \vee \boldsymbol{a}_{m+1} \vee \ldots \vee \boldsymbol{a}_n) \wedge (\boldsymbol{a}_0 \vee \ldots \vee \boldsymbol{a}_m) = [\mathfrak{a}_0 + \ldots + \mathfrak{a}_m]. \tag{7}$$

□

Definition 2. Consider again the situation from Definition 1. Denote by $\boldsymbol{V}$ and $\boldsymbol{W}$ the vector spaces associated with the projective spaces $\boldsymbol{P}_o^n$ and $\boldsymbol{Q}_o^m$, respectively. Let, moreover, $a : \boldsymbol{V} \to \boldsymbol{W}$ and $\sigma : K_1^* \to K_2^*$ be maps such that $a(\mathfrak{o}) = \mathfrak{o}$ and

$$a(\mathfrak{x}\alpha) = a(\mathfrak{x})\sigma(\alpha) \qquad (\mathfrak{x} \in \boldsymbol{V},\ \alpha \in K_1^*). \tag{8}$$

Setting

$$\boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{P}_o^n \longmapsto f(\boldsymbol{x}) := [a(\mathfrak{x})] \in \boldsymbol{Q}_o^m \tag{9}$$

correctly defines a map $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^m$: the *map induced* by a. If there exists a map σ satisfying equation (8) for the non-trivial map a, then it evidently is uniquely determined. □

Example 2. Semi-Linearly Induced Maps. The map $a : \boldsymbol{V} \to \boldsymbol{W}$ considered in Definition 2 is called *semi-linear* if there is an isomorphism of skew fields, $\sigma : K_1 \to K_2$, such that

$$a(\mathfrak{x}\alpha + \mathfrak{y}\beta) = a(\mathfrak{x})\sigma(\alpha) + a(\mathfrak{y})\sigma(\beta) \qquad (\mathfrak{x}, \mathfrak{y} \in \boldsymbol{V},\ \alpha, \beta \in K_1).$$

For brevity, semi-linear maps with the isomorphism σ will often be called σ-linear maps. Obviously, a semi-linear map satisfies (8); the corresponding map f of projective spaces is called *semi-linearly induced.* Of course, linear maps are semi-linear with $K_1 = K_2 = K$ and $\sigma = \mathrm{id}_K$. Another important special case are the *conjugate-linear* maps determined by $K = \mathbf{C}$ and the *conjugation* $\sigma(\alpha) = \bar{\alpha}$ in $\mathbf{C}$. If $K_1 = K_2 = K$ is a skew field and $\boldsymbol{V} = \boldsymbol{W}$ is a right vector space over K, then for every $\mu \in K^*$ the *dilation*

$$d_\mu := \mathfrak{x} \in \boldsymbol{V} \mapsto \mathfrak{x}\mu \in \boldsymbol{V}$$

is semi-linear with respect to the *inner automorphism* $\sigma = \sigma_{\mu^{-1}}$; recall that the inner automorphism σ_ν of a skew field K defined by $\nu \neq 0$ is

$$\sigma_\nu : \alpha \in K \mapsto \sigma_\nu(\alpha) := \nu\alpha\nu^{-1} \in K. \tag{10}$$

σ_ν is different from the identity automorphism for all $\nu \in K^* \setminus Z(K)$. The general properties of semi-linear maps are quite similar to those of the linear maps, so they may be proved by immediately referring to the linear case. Let $a : \boldsymbol{V} \to \boldsymbol{W}$ be σ-linear. Replacing the skew field K_2 of the vector space $\boldsymbol{W}$ by the skew field K_1 using σ leads to a new vector space $\tilde{\boldsymbol{W}}$: As an Abelian group it coincides with $\boldsymbol{W}$, $[\tilde{\boldsymbol{W}}, +] = [\boldsymbol{W}, +]$; the multiplication by elements $\alpha \in K_1$ is introduced via

$$\mathfrak{x}\alpha := \mathfrak{x}\sigma(\alpha) \qquad (\mathfrak{x} \in \boldsymbol{W}, \alpha \in K_1).$$

This construction is called *replacement of scalars*; in Example II.7.1.4 it is defined in greater generality. Obviously, the subspaces of $\tilde{\boldsymbol{W}}$ are nothing but those of $\boldsymbol{W}$, and the dimension of a subspace is not changed by replacing the scalars. Moreover, the relation

$$a(\mathfrak{x}\alpha) = a(\mathfrak{x})\sigma(\alpha) = a(\mathfrak{x})\alpha \qquad (\mathfrak{x} \in \boldsymbol{V}, \alpha \in K_1),$$

implies that, now, i.e. considered as a map from $\boldsymbol{V}$ to $\tilde{\boldsymbol{W}}$, a is even linear. Hence images as well as pre-images of subspaces under a are again subspaces. In particular, the *kernel of* a, $\mathrm{Ker}\, a := a^{-1}(\mathfrak{o}) \subset \boldsymbol{V}$, and the *image of* a, $\mathrm{Im}\, a := a(\boldsymbol{V}) \subset \boldsymbol{W}$ are well-defined subspaces. As for linear maps we define the *rank of* a by $\mathrm{rk}\, a := \dim_{K_2} \mathrm{Im}\, a$. For a finite-dimensional vector space $\boldsymbol{V}$ we then have the same relation as in the case of linear maps:

$$\dim_{K_1} \mathrm{Ker}\, a + \mathrm{rk}\, a = \dim_{K_1} \boldsymbol{V} = n + 1 < \infty. \tag{11}$$

□

Now we will prove that the map f between the associated projective spaces defined, according to (9), by a semi-linear map a is collinear; relation (6) then shows that, in particular, this holds for a central projection p.

Proposition 1. *Let* $a : \boldsymbol{V}^{n+1} \to \boldsymbol{W}^{m+1}$ *be* σ*-linear. Then the map* $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^m$ *defined by* (9) *is collinear, and, moreover,*

$$\operatorname{Dim} f^{-1}(\boldsymbol{o}) + \operatorname{Dim} f(\boldsymbol{P}_o^n) = n - 1. \tag{12}$$

Proof. Since a is semi-linear, we have $a(\boldsymbol{o}) = \boldsymbol{o}$ implying (1). Now take any $\boldsymbol{x} = [\mathfrak{x}]$, $\boldsymbol{y} = [\mathfrak{y}] \in \boldsymbol{P}_o^n$. Then $\boldsymbol{x} \vee \boldsymbol{y} = \pi([\{\mathfrak{x}, \mathfrak{y}\}])$ is determined by the linear hull, and this implies (2):

$$f(\boldsymbol{x} \vee \boldsymbol{y}) = \pi(a([\{\mathfrak{x}, \mathfrak{y}\}])) = \pi([\{a(\mathfrak{x}), a(\mathfrak{y})\}]) = f(\boldsymbol{x}) \vee f(\boldsymbol{y}),$$

where $a([\{\mathfrak{x}, \mathfrak{y}\}]) = [\{a(\mathfrak{x}), a(\mathfrak{y})\}]$ follows from the semi-linearity of a. Hence f is collinear. The dimension formula (12) is a direct consequence of (11). □

As a straightforward consequence of Definition 1 we note the following.

Proposition 2. *The inverse* f^{-1} *of a collineation* f *again is a collineation. If* f *and* g *are collinear, then their composition* $f \circ g$ *is also collinear. The class of projective spaces over arbitrary skew fields with the collinear maps as morphisms form a category.* □

Example 3. The *constant maps* on $\boldsymbol{P}^n$, i.e. the maps $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o$ with $f(\boldsymbol{o}) := \boldsymbol{o}$ and $f(\boldsymbol{x}) := \boldsymbol{y}_0 = \text{const}$ for $\boldsymbol{x} \neq \boldsymbol{o}$, are always collinear. If $\boldsymbol{y}_0 \neq \boldsymbol{o}$, then $\operatorname{Dim} f^{-1}(\boldsymbol{o}) = -1$. In this case, (12) is satisfied if and only if $n = 0$; this is the only situation where f is induced by a semi-linear map (Proposition 1). In case $\boldsymbol{y}_0 = \boldsymbol{o}$ the map f is obviously induced by the zero map. Moreover, an arbitrary bijection of projective lines, $f : \boldsymbol{P}_o^1 \to \boldsymbol{Q}_o^1$, such that $f(\boldsymbol{o}) = \boldsymbol{o}$, is a collineation which, in general, is also not induced by a semi-linear map. □

Exercise 1. Find an example for a collineation $f : \boldsymbol{P}_o^1 \to \boldsymbol{Q}_o^1$ of projective lines over non-isomorphic fields.

Next we want to deduce some properties of collinear maps. Definition 1 and Exercise 1.4 immediately imply the following.

Lemma 3. *Let* $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^m$ *be collinear, and let* $\boldsymbol{A} \subset \boldsymbol{P}_o^n$, $\boldsymbol{B} \subset \boldsymbol{Q}_o^m$ *be projective subspaces. Then* $f(\boldsymbol{A})$ *and* $f^{-1}(\boldsymbol{B})$ *are also projective subspaces; so, in particular,* $f^{-1}(\boldsymbol{o})$ *and* $f(\boldsymbol{P}_o^n)$ *are projective subspaces.* □

Lemma 4. *Let* $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^m$ *be collinear. Then for all* $k \in \mathbf{N}_0$ *and any* $\boldsymbol{x}_i \in \boldsymbol{P}_o^n$

$$f(\boldsymbol{x}_0 \vee \ldots \vee \boldsymbol{x}_k) = f(\boldsymbol{x}_0) \vee \ldots \vee f(\boldsymbol{x}_k). \tag{13}$$

P r o o f. The prove will proceed by induction. For $k = 0, 1$ the statement is trivial or holds by assumption, respectively. Assume that (13) is already proved for k. If $\boldsymbol{x}_{k+1} \in \boldsymbol{H}^r := \boldsymbol{x}_0 \vee \ldots \vee \boldsymbol{x}_k$, then (13) obviously also holds for $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{k+1}\}$. Suppose that $\boldsymbol{x}_{k+1} \notin \boldsymbol{H}^r$, set $\boldsymbol{M}^{r+1} := \boldsymbol{x}_0 \vee \ldots \vee \boldsymbol{x}_{k+1} = \boldsymbol{H}^r \vee \boldsymbol{x}_{k+1}$, and take $\boldsymbol{z} \in \boldsymbol{M}^{r+1}$. Then we will have to prove the following relation:

$$f(\boldsymbol{z}) \in f(\boldsymbol{x}_0) \vee \ldots \vee f(\boldsymbol{x}_{k+1}) = f(\boldsymbol{H}^r) \vee f(\boldsymbol{x}_{k+1}).$$

We may assume $\boldsymbol{z} \neq \boldsymbol{x}_{k+1}$ as well as $\boldsymbol{z} \notin \boldsymbol{H}^r$, since otherwise the claim holds by the induction hypothesis. As $\boldsymbol{H}^r \subset \boldsymbol{M}^{r+1}$ is a hyperplane, the line $\boldsymbol{x}_{k+1} \vee \boldsymbol{z}$ intersects it in precisely one point, $\boldsymbol{y} = (\boldsymbol{x}_{k+1} \vee \boldsymbol{z}) \wedge \boldsymbol{H}^r$. Since f is collinear, $\boldsymbol{z} \in \boldsymbol{y} \vee \boldsymbol{x}_{k+1}$ implies

$$f(\boldsymbol{z}) \in f(\boldsymbol{y} \vee \boldsymbol{x}_{k+1}) = f(\boldsymbol{y}) \vee f(\boldsymbol{x}_{k+1}) \subset f(\boldsymbol{H}^r) \vee f(\boldsymbol{x}_{k+1}),$$

hence

$$f(\boldsymbol{x}_0 \vee \ldots \vee \boldsymbol{x}_{k+1}) \subset f(\boldsymbol{x}_0) \vee \ldots \vee f(\boldsymbol{x}_{k+1}).$$

On the other hand, $\boldsymbol{M}^{r+1}$ is a projective subspace of $\boldsymbol{P}^n_o$, and, by Lemma 3, so is $f(\boldsymbol{M}^{r+1}) \subset \boldsymbol{Q}^m_o$. Since $f(\boldsymbol{x}_i) \in f(\boldsymbol{M}^{r+1})$, $i = 0, \ldots, k+1$, the statement follows from the definition of the join as the smallest of all subspaces containing the one considered. □

Corollary 5. *Let $f : \boldsymbol{P}^n_o \to \boldsymbol{Q}^m_o$ be collinear. Then for any two projective subspaces $\boldsymbol{A}, \boldsymbol{B} \subset \boldsymbol{P}^n_o$ the restriction, $f|\boldsymbol{A} : \boldsymbol{A} \to \boldsymbol{Q}^m_o$, is also collinear, and* $\operatorname{Dim} f(\boldsymbol{A}) \leq \operatorname{Dim} \boldsymbol{A}$. *Moreover,*

$$f(\boldsymbol{A} \vee \boldsymbol{B}) = f(\boldsymbol{A}) \vee f(\boldsymbol{B}). \tag{14}$$

P r o o f. Both claims follow from (13) by considering $\boldsymbol{A}$, $\boldsymbol{B}$ as generated by projectively independent points. □

The following proposition characterizes injective collinear maps:

Proposition 6. *Let $f : \boldsymbol{P}^n_o \to \boldsymbol{Q}^m_o$ be collinear. Then the following statements are equivalent:*
a) *f is injective;*
b) $\operatorname{Dim} f(\boldsymbol{P}^n_o) = n$;
c) $\operatorname{Dim} f(\boldsymbol{A}) = \operatorname{Dim} \boldsymbol{A}$ *for every projective subspace* $\boldsymbol{A}$.

P r o o f. a) implies b): Suppose that $\boldsymbol{B} := f(\boldsymbol{P}^n_o) \neq \boldsymbol{o}$ and $\operatorname{Dim} \boldsymbol{B} = k < n$. Then there are $k+1$ points in general position spanning $\boldsymbol{B}$: $\boldsymbol{B} = \boldsymbol{b}_0 \vee \ldots \vee \boldsymbol{b}_k$. Choose $\boldsymbol{a}_i \in f^{-1}(\boldsymbol{b}_i) \neq \boldsymbol{o}$, $i = 0, \ldots, k$. Setting $\boldsymbol{A} := \boldsymbol{a}_o \vee \ldots \vee \boldsymbol{a}_k$ we conclude $f(\boldsymbol{A}) = \boldsymbol{B} = f(\boldsymbol{P}^n_o)$ and $\operatorname{Dim} \boldsymbol{A} < n$ from (13). Therefore, there is an $\boldsymbol{x} \in \boldsymbol{P}^n_o \setminus \boldsymbol{A}$ for which, nevertheless, $f(\boldsymbol{x}) \in \boldsymbol{B}$ has to hold. Hence f cannot be injective. If $\boldsymbol{B} = \boldsymbol{o}$, then, because of $n \geq 0$ we obtain $f(\boldsymbol{x}) = \boldsymbol{o}$ for all $\boldsymbol{x} \in \boldsymbol{P}^n$, i.e., f is not injective.

b) implies a): Consider a map f that is not injective. Then there are $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{P}^n_o$, $\boldsymbol{x} \neq \boldsymbol{y}$, such that $f(\boldsymbol{x}) = f(\boldsymbol{y})$. If, e.g., $\boldsymbol{x} = \boldsymbol{o}$, then $f(\boldsymbol{y}) = \boldsymbol{o}$ and

$\boldsymbol{y} \neq \boldsymbol{o}$. Set now $\boldsymbol{a}_0 := \boldsymbol{y}$ and choose additional points to obtain a projectively independent sequence $(\boldsymbol{a}_0, \ldots, \boldsymbol{a}_n)$ in $\boldsymbol{P}_o^n$. >From $f(\boldsymbol{a}_0) = \boldsymbol{o}$ we conclude by Lemma 4

$$f(\boldsymbol{P}_o^n) = f(\boldsymbol{a}_1) \vee \ldots \vee f(\boldsymbol{a}_n), \text{ hence } \operatorname{Dim} f(\boldsymbol{P}_o^n) < n. \tag{15}$$

If $\boldsymbol{x}$ and $\boldsymbol{y}$ are different from $\boldsymbol{o}$, denote them by $\boldsymbol{a}_0 := \boldsymbol{x}$, $\boldsymbol{a}_1 := \boldsymbol{y}$, and once more form a sequence of projectively independent points $(\boldsymbol{a}_0, \ldots, \boldsymbol{a}_n)$ in $\boldsymbol{P}_o^n$. Since $f(\boldsymbol{a}_0) = f(\boldsymbol{a}_1)$, this again yields (15).

As for each subspace $\boldsymbol{A} \subset \boldsymbol{P}_o$ the restriction $f|\boldsymbol{A}$ of the map f is also injective, c) immediately follows from the equivalence of a) and b). Since b) is a special case of c), the latter is equivalent to b). □

R e m a r k. Obviously, for each injective collinear map we have $f^{-1}(\boldsymbol{o}) = \boldsymbol{o}$. As the example of a constant map (Example 3) f with $\boldsymbol{y}_0 \neq \boldsymbol{o}$ and $n > 0$ shows, this property does not imply the injectivity of f. Nevertheless, we have the following.

Proposition 7. *Let $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^m$ be collinear and suppose, moreover, that $\operatorname{Dim} f(\boldsymbol{P}_o^n) \geq 1$. The map f is injective if and only if $f^{-1}(\boldsymbol{o}) = \boldsymbol{o}$.*

P r o o f. Without loss of generality we may assume that f is surjective. Let $\boldsymbol{b}_j$, $j = 0, \ldots, m$, be points spanning $\boldsymbol{Q}_o^m$: $\boldsymbol{Q}_o^m = \boldsymbol{b}_0 \vee \ldots \vee \boldsymbol{b}_m$. Choose points $\boldsymbol{a}_j \in f^{-1}(\boldsymbol{b}_j)$, $\boldsymbol{a}_j \neq \boldsymbol{o}$, and form $\boldsymbol{A} := \boldsymbol{a}_0 \vee \ldots \vee \boldsymbol{a}_m$. Relation (13) implies $f(\boldsymbol{A}) = \boldsymbol{Q}_o^m$, and by Corollary 5 we have $\operatorname{Dim} \boldsymbol{A} = m$. According to Proposition 6, $g := f|\boldsymbol{A}$ is bijective. Now the proof proceeds by contradiction. Suppose that $f^{-1}(\boldsymbol{o}) = \boldsymbol{o}$ and that f is not injective. Then there has to be a point $\boldsymbol{c} \in \boldsymbol{P}_o^n \setminus \boldsymbol{A}$ for which there exists a point $\boldsymbol{c}_0 \in \boldsymbol{A}$ satisfying $f(\boldsymbol{c}_0) = f(\boldsymbol{c})$. Since $\boldsymbol{c} \neq \boldsymbol{o}$ and $f^{-1}(\boldsymbol{o}) = \boldsymbol{o}$, we have $f(\boldsymbol{c}) \neq \boldsymbol{o}$ as well as $\boldsymbol{c}_0 \neq \boldsymbol{o}$. Consider the projective subspace $\boldsymbol{H}^{m+1} := \boldsymbol{c} \vee \boldsymbol{A}$ and define the map

$$q := g^{-1} \circ f|\boldsymbol{H}^{m+1} : \boldsymbol{H}^{m+1} \to \boldsymbol{A},$$

which is collinear by Proposition 2. Now $q|\boldsymbol{A} = \mathrm{id}_{\boldsymbol{A}}$ immediately implies $q^2 = q$. For each $\boldsymbol{x} \in \boldsymbol{H}^{m+1}$ there is precisely one $\boldsymbol{x}_0 \in \boldsymbol{A}$, namely $\boldsymbol{x}_0 = q(\boldsymbol{x})$, satisfying $f(\boldsymbol{x}) = f(\boldsymbol{x}_0)$. By definition, $q(\boldsymbol{x}) \in \boldsymbol{A}$, $f(q(\boldsymbol{x})) = (g \circ g^{-1} \circ f)(\boldsymbol{x})$, and $g = f|\boldsymbol{A}$ is injective. Let $\boldsymbol{x} \in \boldsymbol{H} \setminus \boldsymbol{A}$. Then $q(\boldsymbol{x} \vee \boldsymbol{x}_0) = q(\boldsymbol{x}) \vee q(\boldsymbol{x}_0) = \boldsymbol{x}_0$, and hence $\boldsymbol{x} \vee \boldsymbol{x}_0 \subset q^{-1}(\boldsymbol{x}_0)$. We are now going to show $q^{-1}(\boldsymbol{x}_0) = \boldsymbol{x} \vee \boldsymbol{x}_0$. To this end, suppose that $\boldsymbol{x}_1 \notin \boldsymbol{x} \vee \boldsymbol{x}_0$ and $q(\boldsymbol{x}_1) = \boldsymbol{x}_0$. Since $\boldsymbol{x} \neq \boldsymbol{o}$ and $q^{-1}(\boldsymbol{o}) = \boldsymbol{o}$, this implies $\boldsymbol{x}_0 \neq \boldsymbol{o}$ and $\operatorname{Dim} \boldsymbol{x} \vee \boldsymbol{x}_0 \vee \boldsymbol{x}_1 = 2$. The dimension formula then yields

$$\operatorname{Dim}(\boldsymbol{x} \vee \boldsymbol{x}_0 \vee \boldsymbol{x}_1) \wedge \boldsymbol{A} = 2 + m - (m+1) = 1.$$

Because of $q|\boldsymbol{A} = \mathrm{id}_{\boldsymbol{A}}$, this contradicts the relation

$$\boldsymbol{x}_0 = q((\boldsymbol{x} \vee \boldsymbol{x}_0 \vee \boldsymbol{x}_1) \wedge \boldsymbol{A}) = (\boldsymbol{x} \vee \boldsymbol{x}_0 \vee \boldsymbol{x}_1) \wedge \boldsymbol{A}.$$

But by assumption we have $m+1 \geq 2$. Therefore, there has to exist a

$$\boldsymbol{z} \in \boldsymbol{H}^{m+1} \setminus (\boldsymbol{A} \cup (\boldsymbol{x} \vee \boldsymbol{x}_0)).$$

In fact, take any projective frame $(\boldsymbol{a}_0, \ldots, \boldsymbol{a}_{m+1}; e)$ for $\boldsymbol{H}^{m+1}$ such that $\boldsymbol{x} \vee \boldsymbol{x}_0 = \boldsymbol{a}_0 \vee \boldsymbol{a}_{m+1}$ and $\boldsymbol{A} = \boldsymbol{a}_0 \vee \ldots \vee \boldsymbol{a}_m$. Then $\boldsymbol{z} = \boldsymbol{e}$ has the above property. Let $\boldsymbol{z}_0 := q(\boldsymbol{z})$. As shown before, $q^{-1}(\boldsymbol{z}_0) = \boldsymbol{z} \vee \boldsymbol{z}_0$, and hence $\boldsymbol{z}_0 \neq \boldsymbol{x}_0$. Moreover,

$$(\boldsymbol{z} \vee \boldsymbol{z}_0) \wedge (\boldsymbol{x} \vee \boldsymbol{x}_0) = \boldsymbol{o}.$$

Then, $\boldsymbol{y} \iota (\boldsymbol{z} \vee \boldsymbol{z}_0) \wedge (\boldsymbol{x} \vee \boldsymbol{x}_0)$ has to satisfy $q(\boldsymbol{y}) \iota \boldsymbol{x}_0$ as well as $q(\boldsymbol{y}) \iota \boldsymbol{z}_0$. Since $\boldsymbol{x}_0 \neq \boldsymbol{z}_0$, this is only possible for $\boldsymbol{y} = \boldsymbol{o}$. Thus the lines $\boldsymbol{z} \vee \boldsymbol{z}_0$, $\boldsymbol{x} \vee \boldsymbol{x}_0$ are skew, and hence $\boldsymbol{L} := (\boldsymbol{z} \vee \boldsymbol{z}_0) \vee (\boldsymbol{x} \vee \boldsymbol{x}_0)$ is a three-dimensional subspace of $\boldsymbol{H}^{m+1}$. Since $\boldsymbol{H}^{m+1} = \boldsymbol{L} \vee \boldsymbol{A}$, this implies $\operatorname{Dim} \boldsymbol{L} \wedge \boldsymbol{A} = 2$. Because of $q|\boldsymbol{A} = \mathrm{id}_{\boldsymbol{A}}$ we arrive at the contradiction

$$\boldsymbol{A} \wedge \boldsymbol{L} \subset q(\boldsymbol{L}) = q(\boldsymbol{x}) \vee q(\boldsymbol{x}_0) \vee q(\boldsymbol{z}) \vee q(\boldsymbol{z}_0) = \boldsymbol{x}_0 \vee \boldsymbol{z}_0.$$

The converse is trivial. □

Now we can easily deduce the following.

Corollary 8. *Let* $f : \boldsymbol{P}^n_o \to \boldsymbol{Q}^m_o$ *be collinear with* $\operatorname{Dim} f(\boldsymbol{P}^n_o) \geq 1$. *Then* f *can be represented as a composition,* $f = g \circ p$, *of a central projection* p *and an injective collinear map* g *in the following way: For* $\boldsymbol{B} = f^{-1}(\boldsymbol{o})$ *and any subspace* $\boldsymbol{A}$ *complementary to* $\boldsymbol{B}$ *in* $\boldsymbol{P}^n_o$, *let* p *be the central projection* $p : \boldsymbol{P}^n_o \to \boldsymbol{A}$ *with center* $\boldsymbol{B}$. *Then* $g := f|\boldsymbol{A}$ *is injective, and* $f = g \circ p$.

P r o o f. First, by construction $(g \circ p)(\boldsymbol{x}) = f((\boldsymbol{x} \vee \boldsymbol{B}) \wedge \boldsymbol{A}) \iota f(\boldsymbol{x}) \vee f(\boldsymbol{B}) = f(\boldsymbol{x})$. Since $\boldsymbol{A} \wedge \boldsymbol{B} = \boldsymbol{o}$, the relation $(g \circ p)(\boldsymbol{x}) = \boldsymbol{o}$ can only occur for $p(\boldsymbol{x}) = \boldsymbol{o}$, i.e. $\boldsymbol{x} \in \boldsymbol{B}$, and hence $(g \circ p)(\boldsymbol{x}) = f(\boldsymbol{x})$. Moreover, $p(\boldsymbol{x}) \in \boldsymbol{A}$, thus $f(\boldsymbol{P}^n_o) = f(\boldsymbol{A})$. Hence we have $\operatorname{Dim} g(\boldsymbol{A}) \geq 1$, and because of $\boldsymbol{A} \wedge \boldsymbol{B} = \boldsymbol{o}$ we obtain $g^{-1}(\boldsymbol{o}) = \boldsymbol{o}$. By Proposition 7 the map g is injective. □

Proposition 6 implies $\operatorname{Dim} \boldsymbol{A} = \operatorname{Dim} f(\boldsymbol{P}^n_o)$, and hence we immediately arrive at the following property corresponding to (12):

Corollary 9. *Let* $f : \boldsymbol{P}^n_o \to \boldsymbol{Q}^m_o$ *be a collinear map such that* $\operatorname{Dim} f(\boldsymbol{P}^n_o) \geq 1$. *Then, for every projective subspace* $\boldsymbol{H} \subset \boldsymbol{P}^n_o$,

$$\operatorname{Dim} f(\boldsymbol{H}) = \operatorname{Dim} \boldsymbol{H} - \operatorname{Dim} \boldsymbol{H} \wedge f^{-1}(\boldsymbol{o}) - 1. \tag{16}$$

Thus for any collineation f *the dimensions of a projective subspace and its image always coincide,* $\operatorname{Dim} f(\boldsymbol{H}) = \operatorname{Dim} \boldsymbol{H}$.

P r o o f. By Corollary 8, $f = g \circ p$. Hence $f(\boldsymbol{H}) = (g \circ p)(\boldsymbol{H})$, and since g is injective, Proposition 6, c) implies $\operatorname{Dim} f(\boldsymbol{H}) = \operatorname{Dim} p(\boldsymbol{H})$. On the other hand, $p(\boldsymbol{H}) = (\boldsymbol{H} \vee \boldsymbol{B}) \wedge \boldsymbol{A}$. Applying the dimension formula twice, the complementarity of $\boldsymbol{A}$ and $\boldsymbol{B} = f^{-1}(\boldsymbol{o})$ yields:

$$\begin{aligned}\operatorname{Dim} p(\boldsymbol{H}) &= \operatorname{Dim} \boldsymbol{H} + \operatorname{Dim} \boldsymbol{B} - \operatorname{Dim} \boldsymbol{H} \wedge \boldsymbol{B} + \operatorname{Dim} \boldsymbol{A} - n, \\ &= \operatorname{Dim} \boldsymbol{H} - \operatorname{Dim} \boldsymbol{H} \wedge \boldsymbol{B} - 1.\end{aligned}$$

For a collineation f we have by definition $f^{-1}(\boldsymbol{o}) = \boldsymbol{o}$, hence $\operatorname{Dim} \boldsymbol{H} \wedge \boldsymbol{B} = -1$; the case $\operatorname{Dim} f(\boldsymbol{P}_o^n) < 1$ is trivially true for collineations (and in general false for collinear maps, cf. Example 2). □

1.3.2 The Main Theorem of Projective Geometry

In this section we will prove the following *Main Theorem of Projective Geometry*, which also implies the characterization of collinear maps mentioned in the introduction to this chapter (cf. E. Artin [4], J. Dieudonné [36]):

Proposition 10. *Let $\boldsymbol{V}$, $\boldsymbol{W}$ be $(n+1)$-dimensional vector spaces, $n \geq 2$, over the skew fields K_1 and K_2, respectively, with associated projective spaces $\boldsymbol{P}_o^n$ and $\boldsymbol{Q}_o^n$. Let, moreover,*

$$f : \boldsymbol{x} \in \boldsymbol{P}_o^n \longmapsto \boldsymbol{x}' = f(\boldsymbol{x}) \in \boldsymbol{Q}_o^n$$

be a collineation. Then there are an isomorphism of skew fields $\sigma : K_1 \to K_2$ and a σ-linear bijection $a : \boldsymbol{V} \to \boldsymbol{W}$ inducing f, i.e. satisfying $f([\mathfrak{x}]) = [a(\mathfrak{x})]$ for all $\mathfrak{x} \in \boldsymbol{V}$. If $a_1 : \boldsymbol{V} \to \boldsymbol{W}$ is another σ_1-linear bijection inducing f, then there exists $\mu \in K_1^*$ such that

$$a_1(\mathfrak{x}) = a(\mathfrak{x}\mu), \qquad \mathfrak{x} \in \boldsymbol{V}, \tag{17}$$

$$\sigma_1(\xi) = \sigma(\mu^{-1}\xi\mu), \qquad \xi \in K_1. \tag{18}$$

P r o o f. Proposition 6, c) implies that the images of the points in a projective frame $(\boldsymbol{a}_0, \ldots, \boldsymbol{a}_n : \boldsymbol{e})$ of $\boldsymbol{P}_o^n$,

$$\boldsymbol{a}_i' = f(\boldsymbol{a}_i),\ \boldsymbol{e}' = f(\boldsymbol{e}),\ i = 0, \ldots, n, \tag{19}$$

form a projective frame of $\boldsymbol{Q}_o^n$, since any $n+1$ among them are projectively independent. Let $(\mathfrak{a}_i)$ and $(\mathfrak{a}_i')$ be bases of $\boldsymbol{V}$ and $\boldsymbol{W}$, respectively, determining the projective frames $(\boldsymbol{a}_0, \ldots, \boldsymbol{a}_n : \boldsymbol{e})$ and $(\boldsymbol{a}_0', \ldots, \boldsymbol{a}_n' : \boldsymbol{e}')$. By Proposition 2.2 each of them is uniquely defined up to a factor from K_1^* and K_2^*, respectively. We will prove that there is an isomorphism $\sigma : K_1 \to K_2$ for which, in the homogeneous coordinates corresponding to the bases, f is represented by

$$f([\mathfrak{a}_i x^i]) = [\mathfrak{a}_i' \sigma(x^i)]. \tag{20}$$

(Recall the sum convention!), the σ-linear map we are looking for is thus

$$a : \mathfrak{a}_i x^i \in \boldsymbol{V} \mapsto \mathfrak{a}_i' \sigma(x^i) \in \boldsymbol{W}.$$

To see this, consider a coordinate axis $\boldsymbol{a}_0 \vee \boldsymbol{a}_j$, $0 \neq j$. Because of (19) we obtain a bijection $\sigma_j : K_1 \to K_2$ defined by

$$f([\mathfrak{a}_0+\mathfrak{a}_j\xi]) = [\mathfrak{a}'_0+\mathfrak{a}'_j\sigma_j(\xi)],\ \xi\in K_1, \tag{21}$$

(cf. Example 1.3). We prove: *The map σ_j is independent of j.* Consider, say, σ_1 and σ_2. By (19) we know that $\sigma_j(0)=0$. The same formula implies that, in general, f maps the coordinate k-planes into one another. Since f is collinear, it commutes with the join operation $\vee$, and since f is bijective, it furthermore commutes with the section operation $\wedge$ (coinciding with the usual intersection $\cap$). Hence (7) implies that f also maps the unit points of the coordinate k-planes into one another. Applied to the coordinate axes considered this yields $\sigma_j(1)=1$. Let now $\boldsymbol{x}_j(\xi):=[\mathfrak{a}_0+\mathfrak{a}_j\xi]$, $j=1,2$, be the points with parameter values $\xi\neq 0$ on the coordinate axes. It is easy to see that all connecting lines $\boldsymbol{x}_1(\xi)\vee\boldsymbol{x}_2(\xi)$ intersect the coordinate axis $\boldsymbol{a}_1\vee\boldsymbol{a}_2$ in one and the same point $\boldsymbol{c}$: In fact, for fixed ξ, this point satisfies

$$\boldsymbol{c}=[(\mathfrak{a}_0+\mathfrak{a}_1\xi)\alpha+(\mathfrak{a}_0+\mathfrak{a}_2\xi)\beta]=[\mathfrak{a}_1\gamma+\mathfrak{a}_2\delta],$$

which implies $\alpha=-\beta$ and $\gamma=\xi\alpha=-\xi\beta=-\delta$. Hence $\boldsymbol{c}=[\mathfrak{a}_2-\mathfrak{a}_1]$ is independent of ξ.

Remark. Existence as well as uniqueness of the intersection point $\boldsymbol{c}$ are obvious, since the configuration lies in the coordinate plane $\boldsymbol{a}_0\vee\boldsymbol{a}_1\vee\boldsymbol{a}_2$ (cf. Example 1.4). Viewing this plane affinely, with $\boldsymbol{a}_1\vee\boldsymbol{a}_2$ as the line at infinity, this fact can be interpreted by saying that the connecting line of points with equal ordinates on the axes are all parallel, which immediately follows from the intercept theorem.

Consider the image of this configuration under the collineation f. Since lines are mapped into lines by f, which moreover commutes with section and join, in addition taking into account (21), we obtain for the image of the intersection point:

$$\begin{aligned}\boldsymbol{c}' &= (\boldsymbol{x}'_1(\xi)\vee\boldsymbol{x}'_2(\xi))\wedge(\boldsymbol{a}'_1\vee\boldsymbol{a}'_2)\\ &= [(\mathfrak{a}'_0+\mathfrak{a}'_1\sigma_1(\xi))\tilde{\alpha}+(\mathfrak{a}'_0+\mathfrak{a}'_2\sigma_2(\xi))\tilde{\beta}]=[\mathfrak{a}'_1\tilde{\gamma}+\mathfrak{a}'_2\tilde{\delta}].\end{aligned}$$

Comparing coefficients yields

$$\tilde{\alpha}=-\tilde{\beta},\ \tilde{\gamma}=\sigma_1(\xi)\alpha,\ \tilde{\delta}=\sigma_2(\xi)\tilde{\beta}=-\sigma_2(\xi)\tilde{\alpha},$$

and we conclude

$$\boldsymbol{c}'=[\mathfrak{a}'_2\sigma_2(\xi)-\mathfrak{a}'_1\sigma_1(\xi)].$$

Since $\boldsymbol{c}'$ does not depend on ξ, we may set $\xi=1$ and finally obtain from the resulting equation, $\sigma_j(1)=1$, that $\boldsymbol{c}'=[\mathfrak{a}'_2-\mathfrak{a}'_1]$. This together with the last formula implies the assertion $\sigma_1(\xi)=\sigma_2(\xi)$ for all $\xi\in K_1$.

Now we prove (20). First let $\boldsymbol{x}\notin\boldsymbol{H}_0$, where $\boldsymbol{H}_0=\boldsymbol{a}_1\vee\ldots\vee\boldsymbol{a}_n$ denotes the coordinate hyperplane complementary to $\boldsymbol{a}_0$. Since $f(\boldsymbol{H}_0)$ is the coordinate hyperplane $\boldsymbol{H}'_0$ complementary to $\boldsymbol{a}'_0$, we also have $\boldsymbol{x}'=f(\boldsymbol{x})\notin\boldsymbol{H}'_o$. Hence the coordinates x^0, $x'^{\,0}$ of $\boldsymbol{x}$ and $\boldsymbol{x}'$ are different from zero. Normalizing these to have $x^0=1$, $x'^{\,0}=1$ we obtain the normalized representations

$$\boldsymbol{x} = [\mathfrak{a}_0 + \mathfrak{a}_1 x^1 + \ldots \mathfrak{a}_n x^n],\ \boldsymbol{x}' = [\mathfrak{a}'_0 + \mathfrak{a}'_1 x'^1 + \ldots \mathfrak{a}'_n x'^n].$$

Let p_1, p'_1 be the respective projections onto the coordinate axes $\boldsymbol{a}_0 \vee \boldsymbol{a}_1$ and $\boldsymbol{a}'_0 \vee \boldsymbol{a}'_1$. Moreover, denoting by $\boldsymbol{H}_{01}$, $\boldsymbol{H}'_{01}$ the coordinate planes complementary to these axes, we have

$$p_1(\boldsymbol{x}) = (\boldsymbol{x} \vee \boldsymbol{H}_{01}) \wedge (\boldsymbol{a}_0 \vee \boldsymbol{a}_1) = [\mathfrak{a}_0 + \mathfrak{a}_1 x^1]$$
$$p'_1(\boldsymbol{x}') = (\boldsymbol{x}' \vee \boldsymbol{H}'_{01}) \wedge (\boldsymbol{a}'_0 \vee \boldsymbol{a}'_1) = [\mathfrak{a}'_0 + \mathfrak{a}'_1 x'^1].$$

The equations on the right-hand side of these formulas immediately result from Example 1, (6). Now, since f maps the coordinate planes into one another, the fact that the collineation f commutes with $\vee$ and $\wedge$ together with Definition (21) of σ implies:

$$f(p_1(\boldsymbol{x})) = [\mathfrak{a}'_0 + a'_1\sigma(x^1)] = p'_1(f(\boldsymbol{x})) = [\mathfrak{a}'_0 + a'_1 x'^1)],$$

i.e. $x'^1 = \sigma(x^1)$. As corresponding relations hold for all the coordinate axes, we obtain

$$f([\mathfrak{a}_0 + \sum_{j=1}^n \mathfrak{a}_j x^j]) = [\mathfrak{a}'_0 + \sum_{j=1}^n \mathfrak{a}'_j \sigma(x^j)],\ x^j \in K_1, \tag{22}$$

and this is just assertion (20) for the case $\boldsymbol{x} \notin \boldsymbol{H}_0$. If $\boldsymbol{x} \in \boldsymbol{H}_0$, then also $\boldsymbol{x}' = f(\boldsymbol{x}) \in \boldsymbol{H}'_0$, and we have $x^0 = x'^{\,0} = 0$. In this case, $\boldsymbol{x} = [\mathfrak{x}]$ with $\mathfrak{x} = \sum_{j=1}^n \mathfrak{a}_j x^j$ lies on the line $\boldsymbol{a}_0 \vee [\mathfrak{a}_0 + \mathfrak{x}]$. Hence, by (19) and (21), $f(\boldsymbol{x})$ lies on the line $\boldsymbol{a}'_0 \vee [\mathfrak{a}'_0 + \sum_{j=1}^n \mathfrak{a}'_j \sigma(x^j)]$. Thus

$$\boldsymbol{x}' = [\mathfrak{a}'_0 \mu + (\mathfrak{a}'_0 + \sum_{j=1}^n \mathfrak{a}'_j \sigma(x^j))\nu] = [\sum_{j=1}^n \mathfrak{a}'_j y^j)].$$

This implies $\mu = -\nu$ and, since a common factor does not matter, $y^j = \sigma(x^j)$. This again is the asserted equation (20).

Now it is not difficult to show that σ is an isomorphism: The point $\boldsymbol{x} = [\mathfrak{a}_0 + \mathfrak{a}_1(\xi + \eta) + \mathfrak{a}_2]$ lies on the line $[\mathfrak{a}_0 + \mathfrak{a}_1\xi] \vee [\mathfrak{a}_1\eta + \mathfrak{a}_2]$. Consequently, the image point $\boldsymbol{x}' = [\mathfrak{a}'_0 + \mathfrak{a}'_1\sigma(\xi + \eta) + \mathfrak{a}'_2]$ is on the image line $[\mathfrak{a}'_0 + \mathfrak{a}'_1\sigma(\xi)] \vee [\mathfrak{a}'_1\sigma(\eta) + \mathfrak{a}'_2]$. This implies $\sigma(\xi + \eta) = \sigma(\xi) + \sigma(\eta)$. Analogously, the point $\boldsymbol{x} = [\mathfrak{a}_0 + \mathfrak{a}_1\xi\eta + \mathfrak{a}_2\eta]$ lies on the line $\boldsymbol{a}_0 \vee [\mathfrak{a}_1\xi + \mathfrak{a}_2]$, so the image point $\boldsymbol{x}' = [\mathfrak{a}'_0 + \mathfrak{a}'_1\sigma(\xi\eta) + \mathfrak{a}'_2\sigma(\eta)]$ belongs to the image line $\boldsymbol{a}'_0 \vee [\mathfrak{a}'_1\sigma(\xi) + \mathfrak{a}'_2]$. Comparing coefficients in

$$\mathfrak{a}'_0 + \mathfrak{a}'_1\sigma(\xi\eta) + \mathfrak{a}'_2\sigma(\eta) = \mathfrak{a}'_0\nu + (\mathfrak{a}'_1\sigma(\xi) + \mathfrak{a}'_2)\mu$$

yields $\nu = 1$, $\mu = \sigma(\eta)$, $\sigma(\xi\eta) = \sigma(\xi)\sigma(\eta)$. This completes the proof of the existence part of the proposition.

To prove uniqueness consider the map $b := a^{-1} \circ a_1 : \boldsymbol{V} \to \boldsymbol{V}$. It is a semilinear bijection with the isomorphism $\tau := \sigma^{-1} \circ \sigma_1$ of K_1. Since it induces the identity on $\boldsymbol{P}^n_o$, there has to be a function

$$\lambda : \mathfrak{x} \in \boldsymbol{V} \setminus \{\mathfrak{o}\} \mapsto \lambda(\mathfrak{x}) \in K_1^* \text{ with } b(\mathfrak{x}) = \mathfrak{x}\lambda(\mathfrak{x}).$$

In particular, we have $b(\mathfrak{a}_i) = \mathfrak{a}_i\alpha_i$, $\alpha_i \in K_1^*$. From

$$b(\mathfrak{a}_i + \mathfrak{a}_j) = (\mathfrak{a}_i + \mathfrak{a}_j)\lambda(\mathfrak{a}_i + \mathfrak{a}_j) = b(\mathfrak{a}_i) + b(\mathfrak{a}_j) = \mathfrak{a}_i\alpha_i + \mathfrak{a}_j\alpha_j$$

we conclude $\alpha_i = \alpha_j =: \mu$ and hence $b(\mathfrak{a}_j) = \mathfrak{a}_j\mu$ for a certain $\mu \in K_1^*$ and all j. Since $\mathfrak{a}_1$ is an arbitrary vector that is linearly independent of $\mathfrak{a}_0$, and $\dim \boldsymbol{V} \geq 3$, we obtain $\lambda(\mathfrak{x}) = \mu$ for all $\mathfrak{x} \in \boldsymbol{V} \setminus \{\mathfrak{o}\}$. Thus $b(\mathfrak{x}) = \mathfrak{x}\mu$, i.e. $a_1(\mathfrak{x}) = a(\mathfrak{x}\mu)$ for some $\mu \in K_1^*$. Furthermore, we obtain $b(\mathfrak{x}\xi) = \mathfrak{x}\mu\tau(\xi) = \mathfrak{x}\xi\mu$, i.e. $\tau(\xi) = \mu^{-1}\xi\mu$ and $\sigma_1(\xi) = \sigma(\mu^{-1}\xi\mu)$. □

For collinear maps the Main Theorem implies the following.

Corollary 11. *Every collinear map* $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^m$ *such that* $\operatorname{Dim} f(\boldsymbol{P}_o^n) \geq 2$ *is generated by a semi-linear map between the associated vector spaces* $b : \boldsymbol{V} \to \boldsymbol{W}$.

Proof. In the notations of Corollary 8 the map f can be represented in the form $f = g \circ p$; according to (6) the central projection p is induced linearly. Since $g(\boldsymbol{A}) = f(\boldsymbol{P}_o^n)$, the map $g : \boldsymbol{A} \to f(\boldsymbol{P}_o^n)$ is a collineation satisfying the hypotheses of Proposition 10. Hence it is generated by a semi-linear map a. Consequently, the composition $f = g \circ p$ is induced by a semi-linear map $b = a \circ pr$. □

1.3.3 The Group of Auto-Collineations

An important and somewhat surprising partial result of the Main Theorem is the isomorphy between the skew fields of two projective spaces that are mapped onto one another by a collineation. Because of this *invariance of the scalar domains*, considering collinear maps we will always suppose $K_1 = K_2$, i.e., we will be working within the category of projective geometries (or spaces, respectively) over a fixed skew field K. Since, as already mentioned, the notion of a collineation is less rich in substance for $n < 2$, for the description of the automorphism group of projective spaces we will require $n \geq 2$ until the end of this section. The projective lines, for which the cross ratio plays a decisive part, will be studied in the subsequent section. Let us start with the following definition.

Definition 3. For $n \geq 2$, let $\operatorname{Aut} \boldsymbol{P}^n$ denote the group of all collineations $f : \boldsymbol{P}^n \to \boldsymbol{P}^n$ from $\boldsymbol{P}^n$ to itself. The elements $f \in \operatorname{Aut} \boldsymbol{P}^n$ are called the *auto-collineations* of $\boldsymbol{P}^n$. Here and in the sequel we will often omit the index o, when there is no risk of confusion; if necessary, a map $f : \boldsymbol{P}^n \to \boldsymbol{Q}^m$ will be understood to be extended by $f(\boldsymbol{o}) := \boldsymbol{o}$ to a map $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^m$. □

The Main Theorem allows us to describe $\operatorname{Aut} \boldsymbol{P}^n$ algebraically, since each auto-collineation is generated by a semi-linear bijection. First we will prove the following.

Lemma 12. *The group of all the* $f \in \operatorname{Aut} \boldsymbol{P}^n$ *leaving each point of a given projective frame* $(\boldsymbol{a}_i; \boldsymbol{e})$, $i = 0 \dots, n$, *of* $\boldsymbol{P}^n$ *fixed is isomorphic to the* group of automorphisms $\operatorname{Aut} K$ *of the skew field* K *for* $\boldsymbol{P}^n$.

Proof. Let $\sigma \in \operatorname{Aut} K$, and take a basis $(\mathfrak{a}_i)$ of the vector space $\boldsymbol{V}$ associated with $\boldsymbol{P}^n$ determining the frame $(\boldsymbol{a}_i; \boldsymbol{e})$. Then the map $\varphi : \sigma \mapsto f_\sigma$ with

$$f_\sigma([\mathfrak{a}_i x^i]) := [\mathfrak{a}_i \sigma(x^i)] \tag{23}$$

defines a homomorphism from $\operatorname{Aut} K$ to $\operatorname{Aut} \boldsymbol{P}^n$, for which every f_σ belongs to the stationary subgroup of $(\boldsymbol{a}_i; \boldsymbol{e})$, i.e., it satisfies the conditions

$$f(\boldsymbol{a}_i) = \boldsymbol{a}_i,\ f(\boldsymbol{e}) = \boldsymbol{e},\ i = 0, \dots, n, \tag{24}$$

since $\sigma(0) = 0$ and $\sigma(1) = 1$. Now Definition (23) immediately implies $f_{\tau \circ \sigma} = f_\tau \circ f_\sigma$. If, moreover, $f_\sigma = \mathrm{id}_{\boldsymbol{P}^n}$, then in particular,

$$f_\sigma([\mathfrak{a}_0 + \mathfrak{a}_1 \xi]) = [\mathfrak{a}_0 + \mathfrak{a}_1 \sigma(\xi)] = [\mathfrak{a}_0 + \mathfrak{a}_1 \xi],$$

and this implies $\sigma(\xi) = \xi$ for all $\xi \in K$. Hence φ is injective. So we still have to show that each auto-collineation f satisfying (24) has the form (23). First, the equations (24) yield for any semi-linear map a inducing f with corresponding automorphism $\sigma \in \operatorname{Aut} K$:

$$a(\mathfrak{a}_i) = \mathfrak{a}_i \mu_i,\ a(\sum \mathfrak{a}_j) = \sum \mathfrak{a}_j \mu_j = (\sum \mathfrak{a}_j)\mu,$$

i.e. $\mu_i = \mu \neq 0$ for $i = 0, \dots, n$. Hence we have

$$a(\mathfrak{a}_i x^i) = \mathfrak{a}_i \mu \sigma(x^i), \tag{25}$$

$$f([\mathfrak{a}_i x^i]) = [\mathfrak{a}_i \mu \sigma(x^i)] = [\mathfrak{a}_i \mu \sigma(x^i) \mu^{-1}]. \tag{26}$$

Setting thus $\tau := \sigma_\mu \circ \sigma$, where σ_μ is the inner automorphism corresponding to $\mu \in K^*$, implies $f = f_\tau$. □

Lemma 12 leads to a geometric interpretation of the automorphism group of the skew field K as the stationary subgroup of a projective frame for the projective space $\boldsymbol{P}^n$ over K under the action of $\operatorname{Aut} K$. To describe the group $\operatorname{Aut} \boldsymbol{P}^n$ algebraically, we start with an algebraic representation of the *semi-linear group* $\boldsymbol{G} = \boldsymbol{G}(\boldsymbol{V}^{n+1})$, i.e. the group of semi-linear bijections of the vector space $\boldsymbol{V}^{n+1}$. Let $(\mathfrak{a}_i)$ be a fixed basis of $\boldsymbol{V} = \boldsymbol{V}^{n+1}$ and take any $a \in \boldsymbol{G}$. Then we have

$$a(\mathfrak{a}_i x^i) = a(\mathfrak{a}_i)\sigma(x^i) = \mathfrak{a}_j \alpha_i^j \sigma(x^i). \tag{27}$$

Obviously, every semi-linear map a of $\boldsymbol{V}$ can be represented this way, and each of these maps different from the zero map determines the associated automorphism $\sigma \in \operatorname{Aut} K$ uniquely; of course, the zero map is linear. Any $a \in \boldsymbol{G}$ is invertible, and the inverse map a^{-1} is semi-linear with automorphism

σ^{-1}. To determine the matrix of the inverse map we set $a^{-1}(\mathfrak{a}_j) = \mathfrak{a}_k\gamma_j^k$. This implies

$$\mathfrak{a}_i = a^{-1}(a(\mathfrak{a}_i)) = a^{-1}(\mathfrak{a}_j\alpha_i^j) = \mathfrak{a}_k\gamma_j^k\sigma^{-1}(\alpha_i^j),$$

and, by comparing coefficients and applying σ:

$$\delta_i^k = \gamma_j^k\sigma^{-1}(\alpha_i^j) = \sigma(\gamma_j^k)\alpha_i^j.$$

Consequently, the matrix (α_i^j) is invertible, and we obtain a map Φ that is well-defined by (27),

$$\Phi : a \in \boldsymbol{G}(\boldsymbol{V}^{n+1}) \longmapsto \Phi(a) := ((\alpha_i^j),\, \sigma) \in \boldsymbol{GL}(n+1, K) \times \operatorname{Aut} K.$$

Furthermore,

$$\Phi(a^{-1}) = ((\sigma^{-1}(\alpha_i^j))^{-1}, \sigma^{-1}). \tag{28}$$

A straightforward calculation using $\Phi(b) = ((\beta_l^k),\, \tau)$ leads to

$$\Phi(b \circ a) = ((\beta_l^k)(\tau(\alpha_i^l)), \tau \circ \sigma). \tag{29}$$

Defining in the product set

$$\boldsymbol{G}_{n+1}(K) := \boldsymbol{GL}(n+1, K) \times \operatorname{Aut} K$$

a multiplication by

$$((\beta_l^k), \tau) \cdot ((\alpha_i^l), \sigma) := ((\beta_l^k)(\tau(\alpha_i^l)), \tau \circ \sigma), \tag{30}$$

results in a group, the *semi-linear group* $\boldsymbol{G}_{n+1}(K)$ *of order* $n+1$ *over* K, as the semi-direct product of the normal subgroup $\boldsymbol{GL}(n+1) \times \{\mathrm{id}_K\}$ with the subgroup $\{(\delta_j^i)\} \times \operatorname{Aut} K$, cf. Exercise I.3.2.3. Moreover, Φ turns out to be an isomorphism:

$$\boldsymbol{G}(\boldsymbol{V}^{n+1}) \cong \boldsymbol{G}_{n+1}(K),\ n \in \mathbf{N_0}. \tag{31}$$

In Example 2 we introduced the dilation d_μ by $d_\mu(\mathfrak{x}) = \mathfrak{x}\mu$, $\mathfrak{x} \in \boldsymbol{V}$, $\mu \in K^*$; d_μ is semi-linear with automorphism $\sigma_{\mu^{-1}}$. As the following proposition implies, the set of all dilations

$$\boldsymbol{D} := \{d_\mu | \mu \in K^*\} \subset \boldsymbol{G}(\boldsymbol{V})$$

is a normal subgroup of $\boldsymbol{G}(\boldsymbol{V})$. By means of

$$\mu \in K^* \mapsto d_{\mu^{-1}} \in \boldsymbol{D}, \qquad K^* \cong \boldsymbol{D},$$

it is isomorphic to K^*. In the sequel we consider K^* as canonically embedded into the semi-linear group via

$$\mu \in K^* \mapsto \Phi(d_{\mu^{-1}}) = ((\delta_i^j\mu^{-1}),\, \sigma_\mu) \in \boldsymbol{G}_{n+1}(K), \tag{32}$$

and correspondingly write $K^* \subset \boldsymbol{G}_{n+1}(K)$.

Proposition 13. *The map* $F : a \in \boldsymbol{G}(\boldsymbol{V}^{n+1}) \mapsto f_a \in \operatorname{Aut} \boldsymbol{P}^n$, $n \geq 2$, *defined by* $f_a([\mathfrak{x}]) := [a(\mathfrak{x})]$, *is a surjective homomorphism with kernel*

$$\operatorname{Ker} F = \boldsymbol{D}, \tag{33}$$

so that

$$\operatorname{Aut} \boldsymbol{P}^n \cong \boldsymbol{G}_{n+1}(K)/K^*, \quad n \geq 2. \tag{34}$$

Proof. The Main Theorem together with the obvious formula $f_{b \circ a} = f_b \circ f_a$ immediately show that F is a surjective homomorphism. Hence, (34) can be deduced from (33) using (31) and (32). To describe the kernel $\operatorname{Ker} F$, choose a basis for $\boldsymbol{V}^{n+1}$ and consider a map $f_a \in \boldsymbol{G}(\boldsymbol{V}^{n+1})$ determined by $f_a = \Phi^{-1}((\alpha_i^j), \sigma)$, for which $f_a = \mathrm{id}_{\boldsymbol{P}^n}$. Since f_a in particular satisfies (24), we derive relation (26) as in the proof of Lemma 12; we only have to replace μ by μ^{-1}:

$$(\alpha_i^j) = (\delta_i^j \mu^{-1}), \qquad f_a = f_\tau \text{ with } \tau = \sigma_{\mu^{-1}} \circ \sigma,$$

cf. (23). Since the map $\sigma \mapsto f_\sigma$ is injective by Lemma 12, τ has to be the identity of K, i.e. $\sigma = \sigma_\mu$. On the other hand, each $d_\mu \in \boldsymbol{D}$ obviously belongs to $\operatorname{Ker} F$. □

Now we ask whether the sequences $(\boldsymbol{a}_0, \ldots, \boldsymbol{a}_n; \boldsymbol{e})$ we called projective frames actually are *frames in the sense of the group-theoretic view onto geometry*. According to E. Cartan [28] this means that the group, on which the geometry is based, acts simply transitively on this set of frames. As will be shown below this does not hold, in general for the group of auto-collineations; in the important case of the field $K = \mathbf{R}$ of real numbers, however, it is true.

Proposition 14. *Let* $\boldsymbol{P}^n$, $\boldsymbol{Q}^n$ *be* n*-dimensional projective spaces,* $n \geq 1$, *over the skew field* K, *and let* $\sigma \in \operatorname{Aut} K$ *be any automorphism of* K. *Then, for every pair of projective frames,* $(\boldsymbol{a}_j; \boldsymbol{e})$, $(\boldsymbol{a}'_j; \boldsymbol{e}')$ *of* $\boldsymbol{P}^n$ *and* $\boldsymbol{Q}^n$, *respectively, there exists a collineation* $f : \boldsymbol{P}^n \to \boldsymbol{Q}^n$ *generated by a* σ*-linear map such that*

$$f(\boldsymbol{a}_j) = \boldsymbol{a}'_j, \quad f(\boldsymbol{e}) = \boldsymbol{e}'. \tag{35}$$

If K *is commutative,* f *is uniquely determined by* (35) *and* σ. *Conversely, if such a map is uniquely determined by a couple* $(\boldsymbol{a}_j; \boldsymbol{e})$, $(\boldsymbol{a}'_j; \boldsymbol{e}')$ *together with* $\sigma \in \operatorname{Aut} K$, *then* K *is a field.*

Proof. By Proposition 2.2, we find bases $(\mathfrak{a}_j)$ of $\boldsymbol{V}^{n+1}$, $(\mathfrak{a}'_j)$ of $\boldsymbol{W}^{n+1}$ such that $(\boldsymbol{a}_j; \boldsymbol{e})$ and $(\boldsymbol{a}'_j; \boldsymbol{e}')$ are the associated projective frames. Considering (20) as a definition for f, the equations $\sigma(0) = 0$ and $\sigma(1) = 1$ immediately imply (35). Now if g, h are collineations mapping the $(\boldsymbol{a}_j; \boldsymbol{e})$ into the $(\boldsymbol{a}'_j; \boldsymbol{e}')$, both induced by σ-linear maps, then $f = h^{-1} \circ g$ is an auto-collineation of $\boldsymbol{P}^n$ satisfying (24) and, moreover, it is generated by a linear map. The claim of the proposition is an immediate consequence of the following.

Lemma 15. *The scalar domain of $\boldsymbol{P}^n$, $n \in \mathbf{N}$, is a field if and only if there is only one collineation f generated by a linear map $a \in \boldsymbol{GL}(\boldsymbol{V}^{n+1})$ that satisfies* (24)*, namely $f = \mathrm{id}_{\boldsymbol{P}^n}$.*

Proof. As in the proof of Lemma 12, relation (24) implies the condition

$$a(\mathfrak{a}_j) = \mathfrak{a}_j \mu, \qquad \mu \in K^*, \tag{36}$$

for the linear map a we want to find. If K is a field, then each of these maps generates the identity collineation $f = \mathrm{id}_{\boldsymbol{P}^n}$. Consequently, the latter is uniquely determined by requiring (24) as well as linearity. If K is not commutative, there exist μ, $\alpha \in K^*$ with $\mu\alpha\mu^{-1} \neq \alpha$. Since $n \geq 1$, the associated vector space has at least dimension two. Consider the vector $\mathfrak{x} := \mathfrak{a}_0\alpha + \mathfrak{a}_1$ and the linear map a determined by (36) that generates f. Then

$$f([\mathfrak{x}]) = [a(\mathfrak{x})] = [\mathfrak{a}_0\mu\alpha + \mathfrak{a}_1\mu] = [\mathfrak{a}_0\mu\alpha\mu^{-1} + \mathfrak{a}_1] \neq [\mathfrak{a}_0\alpha + \mathfrak{a}_1] = [\mathfrak{x}],$$

since the vectors $\mathfrak{a}_0\alpha + \mathfrak{a}_1$, $\mathfrak{a}_0\mu\alpha\mu^{-1} + \mathfrak{a}_1$ are obviously linearly independent. Thus f is a linearly generated auto-collineation that is different from the identity and satisfies (24). □

Corollary 16. *Let $K = \mathbf{R}$ be the field of real (or rational) numbers. If $f : \boldsymbol{P}^n_o \to \boldsymbol{Q}^m_o$ is a collinear map with $\operatorname{Dim} f(\boldsymbol{P}^n_o) \geq 2$, then f is generated by a linear map between the associated vector spaces. The group $\operatorname{Aut} \boldsymbol{P}^n$, $n \geq 2$, acts simply transitively on the set of projective frames.*

Proof. It can easily be shown, cf. Exercise I.2.1.3, that the the fields of rational and real numbers each have only a single automorphism, the identity map. Hence, in these cases there is only one collineation satisfying (35) by Proposition 14.

Hint. To prove the statement concerning the automorphisms σ of $\mathbf{R}$, first show that for each of these automorphisms the restriction $\sigma|\mathbf{Q}$ to the field $\mathbf{Q}$ of rational numbers is the identity $\mathrm{id}_{\mathbf{Q}}$. Then, starting from $\sigma(\xi^2) = (\sigma(\xi))^2$ prove that σ is monotonous and derive from this its continuity. This, finally, implies $\sigma = \mathrm{id}_{\mathbf{R}}$. □

Exercise 2. Let $f : \boldsymbol{P}^n_o \to \boldsymbol{Q}^m_o$ be a map with the following properties: a) $f(\boldsymbol{o}) = \boldsymbol{o}$; b) $f(\boldsymbol{x} \vee \boldsymbol{y}) \subset f(\boldsymbol{x}) \vee f(\boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{P}^n$; c) $f(\boldsymbol{P}^n_o)$ is a projective subspace of dimension n. Prove that f is an injective collinear map.

Exercise 3. Let $\boldsymbol{P}^n_o$, $\boldsymbol{Q}^m_o$ be projective spaces over the field K, let $(\boldsymbol{a}_j; \boldsymbol{e})$ be a projective frame of $\boldsymbol{P}^n_o$, and let $(\boldsymbol{b}_j; \boldsymbol{b})$, $j = 0, \ldots, n$, be a sequence of points in $\boldsymbol{Q}^m$. Find examples to show that there does not always have to exist a collinear map $f : \boldsymbol{P}^n_o \to \boldsymbol{Q}^m_o$, for which $f(\boldsymbol{a}_j) = \boldsymbol{b}_j$, $j = 0, \ldots, n$, and $f(\boldsymbol{e}) = \boldsymbol{b}$. Prove, moreover, that, in general, these conditions together with the postulate $\sigma = \mathrm{id}_K$ for the automorphism of K associated with f do not determine f uniquely. (Compare this, however, with Proposition I.5.3.4 for affine maps!)

Exercise 4. Let $\boldsymbol{P}_o^n$, $\boldsymbol{Q}_o^m$, $K, (\boldsymbol{a}_j; \boldsymbol{e})$ be as in Exercise 3, and let, in addition, $(\boldsymbol{b}_0, \ldots, \boldsymbol{b}_r; \boldsymbol{b})$, $n \geq r \geq 2$, be a projectively dependent sequence of points from $\boldsymbol{Q}^m$, each $(r+1)$-element subsequence of which is projectively independent. Prove that for every automorphism $\sigma \in \operatorname{Aut} K$ there is precisely one collinear map $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^m$ with the properties: a) $f(\boldsymbol{a}_\rho) = \boldsymbol{b}_\rho$, $\rho = 0, \ldots, r$; b) $f(\boldsymbol{e}) = \boldsymbol{b}$; c) $f(\boldsymbol{a}_{r+1} \vee \ldots \vee \boldsymbol{a}_n) = \boldsymbol{o}$; d) f is generated by a σ-linear map $c : \boldsymbol{V}^{n+1} \to \boldsymbol{W}^{m+1}$ of the associated vector spaces.

Exercise 5. Let $\boldsymbol{V}^{n+1}$, $\boldsymbol{W}$ be right vector spaces over the skew fields K_1 and K_2, respectively, $n \geq 2$, and let $\mathfrak{P}^n = \mathfrak{P}(\boldsymbol{V})$, $\mathfrak{Q} = \mathfrak{P}(\boldsymbol{W})$ be the associated projective geometries (cf. Definition 1.1). A map $F : \mathfrak{P}^n \to \mathfrak{Q}$ is called *monotonous* if for $\boldsymbol{A}, \boldsymbol{B} \in \mathfrak{P}$, $\boldsymbol{A} \subset \boldsymbol{B}$ always implies the relation $F(\boldsymbol{A}) \subset F(\boldsymbol{B})$. Prove: If F is bijective, and F as well as F^{-1} are monotonous, then $f := F|\boldsymbol{P}_o^n$ is a collineation. Conversely, every collineation f generates an isomorphism of the lattices $F : [\mathfrak{P}^n, \subset] \to [\mathfrak{Q}, \subset]$. (In R. Baer [5], § III.1, the projective maps are actually defined as such isomorphisms of lattices.)

Exercise 6. In analogy to Corollary 8, describe all collinear maps $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^m$ with $\operatorname{Dim} f(\boldsymbol{P}_o^n) = 0$.

Exercise 7. Let $f \in \operatorname{Aut} \boldsymbol{P}^n$, $n \geq 2$, and let F be the automorphism of the lattice $[\mathfrak{P}^n, \subset]$ determined by f according to Exercise 5. Prove that f is the identity of $\boldsymbol{P}^n$, if for some k, $0 \leq k < n$, the restriction of F to $\boldsymbol{P}_{n,k}$ is the identity.

Exercise 8. *Coarse Classification of Collinear Maps.* Let $\boldsymbol{P}_o^n, \boldsymbol{Q}_o^m$, $n, m \geq 2$, be projective spaces over the skew field K. Define:

$$\begin{aligned} \mathcal{F} &:= \{f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^m | f \text{ collinear, } \operatorname{Dim} f(\boldsymbol{P}_o^n) \geq 2\}, \\ \boldsymbol{G} &:= \operatorname{Aut} \boldsymbol{Q}^m \times \operatorname{Aut} \boldsymbol{P}^n. \end{aligned}$$

Via $((g,h), f) \in \boldsymbol{G} \times \mathcal{F} \mapsto g \circ f \circ h^{-1} \in \mathcal{F}$ the group $\boldsymbol{G}$ acts on $\mathcal{F}$; denote by $\sim$ the corresponding equivalence relation. Prove that $f_1 \sim f_2$ if and only if $\operatorname{Dim} f_1(\boldsymbol{P}_o^n) = \operatorname{Dim} f_2(\boldsymbol{P}_o^n)$.

Exercise 9. Prove that the group $\operatorname{Aut} \boldsymbol{P}^n$, $n \geq 2$, acts transitively on the *Graßmann manifolds* $\boldsymbol{P}_{n,k}$ of k-planes in n-dimensional projective space. (To this end, consider the realization of k-planes as subsets of point space defined by the canonical map (1.2) as well as the action defined on them via the action of $\operatorname{Aut} \boldsymbol{P}^n$ on $\boldsymbol{P}^n$.) Determine the isotropy group of a k-plane. (The *isotropy group* or *stationary subgroup* of an element x in the transformed space is the set of all those elements g in the transforming group G leaving x fixed, i.e. $gx = x$, cf. I, § 1.4.)

1.4 Cross Ratio and Projective Maps

In affine geometry, the length ratio of two parallel segments or vectors is an important invariant closely related to the affine parameter on a line, cf. Exercise I.4.3.1. In projective geometry, an analogous procedure starting from the projective scales on a line leads to the notion of the cross ratio of four collinear points, that can be interpreted as the quotient of two length ratios, which explains its name[1]. In this section we will first deal with the geometry on the projective line over a skew field K. The cross ratio of four collinear points will be defined by means of a projective scale on the line; nevertheless, it does not depend on the choice of the scale. In general, however, this cross ratio is not invariant under collineations; it is transformed by the automorphism of K associated with the collineation. Requiring invariance of the cross ratio will lead to a special class of collinear maps, the projective maps. These form a category whose objects are the projective geometries (or spaces) over the skew field K, that, in general, is smaller than the category based on all collinear maps. For the field $K = \mathbf{R}$ of real numbers all collinear maps f with $\operatorname{Dim} f(\boldsymbol{P}^n_o) \geq 2$ are projective.

1.4.1 The Group Aut $\boldsymbol{P}^1$

In Example 3.3 we already pointed out that arbitrary bijections between lines are collinear. The Main Theorem of Projective Geometry does not hold for $n = 1$, and the group $\operatorname{Aut}\boldsymbol{P}^1$ is not defined, yet. In order to characterize the projective geometry on a line by a group we will proceed as follows: Consider the Graßmann manifold $\boldsymbol{P}_{n,1}$ of lines in an n-dimensional projective geometry, $n > 1$. The group of auto-collineations of $\boldsymbol{P}^n$ acts transitively over $\boldsymbol{P}_{n,1}$ (Exercise 3.9); denote by H the isotropy group of a fixed line $\boldsymbol{B} \subset \boldsymbol{P}^n$. The elements $h \in H$ leave the line $\boldsymbol{B}$ as a whole fixed, i.e., they transform the points of $\boldsymbol{B}$ among themselves. This way, an action of H over $\boldsymbol{B}$ is defined, that allows to develop a rich projective geometry on the line. To describe this action we choose a projective frame $(\boldsymbol{a}_i; \boldsymbol{e})$ of $\boldsymbol{P}^n$ for which $\boldsymbol{B} = \boldsymbol{a}_0 \vee \boldsymbol{a}_1$ is a coordinate axis. A semi-linear map a inducing the collineation $h \in H \subset \operatorname{Aut}\boldsymbol{P}^n$ then has to satisfy the conditions

$$a(\mathfrak{a}_0) = \mathfrak{a}_0 b^0_0 + \mathfrak{a}_1 b^1_0, \qquad a(\mathfrak{a}_1) = \mathfrak{a}_0 b^0_1 + \mathfrak{a}_1 b^1_1 \tag{1}$$

with respect to a basis $(\mathfrak{a}_i)$, $i = 0, \ldots, n$, determining this frame. Since h is a collineation, the 2×2-matrix (b^α_β) occurring here has to have rank two. Conversely, each semi-linear transformation of the vector space $\boldsymbol{V}^{n+1}$ associated with $\boldsymbol{P}^n$ having property (1) determines a collineation belonging to H. Since the remaining coefficients b^j_i, $i > 1$ or $j > 1$, do not affect the action of h on $\boldsymbol{B}$, we arrive at the same initial situation for the projective line as in the case of higher dimensional projective spaces: Formulas (3.27) to (3.33) retain their validity even for $n = 1$. They describe the action of the restrictions $h|\boldsymbol{B}$ to the line $\boldsymbol{B}$, where, of course, plenty of different elements $h \in H$ yield the same restriction. We obtain a homomorphism F from $\boldsymbol{G}(\boldsymbol{V}^2)$ to the group of all bijections on $\boldsymbol{P}^1 = \boldsymbol{B}$, which is defined as in Proposition 3.13. Setting

[1] Specialize Formula (15) below to the case $K = \mathbf{R}$.

$$\mathrm{Aut}\,\boldsymbol{P}^1 := \{f_a | a \in \boldsymbol{G}(\boldsymbol{V}^2)\} = F(\boldsymbol{G}(\boldsymbol{V}^2)), \tag{2}$$

Formulas (3.33), (3.34) also hold true for $n = 1$. We call $\mathrm{Aut}\,\boldsymbol{P}^1$ the *automorphism group of the projective line.* It is not difficult to see that, formally, the case $n = 0$ is included here; we have $\boldsymbol{G}(\boldsymbol{V}^1) = \boldsymbol{D}$, and $\mathrm{Aut}\,\boldsymbol{P}^0$ just consists of the unit element. *The reader may easily check that* Lemma 3.12 *and* Proposition 3.13 *remain valid for $n = 1$, including their proofs.*

1.4.2 The Cross Ratio

Now we consider the projective line $\boldsymbol{P}^1$ under the action of the group $\mathrm{Aut}\,\boldsymbol{P}^1$ and look for the invariants of this transformation group. By Proposition 3.14, applied to the case $\boldsymbol{Q}^1 = \boldsymbol{P}^1$, the group $\mathrm{Aut}\,\boldsymbol{P}^1$ acts transitively on the triples $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$ formed by pairwise different points on the line. Of course, these triples may always be viewed as the points in a projective frame. Thus, an invariant is first to be expected looking at quadruples $w = (\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d})$ of points on the line $\boldsymbol{P}^1$. The point of departure for the determination of such an invariant is a projective scale on the line, cf. Example 1.3. Let, for the time being, the points $\boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$ of the quadruple w be pairwise different. Then, by Proposition 2.2 we find a basis $(\mathfrak{a}_0, \mathfrak{a}_1)$ of the vector space $\boldsymbol{V}^2$ associated with $\boldsymbol{P}^1$ that is uniquely determined up to a common factor $\mu \in K^*$ and satisfies

$$[\mathfrak{a}_0] = \boldsymbol{c},\ [\mathfrak{a}_1] = \boldsymbol{d},\ [\mathfrak{a}_0 + \mathfrak{a}_1] = \boldsymbol{b}. \tag{3}$$

Let ξ denote the value assigned to $\boldsymbol{a}$ by the projective scale for this basis. If the scalar field K is not commutative, then the factor $\mu \in K^*$, occurring due to the arbitrariness in the choice of the basis, indeed plays a part. By Corollary 2.4, the different scale values, obtained by varying the basis satisfying (3), are related by inner automorphisms of K. Hence, it is not the scale value itself but its conjugacy class that is assigned to the quadruple in a way independent of the chosen basis. We denote this *conjugacy class* by

$$< \xi >:= \{\mu\xi\mu^{-1} | \mu \in K^*\},\ \text{if } \xi \in K,\ < \infty >:= \{\infty\}. \tag{4}$$

The last postulate only serves to include the case $\xi = \infty$ that actually occurs, for $\boldsymbol{a} = \boldsymbol{d}$. If K is a field, we can do without these considerations, since each conjugacy class only consists of the element itself. In general, we agree to set $< \xi >= \xi$, if $< \xi >$ has just a single element. In particular, for a skew field

$$< 0 >= 0,\ < 1 >= 1,\ < \infty >= \infty. \tag{5}$$

Now we summarize the observations above in the following definition, that now refers to the n-dimensional space, since, typically, we will have to deal with lines as embedded objects.

Definition 1. Let $\boldsymbol{P}^n$ be a projective space over the skew field K. A quadruple $w = (\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d})$ of points in $\boldsymbol{P}^n$ is called a *throw* if the points $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$ are collinear, and at least three of them are pairwise different. A throw is called *special* if its last three entries, $\boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$, are pairwise different. If w is a special throw, then its *cross ratio*, abbreviated as CR, is defined by

$$\mathrm{CR}(w) = (\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d}) := < \xi >,$$

where $\xi = \xi(\boldsymbol{a})$ is the scale value of $\boldsymbol{a}$ with respect to a basis for the vector space of the line $\boldsymbol{c} \vee \boldsymbol{d}$ satisfying (3). □

According to the above considerations $\mathrm{CR}(w)$ *does not depend on the choice of the basis satisfying* (3). In particular, because of (3) and (5) we have

$$(\boldsymbol{b}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d}) = 1,\ (\boldsymbol{c}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d}) = 0,\ (\boldsymbol{d}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d}) = \infty. \tag{6}$$

The CR does, however, depend on the order of the points in the throw. In order to remove the asymmetry in Definition 1, and to define the CR for an arbitrary throw we suppose that the four points are mutually different and investigate the way the CR changes under a permutation of the points in the throw. The case that two points of the throw are equal will be considered afterwards. Note that, by (6) this is the case if and only if the CR attains one of the values occurring in (5). This immediately follows from the properties of a projective scale. Since each permutation in the symmetry group S_4 can be represented as a product of the transpositions (1 2), (2 3), (3, 4) (cf. Exercise I.1.2.7), it suffices to consider the action of these transpositions. We will prove:

$$(\boldsymbol{b}, \boldsymbol{a}; \boldsymbol{c}, \boldsymbol{d}) = (\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{d}, \boldsymbol{c}) = (\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d})^{-1}. \tag{7}$$

In this as well as in similar formulas, the algebraic operations are to be applied to all elements in the conjugacy class. Obviously, it suffices to prove the claim for any representative $\xi \in (\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d})$. Using Definition 1 we conclude from (3) the following, where the second line of the formula refers to the computation of $(\boldsymbol{b}, \boldsymbol{a}; \boldsymbol{c}, \boldsymbol{d})$:

$$\begin{aligned} \mathfrak{a} &= \mathfrak{c} + \mathfrak{d}\xi, & \mathfrak{b} &= \mathfrak{c} + \mathfrak{d}, \\ \hat{\mathfrak{a}} &= \hat{\mathfrak{c}} + \hat{\mathfrak{d}}, & \hat{\mathfrak{b}} &= \hat{\mathfrak{c}} + \hat{\mathfrak{d}}\hat{\xi}. \end{aligned} \tag{8}$$

Here, the vectors each have to represent the points:

$$\boldsymbol{a} = [\mathfrak{a}],\ \boldsymbol{b} = [\mathfrak{b}],\ \boldsymbol{c} = [\mathfrak{c}],\ \boldsymbol{d} = [\mathfrak{d}], \tag{9}$$

$$\hat{\mathfrak{a}} = \mathfrak{a}\alpha,\ \hat{\mathfrak{b}} = \mathfrak{b}\beta,\ \hat{\mathfrak{c}} = \mathfrak{c}\gamma,\ \hat{\mathfrak{d}} = \mathfrak{d}\delta, \qquad (\alpha, \beta, \gamma, \delta \in K^*). \tag{10}$$

By (9), equations (8) and (10) hold, if we set

$$\hat{\mathfrak{a}} = \mathfrak{a},\ \hat{\mathfrak{b}} = \mathfrak{b},\ \hat{\mathfrak{c}} = \mathfrak{c},\ \hat{\mathfrak{d}} = \mathfrak{d}\xi.$$

The value of $\hat{\xi}$ has to be computed from

$$\mathfrak{b} = \mathfrak{c} + (\mathfrak{d}\xi)\hat{\xi} = \mathfrak{c} + \mathfrak{d},$$

leading to $\xi\hat{\xi} = 1$, i.e. claim (7) for the permutation of $\boldsymbol{a}$, $\boldsymbol{b}$. The pair $\boldsymbol{c}$, $\boldsymbol{d}$ is treated similarly. The next formula is proved along the same lines (cf. the proof of Proposition 1 below):

$$(\boldsymbol{a}, \boldsymbol{c}; \boldsymbol{b}, \boldsymbol{d}) = 1 - (\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d}). \tag{11}$$

So far, we considered throws consisting of four different points, hence, e.g. $\xi \neq 0$ and $\hat{\xi} = \xi^{-1}$ make sense. If now $\xi = 0$, then $\boldsymbol{a} = \boldsymbol{c}$, and, interchanging $\boldsymbol{c}$, $\boldsymbol{d}$, relation (6) implies

$$0 = (\boldsymbol{c}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d}) \mapsto (\boldsymbol{c}, \boldsymbol{b}; \boldsymbol{d}, \boldsymbol{c}) = \infty.$$

Formally setting $0^{-1} := \infty$ the equality $\hat{\xi} = \xi^{-1}$ remains valid also in this case. Considering in the same way the case $\boldsymbol{a} = \boldsymbol{d}$ we analogously see that $\infty^{-1} := 0$ again yields the correct result. If, finally, $\boldsymbol{a} = \boldsymbol{d}$, the exchange of $\boldsymbol{b}$ and $\boldsymbol{c}$ together with (6) lead to

$$\infty = (\boldsymbol{d}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d}) \mapsto (\boldsymbol{d}, \boldsymbol{c}; \boldsymbol{b}, \boldsymbol{d}) = \infty.$$

Setting $1 - \infty := \infty$ relation (11) also holds in this case. Summarizing we have the following.

Proposition 1. *Let $(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d})$ be a throw containing four different points of the n-dimensional projective space $\boldsymbol{P}^n$ over the skew field K. Then, their cross ratio satisfies the relations* (7) *and* (11). *Moreover,*

$$(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d}) = (\boldsymbol{c}, \boldsymbol{d}; \boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{b}, \boldsymbol{a}; \boldsymbol{d}, \boldsymbol{c}) = (\boldsymbol{d}, \boldsymbol{c}; \boldsymbol{b}, \boldsymbol{a}). \tag{12}$$

Relation (12) *provides a unique way to define the* CR *also for arbitrary throws. Setting*

$$0^{-1} := \infty, \ \infty^{-1} := 0, \ 1 - \infty := \infty, \tag{13}$$

rules (7), (11), (12) *remain true for arbitrary throws. If $\xi \in (\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d})$ is an arbitrary* CR, *then*

$$\xi, \ \xi^{-1}, \ 1 - \xi, \ 1 - \xi^{-1}, \ (1 - \xi)^{-1}, \ \xi(\xi - 1)^{-1} \tag{14}$$

represent all the possible cross ratios for permutations of the four points in the throw.

Proof. We first prove formula (11). Since $\boldsymbol{b}$ and $\boldsymbol{c}$ are interchanged, the second line in formula (8) now reads as

$$\hat{\mathfrak{a}} = \hat{\mathfrak{b}} + \hat{\mathfrak{d}}\hat{\xi}, \qquad \hat{\mathfrak{c}} = \hat{\mathfrak{b}} + \hat{\mathfrak{d}}.$$

Inserting (10) into these equations, using the first line of (8), and comparing coefficients in the representations with respect to the basis $(\mathfrak{c}, \mathfrak{d})$ of the vector space associated with the line yields

$$\alpha = \beta = \gamma = -\delta, \qquad \hat{\xi} = 1 - \xi,$$

and this is just claim (11). For a special throw with only three mutually different points, in all the three cases for $\boldsymbol{a} = \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$, it is straightforward to check (11) taking into account (13). Decomposing the permutation,

$$\begin{pmatrix} 1\ 2\ 3\ 4 \\ 3\ 4\ 1\ 2 \end{pmatrix} = (2\ 3)(3\ 4)(1\ 2)(2\ 3),$$

the first equation of formula (12) follows from (7) and (11). The two remaining equations immediately follow by applying (7) to both pairs of the CR in the resulting equation. If a throw contains only three mutually different points, then, by applying one of the permutations occurring in (12), we can always achieve a situation where the last three points are different from one another. The resulting CR is then correctly defined by extending equation (12) to this case, so that (7), (11), and (12) hold for the CR of an arbitrary throw. Since, according to (12), for each of the CRs corresponding to the 24 throws obtained by permuting four points, there are three others with the same value, altogether at most six different values are possible. These can be obtained as the expressions mentioned in (14) by applying (7) and (11) once or twice. Of course, in particular cases even more of these values may coincide; in the case $K = \mathbf{Z}_2$, e.g., there are only three values at all: $0, 1, \infty$. □

Exercise 1. Let $\boldsymbol{P}^1$ be a projective line over the skew field K, let $\xi : \boldsymbol{P}^1 \to \hat{K}$ be an arbitrary projective scale on $\boldsymbol{P}^1$, and let $\boldsymbol{b}_j \in \boldsymbol{P}^1$, $j = 1, \ldots, 4$, be four different points with scale values $\beta_j := \xi(\boldsymbol{b}_j) \neq \infty$. Prove the following formula for the CR:

$$(\boldsymbol{b}_1, \boldsymbol{b}_2; \boldsymbol{b}_3, \boldsymbol{b}_4) = <(\beta_1 - \beta_3)(\beta_1 - \beta_4)^{-1}[(\beta_2 - \beta_3)(\beta_2 - \beta_4)^{-1}]^{-1}> . \tag{15}$$

(Hint. Apply transformation formula (2.10) from Example 2.2.)

Exercise 2. Let K be a field, let $\boldsymbol{P}^1$ be a projective line over K, and let $(\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3, \boldsymbol{b}_4)$ be a throw in $\boldsymbol{P}^1$. Denote by (x_j, y_j) the homogeneous coordinates of the point $\boldsymbol{b}_j$ with respect to an arbitrary projective frame $(\boldsymbol{a}_0, \boldsymbol{a}_1; \boldsymbol{e})$ of $\boldsymbol{P}^1$, $j = 1, \ldots, 4$ (cf. Corollary 2.3). Prove the following expression for the CR in terms of determinants containing the homogeneous coordinates:

$$(\boldsymbol{b}_1, \boldsymbol{b}_2; \boldsymbol{b}_3, \boldsymbol{b}_4) = \frac{\begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix}}{\begin{vmatrix} x_1 & x_4 \\ y_1 & y_4 \end{vmatrix}} : \frac{\begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix}}{\begin{vmatrix} x_2 & x_4 \\ y_2 & y_4 \end{vmatrix}}. \tag{16}$$

Formulas (15) and (16) are the classical expressions for the CR; they explain the name "cross ratio". The behavior of the CR under collinear maps is described in the following.

Proposition 2. *Let $f : \boldsymbol{P}^n_o \to \boldsymbol{Q}^m_o$ be a collinear map induced by the σ-linear map $g : \boldsymbol{V}^{n+1} \to \boldsymbol{W}^{m+1}$ with $\sigma \in \operatorname{Aut} K$, and let $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d})$ a throw in $\boldsymbol{P}^n$. Then either*

$$\mathrm{Dim}(f(\boldsymbol{a}) \vee f(\boldsymbol{b}) \vee f(\boldsymbol{c}) \vee f(\boldsymbol{d})) < 1,$$

or $(f(\boldsymbol{a}), f(\boldsymbol{b}), f(\boldsymbol{c}), f(\boldsymbol{d}))$ *is a throw in* $\boldsymbol{Q}^m$ *with* CR

$$(f(\boldsymbol{a}), f(\boldsymbol{b}); f(\boldsymbol{c}), f(\boldsymbol{d})) = \sigma((\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d})). \tag{17}$$

In particular, the CR *is invariant under any collineation induced by a semi-linear map with inner automorphism* $\sigma = \sigma_\mu$.

Proof. Consider, say, the case $\boldsymbol{c} \neq \boldsymbol{d}$ and $\boldsymbol{H} = \boldsymbol{c} \vee \boldsymbol{d}$. If $f(\boldsymbol{c}) = f(\boldsymbol{d})$, then $\mathrm{Dim}\, f(\boldsymbol{H}) < 1$, and we are in the first situation. For $f(\boldsymbol{c}) \neq f(\boldsymbol{d})$ we know that $\hat{\boldsymbol{H}} := f(\boldsymbol{H}) = f(\boldsymbol{c}) \vee f(\boldsymbol{d})$ is a line, and, by Proposition 3.6, $f|\boldsymbol{H}$ is a collineation. Hence $(f(\boldsymbol{a}), f(\boldsymbol{b}), f(\boldsymbol{c}), f(\boldsymbol{d}))$ again is a throw. Let $\boldsymbol{U} \subset \boldsymbol{V}$, $\hat{\boldsymbol{U}} \subset \boldsymbol{W}$ be the two-dimensional subspaces associated with $\boldsymbol{H}$ and $\hat{\boldsymbol{H}}$, respectively. Then $g(\boldsymbol{U}) = \hat{\boldsymbol{U}}$. Because of (12) we can suppose that $\boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$ are mutually different, which then has to hold for their images as well. Choosing the basis $(\mathfrak{a}_0, \mathfrak{a}_1)$ for $\boldsymbol{U}$ as in Definition 1, we have, using the notation introduced there,

$$\mathfrak{a} = \mathfrak{a}_0 + \mathfrak{a}_1\xi \text{ und } g(\mathfrak{a}) = g(\mathfrak{a}_0) + g(\mathfrak{a}_1)\sigma(\xi),$$

since g is a σ-linear bijection. As the images of the base points correspond to the image of the throw, $\sigma(\xi)$ corresponds to the CR of this image, proving (17). The last assertion follows from the fact that every inner automorphism leaves each conjugacy class invariant. □

1.4.3 Projective Maps

The observations preceding the definition of the CR show that the CR is a fundamental notion for projective geometry, comparable to the ratio of segment lengths in affine geometry and to the distance in Euclidean geometry. Indeed, in the following section on affine geometry and in Sections 2.5, 2.6 on metric geometries we will see that both these notions can be based on the CR. Hence it should only be reasonable to expect a projective map to leave the CR invariant. Thus excluding the σ-linear maps with an automorphism σ that is not an inner one:

Definition 2. A map $f : \boldsymbol{P}^n_o \to \boldsymbol{Q}^m_o$ is called *projective* if it is collinear and leaves the CR invariant. More precisely, for every throw $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d})$ in $\boldsymbol{P}^n$, such that
$f(\boldsymbol{a}) \vee f(\boldsymbol{b}) \vee f(\boldsymbol{c}) \vee f(\boldsymbol{d})$ is again a line,

$$(f(\boldsymbol{a}), f(\boldsymbol{b}); f(\boldsymbol{c}), f(\boldsymbol{d})) = (\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d}). \tag{18}$$

□

Corollary 3. *The class of spaces over the skew field* K *together with the projective maps as morphisms forms a category.* □

Proposition 4. *Let K be a field, and let $\boldsymbol{P}^n$, $\boldsymbol{Q}^n$ be n-dimensional projective spaces over K, $n > 0$. A map $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^n$ is a projective bijection induced by a linear isomorphism of the associated vector spaces if and only if it has the following properties:*
1. *$f(\boldsymbol{o}) = \boldsymbol{o}$;*
2. *f maps each throw again into a throw;*
3. *relation (18) holds for every throw in $\boldsymbol{P}^n$, i.e., f leaves the CR invariant.*

Proof. 1. and 2. immediately imply that f is injective. To see this, consider a throw $(\boldsymbol{b}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d})$ in $\boldsymbol{P}^n$. Since the image of this throw also is a throw, the three image points have to be mutually different. For each point $\boldsymbol{b} \in \boldsymbol{c} \vee \boldsymbol{d}$, $\boldsymbol{c} \neq \boldsymbol{d}$, we have $f(\boldsymbol{b}) \in f(\boldsymbol{c}) \vee f(\boldsymbol{d})$ for the same reason. Hence $f(\boldsymbol{c} \vee \boldsymbol{d}) \subset f(\boldsymbol{c}) \vee f(\boldsymbol{d})$. Let now $\xi \in \hat{K}$ be arbitrary and take $\hat{\boldsymbol{a}}$ to be the point in $f(\boldsymbol{c}) \vee f(\boldsymbol{d})$ for which

$$\xi = (\hat{\boldsymbol{a}}, f(\boldsymbol{b}); f(\boldsymbol{c}), f(\boldsymbol{d})).$$

Since K is a field, and as the projective scale is bijective, $\hat{\boldsymbol{a}}$ is uniquely determined by $\xi \in \hat{K}$, and vice versa. The same is true for $\xi = (\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d})$ with $\boldsymbol{a} \in \boldsymbol{c} \vee \boldsymbol{d}$. Because of Properties 2 and 3, we thus have $f(\boldsymbol{a}) = \hat{\boldsymbol{a}}$, and hence $f(\boldsymbol{c} \vee \boldsymbol{d}) = f(\boldsymbol{c}) \vee f(\boldsymbol{d})$. So f is collinear, and by Proposition 3.6 it is a collineation. For $n = 1$ the fact just proved means

$$f([\mathfrak{a}_0 + \mathfrak{a}_1\xi]) = [\hat{\mathfrak{a}}_0 + \hat{\mathfrak{a}}_1\xi],$$

where the bases $(\mathfrak{a}_0, \mathfrak{a}_1)$, $(\hat{\mathfrak{a}}_0, \hat{\mathfrak{a}}_1)$ of the vector spaces corresponding to $\boldsymbol{P}^1$, $\boldsymbol{Q}^1$ are chosen to determine the CR as in Definition 1. Thus it is clear that f is induced by the linear map transforming $(\mathfrak{a}_0, \mathfrak{a}_1)$ into $(\hat{\mathfrak{a}}_0, \hat{\mathfrak{a}}_1)$. For $n \geq 2$, the Main Theorem implies that f is induced by a σ-linear bijection of the vector spaces. In the case of a field, the values of the CR belong to $\hat{K}$. As there are no proper inner isomorphisms, because of Proposition 2, (17), condition 3 implies $\sigma = \mathrm{id}_K$, and hence f has to be a linear isomorphism. Conversely, it is clear that each of the maps obtained this way has Properties 1–3. □

Propositions 4 and 3.14 immediately imply the following:

Corollary 5. *Let K be a field, and let $\boldsymbol{P}^n$, $\boldsymbol{Q}^n$ be projective spaces over K. Then for every pair of projective frames $(\boldsymbol{a}_j; \boldsymbol{e})$ of $\boldsymbol{P}^n$, $(\hat{\boldsymbol{a}}_j; \hat{\boldsymbol{e}})$ of $\boldsymbol{Q}^n$, $j = 1, \ldots, n$, there exists a unique projective bijection $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^n$ such that $f(\boldsymbol{a}_j) = \hat{\boldsymbol{a}}_j$, $f(\boldsymbol{e}) = \hat{\boldsymbol{e}}$.* □

Exercise 3. Let K be a field. Prove that assigning to each linear map $a : \boldsymbol{V}^{n+1} \to \boldsymbol{W}^{m+1}$ the projective map it induces defines a covariant functor from the category of linear maps between finite-dimensional vector spaces over K onto the category of projective maps between finite-dimensional projective spaces over K.

Exercise 4. Let K be a skew field, and let $\boldsymbol{P}^n$, $\boldsymbol{Q}^m$ be projective spaces over K. Consider throws $(\boldsymbol{b}_j)$, $(\hat{\boldsymbol{b}}_j)$, $j = 1, \ldots, 4$, in $\boldsymbol{P}^n$ and $\boldsymbol{Q}^m$, respectively. Prove that there is a projective map $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^m$ with $f(\boldsymbol{b}_j) = \hat{\boldsymbol{b}}_j$, $j = 1, \ldots, 4$ if and only if $(\boldsymbol{b}_1, \boldsymbol{b}_2; \boldsymbol{b}_3, \boldsymbol{b}_4) = (\hat{\boldsymbol{b}}_1, \hat{\boldsymbol{b}}_2; \hat{\boldsymbol{b}}_3, \hat{\boldsymbol{b}}_4)$.

Proposition 6. *Let K be a skew field satisfying the following condition:*
I. *Each automorphism $\sigma \in \operatorname{Aut} K$ such that $\sigma(<\xi>) = <\xi>$ for all conjugacy classes $<\xi> \subset K$ is an inner automorphism.*
Then for every projective space $\boldsymbol{P}^n$, $n \geq 2$, over K the group G of projective auto-collineations is isomorphic to the projective linear group $\boldsymbol{PGL}(n+1, K)$ (cf. Definition 2.1). This statement always holds, if K is a field or the skew field $\mathbf{H}$ of quaternions.

P r o o f. According to the Main Theorem, every auto-collineation $f \in \operatorname{Aut} \boldsymbol{P}^n$ is induced by a σ-linear bijection a of the associated vector space $\boldsymbol{V}^{n+1}$. Since f is projective, σ leaves each conjugacy class in K invariant, and so, by condition I, it is an inner automorphism of K. Choose a basis $(\mathfrak{a}_i)$ of $\boldsymbol{V}^{n+1}$, $i = 0, \ldots, n$. With respect to this basis the map a inducing f has the representation

$$a(\mathfrak{a}_i x^i) = \mathfrak{a}_j \alpha_i^j \mu x^i \mu^{-1},$$

where $\mu \in K^*$ determines the associated inner automorphism $\sigma = \sigma_\mu$. But then f is also induced by the linear automorphism of $\boldsymbol{V}^{n+1}$ corresponding to the matrix $(\alpha_i^j \mu)$:

$$f([\mathfrak{a}_i x^i]) = [\mathfrak{a}_j \alpha_i^j \mu x^i]. \tag{19}$$

Hence we already obtain all projective auto-collineations by restricting the homomorphism F from Proposition 3.13 to the subgroup $\boldsymbol{GL}(\boldsymbol{V}^{n+1}) \times \{\mathrm{id}_K\}$ of $\boldsymbol{G}_{n+1}(K)$; denote this restriction by $\hat{F}$. Then by (3.33) we have

$$\begin{aligned} \operatorname{Ker} \hat{F} &= (\operatorname{Ker} F) \cap (\boldsymbol{GL}(\boldsymbol{V}^{n+1}) \times \{\mathrm{id}_K\}) \\ &= \{(\mathrm{id}_{\boldsymbol{V}^{n+1}}\, \mu, \mathrm{id}_K) | \mu \in Z(K^*)\} \cong Z(K^*), \end{aligned} \tag{20}$$

where $Z(K^*)$ denotes the center of K^*. Since $\hat{F}$ is surjective, we deduce from the Homomorphism Theorem for groups (see Proposition I.3.1.6) the claim, cf. Definition 2.1:

$$G \cong \boldsymbol{GL}(n+1, K)/Z(n+1) = \boldsymbol{PGL}(n+1, K). \tag{21}$$

In the case of a field, Property I is satisfied for trivial reasons; the skew field $\mathbf{H}$ of *quaternions* also has it, cf. Example II.8.8.5. □

Nevertheless, there are skew fields not having Property I, cf. G. Köthe [68], R. Baer [5], § III.4. For these skew fields the following definition is geometrically not compelling; it provides, however, a unified foundation for the further elaboration of projective geometry:

Definition 3. Let $\boldsymbol{P}^n$ be the projective space associated with the vector space $\boldsymbol{V}^{n+1}$ over the skew field K. Then a *projectivity* or *projective transformation* of $\boldsymbol{P}^n$ (or from $\boldsymbol{P}^n$ to a projective space $\boldsymbol{Q}^n$ over K) is always understood to be a collineation that can be induced by a linear isomorphism. The group of projectivities from $\boldsymbol{P}^n$ to itself will be denoted by $\boldsymbol{PL}(\boldsymbol{P}^n)$ or, for short $\boldsymbol{PL}_n$; it is called the *projective group* of $\boldsymbol{P}^n$. □

With this definition we now have the general relation

$$\boldsymbol{PL}_n \cong \boldsymbol{GL}(\boldsymbol{V}^{n+1})/Z(\boldsymbol{GL}(\boldsymbol{V}^{n+1})) \cong \boldsymbol{PGL}(n+1, K). \tag{22}$$

This immediately implies: *The projective group $\boldsymbol{PL}_n$ acts transitively on the Graßmann manifolds $\boldsymbol{P}_{n,k}$ of projective k-planes in $\boldsymbol{P}^n$, hence in particular on the projective space $\boldsymbol{P}^n$*, since the linear group acts transitively on the Graßmann manifolds $G_{n+1,k+1}$ of $(k+1)$-dimensional subspaces in the $(n+1)$-dimensional vector space $\boldsymbol{V}^{n+1}$. To see the latter, consider two subspaces $\boldsymbol{U}, \boldsymbol{W} \in G_{n+1,k+1}$ and choose bases $\mathfrak{a}_0, \ldots, \mathfrak{a}_k$ for $\boldsymbol{U}$, $\mathfrak{b}_0, \ldots, \mathfrak{b}_k$ for $\boldsymbol{W}$, and complete both to obtain bases $(\mathfrak{a}_i), (\mathfrak{b}_i)$, $i = 0, \ldots n$, for $\boldsymbol{V}$. Then the linear transformation $g \in \boldsymbol{GL}(\boldsymbol{V}^{n+1})$ defined by $g(\mathfrak{a}_i) = \mathfrak{b}_i$, $i = 0, \ldots n$, satisfies $g\boldsymbol{U} = \boldsymbol{W}$.

Propositions 4 and 6 can be considered as geometric characterizations of the projective group. In accord with the *Erlanger Programm* [65] by F. Klein we want to consider projective geometry as the geometry of the transformation group $[\boldsymbol{PL}_n, \boldsymbol{P}^n]$. Since the group $\boldsymbol{PGL}(n+1, K)$ is the group of coordinate transformations of $\boldsymbol{P}^n$ over K (cf. Definition 2.1), we again obtain the correspondence between geometric and coordinate transformations already described in § I.5.7: The homogeneous coordinate systems of $\boldsymbol{P}^n$ form an atlas adapted to the transformation group $[\boldsymbol{PL}_n, \boldsymbol{P}^n]$ (cf. Definition I.5.7.5 and Exercise I.5.7.6). In the case of a field, the group $\boldsymbol{PL}(\boldsymbol{P}^n)$ is identical to the group of projective bijections of $\boldsymbol{P}^n$ onto itself. For a skew field with Property I this holds by Proposition 6 for $n \geq 2$; in the case $n = 1$ the group of projective bijections is larger, in general.

Example 1. The group $\boldsymbol{PL}_1$ is isomorphic to the *group G_K of fractional linear transformations* (2.10), cf. Exercise 2.4. This fact leads to *models of projective lines $\boldsymbol{P}^1(K)$* as transformation groups $[G_K, \hat{K}]$ that are simple and, in the cases $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$, easy to work with. □

Example 2. By Corollary 3.16, each collineation $f : \boldsymbol{P}^n_o \to \boldsymbol{Q}^n_o$, $n \geq 2$, over the field $K = \mathbf{R}$ of real numbers is a projectivity. In this case, we thus are in the very convenient situation that each bijective map transforming lines again into lines also leaves the CR, and hence all projective properties, invariant. □

Exercise 5. Prove that the only continuous automorphisms of the field $\mathbf{C}$ of complex numbers are the identity and the *conjugation* $\sigma : z \mapsto \bar{z}$. Hint. Prove first that every continuous automorphism τ of $\mathbf{C}$ has to satisfy $\tau(\mathbf{R}) = \mathbf{R}$, cf. Exercise I.2.3.3.

Example 3. If $K = \mathbf{C}$ is the field of complex numbers, then there are infinitely many automorphisms. This follows from Proposition V.6.1 in the algebra book by N. Bourbaki [19]. From the geometric point of view, it is natural to confine oneself to the continuous collinear maps $f : \boldsymbol{P}^n \to \boldsymbol{Q}^m$, $n \geq 2$; here, continuity refers to the topology introduced via inhomogeneous coordinates. If f is continuous, then its restriction to an arbitrary projective line also has to be continuous. Interpreting the CR occurring in (17) as projective scales, the arguments in the proof of Proposition 4 show that $\hat{\xi}(f(\boldsymbol{x})) = \sigma(\xi(\boldsymbol{x}))$, or $\sigma = \hat{\xi} \circ f \circ \xi^{-1}$. Hence σ is continuous. According to Exercise 5, σ is either the identity — then the collineation f is a projectivity — or σ is the conjugation. In this case, we call the collineation f an *anti-projectivity*, cf. also W. Burau [24], § III.20. In analogy to Proposition 4, anti-projectivities are characterized by the property

$$(f(\boldsymbol{a}), f(\boldsymbol{b}); f(\boldsymbol{c}), f(\boldsymbol{d})) = \overline{(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d})}.$$

In the sequel, we will only consider *continuous collinear maps* for complex scalars $K = \mathbf{C}$. This correspondingly also applies to the automorphism group $\operatorname{Aut} \boldsymbol{P}^1$ of the complex projective line. Using a projective scale $z = z(\boldsymbol{x}(\xi))$ (cf. Example 2.2) to describe the group, we see that the projectivities are represented by fractional linear transformations,

$$f(z) = \frac{az + b}{cz + d} \text{ with } ad - bc \neq 0,$$

and the anti-projectivities correspond to fractional conjugate-linear transformations,

$$f(z) = \frac{a\bar{z} + b}{c\bar{z} + d} \text{ with } ad - bc \neq 0.$$

□

Example 4. The sole non-commutative skew field that is important for us is the skew field $\mathbf{H}$ of *quaternions*[2] , cf. § I.2.3. Since by Example II.8.8.5 or Exercise II.8.9.5 every automorphism of $\mathbf{H}$ is an inner one, Propositions 2 and 6 together with the Main Theorem imply: *Each collineation $f : \boldsymbol{P}^n_o \to \boldsymbol{Q}^n_o$, $n \geq 2$, between quaternionic projective spaces can be induced linearly; in addition, it preserves the* CR. More generally, this holds for every collinear map f with $\operatorname{Dim} f(\boldsymbol{P}^n_o) \geq 2$, cf. Corollary 3.11. □

Example 5. Exercise II.7.4.7 immediately implies the rather obvious isomorphy of centers $Z(\boldsymbol{GL}(\boldsymbol{V}^{n+1})) \cong Z(K^*)$, by assigning to $\mu \in Z(K^*)$ the corresponding dilation $d_\mu \in \boldsymbol{GL}(\boldsymbol{V}^{n+1})$. Denoting as usual the *special linear group* by $\boldsymbol{SL}(m, K)$, i.e. the group of square matrices of order m with determinant one over the field K, then (22) and the computation of the kernel for the homomorphism

$$p : g \in \boldsymbol{GL}(n+1, \mathbf{R}) \longmapsto \frac{g}{(\det(g))^{\frac{1}{n+1}}} \in \boldsymbol{SL}(n+1, \mathbf{R}), \qquad n = 2k \in \mathbf{N}_0,$$

[2] An elementary introduction into the quaternions (and the octonions) can be found in I. L. Kantor, A. S. Solodownikow [58] and in many textbooks on algebra.

show the relation

$$\boldsymbol{PL}_n \cong \boldsymbol{SL}(n+1, \mathbf{R}), \qquad n = 2k \in \mathbf{N}_0.$$

Denote, moreover, as usual the unit matrix of order m by I_m. In addition, let $|\boldsymbol{SL}|(m, \mathbf{R})$ be the group of real square matrices of order m whose determinant has absolute value one. Then, for odd n, we analogously obtain

$$\boldsymbol{PL}_n \cong |\boldsymbol{SL}|(n+1, \mathbf{R})/\{\pm I_{n+1}\}, \qquad n = 2k+1 \in \mathbf{N}.$$

For $K = \mathbf{C}$ denote by K_m *the* group of m-th roots of unity, cf. Proposition I.2.3.7. Dividing $g \in \boldsymbol{GL}(\boldsymbol{V}^{n+1})$ by an $(n+1)$-th root of the norm $N(g)$ and multiplying the result by K_{n+1} we obtain

$$\boldsymbol{PL}_n \cong \boldsymbol{SL}(n+1, \mathbf{C})/K_{n+1} \qquad (K = \mathbf{C}).$$

Here, K_{n+1} is embedded into $\boldsymbol{SL}(n+1, \mathbf{C})$ as the set of dilatations d_ϵ, $\epsilon \in K_{n+1}$. □

1.4.4 Harmonic Position

In general, the six CRs corresponding to all permutations of the points in a throw are mutually different according to (14). As already noted, this does not hold in the case (13) that the throw contains two identical points. If the CR of a throw is $(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d}) = -1$, then the throw is called *harmonic*; this is also expressed by saying that the pairs $(\boldsymbol{a}, \boldsymbol{b})$, $(\boldsymbol{c}, \boldsymbol{d})$ are in *harmonic position.* In this case, permuting the points of the throw we only obtain the values $-1, 2, 1/2$ by (14). The following exercise describes a geometric construction for the *fourth point being in harmonic position with three given ones:*

Exercise 6. Let $\boldsymbol{P}^2$ be a projective plane over the skew field K. A *complete quadrangle* is understood to be any configuration consisting of four points, the *vertices* $\boldsymbol{a}_i \in \boldsymbol{P}^2$, $i = 0, \ldots, 3$, no three of which are collinear, together with their six connecting lines, also called the *sides* of the complete quadrangle, cf. Figure 1.8. The sides are arranged into three pairs of *opposite sides*,

$$\{\boldsymbol{a}_0 \vee \boldsymbol{a}_1,\, \boldsymbol{a}_2 \vee \boldsymbol{a}_3\},\ \{\boldsymbol{a}_0 \vee \boldsymbol{a}_2,\, \boldsymbol{a}_1 \vee \boldsymbol{a}_3\},\ \{\boldsymbol{a}_0 \vee \boldsymbol{a}_3,\, \boldsymbol{a}_1 \vee \boldsymbol{a}_2\}, \tag{23}$$

intersecting in the *diagonal points* of the complete quadrangle

$$\begin{aligned} \boldsymbol{d}_0 &:= (\boldsymbol{a}_0 \vee \boldsymbol{a}_3) \wedge (\boldsymbol{a}_1 \vee \boldsymbol{a}_2), \\ \boldsymbol{d}_1 &:= (\boldsymbol{a}_1 \vee \boldsymbol{a}_3) \wedge (\boldsymbol{a}_0 \vee \boldsymbol{a}_2), \\ \boldsymbol{d} = \boldsymbol{d}_2 &:= (\boldsymbol{a}_2 \vee \boldsymbol{a}_3) \wedge (\boldsymbol{a}_0 \vee \boldsymbol{a}_1). \end{aligned} \tag{24}$$

Prove: a) $\boldsymbol{d}_0$, $\boldsymbol{d}_1$, $\boldsymbol{d}_2$ are not collinear if and only if char $K \neq 2$. – b) The connecting lines $\boldsymbol{d}_i \vee \boldsymbol{d}_j$, $i \neq j$, of diagonal points do not pass through the vertices of the quadrangle. (Hint. Compute the homogeneous coordinates of $\boldsymbol{d}_i$ with respect to a basis for the projective frame $(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2; \boldsymbol{e})$ with $\boldsymbol{e} := \boldsymbol{a}_3$.) – c) Each side of the

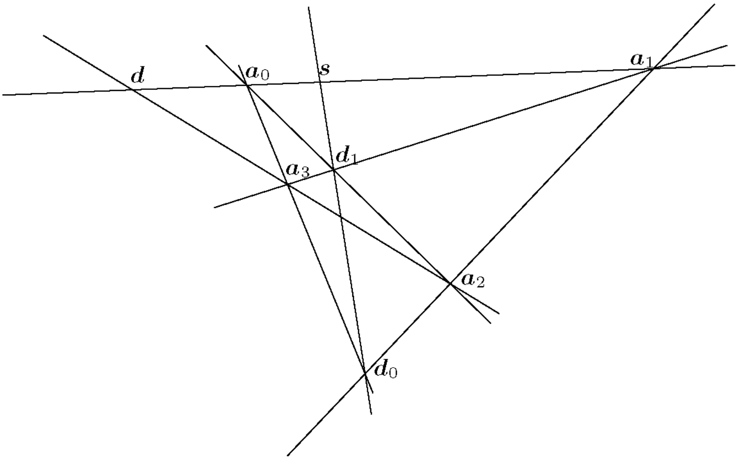


Fig. 1.8. Harmonic Position.

complete quadrangle contains precisely one diagonal point, e.g. $d := d_2 \in a_0 \vee a_1$. Denote by s the point of intersection of the side with the connecting line of the diagonal points not lying on it; e.g., for the side $a_0 \vee a_1$ this point of intersection is

$$s = (a_0 \vee a_1) \wedge (d_0 \vee d_1).$$

Prove that

$$(s, d; a_0, a_1) = -1. \tag{25}$$

d) Consider a complete quadrangle in the real affine plane, in which the sides $a_0 \vee a_1$, $a_2 \vee a_3$ are parallel and prove that s then cuts the segment a_0, a_1 into halves.

Remark. The requirement that there exists a complete quadrangle in P^2 whose diagonal points are not collinear, is known in synthetic projective geometry as the *Fano Postulate*, cf. R. Baer [5], Ch. II. It implies that the characteristic of the synthetically constructed scalar domain is different from two. If the scalar domain has characteristic two, then the relation $1 = -1$ together with (25) imply $s = d$; harmonic position thus is of no interest in this case.

Exercise 7. Let $\operatorname{char} K \neq 2$. Using the notations of Exercise 6 prove that the diagonal points (d_0, d_1) together with the intersection points p, q of the lines $d_0 \vee d_1$ with the sides of the complete quadrangle not passing through d_0, d_1,

$$p := (d_0 \vee d_1) \wedge (a_0 \vee a_1), \qquad q := (d_0 \vee d_1) \wedge (a_2 \vee a_3),$$

form a harmonic throw: $(p, q; d_0, d_1) = -1$.

Exercise 8. Permuting the points in a throw results in at most six different CRs by Proposition 1. Prove that the number of values ξ of the CR under the permutations of the points in a throw is less than six precisely in one of the following cases:

$$a)\ \xi = 1,\ 0,\ \infty; \tag{26}$$

$$b)\ \xi = -1,\ 2,\ \frac{1}{2}; \tag{27}$$

$$c)\ \xi \text{ is a root of the equation } x^2 - x + 1 = 0. \tag{28}$$

In the cases a) and b) the specified values are attained. For $\operatorname{char} K = 2$, a) and b) coincide. If $K = \mathbf{R}$, then only the cases a) and b) are possible. For $K = \mathbf{C}$, in the situation of case c), the throw is called *equi-anharmonic*; in this case, only the two values

$$\xi = (1 \pm \mathrm{i}\,\sqrt{3})/2 \tag{29}$$

can occur. Prove by means of Exercise II.8.9.5 that for the quaternions, $K = \mathbf{H}$, all solutions of (28) are conjugate; there are infinitely many of them. So the value $< \xi >$ of the CR just consists of a single conjugacy class.

Exercise 9. As in Example 1.3 we identify the complex projective line with the Riemann sphere: $\boldsymbol{P}^1(\mathbf{C}) = \hat{\mathbf{C}}$. Prove: If z_0, z_1, z_2 are the vertices of an equilateral triangle in the complex number plane $\mathbf{C}$, and m is its center, then (z_0, z_1, z_2, m) and (z_0, z_1, z_2, ∞) are equi-anharmonic throws.

1.4.5 Staudt's Main Theorem

For the projective geometry over a field K we are now going to prove the following result due to K. G. Chr. v. Staudt (1798–1867) that completes the Main Theorem of Projective Geometry and is therefore sometimes called *Staudt's Main Theorem*:

Proposition 7. *Let $\boldsymbol{P}^1$ be the projective line over a field K with characteristic* $\operatorname{char} K \neq 2$. *A surjective map $f : \boldsymbol{P}^1 \to \boldsymbol{P}^1$ is a projective automorphism if and only if it preserves harmonic position. Thus, for the real projective line f is a projectivity.*

Proof. Obviously, the condition is necessary. To prove the converse, we choose a projective scale ξ in $\boldsymbol{P}^1$ and identify the points $\boldsymbol{x} \in \boldsymbol{P}^1$ with the correspondingly labelled scalars $x = \xi(\boldsymbol{x}) \in \hat{K}$. Since any pair of different points on the line can always be completed to a harmonic throw by Exercise 6, and because of $\operatorname{char} K \neq 2$, the points of a harmonic throw are mutually different, the map f is bijective, and $f(0), f(1), f(\infty)$ are three different points. Let a denote the projectivity satisfying

$$a(f(0)) = 0,\ a(f(1)) = 1,\ a(f(\infty)) = \infty.$$

By Corollary 5 it is uniquely determined. Then, $g := a \circ f$ is a map from $\boldsymbol{P}^1$ to itself preserving harmonic position with fixed points $0 = g(0), 1 = g(1)$, $\infty = g(\infty)$. Now the proof is finished once we have shown that $g|K$ is an automorphism of the field K, since then $f = a^{-1} \circ g$ is a projective automorphism on $\boldsymbol{P}^1$. In the case $K = \mathbf{R}$ this has to be a projectivity according to Exercise I.2.1.3 (note also the hint following Corollary 3.16), because of $g = \mathrm{id}_{\mathbf{R}}$.

Relation (16) implies that for all $x, y \in K$, $x \neq y$:

$$(x, y;\ (x+y)/2, \infty) = (g(x), g(y);\ g((x+y)/2), \infty) = -1, \tag{30}$$

hence

$$g((x+y)/2) = (g(x) + g(y))/2, \qquad x \neq y. \tag{31}$$

Setting here $y = 0$ we conclude from $g(0) = 0$

$$g(x/2) = g(x)/2, \qquad x \in K, \tag{32}$$

and from $x = 2y$ that

$$g(2y) = 2g(y), \qquad y \in K. \tag{33}$$

(31) and (33) imply

$$g(x+y) = g(x) + g(y), \qquad x, y \in K. \tag{34}$$

Once again from (16) we obtain

$$(-z, z;\ 1, z^2) = (g(-z), g(z);\ 1, g(z^2)) = -1, \qquad (z \in K). \tag{35}$$

Since (34) immediately implies $g(-x) = -g(x)$, we deduce from (35) that

$$g(z^2) = g(z)^2, \qquad z \in K. \tag{36}$$

Inserting $z = x + y$ into this last equation and applying (33) and (34) yields

$$g(x \cdot y) = g(x) \cdot g(y), \qquad x, y \in K. \tag{37}$$

Since f and a are bijective, the same holds true for g; hence, by (34) and (37) the map g is an automorphism of the field K. □

For $K = \mathbf{R}$ the assertion can be proved without assuming the surjectivity of f. This is a consequence of two facts: The continuity of each endomorphism of $\mathbf{R}$, and the density of the rational numbers in $\mathbf{R}$.

R e m a r k. The proof of Proposition 7 presented here can also be found in E. Sperner [98], Bd. II, or W. Klingenberg [67]. A more general result for a skew field K with $\operatorname{char} K \neq 2$ contains E. Artin [4], Theorem 2.25, cf. also R. Baer [5], § III.4.

Exercise 10. Let $\boldsymbol{P}^1$ be a projective line over the field K, $\operatorname{char} K \neq 2$, and let, in addition, $f : \boldsymbol{P}^1 \to \boldsymbol{P}^1$ be a surjective map preserving harmonic position and leaving the point $\infty \in \hat{K} \equiv \boldsymbol{P}^1$ fixed. Here we identify $\hat{K}$ with $\boldsymbol{P}^1$ as in the proof of Proposition 7. Prove that there exist $\sigma \in \operatorname{Aut} K$, $\alpha \in K^*$ and $\beta \in K$ such that

$$f(x) = \alpha\sigma(x) + \beta \text{ for } x \in K \equiv \boldsymbol{P}^1 \setminus \{\infty\}.$$

Exercise 11. Let $\operatorname{Int} K$ be the group of inner automorphisms of the skew field K. Then $\operatorname{Int} K$ is a normal subgroup of the group $\operatorname{Aut} K$ of automorphisms of K. Prove that for every projective space $\boldsymbol{P}^n$, $n \geq 2$, over K the group $\boldsymbol{PL}_n$ is a normal subgroup in $\operatorname{Aut} \boldsymbol{P}^n$, and that the following quotients are isomorphic:

$$\operatorname{Aut} \boldsymbol{P}^n / \boldsymbol{PL}_n \cong \operatorname{Aut} K / \operatorname{Int} K.$$

1.4.6 Projective Equivalence of Collinear Maps. Involutions

We conclude this section by some considerations concerning the *classification of collinear maps.* They will be based on the following.

Definition 4. Let $\boldsymbol{P}_o^n$, $\boldsymbol{Q}_o^n$ be projective spaces over the skew field K. Two maps $f_1 : \boldsymbol{P}_o^n \to \boldsymbol{P}_0^n$, $f_2 : \boldsymbol{Q}_o^n \to \boldsymbol{Q}_0^n$ are called *projectively equivalent*, if there exists a projectivity $g : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^n$ such that $f_2 = g \circ f_1 \circ g^{-1}$. □

It is obvious that projectively equivalent collinear maps have to satisfy the condition

$$\operatorname{Dim} f_1(\boldsymbol{P}_o^n) = \operatorname{Dim} f_2(\boldsymbol{Q}_o^n);$$

if the f_j are induced by the σ_j-linear maps a_j, then the σ_j can only differ by an inner automorphism of K. Moreover, concerning the a_j we have: *There is a linear isomorphism $b : \boldsymbol{V}^{n+1} \to \boldsymbol{W}^{n+1}$ as well as a $\mu \in K^*$ such that* $d_\mu \circ a_2 = b \circ a_1 \circ b^{-1}$. Here d_μ denotes the dilation by the factor μ. Obviously, this condition is also sufficient for the projective equivalence of the collinear maps f_j induced by the a_j. This way, the classification problem for collinear maps is essentially reduced to the study of similarity for linear endomorphisms. This again is accomplished for an algebraically closed field K by means of the Jordan normal form, cf. § I.5.8 or § II.7.8.

Exercise 12. Let $f \in \boldsymbol{PL}_n$ be a projectivity of the projective space $\boldsymbol{P}^n$ over the field K. Denote by $P_f := \{\boldsymbol{x} \in \boldsymbol{P}^n | f(\boldsymbol{x}) = \boldsymbol{x}\}$ the set of *fixed points* of f. Prove that P_f can be represented as a union of $r \geq 0$ projective subspaces $\boldsymbol{A}_i \subset \boldsymbol{P}^n$, where

$$a)\ \boldsymbol{A}_i \wedge \bigvee_{j \neq i} \boldsymbol{A}_j = \boldsymbol{o},$$

$$b)\ \sum_{i=1}^{r} (\operatorname{Dim} \boldsymbol{A}_i + 1) \leq n + 1.$$

In the case of a *projective line*, any projectivity $f \neq \mathrm{id}_{\boldsymbol{P}^1}$ can thus have at most $r \leq 2$ fixed points; accordingly, the map f is called *elliptic, parabolic* or *hyperbolic* if $r = 0, 1$, or 2, respectively.

Exercise 13. Let $\boldsymbol{P}^1$ be a projective line over the field K, $\operatorname{char} K \neq 2$. Prove: a) If there are two points $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{P}^1$, $\boldsymbol{a} \neq \boldsymbol{b}$, that are transformed into each other by the projectivity f, $f(\boldsymbol{a}) = \boldsymbol{b}$ and $f(\boldsymbol{b}) = \boldsymbol{a}$, then f is *involutive*, i.e. $f^2 = \mathrm{id}_{\boldsymbol{P}^1}$. – b) If $f \in \boldsymbol{PL}_1$, $f \neq \mathrm{id}_{\boldsymbol{P}^1}$, is involutive, then f either has two different or no fixed point at all. If $\boldsymbol{a}, \boldsymbol{b}$ are two different fixed points of the *involution* f, then we have

$$(\boldsymbol{a}, \boldsymbol{b};\, \boldsymbol{x}, f(\boldsymbol{x})) = -1 \qquad (\boldsymbol{x} \neq \boldsymbol{a}, \boldsymbol{b}). \tag{38}$$

c) Let $\boldsymbol{a}, \boldsymbol{a}', \boldsymbol{b}, \boldsymbol{b}' \in \boldsymbol{P}^1$ be four different points. Prove that there is a unique involutive projectivity with $f(\boldsymbol{a}) = \boldsymbol{a}'$ and $f(\boldsymbol{b}) = \boldsymbol{b}'$.

Exercise 14. Again consider the situation of Exercise 13. Prove that every projectivity $f \in \boldsymbol{PL}_1$ can be represented as the product of two involutions.

Now we want to describe the *involutive collineations* $f \in \operatorname{Aut} \boldsymbol{P}^n$, $f^2 = \mathrm{id}_{\boldsymbol{P}^n}$.

Exercise 15. Let $\boldsymbol{P}^n, n \geq 2$, be a (right) projective space over the skew field K. Consider an involutive and fixed-point free $f \in \operatorname{Aut} \boldsymbol{P}^n$:

$$f^2 = \mathrm{id}_{\boldsymbol{P}^n}, \qquad f(\boldsymbol{x}) \neq \boldsymbol{x} \text{ for all } \boldsymbol{x} \in \boldsymbol{P}^n,$$

Prove that n is odd, and f can be induced by a σ-linear map a of the associated vector space $\boldsymbol{V}^{2k}$, $2k = n+1$, which has the following properties:
a) There is $\lambda \in K^*$ such that $a^2 = d_\lambda$; here the following relations hold

$$\sigma(\lambda) = \lambda \text{ and } \sigma^2 = \sigma_\lambda^{-1}, \tag{39}$$

where $\sigma_\lambda(\xi) = \lambda\xi\lambda^{-1}$ denotes the inner automorphism of K corresponding to λ.
b) There is a basis $(\mathfrak{a}_\kappa, \mathfrak{a}_{k+\kappa})_{\kappa=1,\dots,k}$ of $\boldsymbol{V}^{2k}$ such that, in this basis, a has the following representation:

$$a(\mathfrak{a}_\kappa) = \mathfrak{a}_{k+\kappa}, \quad a(\mathfrak{a}_{k+\kappa}) = \mathfrak{a}_\kappa\lambda. \tag{40}$$

c) There is no $\mu \in K^*$ with

$$\lambda = \mu\sigma(\mu). \tag{41}$$

Conversely, each collineation induced this way is involutive and fixed-point free. Classify all these collineations in the cases $K = \mathbf{R}$ and $K = \mathbf{C}$, σ the identity or the conjugation of $\mathbf{C}$. (Hint. Note that $\mathfrak{x} \neq \mathfrak{o}$ and $a(\mathfrak{x})$ are always linearly independent, since $f(\boldsymbol{x}) \neq \boldsymbol{x} = [\mathfrak{x}]$. Moreover, by $f^2 = \mathrm{id}_{\boldsymbol{P}^n}$ the subspace $\boldsymbol{W} \subset \boldsymbol{V}$ they span is invariant under a.)

The classification result, whose proof we outlined in Exercise 15, may be summarized as follows:

In the case $K = \mathbf{R}$, the fixed-point free involutions $f \in \operatorname{Aut} \boldsymbol{P}^n$ are projectively equivalent to the normal form (40) *with $\lambda = -1$, $\sigma = \mathrm{id}_{\mathbf{R}}$.* Thus, in homogeneous coordinates, f can be represented in the form

$$f([\sum_{\kappa=1}^{k}(\mathfrak{a}_\kappa x^\kappa + \mathfrak{a}_{k+\kappa}x^{k+\kappa})]) = [\sum_{\kappa=1}^{k}(-\mathfrak{a}_\kappa x^{k+\kappa} + \mathfrak{a}_{k+\kappa}x^\kappa)], \tag{42}$$

i.e. f is the simultaneous “rotation through $\pi/2$” in all the two-dimensional subspaces $\boldsymbol{W}_\kappa := \mathfrak{L}(\mathfrak{a}_\kappa, \mathfrak{a}_{k+\kappa})$, $n = 2k$.

In the case $K = \mathbf{C}$, the automorphism σ has to be conjugation, and up to projective equivalence all fixed-point free involutions are of the form

$$f([\sum_{\kappa=1}^{k}(\mathfrak{a}_\kappa x^\kappa + \mathfrak{a}_{k+\kappa}x^{k+\kappa})]) = [\sum_{\kappa=1}^{k}(-\mathfrak{a}_\kappa \bar{x}^{k+\kappa} + \mathfrak{a}_{k+\kappa}\bar{x}^\kappa)]. \tag{43}$$

For $K = \mathbf{C}$, $\sigma = \mathrm{id}_{\mathbf{C}}$, and in the case of the quaternions $K = \mathbf{H}$ there are no fixed-point free involutions at all. This is a consequence of the following.

Exercise 16. Prove: a) If K is an algebraically closed field, then each projectivity $f \in \boldsymbol{PL}_n$ has at least one fixed point. – b) Over the skew field of quaternions $\mathbf{H}$, each collineation $f \in \operatorname{Aut} \boldsymbol{P}^n$, $n \geq 2$, has at least one fixed point.

Exercise 17. Let K be a skew field, char $K \neq 2$, let $\boldsymbol{P}^n$ be a (right) projective space over K, $n \geq 2$, and let $f \in \operatorname{Aut} \boldsymbol{P}^n$ be an involutive collineation having at least one fixed point. Prove that under these conditions f even has at least $n+1$ fixed points in general position. More precisely, prove that there is a basis $(\mathfrak{a}_i)$, $i = 0, \ldots, n$, of the associated vector space $\boldsymbol{V}$ such that a σ-linear map a inducing f has diagonal form in $(\mathfrak{a}_i)$, and

$$a(\mathfrak{a}_i) = \mathfrak{a}_i \beta_i, \; \beta_i \sigma(\beta_i) = 1, \; a^2 = \mathrm{id}_{\boldsymbol{V}}, \; \sigma^2 = \mathrm{id}_K . \tag{44}$$

(Hint. $a^2 = d_\lambda$ and $a(\mathfrak{b}) = \mathfrak{b}\mu$ together imply $\mu\sigma(\mu) = \lambda$. Consider, in addition, the map $a' := d_\mu^{-1} \circ a$ also generating f, and prove that the maximal subspace of $\boldsymbol{V}$ spanned by the eigenvectors of a' is all of $\boldsymbol{V}$.)

Starting from (44) and specializing K as well as σ, it is often possible to obtain a complete classification of the involutions with fixed points. Below we will quote some of the related results, but leave the proofs to the reader.

Example 6. Under the assumptions of Exercise 17 let, in particular, K be a field and $\sigma = \mathrm{id}_K$. Then, with a suitable numbering of the basis elements, (44) implies that for each of these involutions $f \neq \mathrm{id}_{\boldsymbol{P}^n}$ there exists a number $k \in \mathbf{N_0}$, $0 \leq k < n/2$, such that, for an appropriate choice of the basis,

$$\begin{aligned} a(\mathfrak{a}_\alpha) &= -\mathfrak{a}_\alpha, \qquad \alpha = 0, \ldots, k, \\ a(\mathfrak{a}_\kappa) &= \mathfrak{a}_\kappa, \qquad \kappa = k+1, \ldots, n. \end{aligned} \tag{45}$$

Denoting by $\boldsymbol{a}_j := [\mathfrak{a}_j]$ the points corresponding to the basis elements, the k-plane $\boldsymbol{A} := \boldsymbol{a}_0 \vee \ldots \vee \boldsymbol{a}_k$ as well as its complementary $(n-k-1)$-plane $\boldsymbol{B} := \boldsymbol{a}_{k+1} \vee \ldots \vee \boldsymbol{a}_n$ are point-wise fixed; $\boldsymbol{A} \cup \boldsymbol{B}$ is the set of fixed points of f. The map f is called a *reflection* of $\boldsymbol{P}^n$ with respect to the pair of complementary subspaces $\boldsymbol{A}$, $\boldsymbol{B}$. Obviously, for every such pair of subspaces there is precisely one reflection, and two reflections of this kind are projectively equivalent if and only if the dimensions $k, n-k-1$ are the same. Choosing one of the spaces, say $\boldsymbol{B}$, as a subspace of the hyperplane at infinity $\boldsymbol{H}$ of an affine space $\boldsymbol{A}^n = \boldsymbol{P}^n \setminus \boldsymbol{H}$ results in an *affine reflection* in a k-plane $\boldsymbol{A}$ in the direction of the complementary space $\boldsymbol{B}$; compare this also with the results of the subsequent section. □

Example 7. In the case $K = \mathbf{C}$, apart from the cases already treated in Example 6, we now consider the *involutive anti-projectivities* J with the conjugation $\sigma = \tau : z \mapsto \bar{z}$ as inducing automorphism having, in addition, a fixed point. According to Exercise 17 then (44) holds, i.e. $\beta_j \bar{\beta}_j = |\beta_j|^2 = 1$, and hence $\beta_j = \mathrm{e}^{\mathrm{i}\,\varphi_j}$. Transforming the basis according to

$$\mathfrak{b}_j = \mathfrak{a}_j \, \mathrm{e}^{\mathrm{i}\,\varphi_j/2}, \qquad j = 0, \ldots, n,$$

we arrive at $\mathfrak{b}_j = a(\mathfrak{b}_j)$; hence all of these involutions are projectively equivalent. Moreover, in an adapted basis $(\mathfrak{b}_j)$ as just described, they all have the normal form

$$J([\sum_{j=0}^{n} \mathfrak{b}_j x^j]) = [\sum_{j=0}^{n} \mathfrak{b}_j \bar{x}^j]. \tag{46}$$

So the homogeneous coordinates (x^j) of the fixed points satisfy $x^j = \xi^j z$, where $\xi^j \in \mathbf{R}$ and $z \in \mathbf{C}^*$. Obviously, they form a real projective space $\boldsymbol{P}^n(\mathbf{R}) \subset \boldsymbol{P}^n(\mathbf{C})$, for which the points $\boldsymbol{b}_j = [\mathfrak{b}_j]$, $j = 0, \ldots, n$, and $\boldsymbol{e} = [\sum_j \mathfrak{b}_j]$ form a projective frame. This sequence of points is a projective frame for the complex space $\boldsymbol{P}^n(\mathbf{C})$ as well. For this reason, an involution J with normal form (46) is called a *real structure* on $\boldsymbol{P}^n(\mathbf{C})$, cf. Definition II.8.9.5; the set $\boldsymbol{P}^n(\mathbf{R}) \subset \boldsymbol{P}^n(\mathbf{C})$ of fixed points of J is called a *Staudt chain*. Consequently, all real structures, and hence all Staudt chains of a given dimension are projectively equivalent. □

Example 8. Let now $K = \mathbf{H}$ denote the skew field of quaternions. In order to classify the involutions of $\boldsymbol{P}^n(\mathbf{H})$ we could again start from (44). Instead, we prefer to make use of a fact already mentioned in Example 4: Each collineation can be induced by a linear map. According to Exercises 16 and 17 the involution f has at least $n + 1$ fixed points in general position. Thus we find a basis $(\mathfrak{a}_l)$ of the associated vector space $\boldsymbol{V}$ as well as a linear map a inducing f that is in diagonal form with respect to this basis:

$$a(\mathfrak{a}_l) = \mathfrak{a}_l q_l, \qquad q_l \in \mathbf{H}. \tag{47}$$

According to (22), involutivity means that $a^2 = d_\lambda$ has to be a dilation by $\lambda \in Z(\mathbf{H}) = \mathbf{R}$. This again implies:

$$q_l^2 = \lambda \in \mathbf{R}^*, \qquad l = 0, \ldots, n. \tag{48}$$

Turning to the equivalent representation of f by

$$a' := d_\mu \circ a, \qquad \mu := |\lambda|^{-\frac{1}{2}} \tag{49}$$

we can assume without loss of generality that $\lambda = \pm 1$. Recalling how the multiplication is defined in $\mathbf{H}$, cf. II.(8.9.4), it is easy to see that in $\mathbf{H}$ the equation $q^2 = 1$ only has the solutions $q = \pm 1$. So in this case we again arrive at the reflections (45) in a pair of complementary subspaces (Example 6). Now consider $\lambda = -1$. Using II.(8.9.4) we see that the set of solutions of $q^2 = -1$ coincides with the unit sphere in the space of imaginary quaternions $S^2 \subset \mathbf{R}^\perp \subset \mathbf{H}$; here $\mathbf{R}^\perp$ denotes the real subspace of $\mathbf{H}$ spanned by the imaginary units i, j, $\mathrm{k} \in \mathbf{H}$. On this sphere S^2 the group of inner automorphisms of $\mathbf{H}$, which is isomorphic to the special orthogonal group $\boldsymbol{SO}(3)$, acts transitively, cf. Exercise II.8.9.5. Thus, for each of the q_l in (47) we find a $p_l \in \mathbf{H}^*$ such that

$$p_l^{-1} q_l p_l = \mathrm{i}\,. \tag{50}$$

Looking at the basis $\mathfrak{b}_l := \mathfrak{a}_l p_l$, $l = 0, \ldots, n$, we immediately see that

$$a(\mathfrak{b}_l) = \mathfrak{a}_l q_l p_l = \mathfrak{b}_l \,\mathrm{i}\,. \tag{51}$$

Hence any involution of $\boldsymbol{P}^n(\mathbf{H})$, $n \geq 2$, is either a reflection or a projectivity J, which by (51) has the normal form

$$J([\sum_l \mathfrak{b}_l x^l]) = [\sum_l \mathfrak{b}_l \, \mathrm{i} \, x^l]. \tag{52}$$

Note that, according to Exercises 15–17 and the subsequent examples for the cases $K = \mathbf{R}, \mathbf{H}$, the involutive collineations, and for $K = \mathbf{C}$ the continuous involutive collineations are completely classified. □

Exercise 18. Find an example showing that the assertion in Exercise 17 is wrong without the assumption $\operatorname{char} K \neq 2$.

Exercise 19. Let $\boldsymbol{A}, \boldsymbol{B}$ be complementary projective subspaces in $\boldsymbol{P}^n$, both different from $\boldsymbol{o}$, and consider $\boldsymbol{x}_1, \boldsymbol{y}_1 \in \boldsymbol{P}^n \backslash (\boldsymbol{A} \cup \boldsymbol{B})$, $(\boldsymbol{x}_1 \vee \boldsymbol{y}_1) \wedge \boldsymbol{A} = \boldsymbol{a}_1$, $(\boldsymbol{x}_1 \vee \boldsymbol{y}_1) \wedge \boldsymbol{B} = \boldsymbol{b}_1$. a) Prove that there is a unique projectivity $f \in \boldsymbol{PL}_n$ with the properties $f|\boldsymbol{A} = \mathrm{id}_{\boldsymbol{A}}$, $f|\boldsymbol{B} = \mathrm{id}_{\boldsymbol{B}}$, and $f(\boldsymbol{x}_1) = \boldsymbol{y}_1$. – b) For $n \geq 2$, find a geometric way to construct $f(\boldsymbol{x})$ for any point $\boldsymbol{x} \in \boldsymbol{P}^n \setminus (\boldsymbol{A} \cup \boldsymbol{B})$. (Hint. Recall Exercise 2.1.) For $n = 2$, such maps are also called *central collineations*, cf. J. Böhm et al [18], Bd. II. – c) Prove that f is involutive if and only if $(\boldsymbol{x}_1, \boldsymbol{y}_1, \boldsymbol{a}_1, \boldsymbol{b}_1)$ is a harmonic throw.

1.5 Affine Geometry from the Projective Viewpoint

In Section 1 we started from the affine plane and obtained the projective plane by adjoining a "line at infinity". This idea motivated the definition of the projective geometry $\mathfrak{P}^n$ as the lattice of subspaces of an $(n+1)$-dimensional vector space $\boldsymbol{V}^{n+1}$. This way we avoided to pursue the rather awkward process of extending the n-dimensional affine geometry towards the projective one. Now we want to proceed the other way round: Affine geometry is shown to arise from the projective one by distinguishing a projective hyperplane $\boldsymbol{H}$ in $\boldsymbol{P}^n$. All affine notions are derived from the projective ones by relating the latter to this distinguished hyperplane. For simplicity, we will use the symbols introduced in the course of the axiomatic approach to affine geometry in Chapters I.4, I.5 to denote the corresponding projective notions as well. This should not lead to any confusion. Note also that this time we will be more general than there, since now the scalar domain may be a skew field.

Definition 1. Let $\boldsymbol{P}^n$ denote the projective space associated with the $(n+1)$-dimensional right vector space $\boldsymbol{V}^{n+1}$ over the skew field K, and let $\boldsymbol{H} \subset \boldsymbol{P}^n$ be a fixed hyperplane in $\boldsymbol{P}^n$. Then for $n \geq 1$ its complement $\boldsymbol{A}^n := \boldsymbol{P}^n \setminus \boldsymbol{H}$ is called an *n-dimensional affine space* over K. For $n = 0, -1$ we set $\boldsymbol{A}^0 := \boldsymbol{P}^0$, $\boldsymbol{A}^{-1} := \emptyset$. In the case $n \geq 1$, $\boldsymbol{H}$ is called the *hyperplane at infinity* of the affine space $\boldsymbol{A}^n$; it is also called the *absolute hyperplane*, the *ideal hyperplane*, or the *improper hyperplane*. □

Proposition 3.14 immediately implies that the projective linear group acts transitively on the set of hyperplanes: Just assign a frame to each hyperplane in a suitable way and consider the case $\sigma = \mathrm{id}_K$. Hence it does not matter which particular hyperplane is distinguished. To say it lies "at infinity" is suggested by the geometric interpretation described in Section 1. The fundamental property special to affine and not belonging to projective geometry is parallelism:

Definition 2. In the notation of Definition 1, let $\boldsymbol{B}^k$ be a projective k-plane, $k \geq 0$. If $\boldsymbol{B}^k \subset \boldsymbol{H}$, then $\boldsymbol{B}^k$ is called *improper*, and otherwise *proper*. For every proper k-plane we have $\boldsymbol{B}^k \cap \boldsymbol{A}^n \neq \emptyset$; so $\boldsymbol{M}^k := \boldsymbol{B}^k \cap \boldsymbol{A}^n$ is called *the affine k-plane corresponding to* $\boldsymbol{B}^k$. Let $\boldsymbol{B}^k$, $\boldsymbol{C}^l$ be proper projective subspaces of $\boldsymbol{P}^n$, whose dimensions satisfy $1 \leq l \leq k$. Denote, moreover, by $\boldsymbol{D}^l := \boldsymbol{C}^l \cap \boldsymbol{A}^n$ the affine l-plane corresponding to $\boldsymbol{C}^l$. $\boldsymbol{D}^l$ is said to be *parallel* to $\boldsymbol{M}^k$ (or $\boldsymbol{C}^l$ *parallel* to $\boldsymbol{B}^k$, respectively) if

$$\boldsymbol{H} \cap \boldsymbol{C}^l \subset \boldsymbol{H} \cap \boldsymbol{B}^k. \tag{1}$$

Starting, conversely, from an affine k-plane $\boldsymbol{M}^k$, the proper projective k-plane $\boldsymbol{B}^k$ to be defined is uniquely determined; it is called the *projective closure* of $\boldsymbol{M}^k$ and will be denoted by $\bar{\boldsymbol{M}}^k$. It is thus the projective subspace for which $\boldsymbol{M}^k = \bar{\boldsymbol{M}}^k \cap \boldsymbol{A}^n$. □

The dimension formula (Proposition 1.1) yields the following.

Corollary 1. *For every proper projective plane* $\boldsymbol{B}^k$

$$\operatorname{Dim} \boldsymbol{H} \wedge \boldsymbol{B}^k = k - 1. \tag{2}$$

In particular, a proper k-plane $\boldsymbol{B}^k$, $k \geq 1$, is parallel to a proper hyperplane $\boldsymbol{C}^{n-1}$ if and only if $\boldsymbol{B}^k \subset \boldsymbol{C}^{n-1}$ or $\boldsymbol{B}^k \wedge \boldsymbol{C}^{n-1} \subset \boldsymbol{H}$. □

Definition 2 is in accordance with the idea that two different lines in the plane (or, more generally, hyperplanes in $\boldsymbol{P}^n$) are parallel if they intersect "at infinity". It can be shown (cf. Exercise 4 c)), that the k-dimensional vector space associated with an affine k-plane $\boldsymbol{M}^k = \boldsymbol{B}^k \cap \boldsymbol{A}^n$ (cf. Proposition I.4.5.1) is precisely the vector space $\boldsymbol{U} \cap \boldsymbol{W} \subset \boldsymbol{V}^{n+1}$; here $\boldsymbol{U}$ and $\boldsymbol{W}$ denote the vector subspaces determining $\boldsymbol{B}^k$ and $\boldsymbol{H}$, respectively. Hence Definition 2 exactly corresponds to the Definition I.4.5.3 of parallelism in affine geometry. Since parallelism is to be an affinely invariant property, we can only permit those projectivities as *affine transformations* that map $\boldsymbol{H}$ (and hence also $\boldsymbol{A}^n$) into itself: The *affine group* $\mathfrak{A}(n) = \mathfrak{A}(\boldsymbol{A}^n)$ is thus defined to be the stationary subgroup of $\boldsymbol{H}$:

$$\mathfrak{A}(n) := \{f \in \boldsymbol{PL}_n | f(\boldsymbol{H}) = \boldsymbol{H}\}. \tag{3}$$

In order to obtain the coordinate representation of affine transformations in its usual form (cf. I.(5.7.46)) we first have to grasp the notion of an affine coordinate system from the projective point of view. A basis $(\mathfrak{a}_i)$ of the vector space $\boldsymbol{V}^{n+1}$ or a projective frame $(\boldsymbol{a}_i; \boldsymbol{e})$ will be called *affinely adapted* if

$$\boldsymbol{H} = \boldsymbol{a}_1 \vee \ldots \vee \boldsymbol{a}_n \qquad (\boldsymbol{a}_i = [\mathfrak{a}_i]), \tag{4}$$

i.e., if the points $\boldsymbol{a}_i$, $i = 1, \ldots, n$, are improper. The only proper base point $\boldsymbol{a}_0 = [\mathfrak{a}_0]$ is called the *origin* of the *affine coordinate system.* The affine space $\boldsymbol{A}^n$ is then just the domain of definition U_0 for the zero chart φ_0 that is defined by $x^0 \neq 0$ and determined by the inhomogeneous coordinate system belonging to the homogeneous coordinate system for the affinely adapted basis (cf. Lemma 2.6). Hence we will consider the inhomogeneous coordinates defined by (2.8) for $i = 0$ as the *affine coordinates* of a point $\boldsymbol{x} \in \boldsymbol{A}^n$ with respect to the affinely adapted basis $\mathfrak{a}_0, \ldots, \mathfrak{a}_n$: *For each $\boldsymbol{x} \in \boldsymbol{A}^n$ there is precisely one n-tuple $(x^j) \in K^n$, $j = 1, \ldots, n$, such that*

$$\varphi_0(\boldsymbol{x}) = (x^j) \in K^n, \tag{5}$$

$$\textit{hence} \;\; \boldsymbol{x} = [\mathfrak{a}_0 + \sum_{j=1}^{n} \mathfrak{a}_j x^j]. \tag{6}$$

In particular, the unit points $[\mathfrak{a}_0 + \mathfrak{a}_j]$ on the axes $\boldsymbol{a}_0 + [\mathfrak{a}_j]$ coincide with those of the affine geometry, cf. Fig. 1.7 (for $n = 2$). Now let $(\mathfrak{a}_i)$, $i = 0, \ldots, n$, be a fixed affinely adapted frame in $\boldsymbol{P}^n$, and let $f \in \boldsymbol{PL}_n$; take a linear map $b_f \in \boldsymbol{GL}(\boldsymbol{V}^{n+1})$ inducing f. Obviously,

$$f \in \mathfrak{A}(n) \Longleftrightarrow b_f(\mathfrak{a}_j) \in \mathfrak{L}(\mathfrak{a}_1, \ldots, \mathfrak{a}_n), \quad j = 1, \ldots, n. \tag{7}$$

Denote by $\boldsymbol{LA}(n+1, K) \subset \boldsymbol{GL}(n+1, K)$ the set of block matrices of the form

$$\begin{pmatrix} \beta_{00} & \mathfrak{o}' \\ \mathfrak{b} & B \end{pmatrix}, \qquad \beta_{00} \in K^*,\ \mathfrak{b} \in K^n,\ B \in \boldsymbol{GL}(n, K). \tag{8}$$

Then (cf. (4.22))

$$\mathfrak{A}(n) \cong \boldsymbol{LA}(n+1, K)/Z(n+1). \tag{9}$$

Here $Z(n+1)$ denotes the center of $\boldsymbol{GL}(n+1, K)$, i.e. the group of diagonal matrices $I_{n+1}\mu$, μ from the center $Z(K^*)$ of the multiplicative group of K. If K is a field, then we normalize the matrices via multiplication by β_{00}^{-1} to have the form

$$(\mathfrak{b}, B) := \begin{pmatrix} 1 & \mathfrak{o}' \\ \mathfrak{b} & B \end{pmatrix}, \qquad \mathfrak{b} \in K^n,\ B \in \boldsymbol{GL}(n, K). \tag{10}$$

This immediately implies that the recent definition is compatible with the earlier Definition I.5.1.2 of the affine group; since (10) is the coordinate matrix of the affine transformation f, using the normalization $x^0 = 1$, $y^0 = 1$ we obtain for the coordinates of the image point $\boldsymbol{y} = f(\boldsymbol{x})$:

$$\begin{pmatrix} 1 \\ y^j \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \beta^j & \beta^j_i \end{pmatrix} \begin{pmatrix} 1 \\ x^j \end{pmatrix}, \tag{11}$$

thus

$$y^j = \beta^j + \sum_{i=1}^{n} \beta_i^j x^i, \qquad j = 1, \ldots, n. \tag{12}$$

Up to notation, this is the coordinate representation I.(5.7.42) for the affine map f.

Exercise 1. Let K be a field. Taking into account definition (10), prove the multiplication rule in $\mathfrak{A}(n, K)$:

$$(\mathfrak{b}, B) \cdot (\mathfrak{c}, C) = (\mathfrak{b} + B\mathfrak{c}, B \cdot C). \tag{13}$$

Calculate the expression $(\mathfrak{x}, X) = (\mathfrak{b}, B)^{-1}$, and show that $\mathfrak{A}(n, K)$ is the semi-direct product of the normal subgroup

$$\{(\mathfrak{b}, I) | \mathfrak{b} \in K^n\} \cong [K^n, +], \qquad I = (\delta_i^j), \tag{14}$$

with the subgroup

$$\{(\mathfrak{o}, B) | B \in \boldsymbol{GL}(n, K)\} \cong \boldsymbol{GL}(n, K) \tag{15}$$

(cf. § I.5.3).

Exercise 2. Consider the affine space $\boldsymbol{A}^n$ over the field K. a) Let $\boldsymbol{g}^1$ be a proper line, $\boldsymbol{c}_\infty := \boldsymbol{g}^1 \wedge \boldsymbol{H}$, and denote by t the affine parameter on the affine line $\boldsymbol{g}^1 \cap \boldsymbol{A}^n$ with zero point $\boldsymbol{c}_0$ and unit point $\boldsymbol{c}_1 \neq \boldsymbol{c}_0$. Prove the relation

$$\boldsymbol{x} = \boldsymbol{c}_0 + \overrightarrow{\boldsymbol{c}_0\boldsymbol{c}_1}\, t \Longleftrightarrow t = (\boldsymbol{x}, \boldsymbol{c}_1;\, \boldsymbol{c}_0, \boldsymbol{c}_\infty).$$

Thus the affine scales on $\boldsymbol{g}^1$ are just the projective scales assigning the value ∞ to the improper point $\boldsymbol{g}^1 \wedge \boldsymbol{H}$. Moreover, the CR t is equal to the ratio of the parallel vectors $\overrightarrow{\boldsymbol{c}_0\boldsymbol{x}}$ and $\overrightarrow{\boldsymbol{c}_0\boldsymbol{c}_1}$. – b) Let $\operatorname{char} K \neq 2$. Prove for three collinear points $\boldsymbol{c}_0, \boldsymbol{c}_1, \boldsymbol{x} \in \boldsymbol{g}^1 \cap \boldsymbol{A}^n$: $\boldsymbol{x}$ bisects the segment $(\boldsymbol{c}_0, \boldsymbol{c}_1)$ (cf. I.(4.3.30)) if and only if $(\boldsymbol{c}_0, \boldsymbol{c}_1;\, \boldsymbol{c}_\infty, \boldsymbol{x}) = -1$.

Exercise 3. Let again K be a field. Prove: a) The group of *homotheties*[1] of $\boldsymbol{A}^n$ is isomorphic to the group of all those projectivities $f \in \boldsymbol{PL}_n$ for which $f|\boldsymbol{H} = \mathrm{id}_{\boldsymbol{H}}$. – b) Each homothety $g \neq \mathrm{id}_{\boldsymbol{A}^n}$ has at most one fixed point in $\boldsymbol{A}^n$. – c) The homotheties of $\boldsymbol{A}^n$ having no fixed points in $\boldsymbol{A}^n$ together with $\mathrm{id}_{\boldsymbol{A}^n}$ form the group $\boldsymbol{T}(\boldsymbol{A}^n)$ of *translations*, which again is isomorphic to the group (14).

Formulas (11), (12) show that the transformation group $[\mathfrak{A}(\boldsymbol{A}^n), \boldsymbol{A}^n]$ has the same coordinate transformation rules both in the sense of Corollary I.5.1.3 and in the projective perspective; choosing a cartesian as well as an affinely adapted coordinate system yields an equivariant isomorphism onto the transformation group $[\mathfrak{A}(n, K), K^n]$ defined in the matrix algebra over K with the action (11). Therefore, in the sense of the *Erlanger Programm* by F. KLEIN (s. § I.6.5), up to isomorphy, both methods determine one and the same affine geometry. Exercise 3 contains the projective description of translations, which

[1] *Homotheties* are the affine maps for which the image of each affine subspace is parallel to the subspace itself, cf. § I.4.3.

form an n-dimensional vector space over K. If $f \neq \mathrm{id}_{\boldsymbol{A}^n}$ is a translation, then all lines $\boldsymbol{x} \vee f(\boldsymbol{x})$ pass through the same point of $\boldsymbol{H}$, and we have

$$(\boldsymbol{x} \vee f(\boldsymbol{x})) \wedge \boldsymbol{H} = [\mathfrak{b}], \tag{16}$$

if $\boldsymbol{y} = \boldsymbol{x} + \mathfrak{b}$ is the affine representation of f (cf. I. (4.5.3)). According to (16), translations by proportional vectors $\mathfrak{b}, \mathfrak{b}\lambda$, $\lambda \in K^*$, determine the same improper point. Hence the points of the improper hyperplane can be interpreted as *directions* in the affine geometry, i.e. equivalence classes of parallel lines. Viewing translations as vectors it is easy to verify the axioms (I), (II), (III) of affine geometry (§ I.4.3). This completes the projective incorporation of affine geometry.

Exercise 4. Let $(\mathfrak{a}_i)$, $i = 0, \ldots, n$, be an affinely adapted frame; then $\boldsymbol{W}^n := \mathfrak{L}(\mathfrak{a}_1, \ldots, \mathfrak{a}_n) \subset \boldsymbol{V}^{n+1}$ defines the absolute of the affine space $\boldsymbol{A}^n$. a) Prove that $[\boldsymbol{A}^n, \boldsymbol{W}^n, K]$ satisfies axioms (I), (II), (III) of Definition I.4.3.1 for an affine geometry, if the action of $\boldsymbol{W}^n$ on $\boldsymbol{A}^n$ is as follows:

$$t_{\mathfrak{b}} : \boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{A}^n \longmapsto t_{\mathfrak{b}}(\boldsymbol{x}) := [\mathfrak{x} + \mathfrak{b}] \in \boldsymbol{A}^n, \qquad \mathfrak{b} \in \boldsymbol{W}^n.$$

– b) Each $t_{\mathfrak{b}}$ with $\mathfrak{b} \neq \mathfrak{o}$ is a translation in the sense of Exercise 3, i.e., $t_{\mathfrak{b}} = f|\boldsymbol{A}^n$ is the restriction of a projectivity f on $\boldsymbol{P}^n$ with $f|\boldsymbol{H} = \mathrm{id}_{\boldsymbol{H}}$ that has no fixed points in $\boldsymbol{A}^n$. – c) If $\boldsymbol{M}^k = \boldsymbol{A}^n \cap \boldsymbol{B}^k$ is an affine k-plane, and $\boldsymbol{U}^{k+1} \subset \boldsymbol{V}^{n+1}$ denotes the vector subspace associated with $\boldsymbol{B}^k$, then the set of vectors corresponding to the affine k-plane $\boldsymbol{M}^k$ (cf. Proposition I.4.5.1) is

$$V(\boldsymbol{M}^k) = \{\mathfrak{b} \in \boldsymbol{W}^n | t_{\mathfrak{b}}(\boldsymbol{M}^k) = \boldsymbol{M}^k\} = \boldsymbol{W}^n \cap \boldsymbol{U}^{k+1}.$$

Exercise 5. Find an example of two affinely skew planes $\boldsymbol{M}^2, \boldsymbol{B}^2 \subset \boldsymbol{A}^4 \subset \boldsymbol{P}^4$ in the four-dimensional affine space over a field K whose defining projective planes intersect. (Remark: "affinely skew" means that the affine planes are disjoint and not parallel, cf. § I.4.6.)

Exercise 6. Prove (in the notations of Definition 2): If $\boldsymbol{D}^l$ is parallel to $\boldsymbol{M}^k$ and $\boldsymbol{D}^l \cap \boldsymbol{M}^k \neq \emptyset$, then $\boldsymbol{D}^l \subset \boldsymbol{M}^k$.

To conclude this section we want to consider *affine collineations*, i.e. bijective maps $f : \boldsymbol{A}^n \to \hat{\boldsymbol{A}}^m$ between affine spaces over skew fields, not necessarily isomorphic, mapping lines into lines satisfying $f(\boldsymbol{x} \vee \boldsymbol{y}) = f(\boldsymbol{x}) \vee f(\boldsymbol{y})$. Here the operation $\vee$ denotes the restriction of the corresponding projective operation to the affine subspaces; similarly for $\wedge$:

$$\boldsymbol{D}^l \wedge \boldsymbol{M}^k := (\bar{\boldsymbol{D}}^l \wedge \bar{\boldsymbol{M}}^k) \cap \boldsymbol{A}^n, \ \boldsymbol{D}^l \vee \boldsymbol{M}^k := (\bar{\boldsymbol{D}}^l \vee \bar{\boldsymbol{M}}^k) \cap \boldsymbol{A}^n. \tag{17}$$

We will prove that each affine collineation is the restriction of a uniquely determined collineation between the associated projective spaces. Using the Main Theorem of Projective Geometry we can draw some interesting conclusions from this fact.

Lemma 2. *Let* $f : \boldsymbol{A}^n \to \widehat{\boldsymbol{A}}^m$ *be an affine collineation between affine spaces* $\boldsymbol{A}^n \subset \boldsymbol{P}^n$, $\widehat{\boldsymbol{A}}^m \subset \widehat{\boldsymbol{P}}^m$ *over the skew fields* K, $\hat{K}$, $\operatorname{char} K \neq 2$, $\operatorname{char} \hat{K} \neq 2$. *Then there is a unique collineation* $\bar{f} : \boldsymbol{P}^n \to \widehat{\boldsymbol{P}}^m$ *such that* $f = \bar{f}|\boldsymbol{P}^n$.

Proof. For $n = 0, 1$ the statement is trivial. Let now $n \geq 2$. If $\bar{f}$ is a collineation extending f, then we obviously have $\bar{f}(\boldsymbol{H}) = \widehat{\boldsymbol{H}}$ for the improper hyperplanes, and for every affine subspace $\boldsymbol{D} \subset \boldsymbol{A}^n$ the relation

$$\bar{f}(\bar{\boldsymbol{D}} \wedge \boldsymbol{H}) = \overline{f(\boldsymbol{D})} \wedge \widehat{\boldsymbol{H}} \tag{18}$$

has to hold. This allows the following construction for the image $\bar{f}(\boldsymbol{x})$ of any improper point $\boldsymbol{x} \in \boldsymbol{H}$: Choose an affine line $\boldsymbol{h}$ in the direction of $\boldsymbol{x}$, i.e. satisfying $\boldsymbol{x} = \bar{\boldsymbol{h}} \wedge \boldsymbol{H}$, and set $\bar{f}(\boldsymbol{x}) := \overline{f(\boldsymbol{h})} \wedge \widehat{\boldsymbol{H}}$. To justify this definition one has to show that the image $\bar{f}(\boldsymbol{x})$ does not depend on the choice of the affine line $\boldsymbol{h}$ with improper point $\boldsymbol{x}$. In other words, the following lemma has to hold:

Lemma 3. *Under the assumptions of Lemma 2, every affine collineation* $f : \boldsymbol{A}^n \to \widehat{\boldsymbol{A}}^m$ *maps pairs of parallel lines into pairs of parallel lines.*

Proof. Since the closures $\bar{\boldsymbol{h}}_1, \bar{\boldsymbol{h}}_2$ of parallel lines, $\boldsymbol{h}_1 \neq \boldsymbol{h}_2$, intersect in an improper point $\boldsymbol{x} \in \boldsymbol{H}$, they span a plane, i.e., they are not skew. Consider a *parallelogram* $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$ with $\boldsymbol{a} \vee \boldsymbol{d} = \boldsymbol{h}_1$, $\boldsymbol{b} \vee \boldsymbol{c} = \boldsymbol{h}_2$, whose sides $\boldsymbol{a} \vee \boldsymbol{b}, \boldsymbol{c} \vee \boldsymbol{d}$ are parallel, too. The diagonals $\boldsymbol{a} \vee \boldsymbol{c}, \boldsymbol{b} \vee \boldsymbol{d}$ then intersect in a proper point $\boldsymbol{s} = (\boldsymbol{a} \vee \boldsymbol{d}) \wedge (\boldsymbol{b} \vee \boldsymbol{c}) \in \boldsymbol{A}^n$. Of course, this intersection point lies in the plane spanned by the parallels. This plane intersects the improper hyperplane $\boldsymbol{H}$ in the line spanned by the diagonal points of the complete quadrangle $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$ corresponding to the pairs of parallel sides of the parallelogram. Because of $\operatorname{char} K \neq 2$, the diagonal points are not collinear, cf. Exercise 4.6, and this implies $\boldsymbol{s} \in \boldsymbol{A}^n$. Since the affine collineation f is bijective, the image lines intersect in the proper point $f(\boldsymbol{s}) = f(\boldsymbol{a} \vee \boldsymbol{d}) \wedge f(\boldsymbol{b} \vee \boldsymbol{c})$. So the image points $f(\boldsymbol{a}), f(\boldsymbol{b}), f(\boldsymbol{c}), f(\boldsymbol{d})$, and hence also the associated image lines $f(\boldsymbol{h}_1) = f(\boldsymbol{a} \vee \boldsymbol{b}), f(\boldsymbol{h}_2) = f(\boldsymbol{c} \vee \boldsymbol{d})$ lie in a plane. As they are disjoint due to the bijectivity of f, they have to be parallel. □

Now we complete the proof of Lemma 2. Since the map $\bar{f}$ is uniquely determined according to Lemma 3, all we have to show is that it also maps improper lines into improper lines. So let $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \boldsymbol{H}$ be three different collinear points. Consider a projective plane $\boldsymbol{M}^2$ intersecting $\boldsymbol{H}$ in the line $\boldsymbol{x} \vee \boldsymbol{y}$. In this plane we choose three lines $\boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{h}_3$ such that $\boldsymbol{x} = \boldsymbol{H} \wedge \boldsymbol{h}_1$, $\boldsymbol{y} = \boldsymbol{H} \wedge \boldsymbol{h}_2$, $\boldsymbol{z} = \boldsymbol{H} \wedge \boldsymbol{h}_3$. These lines determine a proper triangle in the plane $\boldsymbol{M}^2$. The image points $f(\boldsymbol{a}), f(\boldsymbol{b}), f(\boldsymbol{c})$ of the vertices $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ in this triangle determine a plane $\widehat{\boldsymbol{M}}^2$, and for the images of the improper points we have

$$\bar{f}(\boldsymbol{x}) = \overline{f(\boldsymbol{h}_1)} \cap \widehat{\boldsymbol{H}}, \bar{f}(\boldsymbol{y}) = \overline{f(\boldsymbol{h}_2)} \cap \widehat{\boldsymbol{H}}, \bar{f}(\boldsymbol{z}) = \overline{f(\boldsymbol{h}_3)} \cap \widehat{\boldsymbol{H}}.$$

Because of $f(\boldsymbol{h}_i) \subset \widehat{\boldsymbol{M}}^2$, $i = 1, 2, 3$, these points lie on the line of intersection $\widehat{\boldsymbol{M}}^2 \wedge \widehat{\boldsymbol{H}}$, hence they are collinear. The bijectivity of f follows directly from the bijectivity of $\bar{f}$. Hence $\bar{f}$ is a collineation. □

The Main Theorem of Projective Geometry immediately leads to a series of corollaries:

Corollary 4. *Let $m \geq 2$. Under the assumptions of Lemma 2, the dimensions of the affine spaces coincide, $n = m$, and there is an isomorphism $\sigma : K \to \hat{K}$ such that the affine collineation f has a basis representation of the form*

$$f(\boldsymbol{x}) = f(\boldsymbol{a} + \sum_{i=1}^{n} \mathfrak{a}_i x^i) = f(\boldsymbol{a}) + \sum_{i=1}^{n} \mathfrak{b}_i \sigma(x^i). \tag{19}$$

If K is a field, and f preserves the affine ratio of parallel segments, then f is an affine bijection with $K = \hat{K}$ and $\sigma = \mathrm{id}_K$.

P r o o f. By Corollary 3.9 and Lemma 2 we have $n = m$. Hence $n \geq 2$, and we may apply the Main Theorem of Projective Geometry. For affinely adapted frames we obtain representation (18) starting from the basis representation of the σ-linear map inducing $\bar{f}$. By Exercise 2, the segment ratio is a special case of the cross ratio, and this implies $K = \hat{K}$ as well as $\sigma = \mathrm{id}_K$. □

Corollary 5. *For $n \geq 2$ and a field K with $\operatorname{char} K \neq 2$, the affine group $\mathfrak{A}(n)$ is the group of affine collineations for $\boldsymbol{A}^n$ leaving the ratio of parallel segments invariant.* □

For a later application, we point out the real case $K = \mathbf{R}$. There is only the trivial isomorphism $\sigma = \mathrm{id}_{\mathbf{R}}$, and this leads to a particularly simple description of the affine group:

Proposition 6. *The affine group $\mathfrak{A}(n)$ of the real affine space $\boldsymbol{A}^n$, $n \geq 2$, consists of all bijective maps from $\boldsymbol{A}^n$ to itself mapping lines into lines.* □

Exercise 7. Using the Main Theorem of Projective Geometry prove that the *isometry group*[2] of the Euclidean point space $\boldsymbol{E}^n$, $n \geq 2$, is the Euclidean group. (A different proof of this statement can be found in Exercise I.6.2.13.)

1.6 Duality

In projective geometry, there is a fundamental conceptual symmetry under which the notion pairs join and meet, points and hyperplanes, or, more generally, k-planes and $(n-k-1)$-planes, correspond to one another. This symmetry,

[2] This is understood to be the set of all bijective maps from a metric space to itself preserving distance. Obviously, this is a group with respect to composition, see also Section 2.5.6 below.

called *duality*, allows to produce new geometric results without new proofs, just by formally replacing each notion by its dual counterpart. It reflects the duality of vectors and linear forms, the covectors, cf. § I.5.6; we will apply this relation to study the duality for projective geometries of arbitrary dimension, see 1.6.2 below. Let us start with the simplest interesting case:

1.6.1 Duality in Plane Incidence Geometry

In plane incidence geometry (cf. Example 1.4) there are two basic sets, the set of points and the set of lines, together with a relation, incidence, between points and lines. Performing the *dual replacements*

$$\begin{aligned} &\text{point} \longmapsto \text{line},\\ &\text{line} \longmapsto \text{point},\\ &\text{incidence} \longmapsto \text{incidence}, \end{aligned}$$

in the axiom system P.1–3, Axioms P.1 and P.2 of plane incidence geometry are transformed one into the other:

$$\text{P.1} \longmapsto \text{P.2}, \qquad \text{P.2} \longmapsto \text{P.1},$$

whereas P.3 changes into a new statement P.3$'$:

P.3$'$. *There are four lines no three of which pass through the same point.*

Statement P.3$'$ can, however, be proved using Axioms P.1–3, cf. Exercise 1.7. Adding P.3$'$ as a (superfluous) axiom to the axioms P.1–3 leads to an axiom system of plane incidence geometry that is transformed into itself by the dual replacements. This conceptual symmetry is called the *duality of the projective plane.* Points and lines form a *mutually dual pair*, and the incidence is a notion that is dual to itself[1]. Since all definitions and propositions of plane incidence geometry are exclusively based on the three fundamental notions mentioned together with the quoted self-dual axiom system, dual replacement assigns a corresponding dual notion to every derived notion. Substituting the dual notion for each notion occurring in a proposition (i.e. a proved statement) again produces a valid statement, a proof of which could just proceed by replacing the notions in the original proof of the first one by the dual ones. Hence there is no need to explicate it. In special cases, the dual notion or the dual statement may, of course, coincide with the original one, cf. Exercise 1. The formal rule that for any true proposition in plane incidence geometry there is a dual one also holding is called the *duality principle for the projective plane.* In fact, since in the algebraically defined plane projective geometry over a skew field K the Axioms P.1–3 can be derived as propositions, the plane duality principle can be applied in this case as well. In the subsequent

[1] In fact, *incidence* is defined as *including or to be included*, cf. Definition 1.1.1, and, from the point of view of lattice theory, *to contain* and *to be contained* form a dual pair, see also Equation (2) below.

Section 1.6.2 we will discuss the n-dimensional case, that will establish the plane duality principle once again, this time as a special case. The book by L. HEFFTER [48] commences with a self-dual axiom system for the projective geometry of three-dimensional space, from which the duality principle for this geometry follows.

Exercise 1. Prove that, in the projective plane $\boldsymbol{P}^2$ over a skew field K, the figure in DESARGUES' Theorem (Exercise 1.9, Fig. 1.5) is self-dual. Starting from this, conclude the following self-dual proposition: Let $(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$, $(\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3)$ be two *triangles*, i.e. triples of non-collinear points, and let

$$\begin{array}{l} \boldsymbol{A}_1 = \boldsymbol{a}_2 \vee \boldsymbol{a}_3,\ \boldsymbol{A}_2 = \boldsymbol{a}_3 \vee \boldsymbol{a}_1,\ \boldsymbol{A}_3 = \boldsymbol{a}_1 \vee \boldsymbol{a}_2 \\ \boldsymbol{B}_1 = \boldsymbol{b}_2 \vee \boldsymbol{b}_3,\ \boldsymbol{B}_2 = \boldsymbol{b}_3 \vee \boldsymbol{b}_1,\ \boldsymbol{B}_3 = \boldsymbol{b}_1 \vee \boldsymbol{b}_2 \end{array}$$

be the corresponding *trilaterals*, i.e. triples of non-concentric lines. Then the following holds: The connecting lines of corresponding points, $\boldsymbol{a}_1 \vee \boldsymbol{b}_1$, $\boldsymbol{a}_2 \vee \boldsymbol{b}_2$, $\boldsymbol{a}_3 \vee \boldsymbol{b}_3$, are concentric if and only if the intersection points of corresponding lines, $\boldsymbol{A}_1 \wedge \boldsymbol{B}_1$, $\boldsymbol{A}_2 \wedge \boldsymbol{B}_2$, $\boldsymbol{A}_3 \wedge \boldsymbol{B}_3$, are collinear.

Exercise 2. Dualize PAPPOS' Theorem (Exercise 1.10), and sketch the corresponding situation.

Exercise 3. Let $\boldsymbol{P}^2$ be a projective plane over a skew field K, $\operatorname{char} K \neq 2$. Dualize the notion of a complete quadrangle (Exercise 4.6) to define a *complete quadrilateral.* As the support of the point series the line dually corresponds to the point as the support of the line pencil, i.e. the set of all lines of the plane passing through that point. According to Example 1 one obtains the notions of cross ratio and harmonic position for lines in the pencil by dualizing the corresponding notions defined for the point series. Find a construction for the fourth line that is harmonic to three concentric lines by dualizing Exercise 4.7.

1.6.2 Projective and Algebraic Duality

Introducing projective geometry by starting from linear algebra, the geometric duality results from the algebraic duality between vectors and covectors. In this section, we will only consider finite-dimensional vector spaces $\boldsymbol{V}$ over a skew field K. As in the commutative case, the *dual* $\boldsymbol{V}'$ is defined as the set of linear maps $\mathfrak{u} : \boldsymbol{V} \to K$; these are also called *linear forms* or *covectors*. Defining addition and multiplication by scalars argumentwise, the dual again is a vector space over K of the same dimension. Note that the dual of a right (left) vector space is a left (right) vector space (Proposition II.7.3.7); in fact, in the non-commutative case linearity of the form can be ensured only if scalar multiplication takes place from the opposite side:

$$\mu\mathfrak{u}(\mathfrak{x}\alpha + \mathfrak{y}\beta) = \mu\mathfrak{u}(\mathfrak{x})\alpha + \mu\mathfrak{u}(\mathfrak{y})\beta, \qquad \mu, \alpha, \beta \in K,\ \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V},\ \mathfrak{u} \in \boldsymbol{V}'.$$

As in the commutative case, the dual of the dual space, $(\boldsymbol{V}')'$, again is a vector space that is canonically isomorphic to the original space $\boldsymbol{V}$. As a rule, both spaces are identified via the canonical isomorphism: $(\boldsymbol{V}')' = \boldsymbol{V}$. This can be described most easily using the symmetric notation for the scalar product of a vector and a covector:

$$(\mathfrak{u}, \mathfrak{x}) \in \boldsymbol{V}' \times \boldsymbol{V} \mapsto (\mathfrak{u}|\mathfrak{x}) := \mathfrak{u}(\mathfrak{x}) \in K.$$

The canonical isomorphism $\boldsymbol{V} \to (\boldsymbol{V}')'$ is obtained by fixing the vector $\mathfrak{x} \in \boldsymbol{V}$ in $(\mathfrak{u}|\mathfrak{x})$ and considering $\mathfrak{u} \in \boldsymbol{V}'$ as the variable. It is easy to see that, for any basis $(\mathfrak{a}_i)$ of $\boldsymbol{V}$, the set of projections onto a single component,

$$\mathfrak{a}^j(\mathfrak{a}_i\xi^i) := \xi^j, \qquad \text{i.e. } (\mathfrak{a}^j|\mathfrak{a}_i) = \delta_i^j,$$

uniquely defines a basis $(\mathfrak{a}^j)$ of $\boldsymbol{V}'$, which is called the *dual basis to* $(\mathfrak{a}_i)$; here the sum convention is in action, and δ_i^j denotes the Kronecker symbol.[2] If

$$\mathfrak{x} = \mathfrak{a}_i\xi^i, \qquad \mathfrak{u} = v_i\mathfrak{a}^i$$

are the basis representations of a vector and a covector in mutually dual bases, then their scalar product has the simple expression

$$(\mathfrak{u}|\mathfrak{x}) = v_i\xi^i.$$

It is clear that the propositions of projective geometry hold for right as well as for left vector spaces over the skew field K. In fact, projective geometry is determined by the lattice of subspaces, and, according to Example II.7.2.14, every left vector space over K can be considered as a right vector space over the opposite skew field K_0. So the lattice of subspaces does not change. This is the way to understand the projective geometry $\mathfrak{P}(\boldsymbol{V}')$: The projective geometries of equal dimension over opposite skew fields are isomorphic. Conceptually, projective duality is described by means of the annihilator relation and, in contrast to synthetic projective incidence geometry, accessible to calculations. As is well-known, the *annihilator* (cf. Definition I.5.6.1) is defined as a map $\perp : \mathfrak{P}(\boldsymbol{V}) \to \mathfrak{P}(\boldsymbol{V}')$ in the following way:

$$\boldsymbol{A} \in \mathfrak{P}(\boldsymbol{V}) \mapsto \boldsymbol{A}^\perp := \{\mathfrak{u} \in \boldsymbol{V}' | \, (\mathfrak{u}|\mathfrak{x}) = 0 \text{ for all } \mathfrak{x} \in \boldsymbol{A}\} \in \mathfrak{P}(\boldsymbol{V}'). \tag{1}$$

Here the subspaces $\boldsymbol{A} \in \mathfrak{P}(\boldsymbol{V})$, $\boldsymbol{A}^\perp \in \mathfrak{P}(\boldsymbol{V}')$ are viewed as vector sets; considering them as projective subspaces, the annihilator relation makes sense for both arguments due to the homogeneity of the scalar product; the definition

$$\boldsymbol{u} = [\mathfrak{u}] \in \boldsymbol{A}^\perp :\Longleftrightarrow (\mathfrak{u}|\mathfrak{x}) = 0 \text{ for all } \boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{A}$$

is independent of the chosen representatives $\mathfrak{u}, \mathfrak{x}$ of $\boldsymbol{u}, \boldsymbol{x}$. >From what we proved in linear algebra, cf. § I.5.6 and Proposition II.7.3.7, Exercise II.7.3.7 in the non-commutative case, the following is immediate.

[2] In Chapter II.7, these results are obtained by specializing the propositions concerning free modules over non-commutative rings. An elementary presentation of vector algebra over skew fields can be found in Chapter III of the book by H. Reichardt [91].

Proposition 1. *Let $\mathfrak{P}^n = \mathfrak{P}(\boldsymbol{V})$, $\mathfrak{P}'^n = \mathfrak{P}(\boldsymbol{V}')$ be the projective geometries associated with the $(n+1)$-dimensional vector space $\boldsymbol{V}$ over the skew field K and its dual $\boldsymbol{V}'$, respectively. Then the annihilator relation $\perp : \mathfrak{P}^n \to \mathfrak{P}'^n$ is a bijection of lattices reversing inclusion $\subset$, and hence preserving incidence. In detail:*

$$\boldsymbol{A} \subset \boldsymbol{B} \Longleftrightarrow \boldsymbol{B}^{\perp} \subset \boldsymbol{A}^{\perp}, \tag{2}$$

$$(\boldsymbol{A} \vee \boldsymbol{B})^{\perp} = \boldsymbol{A}^{\perp} \wedge \boldsymbol{B}^{\perp}, \qquad (\boldsymbol{A} \wedge \boldsymbol{B})^{\perp} = \boldsymbol{A}^{\perp} \vee \boldsymbol{B}^{\perp}, \tag{3}$$

$$(\boldsymbol{A}^{\perp})^{\perp} = \boldsymbol{A}, \tag{4}$$

$$\operatorname{Dim} \boldsymbol{A} = k \Longleftrightarrow \operatorname{Dim} \boldsymbol{A}^{\perp} = n - k - 1, \tag{5}$$

$$\boldsymbol{o}^{\perp} = \boldsymbol{P}'^{\,n}, \qquad \boldsymbol{P}^{n\,\perp} = \boldsymbol{o}'. \tag{6}$$

□

According to Proposition 1, *projective duality* can be described as follows: The map $\perp : \mathfrak{P}^n \to \mathfrak{P}'^n$ assigns to each basic notion, more precisely, to each element of the projective geometry $\mathfrak{P}^n$, the dual basic notion for the dual projective geometry $\mathfrak{P}'^n$ so that formulas (2) to (6) hold. Since each derived notion as well as every proposition are obtained starting from these basic notions through finitely many steps, each concept leads to a pair of mutually dual notions, and each result leads to a pair of mutually dual statements, that are either both true or both false. In this connection, recall (4). Since all n-dimensional geometries over the same or the opposite skew fields are isomorphic, according to this *duality principle*, for each proposition there is a proposition dual to it, which, because of (5), as a rule has a different geometric content. As in plane incidence geometry it may, of course, happen that the dual proposition coincides with the original one. Note that, by its very definition, the map $\perp$ always depends on the space it is based on, hence on its dimension, which becomes particularly evident in Formula (5). In the examples to follow we will apply the duality principle.

Example 1. Let $\boldsymbol{u} = [\mathfrak{u}] \in \boldsymbol{P}'^{\,n}$ be a "point" of the dual projective space. Since by (4) the annihilator, viewed as a map $\perp : \mathfrak{P}'^{\,n} \to \mathfrak{P}^n$, is the inverse to (1), $\boldsymbol{u}$ represents the hyperplane

$$\boldsymbol{u}^{\perp} = \{\boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{P}^n | (\mathfrak{u}|\mathfrak{x}) = 0\}. \tag{7}$$

As the map $\perp$ defines a canonical bijection from the dual point space $\boldsymbol{P}'^{\,n}$ onto the Graßmann manifold $\boldsymbol{P}_{n,n-1}$ of hyperplanes in the original space $\boldsymbol{P}^n$, it is reasonable and customary to interpret the elements $\boldsymbol{u} = [\mathfrak{u}] \in \boldsymbol{P}'^{\,n}$ as hyperplanes in $\boldsymbol{P}^n$, even to call them this way. In (7), the hyperplane $\boldsymbol{u}^{\perp}$ is characterized as the set of points whose representatives satisfy the equation $(\mathfrak{u}|\mathfrak{x}) = 0$; fixing, conversely, a point $\boldsymbol{x} = [\mathfrak{x}] \neq \boldsymbol{o}$ of $\boldsymbol{P}^n$ and looking at all hyperplanes $\boldsymbol{u} = [\mathfrak{u}]$ satisfying this equation, the dual configuration $\boldsymbol{x}^{\perp}$ turns out to be the set of all hyperplanes through the point $\boldsymbol{x}$. In general, the Graßmann manifold $\boldsymbol{P}_{n,k}$ of k-planes in the projective space $\boldsymbol{P}^n$ corresponds to the Graßmann manifold $\boldsymbol{P}'_{n,n-k-1}$ of $(n-k-1)$-planes in $\boldsymbol{P}'^n$. Since $\boldsymbol{P}'^n$ itself is a projective point space, dualizing the notions and propositions on projective configurations of points we obtain the dual notions and propositions concerning projective configurations of hyperplanes. If, e.g. $n=3$, then points are dual to planes, and lines are dual to lines. Each line $\boldsymbol{B}$ determines the set of points incident with it; we identified this set via the canonical map π (cf. (1.2)) with the line itself. Without this identification, the object "line" has to be distinguished from this set, which, just for this reason, is sometimes called a *point series*. The associated dual notion is the set of all planes incident with the line, i.e. containing it. The totality of all these planes is called the *plane pencil* with *support* $\boldsymbol{B}$. Let $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{B}$ be two different points on the line; then we have

$$\boldsymbol{B} = \boldsymbol{a} \vee \boldsymbol{b} = (\boldsymbol{a}^\perp \wedge \boldsymbol{b}^\perp)^\perp,$$

i.e., the line appears as the join of two points, and, dually, as the intersection of two planes. Transferring the notion from the point series onto the plane pencil, which, in fact, is nothing but a line in $\boldsymbol{P}'^3$, the latter is thus equipped with a CR. Hence, e.g., there is the notion of harmonic position for four planes in a plane pencil. For $n=2$, the set of all lines through one point dually corresponds to the set of all points in a line. This is called a *line pencil*, and in this set the CR is defined as well. For arbitrary $n>1$, the pencil of all hyperplanes through a fixed $(n-2)$-plane corresponds to a line in the dual space $\boldsymbol{P}'^n$; the latter is again called the support of the hyperplane pencil. Algebraically, viewing a k-plane as the join of points corresponds to its parameter representation, and considering it as the intersection of hyperplanes corresponds to its representation as the solution set of a homogeneous system of equations. □

Exercise 4. a) Dualize the notions collinear, throw, CR of a throw to refer to hyperplanes. – b) Prove for a throw of points in $\boldsymbol{P}^n$, $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d})$, that $(\boldsymbol{a}^\perp, \boldsymbol{b}^\perp, \boldsymbol{c}^\perp, \boldsymbol{d}^\perp)$ is a throw of hyperplanes in the space $\boldsymbol{P}'^n$, and that the CR of these throws coincide. – c) Formulate and prove an analogue of Exercise 4.4 for throws of hyperplanes.

Example 2. In (1.2) we defined a canonical map $\pi : \boldsymbol{V} \to \boldsymbol{P}^n_o$ that, applied to subspaces, allowed to realize the projective subspaces $\boldsymbol{B} \in \mathfrak{P}^n$ as point sets preserving inclusion $\subset$; using the incidence relation this is expressed as ι:

$$\pi(\boldsymbol{B}) = \{\boldsymbol{x} \in \boldsymbol{P}^n | \boldsymbol{x} \ \iota \ \boldsymbol{B}\}. \tag{8}$$

The analogous canonical map $\pi : \boldsymbol{V}' \to \boldsymbol{P}'^n$ can be used to consider each projective subspace $\boldsymbol{U} \subset \boldsymbol{P}'^n$ as a set of points in $\boldsymbol{P}'^n$, hence, by duality, as a set of hyperplanes in $\boldsymbol{P}^n$. By (6), the nopoint $\boldsymbol{o}' = \pi(\mathfrak{o}) \in \boldsymbol{P}'^n$ then corresponds to the whole space $\boldsymbol{P}^n$. Dualizing (8) leads to the following geometric interpretation of the annihilator relation $\perp$:

$$\tau(\boldsymbol{B}) := \{\boldsymbol{u} \in \boldsymbol{P}'^n | \boldsymbol{u} \ \iota \ \boldsymbol{B}\} = \boldsymbol{B}^\perp, \tag{9}$$

since $\tau(\boldsymbol{B})$ by definition is nothing but the incidence in the set of all hyperplanes $\boldsymbol{u}$ containing $\boldsymbol{B}$, for which we thus have $(\boldsymbol{u}|\boldsymbol{B}) = 0$. But this is just $\boldsymbol{B}^{\perp}$; $\tau(\boldsymbol{B})$ is called the *hyperplane pencil with support* $\boldsymbol{B}$. (Still to be included is the whole space $\boldsymbol{P}^n$ corresponding to the nopoint $\boldsymbol{o}'$ that always has to be added to all spaces of hyperplanes. In the sequel we will usually not mention this explicitly.) The canonical map π preserves the relation $\subset$, whereas it is reversed under τ by (2); for this reason, to avoid confusions we will not use τ as an identification like π. □

Exercise 5. Let $(\mathfrak{a}_i)_{i=0,\ldots,n}$ be a basis of $\boldsymbol{V}^{n+1}$, and let $(\mathfrak{a}^j)$ the basis of $\boldsymbol{V}'$ dual to it. The hyperplane $\boldsymbol{H}_j$ corresponding to $\mathfrak{a}^j$ is defined by

$$\boldsymbol{x} \in \boldsymbol{H}_j \Longleftrightarrow \boldsymbol{x} = [\mathfrak{x}] \text{ and } (\mathfrak{a}^j|\mathfrak{x}) = 0;$$

$\boldsymbol{H}_j$ is the face opposite to the point $\boldsymbol{a}_j = [\mathfrak{a}_j]$ in the coordinate simplex for this basis. The object dual to the unit point $\boldsymbol{e} = [\mathfrak{a}_0 + \ldots + \mathfrak{a}_n]$ is the *unit hyperplane*

$$\boldsymbol{E} := [\mathfrak{a}^0 + \ldots + \mathfrak{a}^n] = \{\boldsymbol{x} \in \boldsymbol{P}^n_o | \boldsymbol{x} = [\mathfrak{a}_j x^j] \text{ and } \sum_{j=0}^{n} x^j = 0\}.$$

Choose the unit hyperplane $\boldsymbol{E} \subset \boldsymbol{P}^n$ of a projective frame as the hyperplane at infinity for the affine space $\boldsymbol{A}^n$ and prove

$$\boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{A}^n \Longleftrightarrow \sum_{j=0}^{n} x^j \neq 0;$$

moreover, the normalization

$$\nu^i := x^i (\sum_{j=0}^{n} x^j)^{-1} \qquad (\boldsymbol{x} \in \boldsymbol{A}^n)$$

defines the *barycentric coordinates* of $\boldsymbol{x}$ with respect to the n-simplex $([\mathfrak{a}_0], \ldots, [\mathfrak{a}_n])$, cf. Exercise I.4.5.5.

1.6.3 Projective Pencil Geometries

Let $\mathfrak{P}^n$ be an n-dimensional, e.g. right projective geometry over the skew field K, and let $\boldsymbol{A}^k \in \mathfrak{P}^n$ be an arbitrary, fixed k-plane. The *pencil* of projective subspaces with *support* $\boldsymbol{A}^k$ is defined as the set of all subspaces containing $\boldsymbol{A}^k$:

$$\mathfrak{P}^n/\boldsymbol{A}^k := \{\boldsymbol{B} \in \mathfrak{P}^n | \boldsymbol{A}^k \subset \boldsymbol{B}\}. \tag{10}$$

Now we prove

Proposition 2. *Under the assumptions of Definition* (10), *let* $\boldsymbol{V}^{n+1} \supset \boldsymbol{W}^{k+1}$ *be the vector spaces over the skew field* K *corresponding to* $\boldsymbol{P}^n$ *and* $\boldsymbol{A}^k$, *respectively. Denote by* $\boldsymbol{V}/\boldsymbol{W}$ *the quotient space. Then*

$$p : \boldsymbol{B} \in \mathfrak{P}^n/\boldsymbol{A}^k \longmapsto p(\boldsymbol{B}) := \{b + W | b \in \boldsymbol{B}\} \in \mathfrak{P}(\boldsymbol{V}/\boldsymbol{W}) \tag{11}$$

defines a canonical bijection, which is an isomorphism of lattices. This way, a projective geometry of dimension $n-k-1$ is defined in $\mathfrak{P}^n/\boldsymbol{A}^k$. The space dual to the projective point space $\boldsymbol{P}(\boldsymbol{V}/\boldsymbol{W})$, i.e. the space of its hyperplanes, is canonically isomorphic to $\boldsymbol{A}^{\perp}$.

Proof. For simplicity, let $\boldsymbol{B}$ denote the vector subspace and, at the same time, the projective subspace; depending on the context, it thus has to be interpreted as a set of vectors or points. Considered as a map between vector subspaces, the map p is obviously generated by the canonical projection

$$p : \mathfrak{x} \in \boldsymbol{V} \mapsto \mathfrak{x} + \boldsymbol{W} \in \boldsymbol{V}/\boldsymbol{W},$$

hence it is surjective. Since its kernel is $\boldsymbol{W}$, we have for the zero vector $\hat{0} \in \boldsymbol{V}/\boldsymbol{W}$ the preimage $p^{-1}(\hat{0}) = \boldsymbol{W}$, and this is the nopoint in $\mathfrak{P}^n/\boldsymbol{A}^k$. Thus the map (11) is also injective. Obviously, it preserves the relation $\subset$ and, in addition, satisfies

$$p(\boldsymbol{B}_1) \vee p(\boldsymbol{B}_2) = p(\boldsymbol{B}_1 \vee \boldsymbol{B}_2).$$

Therefore, it is a collineation as well, where the *relative dimension* in the pencil has to be taken into account:

$$\operatorname{Dim}\operatorname{rel}(\boldsymbol{B}) := \operatorname{Dim}\boldsymbol{B} - k - 1. \tag{12}$$

Furthermore, $\boldsymbol{A}^{k\perp}$ is the pencil of hyperplanes in $\boldsymbol{P}^n$ containing $\boldsymbol{A}^k$. For any of these hyperplanes, $\boldsymbol{u} = [\mathfrak{u}]$, we have $\mathfrak{u}|\boldsymbol{W} = 0$, and vice versa. Hence, $\hat{\mathfrak{u}}(\mathfrak{x} + \boldsymbol{W}) := \mathfrak{u}(\mathfrak{x})$ is a well-defined linear form on $\boldsymbol{V}/\boldsymbol{W}$. The map $\mathfrak{u} \in \boldsymbol{W}^{\perp} \mapsto \hat{\mathfrak{u}} \in (\boldsymbol{V}/\boldsymbol{W})'$ is linear and has kernel $\mathfrak{o}$; thus it is injective and, since the dimensions of both spaces coincide, $\dim \boldsymbol{W}^{\perp} = \dim \boldsymbol{V}/\boldsymbol{W} = n - k$, it yields a canonically defined linear isomorphism,

$$\boldsymbol{W}^{\perp} \cong (\boldsymbol{V}/\boldsymbol{W})', \qquad \boldsymbol{W} \subset \boldsymbol{V} \text{ a subspace, } \dim V < \infty. \tag{13}$$

This immediately implies the second assertion. □

Example 3. In order to illustrate Proposition 2, which may appear a bit formal at first glance, we recall the heuristic description for the projective plane from Example 1.1. There we defined the projective plane by means of the pencil of all rays through a point, cf. Fig. 1.1. The relative dimension of a line as an object in the pencil is zero; they are the "points" of the pencil. Intersecting the pencil with support $\boldsymbol{A}$ by a plane $\boldsymbol{H}$ not passing through $\boldsymbol{A}$, the assignment

$$h : \boldsymbol{x} \in \boldsymbol{H} \longmapsto h(\boldsymbol{x}) := \boldsymbol{x} \vee \boldsymbol{A} \in \boldsymbol{P}^n/\boldsymbol{A} \tag{14}$$

is a projectivity; in Example 1.1 we had $n = 3$. In the general case of a pencil in an arbitrary $\boldsymbol{P}^n$ with support $\boldsymbol{A}^k$ one has to choose an $(n-k-1)$-plane complementary to $\boldsymbol{A}^k$ as $\boldsymbol{H}$; $\boldsymbol{H}$ is called a *section* (or a complete system of representatives) for the pencil $\boldsymbol{P}^n/\boldsymbol{A}$. In fact, the map

$$s : \boldsymbol{X} \in \mathfrak{P}^n/\boldsymbol{A} \longmapsto s(\boldsymbol{X}) := \boldsymbol{X} \wedge \boldsymbol{H} \subset \boldsymbol{H} \tag{15}$$

is the inverse of the map between the lattices of projective subspaces defined by h. The map (14) transfers the $(n-k-1)$-dimensional projective geometry from $\boldsymbol{H}$ to the pencil through $\boldsymbol{A}$. Transferring a homogeneous coordinate system from $\boldsymbol{H}$ onto the pencil h can be represented as assigning equal coordinates. Hence it is a projectivity. If $f : \boldsymbol{F} \to \boldsymbol{P}^n/\boldsymbol{A}$ is an analogous projectivity from another $(n-k-1)$-plane $\boldsymbol{F}$ complementary to $\boldsymbol{A}$ onto the pencil with support $\boldsymbol{A}$, then $f^{-1} \circ h : \boldsymbol{H} \to \boldsymbol{F}$ defines a projectivity generalizing the construction of the central projection in Example 1.2, Fig. 1.3. For this reason, we call it the *central projection from* $\boldsymbol{H}$ *onto* $\boldsymbol{F}$ *with center* $\boldsymbol{A}$. Denoting by q the central projection from $\boldsymbol{P}^n$ onto the subspace $\boldsymbol{F}$ with center $\boldsymbol{A}$ defined in Example 3.1, we have $f^{-1} \circ h = q|\boldsymbol{H}$. □

1.6.4 Dual Maps

For a linear map $a : \boldsymbol{V} \to \boldsymbol{W}$ the *associated dual map* $a' : \boldsymbol{W}' \to \boldsymbol{V}'$ is defined by the following well-known formula:

$$(a'(\mathfrak{w})|\mathfrak{x}) := (\mathfrak{w}|a(\mathfrak{x})), \qquad \mathfrak{x} \in \boldsymbol{V},\ \mathfrak{w} \in \boldsymbol{W}',$$

cf. Definition I.5.2, where a' was called the map transposed to a. For the applications in projective geometry a slight generalization of this notion will be needed. Recall, (cf. Definition II.7.2.7), that a map $\sigma : K \to L$ between skew fields is called an *anti-isomorphism* if σ is an isomorphism for the additive groups of the skew fields and, moreover, $\sigma(\alpha\beta) = \sigma(\beta)\sigma(\alpha)$. A map $a : \boldsymbol{V} \to \boldsymbol{W}$ from the right vector space $\boldsymbol{V}$ over K into the left vector space $\boldsymbol{W}$ over L is called σ*-linear* if we have

$$a(\mathfrak{x}\alpha + \mathfrak{y}\beta) = \sigma(\alpha)a(\mathfrak{x}) + \sigma(\beta)a(\mathfrak{y}),\ \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}, \alpha, \beta \in K,$$

where $\sigma : K \to L$ is an anti-isomorphism. It satisfies

$$a(\mathfrak{x}\,\alpha\beta) = \sigma(\beta)\sigma(\alpha)a(\mathfrak{x}),\ \mathfrak{x}, \in \boldsymbol{V}, \alpha, \beta \in K.$$

(There is an analogous definition for maps a from a left into a right vector space.)

Definition 1. Let $\boldsymbol{V}$ be a right vector space over the skew field K, let $\boldsymbol{W}$ be a right (or left, respectively) vector space over the skew field L, let $\sigma : K \to L$ be an isomorphism (or anti-isomorphism, respectively) of the skew fields, and let, finally, $a : \boldsymbol{V} \to \boldsymbol{W}$ be σ-linear. Then the map $a' : \boldsymbol{W}' \to \boldsymbol{V}'$ defined by

$$(a'(\mathfrak{w})|\mathfrak{x}) := \sigma^{-1}((\mathfrak{w}|a(\mathfrak{x}))), \quad \mathfrak{x} \in \boldsymbol{V},\ \mathfrak{w} \in \boldsymbol{W}', \tag{16}$$

$$\text{or } (a'(\mathfrak{w})|\mathfrak{x}) := \sigma^{-1}((a(\mathfrak{x})|\mathfrak{w})) \tag{17}$$

is called the *dual map* to a. □

In Equation (17) we reversed the order in which the vector $\mathfrak{y} \in \boldsymbol{W}$ and the covector $\mathfrak{w} \in \boldsymbol{W}'$ enter the canonical scalar product $(\mathfrak{y}|\mathfrak{w})$, since this way $\boldsymbol{W}'$ becomes a right vector space. The following is straightforward from the definitions.

Lemma 3. *Equations* (16) *or* (17), *respectively, define a* σ^{-1}*-linear map* $a' : \boldsymbol{W}' \to \boldsymbol{V}'$. □

This lemma suggests the following definition:

Definition 2. In the notations of Definition 1, let $\boldsymbol{P}_o^n$ and $\boldsymbol{Q}_o^m$ be the projective spaces associated with the vector spaces $\boldsymbol{V}^{n+1}$ and $\boldsymbol{W}^{m+1}$ over the skew fields K and L, respectively. Moreover, let $f : \boldsymbol{P}_o^n \to \boldsymbol{Q}_o^m$ be a collinear map induced by the σ-linear map $a : \boldsymbol{V} \to \boldsymbol{W}$. Then the collinear map $f' : \boldsymbol{Q}_o'^m \to \boldsymbol{P}_o'^n$ induced by the dual σ^{-1}-linear map a' is called the *dual collinear map associated with the map f.* □

The following proposition contains an interesting application of the dual map: Formula (18) proved below, written a bit differently in the form

$$f^{-1}(\boldsymbol{H}) = f'(\boldsymbol{H}^{\perp})^{\perp}, \qquad \boldsymbol{H} \in \mathfrak{P}(\boldsymbol{W}^{m+1}),$$

shows the possibility to express the inverse images of the collinear map f through the images of f'.

Proposition 4. *Under the assumptions and in the notations of Definition* 2, *the dual collinear map f' is independent of the choice of the map a inducing f, and, moreover,*

$$f^{-1}(\boldsymbol{H})^{\perp} = f'(\boldsymbol{H}^{\perp}), \qquad \boldsymbol{H} \in \mathfrak{P}(\boldsymbol{W}^{m+1}). \tag{18}$$

Proof. Consider the case that $\boldsymbol{W}$ is a left vector space and (17) holds. Let $\boldsymbol{U} \subset \boldsymbol{W}'$ be an arbitrary subspace. We will then prove the equation

$$a^{-1}(\boldsymbol{U}^{\perp}) = a'(\boldsymbol{U})^{\perp}. \tag{19}$$

In fact, $\mathfrak{x} \in a^{-1}(\boldsymbol{U}^{\perp})$ if and only if $a(\mathfrak{x}) \in \boldsymbol{U}^{\perp}$, and this is the case if and only if every $\mathfrak{w} \in \boldsymbol{U}$ satisfies the equation $(a(\mathfrak{x})|\mathfrak{w}) = 0$. By Definition 1, this is equivalent to $(a'(\mathfrak{w})|\mathfrak{x}) = \sigma^{-1}(a(\mathfrak{x})|\mathfrak{w}) = 0$ for all $\mathfrak{w} \in \boldsymbol{U}$, hence $\mathfrak{x} \in a'(\boldsymbol{U})^{\perp}$. Since a induces the map f, we obtain for the projective subspaces

$$f^{-1}(\boldsymbol{U}^{\perp}) = a'(\boldsymbol{U})^{\perp} = f'(\boldsymbol{U})^{\perp},$$

so that f' only depends on f, and is independent of the chosen inducing map a. Setting $\boldsymbol{U} = \boldsymbol{H}^{\perp}$ in the last equation and applying $\perp$ we arrive at Equation (18). □

Proposition 4 justifies the definition of the dual collinear map f'. The following properties are immediate.

$$(f')' = f, \tag{20}$$

$$(f \circ g)' = g' \circ f'. \tag{21}$$

Exercise 6. Let f' be the map dual to $f : \boldsymbol{P}^n_o \to \boldsymbol{Q}^m_o$. Prove: a)

$$\boldsymbol{x} \ \iota \ f'(\boldsymbol{H}) \Longleftrightarrow f(\boldsymbol{x}) \ \iota \ \boldsymbol{H} \qquad (\boldsymbol{x} \in \boldsymbol{P}^n, \ \boldsymbol{H} \in \boldsymbol{Q}'^m), \tag{22}$$

$$f'^{-1}(\boldsymbol{o}') = f(\boldsymbol{P}^n)^\perp = \tau(f(\boldsymbol{P}^n)). \tag{23}$$

b) f determines a collineation $\hat{f}$ from the pencil $\mathfrak{P}^n/f^{-1}(\boldsymbol{o})$ onto the projective geometry of the image $f(\boldsymbol{P}^n)$:

$$\hat{f} : \boldsymbol{X} \in \mathfrak{P}^n/f^{-1}(\boldsymbol{o}) \longmapsto \hat{f}(\boldsymbol{X}) := f(\boldsymbol{X}) \in \mathfrak{P}(f(\boldsymbol{P}^n)), \tag{24}$$

and applying the canonical σ^{-1}-bijection $a(\boldsymbol{V})' \cong \boldsymbol{W}'/a(\boldsymbol{V})^\perp$ yields

$$(\hat{f})' = \widehat{(f')}. \tag{25}$$

Example 4. In the notations of Definition 1 suppose that $a : \boldsymbol{V}^{n+1} \to \boldsymbol{W}^{n+1}$ is bijective. The *contragredient map to* a is defined as

$$a^* := (a^{-1})' : \boldsymbol{V}' \to \boldsymbol{W}'. \tag{26}$$

Obviously, $a^* = (a')^{-1}$ (cf. Exercise II.7.2.12). (18) implies: For all $\boldsymbol{B} \in \mathfrak{P}(\boldsymbol{V})$ we have

$$f^*(\boldsymbol{B}^\perp) = f(\boldsymbol{B})^\perp, \tag{27}$$

where f^* denotes the collineation induced by a^*. Setting $\boldsymbol{M} = \boldsymbol{B}^\perp$ equation (27) can also be written in the form

$$f^*(\boldsymbol{M}) = (f(\boldsymbol{M}^\perp))^\perp.$$

Thus it is clear that the contragredient map does not yield anything essentially new for projective geometry. For all $\boldsymbol{B} \in \mathfrak{P}(\boldsymbol{V})$, the set

$$f(\boldsymbol{B}) = \{f(\boldsymbol{x}) | \boldsymbol{x} \in \boldsymbol{B}\}$$

is the image of $\boldsymbol{B}$ as a point set, and, dually,

$$f^*(\boldsymbol{B}^\perp) = \{f^*(\boldsymbol{H}^{n-1}) | \boldsymbol{H}^{n-1} \supset \boldsymbol{B}\}$$

is the image of $\boldsymbol{B}$ as the support of a pencil of hyperplanes. □

1.7 Correlations

In this chapter we will place special emphasis on the lattice structure of the projective geometry $\mathfrak{P}^n$. According to Exercise 3.5 collineations can be characterized as monotonously increasing bijections $F : \mathfrak{P}^n \to \mathfrak{Q}^n$ between projective geometries whose inverse F^{-1} is monotonously increasing as well. The dualistic structure of projective geometries suggests to consider monotonously decreasing maps likewise.

1.7.1 Definition. Canonical Correlation

Definition 1. Let $F : \mathfrak{P}^n$, $\mathfrak{Q}^m$ be (right or left) projective geometries over the skew fields K and L, respectively. A bijective map $F : \mathfrak{P}^n \to \mathfrak{Q}^m$ is called a *correlation* if it satisfies the following condition:

$$\boldsymbol{A} \subset \boldsymbol{B} \Longleftrightarrow F(\boldsymbol{A}) \supset F(\boldsymbol{B}),\ \boldsymbol{A}, \boldsymbol{B} \in \mathfrak{P}^n, \tag{1}$$

i.e., if F and F^{-1} both are monotonously decreasing. □

Example 1. Forming the annihilator

$$\perp : \mathfrak{P}^n \to \mathfrak{P}'^n,\ \perp : \mathfrak{P}'^n \to \mathfrak{P}^n \tag{2}$$

is a correlation from a right to a left (or from a left to a right) projective geometry that we will call the *canonical correlation*, since it is uniquely determined by the very structure of the projective geometry. For these structural reasons we admit the obvious notational misuse. In fact, since the two maps considered in (2) are mutually inverse, strictly, we would have to distinguish $\perp_{\mathfrak{P}}$ from $\perp_{\mathfrak{P}'}$ and to write $\perp_{\mathfrak{P}'} = \perp_{\mathfrak{P}}^{-1}$. □

As the correlations are anti-isomorphisms for the lattice structures, the characterization of $\bigwedge$ as the infimum and that of $\bigvee$ as the supremum lead to the following rules:

$$F(\bigwedge \boldsymbol{A}_\iota) = \bigvee F(\boldsymbol{A}_\iota),\ F(\bigvee \boldsymbol{A}_\iota) = \bigwedge F(\boldsymbol{A}_\iota),\ \boldsymbol{A}_\iota \in \mathfrak{P}^n. \tag{3}$$

Now we ask under which conditions there exists a correlation between two projective geometries. Since for $n = 1$ the order $\subset$ is rather useless — each bijection $F : \mathfrak{P}^1 \to \mathfrak{Q}^1$ with $F(\boldsymbol{o}) = \boldsymbol{Q}^1$ and $F(\boldsymbol{P}^1) = \boldsymbol{o}$ is a correlation — we will suppose $n \geq 2$, as a rule. First we note two facts immediately following from the definition.

Corollary 1. *The composition $F \circ G$ of two correlations is a collineation. The composition of a correlation and a collineation is a correlation.* □

In order to discuss the existence question, consider two right projective geometries; if one of the geometries is a left one, then by Example II.7.2.14 we may pass to the opposite skew field and vector space, and this way reduce everything to the case of right geometries.

Proposition 2. *Let $\mathfrak{P}^n$, $\mathfrak{Q}^m$ be right projective geometries over the skew fields K and L, respectively, $n \geq 2$. There exists a correlation $F : \mathfrak{P}^n \to \mathfrak{Q}^m$ if and only if $n = m$ and if, in addition, there is an anti-isomorphism $\sigma : K \to L$. If both these conditions are satisfied, then each correlation $F : \mathfrak{P}^n \to \mathfrak{Q}^n$ has the form $F = \perp \circ G$, where $G : \mathfrak{P}^n \to \mathfrak{Q}'^n$ is a collineation induced by a σ-linear bijection, where σ runs through the set of all anti-isomorphisms from K onto L.*

P r o o f. Let $F : \mathfrak{P}^n \to \mathfrak{Q}^m$ be any correlation. According to Example 1 and Corollary 1, $G := \perp \circ F$ is a collineation from the right projective geometry $\mathfrak{P}^n$ over K onto the left projective geometry $\mathfrak{Q}'^m$ over L. Let $\hat{L}$ denote the skew field opposite to L, and denote by $\hat{\mathfrak{Q}}'$ the right projective geometry that arises by passing to the opposite module structure $\hat{\boldsymbol{W}}'$ over the vector space $\boldsymbol{W}'$ associated with $\mathfrak{Q}'$. Since the lattices $\hat{\mathfrak{Q}}'$ and $\mathfrak{Q}'$ coincide, $G : \mathfrak{P} \to \hat{\mathfrak{Q}}'$ is a collineation of right projective geometries. Hence, by Proposition 3.6 we have $n = m$, and from the Main Theorem 3.10 we conclude the existence of an isomorphism $\sigma : K \to \hat{L}$ and a σ-linear map $a : \boldsymbol{V} \to \hat{\boldsymbol{W}}'$ inducing G due to $G([\mathfrak{x}]) = [a(\mathfrak{x})]$. Returning now to the original structures, because of $\sigma(\alpha \cdot \beta) = \sigma(\alpha) \times \sigma(\beta) = \sigma(\beta) \cdot \sigma(\alpha)$, the map $\sigma : K \to L$ is an anti-isomorphism, $a : \boldsymbol{V} \to \boldsymbol{W}'$ a σ-linear bijection, and, moreover, $F = \perp \circ G$. The remaining assertions are obvious. □

Note that the map F from Proposition 2 is also induced by a σ_1-linear map a_1:

$$a_1(\mathfrak{x}) = a(\mathfrak{x}\mu), \; \sigma_1(\xi) = \sigma(\mu^{-1}\xi\mu), \; \mu \in K^*, \tag{4}$$

and that σ is only determined up to an inner automorphism of K.

Exercise 1. Formulate and prove a result similar to Proposition 2 for the case that $\mathfrak{P}^n$ is a right, and $\mathfrak{Q}^m$ is a left projective geometry. In the case of the canonical correlation $F = \perp$ we have $G = \perp \circ \perp = \mathrm{id}_{\mathfrak{P}}$.

1.7.2 Correlative Maps

In analogy with collinear maps we may also introduce correlative maps:

Definition 2. Let $\mathfrak{P}^n = \mathfrak{P}(\boldsymbol{V})$, $\mathfrak{Q}^m = \mathfrak{Q}(\boldsymbol{W})$ be projective geometries over the skew fields K and L, respectively. A map $F : \mathfrak{P} \to \mathfrak{Q}$ is called *correlative* if there is a σ-linear map $a : \boldsymbol{V} \to \boldsymbol{W}'$ such that

$$F(\boldsymbol{A}) = a(\boldsymbol{A})^{\perp}, \; \boldsymbol{A} \in \mathfrak{P}^n; \tag{5}$$

a is called a σ-linear map *inducing* F. $\sigma : K \to L$ is an isomorphism or an anti-isomorphism depending on whether $\boldsymbol{V}$ and $\boldsymbol{W}$ are opposite or like vector spaces. By Proposition 2 the map F is a *correlation* if a is a σ-linear bijection (in the case $n \geq 2$); in case $n = 1$, we always suppose that correlations are correlative bijections. □

For brevity, we only formulated an algebraic definition for correlative maps; it is clear that, just as collinear maps (cf. Section 1.3.1), they should also have a geometric definition. Composing them, instead, with the canonical correlation and, this way, reducing the situation to the case of collinear maps turns out to be easier. By Proposition 2, every correlation is a correlative map as well. Obviously, each correlative map is monotonously decreasing:

$$\boldsymbol{A} \subset \boldsymbol{B} \text{ implies } F(\boldsymbol{A}) \supset F(\boldsymbol{B}), \; \boldsymbol{A}, \boldsymbol{B} \in \mathfrak{P}^n. \tag{6}$$

Exercise 2. Let $F : \mathfrak{P}^n \to \mathfrak{Q}^m$ be the correlative map induced by the σ-linear map $a : \boldsymbol{V} \to \boldsymbol{W}'$. Prove that if $a_1 : \boldsymbol{V} \to \boldsymbol{W}'$ is a σ_1-linear map also inducing F, then there is $\mu \in K^*$ such that (4) holds.

Example 2. Let $F : \mathfrak{P}^n \to \mathfrak{Q}^n$ be a correlation. Proposition 2 with $G := \perp \circ F$ implies for each k-plane $\boldsymbol{H} \in \mathfrak{P}^n$:

$$\operatorname{Dim} F(\boldsymbol{H}^k) = \operatorname{Dim}(G(\boldsymbol{H}^k)^{\perp}) = n - k - 1.$$

In particular, the image of a line $\boldsymbol{H}^1$ is an $(n-2)$-plane, and the set of points incident with $\boldsymbol{H}^1$ is transformed into the pencil of hyperplanes incident with $F(\boldsymbol{H}^1)$. By Exercise 6.4, these hyperplanes are mapped into the "points" of a line L, on which, consequently, a CR is defined. For a correlation to be *projective*, i.e. to preserve the CR, it is necessary that $K = L$ and $\sigma = \mathrm{id}_K$; moreover, if $\mathfrak{P}^n$ and $\mathfrak{Q}^n$ both are right projective geometries, then K has to be a field, in addition. □

Example 3. Let $K = L = \mathbf{R}$ be the field of real numbers. Then there is just the isomorphism $\sigma = \mathrm{id}_{\mathbf{R}}$, cf. Exercise I.2.1.3). Each correlative map $F : \mathfrak{P}^n \to \mathfrak{Q}^m$ is thus induced by a linear map $a : \boldsymbol{V} \to \boldsymbol{W}'$. □

Example 4. In the case of complex numbers, $K = L = \mathbf{C}$, as for collineations, we confine ourselves to the continuous automorphisms of $\mathbf{C}$, cf. Example 4.3. Accordingly, there are two types of correlative maps: Taking $\sigma = \mathrm{id}_{\mathbf{C}}$ we obtain the *(proper) correlative maps*, and choosing the conjugation as σ leads to the *anti-correlative maps.* □

Example 5. Let $K = L = \mathbf{H}$ be the skew field of quaternions, and let $\mathfrak{P}^n$ be a right projective geometry over $\mathbf{H}$. If $\mathfrak{Q}^m$ is another right projective geometry over $\mathbf{H}$, and the correlative map $F : \mathfrak{P}^n \to \mathfrak{Q}^m$ is induced by a σ-linear map, then we may always renormalize according to (4) and arrive at the conjugation $\tau : q \mapsto \bar{q}$ of $\mathbf{H}$ as σ_1 and an *anti-linear* map a_1, i.e.

$$a_1(\mathfrak{x}\lambda) = \bar{\lambda} a_1(\mathfrak{x}),\ \ \lambda \in \mathbf{H};$$

recall Example 4.4 and the fact that $\tau \circ \sigma$ is an isomorphism of $\mathbf{H}$. Similarly, the same example implies that F may be induced by a linear map $a : \boldsymbol{V} \to \boldsymbol{W}'$, if $\mathfrak{Q}^m$ is a left projective geometry over $\mathbf{H}$. □

If $F : \mathfrak{P}^n \to \mathfrak{Q}^m$ is correlative, then $F(\boldsymbol{o}) = \boldsymbol{Q}^m$. For a point $\boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{P}^n$ we have

$$F(\boldsymbol{x}) = \boldsymbol{Q}^m \Longleftrightarrow \mathfrak{x} \in \operatorname{Ker} a; \tag{7}$$

here, F is induced by a. Moreover,

$$\operatorname{Dim} F(\boldsymbol{x}) = m - 1 \Longleftrightarrow \boldsymbol{x} = [\mathfrak{x}] \text{ and } \mathfrak{x} \notin \operatorname{Ker} a. \tag{8}$$

More generally, *the kernel of* F defined by $\operatorname{Ker} F := \operatorname{Ker} a$ is a projective subspace of dimension

$$\operatorname{Dim} \operatorname{Ker} F = n - \operatorname{rk} a. \tag{9}$$

If $\operatorname{rk} a = 0$, i.e. $\operatorname{Ker} F = \boldsymbol{P}^n$, then F is called the *correlative null map*; it is characterized by $F(\boldsymbol{A}) = \boldsymbol{Q}^m$ for all $\boldsymbol{A} \in \mathfrak{P}^n$. For any correlative map F and each projective subspace $\boldsymbol{H} \in \mathfrak{P}^n$ we have

$$\operatorname{Dim} F(\boldsymbol{H}) = m - \operatorname{Dim} \boldsymbol{H} + \operatorname{Dim}(\boldsymbol{H} \wedge \operatorname{Ker} F). \tag{10}$$

Note that, in this relation, $F(\boldsymbol{H})$ appears as the support of the hyperplane pencil

$$\tau(F(\boldsymbol{H})) = \{F(\boldsymbol{x}) | \boldsymbol{x} \in \boldsymbol{H}\};$$

in particular, F is surjective if and only if $F(\boldsymbol{P}^n) = \boldsymbol{o}$.

Definition 3. Let $F : \mathfrak{P}^n \to \mathfrak{Q}^m$ be a correlative map induced by the σ-linear map $a : \boldsymbol{V} \to \boldsymbol{W}'$. Then the correlative map induced by the dual σ^{-1}-linear map $a' : \boldsymbol{W} \to \boldsymbol{V}'$ is called the map *dual to* F; it is denoted by

$$F' : \boldsymbol{X} \in \mathfrak{Q}^m \mapsto F'(\boldsymbol{X}) := a'(\boldsymbol{X})^{\perp} \in \mathfrak{P}^n. \tag{11}$$

□

Proposition 3. *Let* $F : \mathfrak{P}^n \to \mathfrak{Q}^m$ *be correlative. Then the dual map* F' *does not depend on the choice of the* σ*-linear map* a *inducing* F*. Moreover,*

$$F'(\boldsymbol{Q}^m) = \operatorname{Ker} F, \tag{12}$$

$$F \circ F'(\boldsymbol{Y}) \supset \boldsymbol{Y}, \; F' \circ F(\boldsymbol{X}) \supset \boldsymbol{X}, \; \textit{for } \boldsymbol{Y} \in \mathfrak{Q}^m, \, \boldsymbol{X} \in \mathfrak{P}^n, \tag{13}$$

$$(F')' = F. \tag{14}$$

If F *is a correlation, then*

$$F' = F^{-1}. \tag{15}$$

Proof. The first statement follows from Proposition 6.4. Furthermore, by (11) with $\boldsymbol{X} = \boldsymbol{Q}^m$ and (6.19) with $\boldsymbol{U} = \boldsymbol{W}$ we have

$$F'(\boldsymbol{Q}^m) = a'(\boldsymbol{W})^{\perp} = \operatorname{Ker} a = \operatorname{Ker} F.$$

The first relation in (13) also follows from the definitions using (6.19),

$$F(F'(\boldsymbol{Y})) = (a(a'(\boldsymbol{Y})^{\perp}))^{\perp} = (a(a^{-1}(\boldsymbol{Y}^{\perp})))^{\perp} \supset \boldsymbol{Y},$$

since $a(a^{-1}(\boldsymbol{Y}^{\perp})) \subset \boldsymbol{Y}^{\perp}$. The second part of (13) is a consequence of the first combined with equation (14), which immediately follows from Definition 3. If F is a correlation, then this is also true of F', and, for dimensional reasons, equality has to hold in both parts of (13), implying (15). □

1.7.3 F-Correspondences and σ-Biforms

Now we want to express correlative maps in terms of the points $\boldsymbol{x} \in \boldsymbol{P}_o^n$, $\boldsymbol{y} \in \boldsymbol{Q}_o^m$. To this end, we start with the following definition.

Definition 4. Let $F : \mathfrak{P}^n \to \mathfrak{Q}^m$ be a correlative map. Then the pair $(\boldsymbol{x}, \boldsymbol{y}) \in \boldsymbol{P}_o^n \times \boldsymbol{Q}_o^m$ is said to *correspond under the map* F, or to *belong to the* F*-correspondence* if $\boldsymbol{y} \,\iota F(\boldsymbol{x})$. □

Corollary 4. *F-correspondences have the following properties:*
a) The set $\{\boldsymbol{y} \in \boldsymbol{Q}_o^m | \boldsymbol{y} \,\iota F(\boldsymbol{x})\}$ is the hyperplane $F(\boldsymbol{x}) \subset \boldsymbol{Q}_o^m$ considered as a point set, if $\boldsymbol{x} \notin \operatorname{Ker} F$, and it is the whole space $\boldsymbol{Q}_o^m$, if $\boldsymbol{x} \in \operatorname{Ker} F$.
b) In general,

$$\boldsymbol{y} \,\iota F(\boldsymbol{x}) \iff \boldsymbol{x} \,\iota F'(\boldsymbol{y}); \tag{16}$$

hence for fixed $\boldsymbol{y}$ the set $\{\boldsymbol{x} \in \boldsymbol{P}_o^n | \boldsymbol{y} \,\iota F(\boldsymbol{x})\}$ is the hyperplane $F'(\boldsymbol{y}) \subset \boldsymbol{P}_o^n$ considered as a point set, if $\boldsymbol{y} \notin F(\boldsymbol{P}_o^n)$, and all of $\boldsymbol{P}_o^n$, otherwise. □

$>$From now on, we will again suppose that $K = L$ is an arbitrary skew field and σ an anti-automorphism of K. Let $\mathfrak{P}^n = \mathfrak{P}(\boldsymbol{V}^{n+1})$, $\mathfrak{Q}^m = \mathfrak{P}(\boldsymbol{W}^{m+1})$ be projective geometries associated with right vector spaces over K, and let $F : \mathfrak{P}^n \to \mathfrak{Q}^m$ be a correlative map induced by the σ-linear map $a : \boldsymbol{V} \to \boldsymbol{W}'$. Then we define

$$b(\mathfrak{x}, \mathfrak{y}) := (a(\mathfrak{x})|\mathfrak{y}), \ (\mathfrak{x}, \mathfrak{y}) \in \boldsymbol{V} \times \boldsymbol{W}. \tag{17}$$

Lemma 5. *The map $b : \boldsymbol{V} \times \boldsymbol{W} \to K$ defined by* (17) *has the following properties:*
a) b is linear in $\mathfrak{y}$, i.e.

$$b(\mathfrak{x}, \mathfrak{y}\alpha + \mathfrak{z}\beta) = b(\mathfrak{x}, \mathfrak{y})\alpha + b(\mathfrak{x}, \mathfrak{z})\beta,$$

b) b is σ-linear in $\mathfrak{x}$, i.e.

$$b(\mathfrak{x}\alpha + \mathfrak{u}\beta, \mathfrak{y}) = \sigma(\alpha)b(\mathfrak{x}, \mathfrak{y}) + \sigma(\beta)b(\mathfrak{u}, \mathfrak{y}). \tag{18}$$

□

Definition 5. Let $\boldsymbol{V}$, $\boldsymbol{W}$ be right vector spaces over the skew field K. The map
$b : \boldsymbol{V} \times \boldsymbol{W} \to K$ is called a *biform* or, more precisely, a σ*-biform* if it has properties a) and b) from Lemma 5. □

Bilinear or hermitian forms are examples of σ-biforms. In the literature, instead of σ-biforms the term *sesqui-linear form* is frequently used. If $a_1 : \boldsymbol{V} \to \boldsymbol{W}'$ is a map that also induces the correlative map F, then (4) holds (cf. Exercise 2), and a_1 determines the σ_1-biform

$$b_1(\mathfrak{x}, \mathfrak{y}) := (a_1(\mathfrak{x})|\mathfrak{y}) = \sigma(\mu)b(\mathfrak{x}, \mathfrak{y}). \tag{19}$$

Two biforms b, b_1 differing only in a factor $\nu \in K^*$ will be called *proportional.* If b is a σ-biform, and $b_1 = \nu b$ is proportional to b, then b_1 is a σ_1-biform with $\sigma_1(\alpha) = \nu\sigma(\alpha)\nu^{-1}$. Obviously, each biform $b \neq 0$ uniquely determines the corresponding anti-automorphism of K. Biforms are particularly well suited to describe correlative maps as correspondences:

Proposition 6. *For each correlative map $F : \mathfrak{P}^n \to \mathfrak{Q}^m$ between right projective geometries over the same skew field K there exists a biform $b : \boldsymbol{V} \times \boldsymbol{W} \to K$, which is uniquely determined up to proportionality, such that*

$$[\mathfrak{y}]\,\iota F([\mathfrak{x}]) \Longleftrightarrow b(\mathfrak{x}, \mathfrak{y}) = 0,\ (\mathfrak{x}, \mathfrak{y}) \in \boldsymbol{V} \times \boldsymbol{W}. \tag{20}$$

Conversely, according to (20), *to each class of mutually proportional biforms there corresponds a correlative map $F : \mathfrak{P}^n \to \mathfrak{Q}^m$.*

Proof. The first part of the assertion follows from Lemma 5 combined with the argument leading to (19). If, conversely, b is a biform, then it defines a σ-linear map a inducing F by

$$\mathfrak{x} \in \boldsymbol{V} \mapsto a(\mathfrak{x}) \in \boldsymbol{W}',\ a(\mathfrak{x})(\mathfrak{y}) := b(\mathfrak{x}, \mathfrak{y}),\ \mathfrak{y} \in \boldsymbol{W}. \tag{21}$$

But proportional biforms define the same correlative map. □

In the above notations, Corollary (4) and (20) immediately imply

$$\boldsymbol{y}\,\iota F(\boldsymbol{x}) \Longleftrightarrow b(\mathfrak{x}, \mathfrak{y}) = 0 \Longleftrightarrow \boldsymbol{x}\,\iota F'(\boldsymbol{y}), \tag{22}$$

so that we can also describe the dual correlative map by b. On the other hand, applying (17) to the dual map a' we see that, because of (6.16), the *transposed biform* corresponds to F',

$$b'(\mathfrak{y}, \mathfrak{x}) := \sigma^{-1}(b(\mathfrak{x}, \mathfrak{y})),\ (\mathfrak{x}, \mathfrak{y}) \in \boldsymbol{V} \times \boldsymbol{W}. \tag{23}$$

In fact, we have $(a'(\mathfrak{y})|\mathfrak{x}) = \sigma^{-1}((a(\mathfrak{x})|\mathfrak{y}))$. Obviously, b' is σ^{-1}-linear in $\mathfrak{y}$ and linear in $\mathfrak{x}$.

Next we want to describe the representation of correlative maps in homogeneous coordinates. Consider a pair of dual bases,

$$(\mathfrak{a}_i),\ (\mathfrak{a}^j),\ (\mathfrak{a}_i|\mathfrak{a}^j) = \delta_i^j,\quad i, j = 0, \dots, n,$$
$$(\mathfrak{b}_\alpha),\ (\mathfrak{b}^\beta),\ (\mathfrak{b}_\alpha|\mathfrak{b}^\beta) = \delta_\alpha^\beta,\quad \alpha, \beta = 0, \dots, m,$$

determining coordinate systems on $\boldsymbol{P}^n$ and $\boldsymbol{Q}^m$, respectively. Here we will write the coordinates (x^i) of a vector $\mathfrak{x}$ as a column and those of a dual vector as a row; the sum convention will again be used. The matrix $(a_{i\,\alpha})$ describing the σ-linear map $a : \boldsymbol{V} \to \boldsymbol{W}'$ is defined by

$$a(\mathfrak{a}_i) = a_{i\,\alpha}\mathfrak{b}^\alpha,\ i = 0, \dots, n,\ \alpha = 0, \dots, m. \tag{24}$$

If the coordinates of $\mathfrak{x} \in \boldsymbol{V}$ and $\mathfrak{u} = a(\mathfrak{x}) \in \boldsymbol{W}'$ are (x^i) and (u_α), respectively, then

$$u_\alpha = \sigma(x^i)a_{i\,\alpha}; \tag{25}$$

(u_α) are the homogeneous coordinates for the hyperplane $F(\boldsymbol{x})$ corresponding to the point $\boldsymbol{x} = [\mathfrak{x}]$ with homogeneous coordinates (x^i). At the same time, $(a_{i\alpha}) \in M_{n+1,m+1}(K)$ is the matrix of the σ-biform b corresponding to F:

$$b(\mathfrak{x}, \mathfrak{y}) = (a(\mathfrak{x})|\mathfrak{y}) = \sigma(x^i)a_{i\alpha}y^\alpha, \; a_{i\alpha} = b(\mathfrak{a}_i, \mathfrak{b}_\alpha). \tag{26}$$

Here, (y^α) are the coordinates of $\mathfrak{y}$. For the map F' dual to F a straightforward calculation using (24) and (6.16) yields

$$a'(\mathfrak{b}_\alpha) = a'_{\alpha i}\mathfrak{a}^i = \sigma^{-1}(a_{i\alpha})\mathfrak{a}^i, \tag{27}$$

i.e., the matrix $(a'_{\alpha i})$ of a' coincides with the transposed of the matrix obtained by transforming that of a by means of σ^{-1}:

$$(a'_{\alpha i}) = (\sigma^{-1}(a_{i\alpha}))'. \tag{28}$$

Exercise 3. Let $(a_{i\alpha})$ be the matrix of the correlative map $F : \mathfrak{P}^n \to \mathfrak{Q}^m$ according to (24). Prove the relation

$$\operatorname{Dim}\operatorname{Ker} F = n - r, \; \operatorname{Dim}\operatorname{Ker} F' = m - r, \tag{29}$$

where r is the rank of $(a_{i\alpha})$ (cf. Exercise II.7.3.7). This implies that r does not depend on the chosen coordinates; the integer r is called the *rank* of F.

Exercise 4. Let $\mathfrak{B} := \mathfrak{B}(\mathfrak{P}^n, \mathfrak{Q}^m)$ denote the set of all correlative maps from $\mathfrak{P}^n$ to $\mathfrak{Q}^m$. Consider the action of the group $\boldsymbol{G} := \boldsymbol{PL}(\boldsymbol{P}^n) \times \boldsymbol{PL}(\boldsymbol{Q}^m)$ of pairs of projectivities on $\mathfrak{B}$ defined by

$$(g_1, g_2) \in \boldsymbol{G}, \; F \in \mathfrak{B} \mapsto g_2 \circ F \circ g_1^{-1} \in \mathfrak{B}. \tag{30}$$

Prove that two correlative maps F, $\tilde{F} \in \mathfrak{B}$ are equivalent under this action of $\boldsymbol{G}$ if and only if they have equal rank and determine the same class $\hat{\sigma}$ in the group $\operatorname{Aut}(K)/\operatorname{Int}(K)$, cf. Exercise 4.11 and (4).

The following proposition will allow to define the *restriction* $F|\boldsymbol{A}$ of an auto-correlative map $F : \mathfrak{P}^n \to \mathfrak{P}^n$ to a subspace $\boldsymbol{A} \subset \boldsymbol{P}^n$:

Proposition 7. *Consider* $\mathfrak{P}^n = \mathfrak{P}(\boldsymbol{V}^{n+1})$, $\boldsymbol{A} \in \mathfrak{P}^n$ *with* $\boldsymbol{A} = \pi(\boldsymbol{W})$ *for a subspace* $\boldsymbol{W} \subset \boldsymbol{V}$ *of the vector space* $\boldsymbol{V}$. *Let* $F : \mathfrak{P}^n \to \mathfrak{P}^n$ *be an auto-correlative map induced by the* σ*-linear map* $a : \boldsymbol{V} \to \boldsymbol{V}'$. *Then*

$$\boldsymbol{x} \in \boldsymbol{A} \mapsto F|\boldsymbol{A}(\boldsymbol{x}) := F(\boldsymbol{x}) \wedge \boldsymbol{A} \subset \boldsymbol{A} \tag{31}$$

defines an auto-correlative map $F|\boldsymbol{A}$ *that is induced by the* σ*-linear map*

$$\mathfrak{x} \in \boldsymbol{W} \mapsto a(\mathfrak{x})|\boldsymbol{W} \in \boldsymbol{W}'. \tag{32}$$

If b *is a* σ*-biform describing* F, *then the restricted* σ*-biform*

$$b_{\boldsymbol{A}} := b|\boldsymbol{W} \times \boldsymbol{W} \tag{33}$$

describes the restriction $F|\boldsymbol{A}$. □

The proof immediately follows from the above definitions. Looking at examples it is obvious that $F|\boldsymbol{A}$ may well be the correlative zero map, even for $\operatorname{Ker} F = \boldsymbol{o}$ (cf. also Section 1.10 below).

A notion dual to itself is that of a mixed throw:

Definition 6. Let $\mathfrak{P}^n$, $n \geq 2$, be a projective geometry over the skew field K, let $\boldsymbol{a}, \boldsymbol{b} \in \mathfrak{P}^n$ be points, and let $\boldsymbol{C}, \boldsymbol{D} \in \mathfrak{P}^n$ be hyperplanes, $\boldsymbol{a} \neq \boldsymbol{b}$, $\boldsymbol{C} \neq \boldsymbol{D}$. Set

$$\boldsymbol{c} := \boldsymbol{C} \wedge (\boldsymbol{a} \vee \boldsymbol{b}),\ \boldsymbol{d} := \boldsymbol{D} \wedge (\boldsymbol{a} \vee \boldsymbol{b}).$$

The quadruple $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{C}, \boldsymbol{D})$ is called a *mixed throw* if $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d})$ is a throw. The *cross ratio of the mixed throw* is defined as (cf. L. Heffter [48], §18)

$$(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{C}, \boldsymbol{D}) := (\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{c}, \boldsymbol{d}). \tag{34}$$

□

Exercise 5. Under the assumptions of Definition 6 show the following: A quadruple $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{C}, \boldsymbol{D})$ is a mixed throw if and only if $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D})$ is a throw of hyperplanes, where

$$\boldsymbol{A} := \boldsymbol{a} \vee (\boldsymbol{C} \wedge \boldsymbol{D}),\ \boldsymbol{B} := \boldsymbol{b} \vee (\boldsymbol{C} \wedge \boldsymbol{D})$$

(cf. Exercises 6.3, 6.4). In addition, for the CR we have

$$(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{C}, \boldsymbol{D}) = (\boldsymbol{A}, \boldsymbol{B}; \boldsymbol{C}, \boldsymbol{D}). \tag{35}$$

Exercise 6. Let $\mathfrak{P}^n$, $\mathfrak{Q}^m$, $n \geq m \geq 2$, be right projective geometries over the skew field K; moreover, let $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{C}, \boldsymbol{D})$ and $(\boldsymbol{a}_1, \boldsymbol{b}_1, \boldsymbol{C}_1, \boldsymbol{D}_1)$ be mixed throws in $\mathfrak{P}^n$ and $\mathfrak{Q}^m$, respectively. Prove that there is a projective map $F : \mathfrak{P}^n \to \mathfrak{Q}^m$ with

$$F(\boldsymbol{a}) = \boldsymbol{a}_1,\ F(\boldsymbol{b}) = \boldsymbol{b}_1,\ F(\boldsymbol{C}) = \boldsymbol{C}_1,\ F(\boldsymbol{D}) = \boldsymbol{D}_1,$$

if and only if their CRs coincide:

$$(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{C}, \boldsymbol{D}) = (\boldsymbol{a}_1, \boldsymbol{b}_1; \boldsymbol{C}_1, \boldsymbol{D}_1). \tag{36}$$

Exercise 7. In the notations of Exercise 6 with the additional assumptions that $n = m$ and K is a field, show the following: There is a projective correlation $F : \mathfrak{P}^n \to \mathfrak{Q}^n$ with

$$F(\boldsymbol{a}) = \boldsymbol{C}_1,\ F(\boldsymbol{b}) = \boldsymbol{D}_1,\ F(\boldsymbol{C}) = \boldsymbol{a}_1,\ F(\boldsymbol{D}) = \boldsymbol{b}_1,$$

if and only if (36) holds.

1.8 Symmetric Auto-Correlative Maps

In this section we will study auto-correlative maps $F : \mathfrak{P}^n \to \mathfrak{P}^n$. Such a map is called *symmetric* if it coincides with its dual, i.e., if

$$F = F'. \tag{1}$$

This name is suggested by the fact that the associated F-correspondence is a symmetric relation over $\boldsymbol{P}_o^n$: By (7.16) F is symmetric if and only if

$$\boldsymbol{x} \;\iota F(\boldsymbol{y}) \Longleftrightarrow \boldsymbol{y} \;\iota F(\boldsymbol{x}),\ \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{P}_o^n. \tag{2}$$

There are two types of symmetric auto-correlative maps: the null systems and the polar maps. Now we are going to obtain a projective classification of these maps; for the null systems this will be completely accomplished in this paragraph, whereas for the polar maps we will achieve a classification of important special cases in the next paragraph. Null systems form the basis for symplectic geometry; we will also interpret them geometrically as linear line complexes.

Remarks concerning the terminology. The obvious name "symmetric" for property (1) appears to have rarely (or even never?) been used until now. The reader may forgive us, that we add still another one to the various names already to be found in the literature. The name "reflexive" chosen by J. Dieudonné [36], § I.6, does not seem very fortunate to us, since the associated F-correspondence is reflexive only for the null systems (compare Definition I.0.2.7, Property 1, with Definition 1 below). W. Burau [24] calls the correlative maps satisfying (1) "involutory", even though by 7.3, (7.15), only the symmetric correlations share the property $F \circ F = \mathrm{id}_{\mathfrak{P}^n}$ characterizing involutions; these symmetric correlations are called "polar maps" by R.Baer [5], thereby possibly making the notational confusion complete. On the other hand, R. Baer uses the name "dual map" instead of correlation, and "self-dual map" instead of auto-correlation, so that we preferred to avoid the obvious name "auto-dual" for (1).

1.8.1 Null Systems and Polar Maps

Definition 1. Let $F : \mathfrak{P}^n \to \mathfrak{P}^n$ be a symmetric correlative map. The points in the kernel, $\boldsymbol{x} \in \operatorname{Ker} F$, are called *singular* and the remaining ones, $\boldsymbol{x} \notin \operatorname{Ker} F$, *regular* points of F; a point $\boldsymbol{x} \notin \operatorname{Ker} F$ is called a *pole* of the hyperplane $F(\boldsymbol{x})$, and $F(\boldsymbol{x})$ the *polar* of $\boldsymbol{x}$ for the correlative map F. An element $\boldsymbol{X} \in \mathfrak{P}^n$ is called *auto-polar* if $\boldsymbol{X} \;\iota F(\boldsymbol{X})$. Let $Q_F \subset \boldsymbol{P}^n$ denote the set of auto-polar points of F. If $Q_F = \boldsymbol{P}^n$, then F is called a *null system*; otherwise, F is a *polar map*. Moreover, we will call F *non-degenerate* if $\operatorname{Ker} F = \boldsymbol{o}$, i.e., if F is a correlation; a polar correlation is a *polarity*. □

Note that, according to Definition 7.2, every auto-correlative map F is induced by a semi-linear map $a : \boldsymbol{V} \to \boldsymbol{V}'$. Obviously, the polar of a regular point is uniquely determined; conversely, for the pole this only holds in the case of a non-degenerate F: If $\boldsymbol{x} \notin \operatorname{Ker} F$, then every point $\boldsymbol{y} \in (\boldsymbol{x} \vee \operatorname{Ker} F) \setminus \operatorname{Ker} F$ is a pole of $F(\boldsymbol{x})$; in fact, $\boldsymbol{P}^n \neq F(\boldsymbol{y}) \supset F(\boldsymbol{x}) \wedge F(\operatorname{Ker} F) = F(\boldsymbol{x})$.

Corollary 1. *An auto-correlation $F : \mathfrak{P}^n \to \mathfrak{P}^n$ is symmetric if and only if it is involutive:*

$$F^2 = \mathrm{id}_{\mathfrak{P}^n}\,.$$

Proof. The assertion immediately follows from Definition (1) and (7.15): $F = F' = F^{-1}$. □

The *symmetric correlations* F thus are the involutions in projective geometry $\mathfrak{P}^n$: The equality $\boldsymbol{A} = F(\boldsymbol{B})$ holds if and only if $\boldsymbol{B} = F(\boldsymbol{A})$; in this case, $\boldsymbol{A}, \boldsymbol{B}$ *are called mutually polar under* F. For the dimensions of a polar pair we have (cf. Example 7.2)

$$\operatorname{Dim} \boldsymbol{A} = k \Longleftrightarrow \operatorname{Dim} F(\boldsymbol{A}) = n - k - 1. \tag{3}$$

By means of the following proposition, the symmetric correlative maps will be reduced to the symmetric correlations:

Proposition 2. *Let F be a symmetric correlative map of the projective geometry $\mathfrak{P}^n$. Then*

$$\operatorname{Ker} F = F(\boldsymbol{P}^n), \tag{4}$$

and, moreover,

$$\operatorname{Ker} F \subset F(\boldsymbol{A}) \text{ for all } \boldsymbol{A} \in \mathfrak{P}^n. \tag{5}$$

F determines the symmetric correlation of the bundle of subspaces through $\operatorname{Ker} F$*:*

$$\hat{F} : \boldsymbol{X} \in \mathfrak{P}^n / \operatorname{Ker} F \longmapsto F(\boldsymbol{X}) \in \mathfrak{P}^n / \operatorname{Ker} F. \tag{6}$$

Proof. By Proposition 7.3, (7.12), (4) holds because of $F = F'$. By (7.6), Equation (5) for $\boldsymbol{A} \subset \boldsymbol{P}^n$ follows from (4). Now recall the definition of the projective bundle geometries, Section 1.6.3. Because of (5), the map $\hat{F}$ is well-defined. Let F be induced by the σ-linear map $a : \boldsymbol{V} \to \boldsymbol{V}'$ with $\operatorname{Ker} F = \operatorname{Ker} a =: \boldsymbol{W}$; then $\hat{F}$ is induced by the canonically associated, σ-linear bijection $\hat{a} : \boldsymbol{V}/\boldsymbol{W} \to (\boldsymbol{V}/\boldsymbol{W})' \cong \boldsymbol{W}^\perp$ that is defined by

$$(\hat{a}(\mathfrak{x} + \boldsymbol{W}) | \mathfrak{y} + \boldsymbol{W}) := (a(\mathfrak{x}) | \mathfrak{y}). \tag{7}$$

We still have to show that $\hat{F}$ is symmetric, i.e. $(\hat{F})' = \hat{F}$. By definition of the dual map, 6.1, (6.16), it is induced by the dual semi-linear map, which is σ^{-1}-linear:

$$((\hat{a})'(\mathfrak{x} + \boldsymbol{W}) | \mathfrak{y} + \boldsymbol{W}) = \sigma^{-1}((\hat{a}(\mathfrak{y} + \boldsymbol{W}) | \mathfrak{x} + \boldsymbol{W})) = \sigma^{-1}((a(\mathfrak{y}) | \mathfrak{x})) = (a'(\mathfrak{x}) | \mathfrak{y}).$$

Since F is symmetric, the dual semi-linear map a' also induces F. Hence, according to the uniqueness statement in the Main Theorem 3.10, (3.17), there is $\mu \in K^*$ such that $a'(\mathfrak{x}) = a(\mathfrak{x}\mu)$. Inserting this into the last equation yields

$$((\hat{a})'(\mathfrak{x} + \boldsymbol{W})|\mathfrak{y} + \boldsymbol{W}) = (a(\mathfrak{x}\mu)|\mathfrak{y}) = (\hat{a}(\mathfrak{x}\mu + \boldsymbol{W})|\mathfrak{y} + \boldsymbol{W}),$$

i.e., $(\hat{a})'(\mathfrak{x} + \boldsymbol{W}) = \hat{a}((\mathfrak{x} + \boldsymbol{W})\mu)$, implying the symmetry of $\hat{F}$. □

1.8.2 Equivalence of Auto-Correlative Maps

Now we define the projective equivalence of two auto-correlative maps:

Definition 2. Two auto-correlative maps F_α, $\alpha = 1, 2$, of the projective geometries $\mathfrak{P}^n_\alpha$ are called *projectively equivalent*, $F_1 \sim F_2$, if there is a projectivity $g : \mathfrak{P}^n_1 \to \mathfrak{P}^n_2$ such that

$$F_2 = g \circ F_1 \circ g^{-1}. \tag{8}$$

□

Again, we may and will consider only the case $\mathfrak{P}^n_1 = \mathfrak{P}^n_2 = \mathfrak{P}(\boldsymbol{V})$. Let $a_1 : \boldsymbol{V} \to \boldsymbol{V}'$ be a σ_1-linear map inducing F_1, and let the linear transformation $c \in \mathbf{GL}(\boldsymbol{V})$ induce the projectivity g; by $c^* := c'^{-1}$ we denote the transformation contragredient to c. Since

$$(c^*\mathfrak{y}|c\mathfrak{x}) = (\mathfrak{y}|\mathfrak{x}), \tag{9}$$

we obtain

$$c(\boldsymbol{X})^\perp = c^*(\boldsymbol{X}^\perp),\ c(\boldsymbol{Y}^\perp) = c^*(\boldsymbol{Y})^\perp,\ \boldsymbol{X} \subset \boldsymbol{V},\ \boldsymbol{Y} \subset \boldsymbol{V}'. \tag{10}$$

Because of (8) and (10) the likewise σ_1-linear map $c^* \circ a_1 \circ c^{-1}$ induces the map F_2; in fact,

$$c^* \circ a_1 \circ c^{-1}(\boldsymbol{X})^\perp = c^*(a_1(c^{-1}(\boldsymbol{X})))^\perp = c(a_1(c^{-1}(\boldsymbol{X}))^\perp) = g \circ F \circ g^{-1}(\boldsymbol{X}),$$

cf. (7.5). According to the uniqueness statement in the Main Theorem 3.10 we know: If a_2 is any σ_2-linear map inducing F_2, then $F_1 \sim F_2$ if and only if there are $c \in \mathbf{GL}(\boldsymbol{V})$ and $\mu \in K^*$ such that

$$a_2 = c^* \circ a_1 \circ c^{-1} \circ d_\mu \text{ and } \sigma_2 = \sigma_1 \circ \sigma_\mu^{-1}; \tag{11}$$

here d_μ denotes the dilation by the factor μ in $\boldsymbol{V}$, and σ_μ is the inner automorphism of K associated with μ, cf. Exercise 7.2. From (11) and Proposition 7.6 we immediately conclude: If b_α are biforms inducing the correlative maps F_α, $\alpha = 1, 2$, then $F_1 \sim F_2$ if and only if there exist $\kappa \in K^*$ as well as $c \in \mathbf{GL}(\boldsymbol{V})$ such that

$$b_2(c^{-1}(\mathfrak{x}), c^{-1}(\mathfrak{y})) = \kappa b_1(\mathfrak{x}, \mathfrak{y}),\ \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}; \tag{12}$$

hence there has to be a projectivity transforming the associated correspondences into one another. Choosing a basis $(\mathfrak{a}_i)$ in $\boldsymbol{V}$ and denoting by $(\gamma^i{}_j)$, (a_{ij}), and (b_{ij}), respectively, the matrices of c^{-1}, b_1 and b_2 with respect to this basis, then (12) and (7.26) immediately imply the corresponding properties of these matrices (Note again the sum convention!):

$$\begin{aligned} \kappa a_{ij} &= \sigma_2(\gamma^k{}_i) b_{kl} \gamma^l{}_j, \\ \kappa(a_{ij}) &= (\sigma_2(\gamma^k{}_i))'(b_{kl})(\gamma^l{}_j). \end{aligned} \tag{13}$$

Formulas (12) and (13) are generalizations of the transformation rules (I.5.9.45) and (I.5.9.81) for bilinear and Hermitean forms, respectively.

These considerations apply to an arbitrary correlative map F from $\mathfrak{P}^n$ to itself; by Proposition 7.6, (7.23), the symmetric correlative maps are characterized by the condition

$$\kappa b(\mathfrak{y}, \mathfrak{x}) = \sigma^{-1}(b(\mathfrak{x}, \mathfrak{y})), \tag{14}$$

or, in matrix form, by

$$\kappa(a_{ij}) = (\sigma^{-1}(a_{ij}))'. \tag{15}$$

Apart from the symmetric bilinear forms, the Hermitean biforms and matrices, respectively, are special cases, cf. (I.5.9.77). The correlative map F is non-degenerate if there is no $\mathfrak{x} \in \boldsymbol{V}$, $\mathfrak{x} \neq \mathfrak{o}$, such that $b(\mathfrak{x}, \mathfrak{y}) = 0$ for all $\mathfrak{y} \in \boldsymbol{V}$; biforms with this property are also called *non-degenerate*. Exercises 7.3 and 7.4 immediately yield the following necessary condition for projective equivalence:

Corollary 3. *Projectively equivalent correlative maps from $\mathfrak{P}^n$ to itself have equal rank. Such a map is a correlation if and only if it is non-degenerate, i.e. its rank is $n+1$.* □

Exercise 1. Let $F_\alpha, \alpha = 1, 2$, be symmetric, auto-correlative maps of the projective geometries $\mathfrak{P}^n_\alpha$. Prove that $F_1 \sim F_2$ if and only if the correlations $\hat{F}_1, \hat{F}_2$ associated with them according to Proposition 2 are projectively equivalent. (Hint. Consider subspaces $\boldsymbol{A}_\alpha \in \mathfrak{P}^n_\alpha$ complementary to $\operatorname{Ker} F_\alpha$; identify $\mathfrak{P}(\boldsymbol{A}_\alpha)$ with $\mathfrak{P}^n_\alpha / \operatorname{Ker} F_\alpha$ by means of

$$\boldsymbol{B} \in \mathfrak{P}(\boldsymbol{A}_\alpha) \longmapsto \boldsymbol{B} \vee \operatorname{Ker} F_\alpha \in \mathfrak{P}^n_\alpha / \operatorname{Ker} F_\alpha,$$

and realize the $\hat{F}_\alpha$ as correlations of $\mathfrak{P}(\boldsymbol{A}_\alpha)$.)

1.8.3 Classification of Null Systems

Let now F be a null system. If $b \neq 0$ is a σ-biform defining F, then for all $\mathfrak{x} \in \boldsymbol{V}$ the relation

$$b(\mathfrak{x}, \mathfrak{x}) = 0 \tag{16}$$

has to hold, from which, by inserting $\mathfrak{x} + \mathfrak{y}$ for $\mathfrak{x}$, we immediately obtain

$$b(\mathfrak{x}, \mathfrak{y}) + b(\mathfrak{y}, \mathfrak{x}) = 0. \tag{17}$$

Since $b \neq 0$, by the linearity in $\mathfrak{y}$ we find a pair $(\mathfrak{x}_o, \mathfrak{y}_o) \in \boldsymbol{V} \times \boldsymbol{V}$ with

$$b(\mathfrak{x}_o, \mathfrak{y}_o) = 1. \tag{18}$$

As b is a σ-biform, (17) and (18) imply

$$b(\mathfrak{x}_o \xi, \mathfrak{y}_o) = \sigma(\xi) = -b(\mathfrak{y}_o, \mathfrak{x}_o \xi) = \xi$$

for all $\xi \in K$. Hence $\sigma = \mathrm{id}_K$, K is a field, and b has to be an alternating bilinear form over $\boldsymbol{V}$, cf. Proposition 7.6 and Definition I.5.9.3. Thus we arrive at (cf. also Exercise I.5.9.13)

Lemma 4. *An alternating bilinear form on a finite-dimensional vector space* $\boldsymbol{V}$ *over the field* K *always has even rank* $2r$. *There is a basis* $(\mathfrak{a}_i)$ *of* $\boldsymbol{V}$ *with respect to which the matrix* (b_{ij}) *of* b *has quasi-diagonal form with* r *block matrices*

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

along the principal diagonal, and zeroes everywhere else (cf. (I.5.9.82)*):*

$$(b_{ij}) = \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & O & \dots & \dots & O \\ O & \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & O & \dots & O \\ & & \dots & & \\ & & \dots & & \\ & & & \dots & \\ O & \dots & O & \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & O \\ O & \dots & \dots & \dots & O \end{pmatrix}. \tag{19}$$

Proof. Because of (16), the assertion is trivial in the case $\dim \boldsymbol{V} = 1$: Since $b = 0$, we have $\operatorname{rk} b = 0$. For $\dim \boldsymbol{V} = 2$ the assertion follows from (18): In case $b = 0$ we have $r = 0$, and for $b \neq 0$ we set $\mathfrak{a}_1 = \mathfrak{x}_o$, $\mathfrak{a}_2 = \mathfrak{y}_o$. Suppose that we already found vectors $\mathfrak{a}_1, \dots, \mathfrak{a}_{2k} \in \boldsymbol{V}$ such that

$$b(\mathfrak{a}_{2j-1}, \mathfrak{a}_{2j}) = -b(\mathfrak{a}_{2j}, \mathfrak{a}_{2j-1}) = 1 \quad \text{for} \quad j = 1, \dots, k,$$
$$b(\mathfrak{a}_i, \mathfrak{a}_l) = 0 \quad \text{for} \quad i, l = 1, \dots, 2k, |i - l| \neq 1.$$

Consider the linear span $\boldsymbol{W}^{2k} := \mathfrak{L}(\mathfrak{a}_1, \dots, \mathfrak{a}_{2k})$ and its orthogonal complement

$$\hat{\boldsymbol{W}} := \{\mathfrak{x} \in \boldsymbol{V} | b(\mathfrak{y}, \mathfrak{x}) = 0 \text{ for all } \mathfrak{y} \in \boldsymbol{W}\}.$$

Since $b|\boldsymbol{W} \times \boldsymbol{W}$ is non-degenerate, we have $\boldsymbol{W} \cap \hat{\boldsymbol{W}} = \{\mathfrak{o}\}$, and a straightforward consideration of ranks implies $\dim \hat{\boldsymbol{W}} = \dim \boldsymbol{V} - 2k$. Hence $\boldsymbol{V} = \boldsymbol{W} \oplus \hat{\boldsymbol{W}}$. If $b|\hat{\boldsymbol{W}} \times \hat{\boldsymbol{W}} = 0$, then we complete $\mathfrak{a}_1, \dots, \mathfrak{a}_{2k}$ by means of an arbitrary basis of $\hat{\boldsymbol{W}}$ to obtain a basis for $\boldsymbol{V}$. This leads to the normal form (19) with $k = r$.

Otherwise, we find vectors $\mathfrak{a}_{2k+1}, \mathfrak{a}_{2k+2} \in \hat{\boldsymbol{W}}$ with $b(\mathfrak{a}_{2k+1}, \mathfrak{a}_{2k+2}) = 1$ and proceed like before. □

Lemma 4 and what was said before immediately lead to the projective classification of all null systems:

Proposition 5. *Let $\mathfrak{P}^n$ be an n-dimensional projective geometry over the skew field K, $n \geq 2$. There exists a null system F of $\mathfrak{P}^n$ with $\operatorname{rk} F > 0$ if and only if K is a field. The biforms determining a null system F are bilinear, alternating and of even rank. Two null systems are projectively equivalent if and only if they have the same scalar field K and equal rank. Non-degenerate null systems only exist for projective geometries of odd dimension.*

Proof. Apart from what was just said it suffices to note that null systems of equal rank also have coinciding normal forms (19). Choosing bases for F_1, F_2 in which each has normal form, and then assigning corresponding vectors defines a linear isomorphism mapping the null systems into one another. Hence it provides the projective equivalence g in (8). The converse is obvious by Corollary 3. □

1.8.4 Linear Line Complexes

A geometric interpretation of null systems can be obtained by considering the Graßmann manifold $\boldsymbol{P}_{n,1}$ of lines in $\mathfrak{P}^n$. Let F be a null system of rank $2r > 0$. Then the set of lines

$$\mathfrak{K}_F := \{\boldsymbol{H} \in \boldsymbol{P}_{n,1} | \, \boldsymbol{H} \;\iota\; F(\boldsymbol{H})\} \tag{20}$$

is called a *linear line complex*, if $n \geq 3$. Concerning the case $n = 2$ note Exercise 2.

Exercise 2. If F is a null system of rank $2r > 0$ in the projective plane $\mathfrak{P}^2$, then $r = 1$, $\operatorname{Ker} F = \boldsymbol{z}_o$ is a point, and $\mathfrak{K}_F = \tau(\operatorname{Ker} F)$ is the pencil of lines through $\boldsymbol{z}_o$.

No we will show that $\mathfrak{K}_F$ is the set of all connecting lines $\boldsymbol{x} \vee \boldsymbol{y}$ for which $\boldsymbol{x}$ is incident with $F(\boldsymbol{y})$.

Lemma 6. *If F is a null system in $\mathfrak{P}^n$, $n \geq 3$, then*

$$\mathfrak{K}_F = \{\boldsymbol{x} \vee \boldsymbol{y} | \, \boldsymbol{x}, \boldsymbol{y} \in \mathfrak{P}^n, \, \boldsymbol{x} \neq \boldsymbol{y}, \boldsymbol{x} \;\iota\; F(\boldsymbol{y})\}. \tag{21}$$

Proof. For $\boldsymbol{H} = \boldsymbol{x} \vee \boldsymbol{y} \in \mathfrak{K}_F$ (20) implies $\boldsymbol{x} \vee \boldsymbol{y} \;\iota\; F(\boldsymbol{x} \vee \boldsymbol{y}) \subset F(\boldsymbol{x}) \wedge F(\boldsymbol{y})$; in fact, F is monotonously decreasing. Hence, in particular, $\boldsymbol{x} \;\iota\; F(\boldsymbol{y})$. Conversely, assume that $\boldsymbol{x} \;\iota\; F(\boldsymbol{y})$. Since F is a null system, $\boldsymbol{y} \;\iota\; F(\boldsymbol{x})$, $\boldsymbol{x} \;\iota F(\boldsymbol{x})$, and $\boldsymbol{y} \;\iota\; F(\boldsymbol{y})$. Let now $a : \boldsymbol{V} \to \boldsymbol{V}'$ be a linear map inducing F; then $\boldsymbol{x} = [\mathfrak{x}]$ and $\boldsymbol{y} = [\mathfrak{y}]$ imply that the linear span is $\mathfrak{L}(\mathfrak{x}, \mathfrak{y}) = \boldsymbol{H}$. Thus $F(\boldsymbol{H}) = (a(\mathfrak{L}(\mathfrak{x}, \mathfrak{y})))^{\perp} = \mathfrak{L}(a(\mathfrak{x}), a(\mathfrak{y}))^{\perp}$ is characterized by

$$\boldsymbol{z} = [\mathfrak{z}] \in F(\boldsymbol{H}) \iff (a\mathfrak{x}|\mathfrak{z}) = 0 \text{ und } (a\mathfrak{y}|\mathfrak{z}) = 0.$$

From the incidences above we conclude

$$(a\mathfrak{x}|\mathfrak{x}) = (a\mathfrak{y}|\mathfrak{x}) = (a\mathfrak{x}|\mathfrak{y}) = (a\mathfrak{y}|\mathfrak{y}) = 0.$$

Hence $\boldsymbol{x} \,\iota\, F(\boldsymbol{H})$, $\boldsymbol{y} \,\iota\, F(\boldsymbol{H})$, and, since $F(\boldsymbol{H})$ is a projective subspace, $\boldsymbol{x} \vee \boldsymbol{y} \,\iota\, F(\boldsymbol{H})$, i.e. $\boldsymbol{H} \in \mathfrak{K}_F$. □

For a better understanding of the concept of a linear line complex we need an embedding of the Graßmann manifold of all lines $\boldsymbol{P}_{n,1} = \boldsymbol{G}_{n+1,2}$, cf. Definition 1.1, as an algebraic submanifold of a certain higher-dimensional projective space. To obtain such an embedding we need some simple tools from exterior algebra, see e.g. II.8.3. Let $\Lambda^2\boldsymbol{V}$ denote the space of all alternating contravariant tensors of degree 2 associated with V, also called bivectors. It is generated by the *splitting bivectors*

$$\mathfrak{x} \wedge \mathfrak{y} := \mathfrak{x} \otimes \mathfrak{y} - \mathfrak{y} \otimes \mathfrak{x}; \quad \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}.$$

If $(\mathfrak{a}_i)$, $i = 0, \dots, n$, is a basis for $\boldsymbol{V}$, then the set of all bivectors $\mathfrak{a}_i \wedge \mathfrak{a}_j$ with $0 \le i < j \le n$ is a basis for $\Lambda^2\boldsymbol{V}$; hence

$$\dim \Lambda^2\boldsymbol{V} = (n+1)n/2.$$

Thus the associated projective space $\boldsymbol{P}^N(\Lambda^2\boldsymbol{V})$ has dimension

$$N = \frac{(n+1)n}{2} - 1.$$

Now let $\mathfrak{x} = \mathfrak{a}_i x^i$, $\mathfrak{y} = \mathfrak{a}_i y^i$ be the basis representations of two vectors $\mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}$. Then we have

$$\mathfrak{x} \wedge \mathfrak{y} = \sum_{i<j} \mathfrak{a}_i \wedge \mathfrak{a}_j p^{ij} \text{ with } p^{ij} = x^i y^j - x^j y^i. \tag{22}$$

Obviously, $\mathfrak{x} \wedge \mathfrak{y} \neq 0$ if and only if the vectors are linearly independent. Set $\boldsymbol{H} = x \vee x \in \boldsymbol{P}_{n,1}$ for $x = [\mathfrak{x}]$, $y = [\mathfrak{y}]$. Then the coordinates p^{ij} of $\mathfrak{x} \wedge \mathfrak{y}$ are called the *Plücker coordinates* of the line $\boldsymbol{H}$. If $\boldsymbol{H} = a \vee b$, $a = [\mathfrak{a}]$, $b = [\mathfrak{b}]$ is another representation of $\boldsymbol{H}$, then $\mathfrak{a} \vee \mathfrak{b} = \mathfrak{x} \vee \mathfrak{y}\kappa$ for $\kappa \in K^*$. Hence the map

$$h : \boldsymbol{H} = x \vee y \in \boldsymbol{P}_{n,1} \to h(\boldsymbol{H}) := [\mathfrak{x} \wedge \mathfrak{y}] \in \boldsymbol{P}^N(\Lambda^2\boldsymbol{V}) \tag{23}$$

is correctly defined. Since each splitting bivector $\mathfrak{x} \wedge \mathfrak{y} \neq 0$ uniquely defines the line:

$$z = [\mathfrak{z}] \in \boldsymbol{H} \quad \text{if and only if} \quad \mathfrak{x} \vee \mathfrak{y} \vee \mathfrak{z} = 0,$$

the map h is injective and defines the embedding of $\boldsymbol{P}^{n,1}$ as a manifold of points in $\boldsymbol{P}^N(\Lambda^2\boldsymbol{V})$ we were looking for. In the case $n = 2$ the map h can be shown to be even surjective. For $n \ge 3$ this is not true. A general bivector

$T = \frac{1}{2}\mathfrak{a}_i \wedge \mathfrak{a}_j p^{ij}$, $p^{ij} + p^{ji} = 0$, is splitting, i.e. has the form $T = \mathfrak{a} \vee \mathfrak{b}$ for some $\mathfrak{a}, \mathfrak{b} \in \boldsymbol{V}$ if and only if its coordinates satisfy the *Plücker relations*

$$p^{i_1 i_2} p^{i_3 l} + p^{i_2 i_3} p^{i_1 l} + p^{i_3 i_1} p^{i_2 l} = 0, 0 \leq i_1 < i_1 < i_3 \leq n, l = 0, \ldots, n. \quad (24)$$

Formula (24) is a special case of II.8.3.20. The equations (24) are not independent. For $n = 3$ they reduce to a simple quadratic equation defining the *Plücker quadric* $Q \subset \boldsymbol{P}^5(\Lambda^2 \boldsymbol{V}^4)$:

$$p^{01} p^{23} + p^{12} p^{03} + p^{20} p^{13} = 0, \quad (25)$$

(see the next section).

According to Proposition 5, the null systems F of $\mathfrak{P}^n$ correspond to the alternating bilinear forms $b \in \bigwedge^2 \boldsymbol{V}'$. The latter are uniquely determined by F up to a factor $\kappa \in K^*$; in other words: The null systems of $\mathfrak{P}^n$ are described bijectively by the points of the projective space $\boldsymbol{P}(\bigwedge^2 \boldsymbol{V}')$. Let a_{ij} be the homogeneous coordinates of F (or b); then we have

$$a_{ij} + a_{ji} = 0, \; i, j = 0, \ldots, n, \quad (26)$$

and this implies:

Lemma 7. *Let F be a null system corresponding to the alternating 2-form with coordinates a_{ij}. Then $\mathfrak{K}_F$ consists of all lines $\boldsymbol{H} = \boldsymbol{x} \vee \boldsymbol{y}$ whose Plücker coordinates* (22) *satisfy the linear equation*

$$\sum_{i<j} a_{ij} p^{ij} = 0. \quad (27)$$

□

In fact, according to Lemma 6 we have $a_{ij} x y^j = 0$, and (25) follows by a straightforward computation from (22) and (26). For $b \neq 0$ it is obvious from (25) that a linear line complex is the intersection of the Graßmann manifold, according to (23) embedded into $\boldsymbol{P}^N$, with a hyperplane of $\boldsymbol{P}^N$. More generally, apart from the linear line complexes there are more general line manifolds, whose definition is not based on the single linear equation (25), but on a system of algebraic or even just differentiable equations. In the older literature, the geometry of the line space (for different transformation groups) is also called *line geometry*, and, accordingly, the terms line coordinates, line, or ray complexes etc. are used.

Exercise 3. Prove that under the canonical embedding (23) each hyperplane $\boldsymbol{A} \subset \boldsymbol{P}^N(\bigwedge^2 \boldsymbol{V})$ intersects the Graßmann manifold (23) in a proper subset $\mathfrak{K} := \boldsymbol{A} \cap h(\boldsymbol{P}_{n,1}) \neq \emptyset$, and that $\mathfrak{K} = \mathfrak{K}_F$ for a suitable null system F of $\mathfrak{P}^n$.

Exercise 4. Prove that for a symmetric auto-correlative map F of $\mathfrak{P}^n$ and any subspace $\boldsymbol{A} \subset \boldsymbol{P}^n$ the restriction $F|\boldsymbol{A}$ (cf. Proposition 7.7) is a symmetric auto-correlative map in $\mathfrak{P}(\boldsymbol{A})$. In particular, the restriction of a null system is again a null system.

1.9 Polarities and Quadrics

In this section we consider polar maps $F : \mathfrak{P}^n \to \mathfrak{P}^n$ (Definition 8.1). First we will show that the associated correspondences can be described by σ-Hermitean biforms b, which are generalizations both of the symmetric bilinear and the Hermitean forms. According to Definition 8.1, the set $Q_F \subset \boldsymbol{P}^n$ of auto-polar points of F is a proper subset of projective space, $Q_F \neq \boldsymbol{P}^n$; this set Q_F will be called the *quadric* in $\boldsymbol{P}^n$ determined by F. If K is a field and b is bilinear, this definition coincides with the classical one considered in § I.5.9. In this case, a quadric is the set of points whose coordinates satisfy a quadratic equation of the form (I.5.9.1). Since projective geometry is described by means of homogeneous coordinates, the homogeneous equations occurring are of second degree: If a point with coordinates (x^i) is a solution, then for any $\xi \in K$ the point with coordinates $(x^i)\xi$ is a solution as well. Hence, in a projective geometry over the field K with $\sigma = \mathrm{id}_K$ the equation for a quadric in $\boldsymbol{P}^n$ corresponding to (I.5.9.1) has the form

$$\sum_{i,j=0}^{n} a_{ij} x^i x^j = 0,$$

where not all coefficients a_{ij} are equal to 0. Note that the upper i, j are coordinate indices and no powers. If $\operatorname{char} K \neq 2$, then one may additionally assume that the coefficient matrix is symmetric; in fact, the coefficients a_{ij} may be replaced by $\hat{a}_{ij} := (a_{ij} + a_{ji})/2$ without changing the solution set. Of course, this set may well be empty. In the general case, the anti-automorphism σ also comes into play, cf. (1) below. The aim of this section is to obtain a classification of polar maps and, consequently, for the corresponding quadrics in projective geometries whose scalar field are the real, the complex, or the quaternionic numbers. Starting from this classification we will introduce the most important Cayley-Klein geometries in the next chapter.

1.9.1 σ-Hermitean Biforms

Let b be a σ-biform determining F, and let $(a_{ij}) \in M_{n+1}(K)$ be its matrix in a homogeneous coordinate system defined by the basis $(\mathfrak{a}_i)$ of $\boldsymbol{V}^{n+1}$. Then we obtain the equation for the associated quadric $Q_F \subset \boldsymbol{P}^n$ from (7.22), (7.16) (Note that the sum convention is in action!):

$$\boldsymbol{x} = [\mathfrak{x}] \in Q_F \Longleftrightarrow b(\mathfrak{x}, \mathfrak{x}) = 0 \Longleftrightarrow \sigma(x^i) a_{ij} x^j = 0, \ (\boldsymbol{x} \neq \boldsymbol{o}); \tag{1}$$

here x^i denote the homogeneous coordinates of the point $\boldsymbol{x}$ with respect to the basis $(\mathfrak{a}_i)$, $i = 0, \ldots, n$.

Definition 1. Let $\boldsymbol{V}$ be a vector space over the skew field K, and let σ be an anti-automorphism of K. A σ-biform $b : \boldsymbol{V} \times \boldsymbol{V} \to K$ is called *σ-Hermitean* if it satisfies

$$b(\mathfrak{x}, \mathfrak{y}) = \sigma(b(\mathfrak{y}, \mathfrak{x})),\ \mathfrak{x},\ \mathfrak{y} \in \boldsymbol{V}. \tag{2}$$

□

Proposition 1. *Let $\mathfrak{P}^n = \mathfrak{P}^n(\boldsymbol{V})$ be a projective geometry over the skew field K. Then:*
a) For every polar map $F : \mathfrak{P}^n \to \mathfrak{P}^n$ there are an anti-automorphism σ of K, a σ-Hermitean biform b over $\boldsymbol{V}$ determining F, and an $\mathfrak{x}_o \in \boldsymbol{V}$ such that

$$b(\mathfrak{x}_o, \mathfrak{x}_o) = 1. \tag{3}$$

Conversely, each σ-Hermitean biform of this kind determines a polar map F of $\mathfrak{P}^n$.
b) The anti-automorphism σ corresponding to a σ-Hermitean biform $b \neq 0$ is involutive, i.e.

$$\sigma \circ \sigma = \mathrm{id}_K\,. \tag{4}$$

Because of (2), *in particular,*

$$\sigma(b(\mathfrak{x}, \mathfrak{x})) = b(\mathfrak{x}, \mathfrak{x}),\ \mathfrak{x} \in \boldsymbol{V}. \tag{5}$$

Proof. Let b_o be any biform defining F. Since F is polar, there is an $\boldsymbol{x}_o = [\mathfrak{x}_o] \in \boldsymbol{P}^n$ with $\boldsymbol{x}_o \notin F(\boldsymbol{x}_o)$, hence $\lambda := b_o(\mathfrak{x}_o, \mathfrak{x}_0) \neq 0$. Setting $b := \lambda^{-1} b_0$ we obtain (3). According to Proposition 7.6, the biform b also defines F; let σ be the anti-automorphism corresponding to b. Since F is polar, and hence symmetric, (8.14) holds for $\kappa \in K^*$. Inserting $\mathfrak{x} = \mathfrak{y} = \mathfrak{x}_o$ together with (3) leads to:

$$\kappa = \kappa b(\mathfrak{x}_o, \mathfrak{x}_o) = \sigma^{-1}(b(\mathfrak{x}_o, \mathfrak{x}_o)) = 1.$$

Thus (2) immediately follows from (8.14). Conversely, (2) is a special case of (8.14). Hence the correlation determined by b is symmetric, and due to (3) a polar map (Definition 8.1).
b) Since b is non-zero and linear in $\mathfrak{y}$, for every $\xi \in K$ one can find a pair $(\mathfrak{x}, \mathfrak{y}) \in \boldsymbol{V} \times \boldsymbol{V}$ with $b(\mathfrak{x}, \mathfrak{y}) = \xi$. From (2) we now conclude (4):

$$\xi = \sigma(b(\mathfrak{y}, \mathfrak{x})) = \sigma(\sigma(b(\mathfrak{x}, \mathfrak{y}))) = \sigma \circ \sigma(\xi).$$

□

Exercise 1. Let $F, b, \sigma, \mathfrak{x}_o$ be as in Proposition 1. If b_1 is a σ_1-biform also defining F, then $b_1 = \kappa b$, where $\kappa \in K$ is a scalar for which there is an $\mathfrak{y} \in \boldsymbol{V}$ with $\kappa = b(\mathfrak{y}, \mathfrak{y}) \neq 0$; b_1 is σ_1-Hermitean, $\sigma_1 = \sigma_\kappa \circ \sigma$, and $\sigma_\kappa(\xi) = \kappa \xi \kappa^{-1}$. Conversely, if $\kappa = b(\mathfrak{y}, \mathfrak{y}) \neq 0$ holds for some $\mathfrak{y} \in \boldsymbol{V}$, then $b_1 = \kappa b$ is a σ_1-Hermitean biform also defining F.

Example 1. Let K be a field of characteristic 2 and consider $\boldsymbol{V} = K^n$. Then

$$b(\mathfrak{x}, \mathfrak{y}) := x^1 y^2 + x^2 y^1$$

is a symmetric bilinear form of rank 2 for which $b(\mathfrak{x},\mathfrak{x}) = 0$ holds for all $\mathfrak{x} \in \boldsymbol{V}$; hence it does not determine a polar map in the associated projective geometry. The following lemma is obtained in a way similar to the argument in the proof of Proposition 1 showing, in the case $\operatorname{char} K \neq 2$, that every σ-Hermitean biform $b \neq 0$ defines a polar map. □

Lemma 2. *Let K be a skew field,* $\operatorname{char} K \neq 2$, *and let $b \neq 0$ be a σ-Hermitean biform over $\boldsymbol{V}$. Then there is an $\mathfrak{x} \in \boldsymbol{V}$ with $b(\mathfrak{x},\mathfrak{x}) \neq 0$.*

Proof. Assume that $b(\mathfrak{x},\mathfrak{x}) = 0$ for all $\mathfrak{x} \in \boldsymbol{V}$. Because of $b \neq 0$, there are $\mathfrak{x},\mathfrak{y} \in \boldsymbol{V}$ such that $b(\mathfrak{x},\mathfrak{y}) \neq 0$; due to the linearity of b in $\mathfrak{y}$ we may suppose $b(\mathfrak{x},\mathfrak{y}) = 1$. Then for every $\lambda \in K$

$$0 = b(\mathfrak{x}+\mathfrak{y}\lambda, \mathfrak{x}+\mathfrak{y}\lambda) = b(\mathfrak{x},\mathfrak{y}\lambda) + b(\mathfrak{y}\lambda,\mathfrak{x}) = \lambda + \sigma(\lambda).$$

Thus $\sigma(\lambda) = -\lambda$ for all $\lambda \in K$, hence, in particular, $\sigma(1) = 1 = -1$. But this implies $\operatorname{char} K = 2$. □

Definition 2. Let $b : \boldsymbol{V} \times \boldsymbol{V} \to K$ be σ-Hermitean. The *defect subspace* of b is the subspace

$$\boldsymbol{W}_b := \{\mathfrak{x} \in \boldsymbol{V} \,|\, b(\mathfrak{x},\mathfrak{y}) = 0 \text{ for all } \mathfrak{y} \in \boldsymbol{V}\}. \tag{6}$$

The *defect* of b is defined by

$$\operatorname{def}(b) := \dim \boldsymbol{W}_b. \tag{7}$$

If a is the σ-linear map from$\boldsymbol{V}$ to $\boldsymbol{V}'$ defined by b, and if, moreover, $\dim \boldsymbol{V} < \infty$, then the *rank* of b is defined by the equation

$$\operatorname{rk}(b) := \dim \boldsymbol{V} - \operatorname{def}(b). \tag{8}$$

Obviously, *the defect subspace of b is equal to* $\boldsymbol{W}_b = \operatorname{Ker} a = \operatorname{Ker} F$, if F denotes the polar map defined by b, cf. (7.9). Hence we have $\operatorname{rk}(b) = \operatorname{rk}(a)$. □

1.9.2 Classification of Polar Maps

The following proposition can be proved in a way similar to Proposition I.5.9.5:

Proposition 3. *Let b be a σ-Hermitean biform on the $(n+1)$-dimensional vector space $\boldsymbol{V}$ over the skew field K with* $\operatorname{char} K \neq 2$. *Then there is a basis $(\mathfrak{a}_i)$ for $\boldsymbol{V}$ in which the matrix of b has diagonal form:*

$$b(\mathfrak{a}_i, \mathfrak{a}_j) = \begin{pmatrix} \beta_0 & 0 & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & & & & & & & & \cdot \\ \cdot & & \cdot & & & & & & & \cdot \\ \cdot & & & \cdot & & & & & & \cdot \\ \cdot & & & & \beta_{r-1} & & & & & \cdot \\ \cdot & & & & & 0 & & & & \cdot \\ \cdot & & & & & & \cdot & & & \cdot \\ \cdot & & & & & & & \cdot & & \cdot \\ \cdot & & & & & & & & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & 0 \end{pmatrix}; \tag{9}$$

here

$$\beta_h = \sigma(\beta_h) \neq 0 \textit{ for } h = 0, \ldots, r-1; r = \operatorname{rk}(b). \tag{10}$$

□

Of course, the numbers $\beta_h \in K$ in (9) are not uniquely determined by b. Replacing $\mathfrak{a}_h$ by $\hat{\mathfrak{a}}_h = \mathfrak{a}_h \lambda_h$, $\lambda_h \in K^*$, relations (9), (10) remain valid, and hence

$$\hat{\beta}_h := b(\hat{\mathfrak{a}}_h, \hat{\mathfrak{a}}_h) = \sigma(\lambda_h)\beta_h\lambda_h. \tag{11}$$

The transformation (11) can be used to obtain normal forms for polar maps and σ-biforms, respectively. These normal forms depend upon the properties of both the skew field K and the anti-automorphism σ. For the classification one first has to describe all the involutive anti-automorphisms σ of K, and then to determine, for each σ, corresponding normal forms for the σ-Hermitean biforms b. In the subsections to follow we will discuss the cases $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$ separately.

Exercise 2. Let F be a polar map of the projective geometry $\mathfrak{P}^n$, $n \geq 2$ and $\operatorname{char} K \neq 2$. Prove: a) There are $n+1$ points $\boldsymbol{a}_i \in \boldsymbol{P}^n, i = 0, \ldots, n$, such that $\boldsymbol{a}_i \in F(\boldsymbol{a}_j)$ for $i \neq j$ and $\boldsymbol{a}_i \notin F(\boldsymbol{a}_i)$ if and only if F is a polarity. – b) Formulate the dual of Statement a). – c) The points $\boldsymbol{a}_i$ with the properties stated in a) are in general position; choosing a projective frame $(\boldsymbol{a}_i; \boldsymbol{e})$ with $\boldsymbol{a}_i$ as base points, the matrix of F has diagonal form (9) with $r = n+1$. A simplex (a_i), $i = 0, \ldots, n$, with the properties from a) is called a *polar simplex* for the polarity F.

Exercise 3. Let F be a polarity satisfying the assumptions from Exercise 2, and let $\boldsymbol{a}_i \in \boldsymbol{P}^n$, $i = 0, \ldots, n$, be a sequence of points in general position. Prove that the following statements are equivalent:

a) $(\boldsymbol{a}_i)$ is a polar simplex;

b) $F(\boldsymbol{a}_i) = \bigvee_{j \neq i} \boldsymbol{a}_j$;

c) $\boldsymbol{a}_i = \bigwedge_{j \neq i} F(\boldsymbol{a}_j)$.

1.9.3 The Real Polar Maps

As is well-known, the only anti-automorphism of the field of real numbers is the identity $\sigma = \mathrm{id}_{\mathbf{R}}$. Hence, by Sylvester's Law of Inertia, Proposition I.5.9.6, polar maps are classified by their rank and index. Its projective formulation is

Proposition 4. *Let $K = \mathbf{R}$ be the field of real numbers, and let $\mathfrak{P}^n$ be an n-dimensional projective geometry over $\mathbf{R}$, $n \geq 2$. Choose any basis and denote by $F_{r,l}$ the polar map of $\mathfrak{P}^n$ whose matrix in this basis has normal form* (9) *(cf.* (I.5.9.49)*), where*[1]

$$\beta_\lambda = -1,\ \lambda = 0, \ldots, l-1;\ \beta_\mu = 1,\ \mu = l, \ldots, r-1, 0 \leq l \leq [r/2],\ 0 < r \leq n+1. \tag{12}$$

Denoting by x^i the homogeneous point coordinates, and by y_i the homogeneous hyperplane coordinates dual to them, the polar map $\boldsymbol{y} = F_{r,l}(\boldsymbol{x})$ has the following normal form:

$$\begin{array}{rl} y_j = -x^j & \text{for } j = 0, \ldots, l-1; \\ y_k = x^k & \text{for } k = l, \ldots, r-1,\ (0 < r \leq n+1); \\ y_h = 0 & \text{for } h = r, \ldots, n. \end{array} \tag{13}$$

Each polar map F of $\mathfrak{P}^n$ is projectively equivalent to one of the $F_{r,l}$, and no two of the latter are projectively equivalent.

Proof. According to Exercise I.2.1.3, cf. also Corollary 3.16, $\sigma = \mathrm{id}_{\mathbf{R}}$, and the biform b determining F has to be bilinear; moreover, because of (2) it is also symmetric. In (9), the $\beta_h \neq 0$ are arbitrary. Choosing the transformation (11) with $\lambda_h := |\beta_h|^{-1/2}$ and renumbering the basis elements in such a way that the first l diagonal entries are equal to -1 followed by $r - l$ times the entry 1, we arrive at Sylvester's Law of Inertia I.5.9.6; the matrix of b has the normal form (9) with $0 \leq l \leq r$. For $l > [r/2]$ the form (12) is obtained via multiplication of b by $\kappa = -1$ and a suitable numbering of the basis elements. Choosing again the indices in the range $i = 0, \ldots, n$ implies that each polar map of $\mathfrak{P}^n$ is projectively equivalent to one of the $F_{r,l}$. By (8.13) and Proposition I.5.9.6, no two of the $F_{r,l}$ are projectively equivalent. □

1.9.4 The Complex Polar Maps

As in the case of collinear maps (Example 4.3), the continuity of correlative maps implies that we only have to consider continuous automorphisms of the field $\mathbf{C}$ of complex numbers, cf. also Example 7.4. Hence there are two types of polar maps to be classified: those corresponding to $\sigma = \mathrm{id}_{\mathbf{C}}$ and the ones associated with the conjugation $\tau : z \mapsto \bar{z}$. For $K = \mathbf{C}$ we agree that the

[1] As usual, for $\xi \in \mathbf{R}$ the greatest integer less than or equal to ξ is denoted by $[\xi]$.

term *polar map* always refers to $\sigma = \mathrm{id}_{\mathbf{C}}$. On the other hand, the polar maps associated with $\sigma = \tau$ will be called *antipolar*. Starting from (9) the proof of the analogue to Proposition 4 is straightforward.

Proposition 5. *Two polar maps of a complex projective geometry* $\mathfrak{P}^n$, $n \geq 2$, *with* $\sigma = \mathrm{id}_{\mathbf{C}}$ *are projectively equivalent if and only if they have equal rank* r, $1 \leq r \leq n+1$. *Hence the coordinate representations*

$$F_r : y_j = x^j \text{ for } j = 0, \ldots, r-1,\ y_k = 0 \text{ for } k = r, \ldots, n, \tag{14}$$

provide a complete set of normal forms for these polar maps. □

Corollary 6. *All polarities of the complex projective geometry* $\mathfrak{P}^n$, $n \geq 2$, *with* $\sigma = \mathrm{id}_{\mathbf{C}}$ *are projectively equivalent.* □

The following proposition dealing with τ-Hermitean biforms and the corresponding anti-polar maps is proved along the same lines as the Law of Inertia (cf. Proposition I.9.13). As the proof works without any change also in the case of quaternionic projective geometries, we simply combine both situations into one statement.

Proposition 7. *Let* τ *be the conjugation in the field* $\mathbf{C}$ *of complex numbers or in the skew field* $\mathbf{H}$ *of quaternions. Denote by* $F_{r,l}$ *the polar map of the projective geometry* $\mathfrak{P}^n$ *that, in a fixed coordinate system, has the normal form*

$$\begin{aligned} y_j &= -\bar{x}^j \ \text{ for } \ j = 0, \ldots, l-1; \\ y_k &= \bar{x}^k \ \text{ for } \ k = l, \ldots, r-1,\ (0 < r \leq n+1); \\ y_h &= 0 \ \text{ for } \ h = r, \ldots, n, \end{aligned} \tag{15}$$

where the matrix (9) *satisfies* (12). *Every polar map of the projective geometry* $\mathfrak{P}^n$ *determined by a* τ*-Hermitean biform is projectively equivalent to one of the maps* $F_{r,l}$, *and no two of the latter are projectively equivalent.*

P r o o f. According to Proposition 3 we may start from normal form (9) for the matrix of F. Because of (10), we have $\beta_h \in \mathbf{R}^*$. As in the proof of Proposition 4 we obtain the diagonal form for the matrix of F with property (12). The following lemma characterizes the number l in an invariant way:

Lemma 8. *Let* $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$. *Take* σ *to be the identity in the case of the real numbers, and take conjugation as* σ *for* $K = \mathbf{C}, \mathbf{H}$. *Consider a* σ*-Hermitean form* b *whose matrix has the diagonal form* (9) *in a suitable basis* $(\mathfrak{a}_i)$. *Then the number* l *of negative values among the diagonal entries* β_i *is equal to the maximal dimension of a subspace* $\boldsymbol{W}^- \subset \boldsymbol{V}^n$ *for which the restriction* $b|\boldsymbol{W}^- \times \boldsymbol{W}^-$ *is negative-definite, i.e.,* $b(\mathfrak{x}, \mathfrak{x}) < 0$ *for all* $\mathfrak{x} \in \boldsymbol{W}^-, \mathfrak{x} \neq \mathfrak{o}$.

Proof. Because of (2), even in the case $K = \mathbf{H}$ we immediately have $\beta_h \in \mathbf{R}$. Take $\beta_k < 0$ for $k = 0, \dots, l-1$, and $\beta_h \geq 0$ otherwise. Obviously, the restriction of b to the subspace $\boldsymbol{U}^l := \mathfrak{L}(\mathfrak{a}_0, \dots, \mathfrak{a}_{l-1})$ spanned by $\mathfrak{a}_0, \dots, \mathfrak{a}_{l-1}$ is negative-definite. Consider the complementary subspace $\boldsymbol{B} := \mathfrak{L}(\mathfrak{a}_l, \dots, \mathfrak{a}_n)$. For every subspace $\boldsymbol{W} \subset \boldsymbol{V}$ whose dimension satisfies $\dim W > l$ there is a vector $\mathfrak{x} \in \boldsymbol{B} \cap \boldsymbol{W}$ such that $\mathfrak{x} \neq \mathfrak{o}$ by the dimension formula. Since we obviously have $b(\mathfrak{x}, \mathfrak{x}) \geq 0$ for all $\mathfrak{x} \in \boldsymbol{B}$, b cannot be negative-definite on $\boldsymbol{W}$. □

Lemma 8 implies that the *index* l of b, defined to be the number of negative diagonal entries in (9), does not depend on the chosen basis in which b has this diagonal form. The same holds for the dimension $n - r$ of the defect subspace of b, hence also for the rank r. Summarizing, we arrive at the analogue of Sylvester's Law of Inertia I.5.9.6 for τ-Hermitean biforms in the cases $K = \mathbf{C}, \mathbf{H}$, and complete the proof just as we did proving Proposition 4. □

Exercise 4. Let $\boldsymbol{V}^{n+1}$ be a complex vector space, and let τ be conjugation in $\mathbf{C}$. A τ-biform b is called *skew Hermitean* if it satisfies the condition

$$b(\mathfrak{y}, \mathfrak{x}) = -\overline{b(\mathfrak{x}, \mathfrak{y})}, \ \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}. \tag{16}$$

Prove: a) Each skew Hermitean τ-biform determines an anti-polar map of $\mathfrak{P}^n(\boldsymbol{V})$, that can also be described by a τ-Hermitean biform. – b) Classify the skew Hermitean biforms with respect to the action (8.12) with $\kappa = 1$ of the linear group of $\boldsymbol{V}^{n+1}$. To this end, prove that $b(\mathfrak{x}, \mathfrak{y})$ is skew Hermitean (or Hermitean, respectively) if and only if $\mathrm{i}\, b(\mathfrak{x}, \mathfrak{y})$ (with $\mathrm{i}^2 = -1$) is Hermitean (or skew Hermitean, respectively).

1.9.5 The Quaternionic Polar Maps

It is well-known that the center of the skew field $\mathbf{H}$ of quaternions is the real subfield $\mathbf{R}$. Denote by $\mathbf{R}^\perp$ the real subspace of $\mathbf{H}$ spanned by the imaginary units $\mathrm{i}, \mathrm{j}, \mathrm{k} \in \mathbf{H}$. First we prove:

Lemma 9. *The only involutive anti-automorphisms σ of the skew field $\mathbf{H}$ of quaternions are the conjugation $\tau : q \mapsto \bar{q}$ and the anti-automorphisms defined by*

$$\tau_q(\lambda) := -q\bar{\lambda}q, \ q \in \mathbf{R}^\perp, |q| = 1. \tag{17}$$

Proof. Let σ be an arbitrary anti-automorphism of $\mathbf{H}$. Then, according to Example II.8.8.5, $\tau \circ \sigma$ has to be an inner automorphism of $\mathbf{H}$. Hence there is $q \in \mathbf{H}^*$, $|q| = 1$, such that $\tau \circ \sigma = \sigma_q$. This implies $\sigma = \tau \circ \sigma_q$. Since σ is involutive, we obtain for all $\lambda \in \mathbf{H}$

$$q^{-1}\bar{q}\lambda = \lambda q^{-1}\bar{q}.$$

Consequently, $q^{-1}\bar{q}$ must belong to the center $\mathbf{R}$ of $\mathbf{H}$. Now $|q| = 1$ implies $|q^{-1}\bar{q}| = 1$, hence $q^{-1}\bar{q} = \pm 1$, i.e. $\bar{q} = \pm q$. If $\bar{q} = q$, then $q = \pm 1 \in \mathbf{R}$, and $\sigma_q = \mathrm{id}_{\mathbf{H}}$, thus $\sigma = \tau$. On the other hand, $\bar{q} = -q$ implies $q \in \mathbf{R}^\perp = \mathfrak{L}_{\mathbf{R}}(\mathrm{i}, \mathrm{j}, \mathrm{k})$, and $\sigma(\lambda) = -q\bar{\lambda}q$. A straightforward computation shows that each of these anti-automorphisms is involutive. □

The polar maps corresponding to τ-Hermitean biforms were already classified in Proposition 7. For later application, we write out the formula for such a biform over an n-dimensional vector space $\boldsymbol{V}^n$ below; according to Sylvester's Law of Inertia this is characterized by its index l and rank r and thus, in a suitable basis $(\mathfrak{a}_i), i = 1, \dots n$, has the form:

$$b(\mathfrak{x}, \mathfrak{y}) = -\sum_{a=1}^{l} \bar{x}^a y^a + \sum_{c=l+1}^{r} \bar{x}^c y^c,\ 0 \leq l \leq r. \tag{18}$$

Here $x^i, y^i, i = 1, \dots, n$, are the vector coordinates of $\mathfrak{x}, \mathfrak{y}$ in this basis. Now we turn to the τ_q-Hermitean biforms and start by noting

Lemma 10. *Let τ_q be defined by* (17). *Then a biform $b : \boldsymbol{V} \times \boldsymbol{V} \to \mathbf{H}$ is τ_q-Hermitean if and only if $\hat{b} = qb$ is skew Hermitean in the sense of* (16). □

The proof is nothing but a straightforward calculation. Since $\tau_q = \tau_{\bar{q}} = \tau_{-q}$, we immediately obtain

Corollary 11. *If b is a τ_q-Hermitean biform over $\boldsymbol{V}$, then $\tilde{b} := \pm\,\mathrm{i}\,qb$ is a τ_{i}-Hermitean biform over $\boldsymbol{V}$.* □

First we consider the τ_{i}-Hermitean biforms. Later, the general case will be reduced to this situation. A simple calculation shows: *The set of fixed elements of τ_{i} is the real subspace of* $\mathbf{H}$ *spanned by* $1, \mathrm{j}, \mathrm{k}$*:*

$$\{\beta \in \mathbf{H} |\, \tau_{\mathrm{i}}(\beta) = \beta\} = \mathfrak{L}_{\mathbf{R}}(\mathbf{1}, \mathrm{j}, \mathrm{k}). \tag{19}$$

Let now b be a τ_{i}-Hermitean biform, and let $(\mathfrak{a}_j), j = 1, \dots, n$, be a basis in which the matrix of b has the form (9). In this case, transformation (11) takes the form

$$\hat{\beta} = -\,\mathrm{i}\,\bar{\lambda}\,\mathrm{i}\,\beta\lambda. \tag{20}$$

Choosing first $\lambda = |\beta|^{-1/2}$, which is possible by $\beta \neq 0$, we obtain $|\hat{\beta}| = 1$. Thus we may already assume $|\beta| = 1$ and henceforth only consider transformations with$|\lambda|^2 = \lambda\bar{\lambda} = 1$, i.e. $\bar{\lambda} = \lambda^{-1}$. Because of (19), the element $\mathrm{i}\,\beta$ lies on the unit sphere in $\mathbf{R}^{\perp}$. According to Exercise II.8.9.5, each element of the special orthogonal group $\mathbf{SO}(\mathbf{R}^{\perp})$ can be represented as an inner automorphism σ_λ, $|\lambda| = 1$. As is well-known, $\mathbf{SO}(\mathbf{R}^{\perp}) \cong \mathbf{SO}(\mathbf{3})$ acts transitively on the unit sphere; hence we can find λ such that $\bar{\lambda}\,\mathrm{i}\,\beta\lambda = \mathrm{i}$, i.e. $\hat{\beta} = 1$. Performing these transformations for each of the diagonal entries β_h in (9), Proposition 3 together with Corollary 11 implies

Proposition 12. *Let $\boldsymbol{V}$ be an n-dimensional vector space over the skew field* $\mathbf{H}$ *of quaternions, and let τ_q be the anti-automorphism defined by* (17). *If b is* *a τ_q-Hermitean biform, then there exists a basis $(\mathfrak{a}_\nu), \nu = 1, \ldots, n$, of* $\mathbf{V}$ *such that b has the following normal form:*

$$b(\mathfrak{x}, \mathfrak{y}) = \sum_{a=1}^{r} \tau_q(x^a)\bar{q}\,\mathrm{i}\,y^a = -q\sum_{a=1}^{r} \bar{x}^a\,\mathrm{i}\,y^a. \tag{21}$$

Here r is the rank of b, x^ν and y^ν are the coordinates of $\mathfrak{x}$ and $\mathfrak{y}$ with respect to the basis $(\mathfrak{a}_\nu)$.

Proof. By the above, for $q = \mathrm{i}$ there is a basis with $\beta_a = \bar{q}\,\mathrm{i} = 1$ for $a = 1, \ldots, r$, hence (21). If b is a τ_q-Hermitean form, then Corollary 11 implies $b = -q\,\mathrm{i}\,\tilde{b}$, where $\tilde{b}$ is a τ_{i}-Hermitean form. Bringing this into normal form we again obtain (21) for arbitrary $q \in \mathbf{H}^\perp$, $|q| = 1$. □

In the notations of Proposition 12 we have the following

Corollary 13. *If $\hat{b}$ is a skew Hermitean biform of rank r over $\boldsymbol{V}$, then there is a basis $(\mathfrak{a}_\nu)$ of $\boldsymbol{V}$ in which $\hat{b}$ has the normal form*

$$\hat{b}(\mathfrak{x}, \mathfrak{y}) = \sum_{a=1}^{r} \bar{x}^a\,\mathrm{i}\,y^a. \tag{22}$$

□

The proof is a consequence of (21) and Lemma 10. Note that (21) and (22) provide sets of normal forms with respect to the action (8.12) with $\kappa = 1$ of the quaternionic linear group for the τ_q-Hermitean and the skew Hermitean biforms, respectively; no two of these normal forms are equivalent. In fact, the rank of b is obviously invariant, and to be τ_q-Hermitean or skew Hermitean, respectively, is a property preserved by any linear transformation. Combining all these results immediately leads to the classification of polar maps:

Proposition 14. *Let F be a polar map in the n-dimensional projective geometry $\mathfrak{P}^n$ over the skew field* $\mathbf{H}$ *of quaternions, $n \geq 2$. Then F is projectively equivalent to one of the maps $F_{r,l}$ by Proposition 7, or F is projectively equivalent to one of the maps J_r with normal form*

$$J_r : \ y_a = \bar{x}^a\,\mathrm{i} \ \textit{for}\ a = 0, \ldots, r-1;\ y_c = 0 \ \textit{for}\ c = r, \ldots, n;\ 0 < r \leq n. \tag{23}$$

In both cases, r is the rank of F, and the corresponding anti-automorphism is the conjugation τ in $\mathbf{H}$. *No two of these normal forms are projectively equivalent.*

Proof. According to Proposition 7.6 F can be described by a σ-biform (7.20). Proposition 1 implies that σ is an involutive anti-automorphism of $\mathbf{H}$. By Lemma 9 the anti-automorphism σ has to be either conjugation, $\sigma = \tau$, or $\sigma = \tau_q$ as defined by (17). Corollary 11 implies that after multiplication by $q\,\mathrm{i}$ the biform describing the same correspondence becomes τ_{i}-Hermitean. Hence we can always require this to hold. From Proposition 12 we derive the following set of normal forms for the polar maps with $\sigma = \tau_{\mathrm{i}}$:

$$y_a = \tau_{\mathrm{i}}(x^a) = -\,\mathrm{i}\,\bar{x}^a\,\mathrm{i} \text{ for } a = 0, \ldots, r-1;\; y_c = 0 \text{ for } c = r, \ldots, n;\; 0 < r \leq n. \tag{24}$$

The transition to $\hat{b} = \mathrm{i}\,b$ discussed in Lemma 10 leads to the skew Hermitean normal form (22) and this again to (23). No two of the normal forms (15) (with (12)) and (23) (or (24) instead) are projectively equivalent: According to Proposition 7, and because of the invariance of the rank, it suffices to show that no τ_{i}-Hermitean biform can be transformed into a τ-Hermitean biform $b \neq 0$ via multiplication by $h \in \mathbf{H}^*$, and vice versa. Let $b_1 = hb$ be a τ_{i}-biform. Since b is a τ-biform, b_1 is a biform with respect to the anti-automorphism $\xi \mapsto h\bar{\xi}h^{-1}$. If this were equal to τ_{i}, then, up to an irrelevant real factor, we had $h = \mathrm{i}$. But this again implied

$$b_1(\mathfrak{y}, \mathfrak{x}) = \mathrm{i}\,b(\mathfrak{y}, \mathfrak{x}) = \mathrm{i}\,\overline{b(\mathfrak{x}, \mathfrak{y})} = \mathrm{i}\,\mathrm{i}\,\overline{b(\mathfrak{x}, \mathfrak{y})}\,\mathrm{i},$$

hence $b_1(\mathfrak{y}, \mathfrak{x}) = -\tau_{\mathrm{i}}(b_1(\mathfrak{x}, \mathfrak{y}))$. However, this can only be τ_{i}-Hermitean if and only if $b = b_1 = 0$. □

1.9.6 Quadrics

Because of Equation (1), the projective classification of polar maps immediately leads to a corresponding classification of quadrics. Allowing for some incorrectness, we will call two quadrics ***projectively equivalent*** if the polar maps defining them are related this way. It is easy to prove that equivalence of the polar maps implies equivalence of the quadrics as sets; more precisely: If F_0, F_1 are two equivalent polar maps, and if g is a polarity realizing their equivalence, i.e. $F_1 = gF_0g^{-1}$, then for the associated quadrics Q_0, Q_1 we have the relation $Q_1 = g(Q_0)$. The converse of this statement, however, is, in general, not true. Consider, for example, the projective plane over the field $\mathbf{Q}$ of rational numbers. For every prime number $p \in \mathbf{Z}$ the quadrics determined by the equations

$$(x^0)^2 + p(x^1)^2 = 0 \text{ and } (x^0)^2 - p(x^1)^2 = 0$$

are equal. In fact, they only contain the point satisfying $x^0 = x^1 = 0$, as can easily be seen using the prime decomposition in $\boldsymbol{Q}$. The polar maps defining them, however, cannot be equivalent: Any polarity over the rational numbers realizing the equivalence would also be an equivalence over the real numbers, but this contradicted the Theorem of Inertia. Considering the equations over the real numbers, their solution sets are obviously not projectively equivalent: One consists of a single point, and the other is the union of two lines. Over the complex numbers the defining polar maps, and hence their solution sets as well, projectively equivalent.

Remark. In the case that the scalar domain K is an *algebraically closed field,* e.g. the field $\mathbf{C}$ of complex numbers, and if the polar maps and quadrics un question are defined by a *symmetric bilinear form,* then the converse holds in general: *If, under the assumptions above, the quadrics $Q, \hat{Q}$ corresponding to the symmetric bilinear forms $b, \hat{b}$ are equal, then the b and $\hat{b}$ are proportional.* To prove this we show that the quadrics are different, if the forms are not proportional. So take a point $x = [\mathfrak{x}]$ with $b(\mathfrak{x}, \mathfrak{x}) = 1$ (cf. (3)). If $\hat{b}(\mathfrak{x}, \mathfrak{x}) = 0$, then $x \in \hat{Q} \setminus Q$, i.e. the quadrics are different. Hence, normalizing $\hat{b}$ by a suitable factor $\kappa \in K^*$ we may assume that $\hat{b}(\mathfrak{x}, \mathfrak{x}) = 1$ as well. Since the forms are supposed to be not proportional, there is a vector $\mathfrak{y}$ such that $b(\mathfrak{y}, \mathfrak{y}) \neq \hat{b}(\mathfrak{y}, \mathfrak{y})$. Now consider the line joining x to the corresponding point $y = [\mathfrak{y}]$. It has the parametrization

$$z(t) = [\mathfrak{z}(t)] \text{ with } \mathfrak{z}(t) = \mathfrak{x} + \mathfrak{y}t.$$

Its intersection with the quadric Q is determined by the roots t_1, t_2 of the quadratic equation

$$b(\mathfrak{z}(t), \mathfrak{z}(t)) = t^2 + 2tb(\mathfrak{x}, \mathfrak{y}) + b(\mathfrak{y}, \mathfrak{y}) = 0,$$

for which $t_1 t_2 = b(\mathfrak{y}, \mathfrak{y})$. Considering the intersection $\hat{Q} \cap x \vee y$ we analogously obtain two parameter values $\hat{t}_1$, $\hat{t}_2$ with $\hat{t}_1 \hat{t}_2 = \hat{b}(\mathfrak{y}, \mathfrak{y})$. These roots exist, since the scalar domain is an algebraically closed field. Because of $b(\mathfrak{y}, \mathfrak{y}) \neq \hat{b}(\mathfrak{y}, \mathfrak{y})$, they have to be different. Hence the intersections with the line $x \vee y$, and therefore the quadrics, cannot be equal. □

The *affine classification of quadrics* can be derived by first distinguishing a hyperplane at infinity and then discussing the various position possibilities for the quadric with respect to this hyperplane. Conversely, one might start from the affine classification (cf. Proposition I.5.9.11 for the case $K = \mathbf{R}$) and then, after having passed to homogeneous coordinates, decide, which among the affine types are projectively equivalent.

Example 2. Let $K = \mathbf{R}$. We denote by $Q_{r,l}$ the quadric corresponding to the normal form $F_{r,l}$. Then the equation of the quadric $Q_{r,0}$ is

$$\sum_{i=0}^{r-1} (x^i)^2 = 0, \ 0 < r \leq n+1; \tag{25}$$

hence $Q_{r,0} = \operatorname{Ker} F_{r,0}$. Note that we never consider the nopoint $\boldsymbol{o}$ as an element of any quadric. In particular, $Q_{n+1,0} = \emptyset$ is the empty set; this is expressed by calling it the *empty quadric.* The set of base points for the related coordinate system $\boldsymbol{a}_k = [\mathfrak{a}_k], k = r, \ldots, n$, span the kernel $\operatorname{Ker} F_{r,0} = \boldsymbol{a}_r \vee \ldots \vee \boldsymbol{a}_n$, which thus, for $r \leq n$, always lies in the coordinate hyperplane $x^0 = 0$. Since this projective $(n-r)$-plane is now defined by the quadratic equation (25), it is usually called a *double counted* projective $(n-r)$-plane. The reason for this term is the following: Consider the family of quadrics defined by

$$(x^0)^2 - t^2(x^1)^2 = 0$$

depending on the parameter $t \in \mathbf{R}$. For $t \neq 0$ each of these quadrics "decomposes" into the pair of hyperplanes

$$x^0 - tx^1 = 0, \; x^0 + tx^1 = 0.$$

Letting then t tend to zero, both the hyperplanes move towards one and the same hyperplane $(x^0)^2 = 0$, which, for this reason, has to be counted twice. Let us now return to equation (25). As in Section 5 we choose the hyperplane $x^0 = 0$ as the hyperplane at infinity for an affine space $\boldsymbol{A}^n \subset \boldsymbol{P}^n$ and consider, as usual, only the part $AQ_{r,l} := \boldsymbol{A}^n \cap Q_{r,l}$ as the affine quadric. In the inhomogeneous coordinates, the affine points $x^0 = 1$ from (25) satisfy the equation for $AQ_{r,0}$:

$$\sum_{i=1}^{r-1}(x^i)^2 = -1.$$

This, obviously, has no real solution, so that *for all* r the affine quadric $AQ_{r,0}$ is the empty set: The kernel $\operatorname{Ker} F_{r,0}$ completely lies in the hyperplane at infinity. A similar consideration shows: The quadric $Q_{n+1,1}$ has the equation

$$-(x^0)^2 + \sum_{i=1}^{n}(x^i)^2 = 0, \tag{26}$$

its section with the hyperplane at infinity $x^0 = 0$ is the empty quadric $Q_{n,0}$, and hence it is the *hyperellipsoid* $AQ_{n+1,1} = Q_{n+1,1}$, whose equation has the following normal form in inhomogeneous affine coordinates

$$\sum_{i=1}^{n}(x^i)^2 = 1, \quad \text{(cf. Proposition I.5.9.11)} \tag{27}$$

Interpreting the x^i as the orthonormal coordinate of an n-dimensional Euclidean space (27) becomes the equation of a hypersphere with radius 1 and the origin of the coordinate system as its center; for $n = 3$ Figure 1.9 shows an *ellipsoid,* which is projectively, even affinely, equivalent to it. It is not difficult to prove that for $n > 1$ the hyperellipsoids are the only non-empty real quadrics for which there is a hyperplane not intersecting them. Hence the hyperellipsoids are the only real projective quadrics that can be represented as closed surfaces in a Euclidean space; any other non-empty quadric intersects the hyperplane at infinity, if it is considered as affine or Euclidean. For topological reasons they are, however, always compact subsets of the ambient projective space. Let now $Q = Q_{n+1,l}$ be a non-degenerate quadric of the n-dimensional projective space; the degenerate quadrics may always be viewed as cones over non-degenerate ones, cf. Example 4 below. We suppose that its index l is greater than one and write its equation in the form

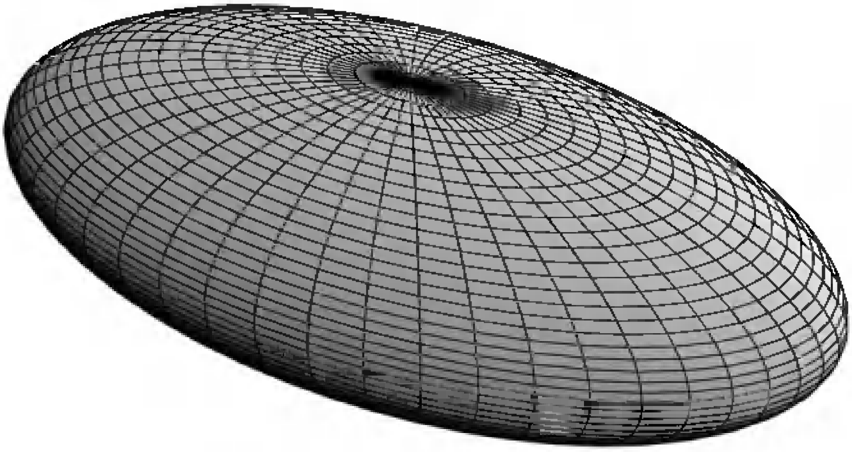


Fig. 1.9. An Ellipsoid

$$|\mathfrak{x}_0|^2 = \Sigma_{\alpha=0}^{l-1}(x^\alpha)^2 = \Sigma_{\kappa=l}^{n}(x^\kappa)^2 = |\mathfrak{x}_1|^2.$$

Here $|\mathfrak{x}|$ denotes the Euclidean norm in the vector space $\boldsymbol{V}^{n+1}$ for which the x^i, $i = 0, \ldots, n$, are orthonormal coordinates. Since these coordinates are, at the same time, also homogeneous point coordinates for the associated projective space, and as the nopoint does not belong to the quadric, for arbitrary $r > 0$ the solution set for this equation is the product of two spheres of dimensions $l - 1$ and $n - l$ of radius r with the origin as their center lying in orthogonal coordinate planes:

$$S_0^{l-1}(r) \times S_1^{n-l}(r) =$$

$$\{(\mathfrak{x}_0, \mathfrak{x}_1) | \mathfrak{x}_0 = (x^0, \ldots, x^{l-1}), \mathfrak{x}_1 = (x^l, \ldots, x^n), |\mathfrak{x}_0| = |\mathfrak{x}_1| = r\}.$$

Since each point $\boldsymbol{x} = [(\mathfrak{x}_0, \mathfrak{x}_1)] \in Q$ has precisely two representatives in this solution set, differing in the factor ± 1, we still have to identify opposite vectors in the solution set; assuming $r = 1$ leads to: *The real, non-degenerate projective quadric $Q_{n+1,l} \subset \boldsymbol{P}^n$ of index $l \geq 1$ is bijectively represented by the quotient set*

$$Q_{n+1,l} \cong S^{l-1} \times S^{n-l}/\{\pm 1\},\ l \geq 1.$$

Applying straightforward topological or differential geometric considerations the bijection in question can be proved to be a homeomorphism or diffeomorphism with respect to the canonically induced structures. In the case $l = 1$ already considered, the first factor is $S^0 = \{1, -1\}$; choosing, say, 1 as the representative for $\mathfrak{x}_0$ the representative of the point is already uniquely determined, and we again obtain the representation of the quadric as the hypersphere (27). Fixing a homogeneous coordinate system the points $\boldsymbol{x} \in \boldsymbol{P}^n$ of the real projective space are bijectively represented by unit vectors whose first coordinate is positive; this way, the projective space itself is represented as the unit hypersphere of the associated vector space with identified opposite points (cf. also Example 2.2.1):

$$\boldsymbol{P}^n \cong S^n/\{\pm 1\},\ (K = \mathbf{R}).$$

Since now, in the representation of Q just described, already the first vector $\mathfrak{x}_0$ has to be non-zero, we conclude for the *representation of the real non-degenerate quadrics* in general

$$Q_{n+1,l} \cong \boldsymbol{P}^{l-1} \times S^{n-l},\ 1 \leq l \leq (n+1)/2.$$

For $n = 3$ we obtain apart from the ellipsoid ($l = 1$) also the *hyperboloid* $Q_{4,2}$, whose affine image (cf. Fig. 2.18) has to be completed by an ellipse in the plane at infinity $x^0 = 0$. Since the projective line itself is homeomorphic to a circle, the above representation amounts to the homeomorphy of the hyperboloid with the product $S^1 \times S^1$, which is also called a *torus*. Figure 2.53 shows a torus embedded into the Euclidean space $\boldsymbol{E}^3$ as the surface with the parametrization described in (2.7.56). Obviously, the hyperboloid itself is not embedded in this way into the Euclidean (or projective) space; note that each plane, hence that at infinity as well, intersects the hyperboloid in a non-empty set. $\square$

Exercise 5. Prove for $K = \mathbf{R}$ and $n = 2$: If $Q \subset \boldsymbol{A}^2$ is one of the affine quadrics ellipse, hyperbola, or parabola, then there is a unique projective quadric $\hat{Q} \subset \boldsymbol{P}^2$ satisfying $Q = \hat{Q} \cap \boldsymbol{A}^2$. Discuss the possible positions of these quadrics relative to the line at infinity, and prove that their projective extensions $\hat{Q}$ are projectively equivalent quadrics. More generally, investigate analogous questions for the affine quadrics classified in Proposition I.5.9.11. In particular, for $n = 3$, the hyperbolic paraboloid and the one-sheeted hyperboloid both are completed to become projective hyperboloids; hence they are projectively equivalent. The two-sheeted hyperboloid is completed to a projective ellipsoid.

Exercise 6. For $K = \mathbf{R}$ consider the empty quadric $F = F_{n+1,0}$, cf. (25). By Exercise 2, the points $\boldsymbol{a}_i = [\mathfrak{a}_i]$ form a polar simplex for F; choose $\boldsymbol{a}_0$ as its origin and $\boldsymbol{H} := F(\boldsymbol{a}_0) = \boldsymbol{a}_1 \vee \ldots \vee \boldsymbol{a}_n$ as the hyperplane at infinity. Then the homogeneous coordinates normalized by $x^0 = 1$ yield an orthonormal Euclidean coordinate system x^i, $i = 1, \ldots, n$, in the affine space $\boldsymbol{A}^n$ with respect to the restriction of the symmetric bilinear form on $\boldsymbol{H}$ corresponding to F:

$$b_{\boldsymbol{H}}(\mathfrak{x}, \mathfrak{y}) = \sum_{i=1}^{n} x^i y^i =: \langle \mathfrak{x}, \mathfrak{y} \rangle.$$

In fact, according to Exercise 5.4 we may also identify the vector space $\boldsymbol{W}^n := \mathfrak{L}(\mathfrak{a}_1, \ldots, \mathfrak{a}_n)$ of the hyperplane at infinity with the vector space of the affine geometry $\boldsymbol{A}^n$. Prove that for $\mathfrak{x} \neq \mathfrak{o}$ the polar $F(\boldsymbol{x})$ of the point $\boldsymbol{x} := [\mathfrak{a}_0 + \mathfrak{x}]$ is the hyperplane with Hessian normal form (cf. (I.6.3.37))

$$z = [\mathfrak{a}_0 + \mathfrak{z}] \in F(x) \Longleftrightarrow \langle \mathfrak{z}, \frac{-\mathfrak{x}}{||\mathfrak{x}||} \rangle = \frac{1}{||\mathfrak{x}||}.$$

For $\mathfrak{x}$ tending to $\mathfrak{o}$, the point $\boldsymbol{x}$ tends to $\boldsymbol{a}_0$, and $F(\boldsymbol{x})$ converges to the hyperplane at infinity $F(\boldsymbol{a}_o) = \boldsymbol{H}$.

Example 3. Obviously, for $K = \mathbf{C}$ the quadric Q_{n+1} corresponding to F_{n+1} (Equation (25) with $r = n+1$) is not empty; e.g., the point with coordinates $(1, \mathrm{i}, 0, \ldots, 0)$ belongs to Q_{n+1}. Again denoting by $\boldsymbol{a}_j = [\mathfrak{a}_j], j = 0, \ldots, n$, the vertices of a polar simplex for F_{n+1} we obtain

$$\boldsymbol{c}_\rho := [\mathfrak{a}_{2\rho} + \mathfrak{a}_{2\rho+1}\,\mathrm{i}] \in Q_{n+1},\ \ \rho = 0, \ldots, k := [(n-1)/2], \tag{28}$$

$$\boldsymbol{M}^k := \boldsymbol{c}_0 \vee \ldots \vee \boldsymbol{c}_k \subset Q_{n+1}. \tag{29}$$

Note that the restriction of the bilinear form associated with F_{n+1} to $\boldsymbol{M}^k$ vanishes identically. □

Definition 3. Let b be a σ-Hermitean, skew Hermitean, or alternating biform over the vector space $\boldsymbol{V}$ with the skew field K as scalar domain. A subspace $\boldsymbol{W} \subset \boldsymbol{V}$ is called *isotropic (with respect to b)* if there exists $\mathfrak{x} \in \boldsymbol{W}, \mathfrak{x} \neq \mathfrak{o}$, such that for all $\mathfrak{y} \in \boldsymbol{W}$ the equation $b(\mathfrak{x}, \mathfrak{y}) = 0$ holds; $\boldsymbol{W}$ is called *totally isotropic* if the restriction of b vanishes, $b|\boldsymbol{W} \times \boldsymbol{W} = 0$. □

Corollary 15. *Let* char $K \neq 2$, *and let b be a σ-biform determining the polar map F of the projective geometry $\mathfrak{P}^n$. Then a projective subspace $\boldsymbol{A}$ is contained in the quadric Q_F corresponding to F if and only if the restriction of b to it vanishes, $b_{\boldsymbol{A}} = 0$, i.e., if the subspace $\boldsymbol{W} \subset \boldsymbol{V}$ corresponding to $\boldsymbol{A}$ is totally isotropic (*cf. (7.33).

Proof. If $\boldsymbol{A} \subset Q_F$, then because of $\boldsymbol{x} = [\mathfrak{x}] \ \iota\ F(\boldsymbol{x})$ for all $\mathfrak{x} \in \boldsymbol{W}$, we have $b(\mathfrak{x}, \mathfrak{x}) = 0$; hence by Lemma 2 $b_{\boldsymbol{A}} = 0$. The converse is trivial. □

Corollary 16. *Let* char $K \neq 2$. *If $\boldsymbol{A}$ is a subspace contained in the quadric Q_F of the polar map F, then $\boldsymbol{A} \vee$* Ker F *also lies in Q_F; in particular,* Ker $F \subset Q_F$.

Proof. If $\boldsymbol{W}$ is the vector space corresponding to $\boldsymbol{A}$, then $\boldsymbol{W}$ is totally isotropic. Since the defect subspace (6) is also totally isotropic, and, in addition, $b(\boldsymbol{W}_b, \boldsymbol{W}) = 0$, the sum $\boldsymbol{W} + \boldsymbol{W}_b$ again is totally isotropic. As the defect subspace corresponds to the kernel of F, this implies the assertion of Corollary 16. □

Example 4. Let Q_F be a quadric with Ker $F \neq \boldsymbol{o}$. Then Ker F is called the *vertex space* of the quadric. For char $K \neq 2$, according to (16) we have $\boldsymbol{x} \vee$ Ker $F \subset Q_F$ for each $\boldsymbol{x} \in Q_F$. For this reason, quadrics corresponding to degenerate polar maps are also called *projective cones*, and the maximal projective subspace contained in Q_F is the *generator* of the cone. □

Exercise 7. Let $\operatorname{char} K \neq 2$, let Q_F be the projective cone determined by the σ-biform b, and let $\boldsymbol{A}$ be a subspace complementary to $\operatorname{Ker} F$. a) Prove:

$$Q_F = \bigcup_{\boldsymbol{x} \in Q_F \cap \boldsymbol{A}} x \vee \operatorname{Ker} F,$$

$$(\boldsymbol{x} \vee \operatorname{Ker} F) \wedge (\boldsymbol{y} \vee \operatorname{Ker} F) = \operatorname{Ker} F \text{ for } \boldsymbol{x}, \boldsymbol{y} \in Q_F \cap \boldsymbol{A}, \boldsymbol{x} \neq \boldsymbol{y}.$$

b) The restriction (cf. Proposition 7.7) $F|\boldsymbol{A}$ is a polarity described by the biform $b|\boldsymbol{W} \times \boldsymbol{W}$, where $\boldsymbol{W}$ denotes the vector space associated with $\boldsymbol{A}$. – c) If $\boldsymbol{A}_1, \boldsymbol{A}_2$ both are complementary subspaces to $\operatorname{Ker} F$, then the polarities $F|\boldsymbol{A}_1$, $F|\boldsymbol{A}_2$ are projectively equivalent. Using the elementary definition of cone sections (Exercise I.5.9.3) this leads to a new proof for the first part of Exercise 5.— Hint. Recall, in this connection, (7.33) and Proposition 8.2.

1.9.7 Polar Maps of Projective Lines

As a preparation for the study of tangents and tangent subspaces of a quadric, we first want to discuss polar maps F of a projective line $\mathfrak{P}^1$. For the rank r of F the only possible values are $r = 1, 2$ by (8). The first case is the subject of the following simple proposition:

Proposition 17. *Let F be a polar map of the projective line $\mathfrak{P}^1$ over the skew field K, $\operatorname{char} K \neq 2$. Then the rank r of F is equal to 1 if and only if all points $\boldsymbol{x} \notin \operatorname{Ker} F$ have the same image $F(\boldsymbol{x}) = \boldsymbol{a}$; in this case,*

$$Q_F = \operatorname{Ker} F = \boldsymbol{a}.$$

Proof. Let b be the σ-Hermitean form defining F. Now F has rank 1 if and only if $\operatorname{Ker} F$ just consists of a single point, say $\boldsymbol{a} = \operatorname{Ker} F$. Obviously, $\boldsymbol{a} \in Q_F$. According to Proposition 3, we may suppose that the matrix of b has diagonal form; hence the equation for Q_F is

$$\sigma(x^0)\beta_0 x^0 = 0 \text{ with } \beta_0 \neq 0.$$

This implies $x^0 = 0$. Thus $Q_F = \boldsymbol{a}_1 = [\mathfrak{a}_1]$ is the first base point of the corresponding projective frame. Hence Q_F contains just one element, and this implies $\boldsymbol{a} = \boldsymbol{a}_1$. The coordinates of the image point $F(\boldsymbol{x})$, $\boldsymbol{x} \neq \boldsymbol{a}$, can be computed using the annihilator of the covector with components

$$u_0 = \sigma(x^0)\beta_0,\ u_1 = 0,$$

yielding $y^0 = 0$, $y^1 = 1$. Thus $F(\boldsymbol{x}) = \boldsymbol{a}$. The converse immediately follows from the definition of the rank. □

The possibilities occurring in the case $r = 2$ again depend both on the field K and on the type of the quadric. According to Corollary 8.3, F is now bijective and thus defines an involution of $\boldsymbol{P}^1$.

Example 5. Let K be a skew field with $\operatorname{char} K \neq 2$. By Proposition 3, any polarity of $\mathfrak{P}^1$ can be described by a σ-biform

$$b(\mathfrak{x}, \mathfrak{y}) = \sigma(x^0)\beta_0 y^0 + \sigma(x^1)\beta_1 y^1 = 0. \tag{30}$$

Because of $b(\mathfrak{a}_i, \mathfrak{a}_i) = \beta_i \neq 0$, $i = 0, 1$, the base points $\boldsymbol{a}_i = [\mathfrak{a}_i]$ do not lie on the quadric Q_F. For every point $\boldsymbol{x} \in Q_F$ we may thus take the coordinates to be $(x^0, x^1) = (1, \xi)$ with $\xi \in K^*$, and then write $\xi = \xi(\boldsymbol{x})$. This implies

$$\boldsymbol{x} \in Q_F \iff \sigma(\xi)\beta_1\xi = -\beta_0. \tag{31}$$

If K is a field, and $\sigma = \mathrm{id}_K$, then (31) is equivalent to

$$\xi^2 = -\beta_0/\beta_1. \tag{32}$$

So there are only two possibilities, Q_F may either be empty or else contain two points. This coincides with the result of Exercise 4.13 b); in fact, in this case F is, at the same time, a projectivity defined by the linear isomorphism

$$y^0 = -\beta_1 x^1,\ y^1 = \beta_0 x^0. \tag{33}$$

Relation (30) immediately implies (even in the case of an arbitrary skew field K) that the points of the quadric Q_F are nothing but the fixed points of the involution F. For $K = \mathbf{R}$ both possibilities occur according to Proposition 4: $Q_F = \emptyset$ for $F = F_{2,0}$, and Q_F consists of two points for $F = F_{2,1}$. If K is algebraically closed, e.g. $K = \mathbf{C}$, and $\sigma = \mathrm{id}_K$, then it is just the second case that occurs. □

Example 6. Consider the Hermitean normal forms F from Proposition 7 for $K = \mathbf{C}, \mathbf{H}$, $n = 1$, $r = 2$. The equation of this quadric is

$$\epsilon \bar{z}^0 z^0 + \bar{z}^1 z^1 = 0,\ \epsilon = \pm 1. \tag{34}$$

For $\epsilon = +1$ equation (34) has no solution in $\boldsymbol{P}^1$. In case $\epsilon = -1$, each solution of (34) satisfies $x^0 \neq 0, x^1 \neq 0$; hence we can take the coordinates of the solution point to be $(z^0, z^1) = (1, \xi)$; $\xi \in K^*$. Then

$$\bar{\xi}\xi = |\xi|^2 = 1\ (\epsilon = -1), \tag{35}$$

i.e., the set of solutions is the unit circle for $K = \mathbf{C}$, and the unit hypersphere S^3 in case $K = \mathbf{H}$. Let now $\boldsymbol{x} \in \boldsymbol{P}^1$ be a variable pre-image point with coordinates $(1, \xi)$, $\xi \in K^*$. For the homogeneous coordinates of the image point $\boldsymbol{y} = F(\boldsymbol{x})$ we then obtain from (30):

$$\epsilon y^0 + \bar{\xi} y^1 = 0.$$

Because of $\xi \in K^*$, necessarily $y^0 \neq 0$; again setting $(y^0, y^1) = (1, \eta)$ this yields

$$\eta = -\epsilon\bar{\xi}^{-1} = -\epsilon\xi/|\xi|^2 \ (\xi \in K^*). \tag{36}$$

In case $\epsilon = -1$ this is nothing but the formula for the *inversion* (or the *map by reciprocal radii*) in the unit circle $S^1 \subset \mathbf{C}$ or in the unit hypersphere $S^3 \subset \mathbf{H}$, respectively. For $\epsilon = +1$ this inversion still has to be composed with the point reflection in 0. In each case these involutions interchange the points $0, \infty \in \hat{K}$. □

Example 7. Let now $K = \mathbf{H}$. Consider the involution J_2 on the quaternionic projective line $\boldsymbol{P}^1$, cf. (23). The associated quadric Q has the equation

$$\bar{x}^0\,\mathrm{i}\,x^0 + \bar{x}^1\,\mathrm{i}\,x^1 = 0. \tag{37}$$

Its solutions are all points $\boldsymbol{x}$ with coordinates $(1, \xi)$ satisfying

$$\xi = \mathrm{j}\cos\alpha + \mathrm{k}\sin\alpha,\ \ \alpha \in \mathbf{R}, \tag{38}$$

i.e., the unit circle in the j, k-plane of $\mathbf{H}$. The image $\boldsymbol{y} = F(\boldsymbol{x})$ of the point $\boldsymbol{x} \in \boldsymbol{P}^1$ with coordinates $(1, \xi)$, $\xi \in \mathbf{H}^*$, has coordinates $(1, \eta)$ satisfying

$$\eta = \mathrm{i}\,\xi\,\mathrm{i}\,/|\xi|^2. \tag{39}$$

In $\mathbf{H} = \mathbf{R}^4$ this is a reflection in the j, k-plane composed with the involution in the unit hypersphere $S^3 \subset \mathbf{H}$; again the points 0 and ∞ of $\hat{\mathbf{H}}$ are interchanged by F. □

1.9.8 Tangents and Tangent Subspaces

In this section we want to describe the possible positions of a line relative to a quadric. This will lead to the notions of tangents and, more generally, tangent subspaces of a quadric.

Example 8. Let again $\operatorname{char} K \neq 2$, let F denote a polar map, and let Q_F be the associated quadric in $\boldsymbol{P}^n$. For $\boldsymbol{x} = [\mathfrak{x}] \in Q_F$ consider the line $\boldsymbol{h} = \boldsymbol{x} \vee \boldsymbol{y}$, where $\boldsymbol{y} = [\mathfrak{y}] \neq \boldsymbol{x}$ is an arbitrary point in $\boldsymbol{P}^n$. Each point $\boldsymbol{z} \in \boldsymbol{h}$, $\boldsymbol{z} \neq \boldsymbol{x}$, can be represented in the form $\boldsymbol{z} = [\mathfrak{x}\xi + \mathfrak{y}]$ with uniquely determined $\xi \in K$. Because of $b(\mathfrak{x}, \mathfrak{x}) = 0$ we have: *$\boldsymbol{z} \in Q_F \cap \boldsymbol{h}$ if and only if $\boldsymbol{z} = \boldsymbol{x}$ or if the following equation is satisfied:*

$$\sigma(\xi)b(\mathfrak{x}, \mathfrak{y}) + b(\mathfrak{y}, \mathfrak{x})\xi + b(\mathfrak{y}, \mathfrak{y}) = 0. \tag{40}$$

Next we will distinguish the following cases:

a) $\operatorname{rk}(F|\boldsymbol{h}) = 0$: This holds if and only if $b(\mathfrak{x}, \mathfrak{y}) = b(\mathfrak{y}, \mathfrak{y}) = 0$; by Corollary 15 it is thus equivalent to $\boldsymbol{x} \vee \boldsymbol{y} \subset Q_F$.

b) $\operatorname{rk}(F|\boldsymbol{h}) = 1$: Then, according to Proposition 17 and Example 4, $Q_{F|\boldsymbol{h}} = Q_F \cap \boldsymbol{h}$ consists of a single point, hence $Q_F \cap \boldsymbol{H} = \{\boldsymbol{x}\}$. This is the case if and only if $b(\mathfrak{x}, \mathfrak{y}) = 0$ as well as $b(\mathfrak{y}, \mathfrak{y}) \neq 0$ in (40). If $\boldsymbol{x}$ is singular, i.e. $\boldsymbol{x} \in \operatorname{Ker} F$, then $b(\mathfrak{x}, \mathfrak{y}) = 0$ for all $\boldsymbol{y} \in \boldsymbol{P}^n$. Thus $(\boldsymbol{x} \vee \boldsymbol{y}) \cap Q_F = \{\boldsymbol{x}\}$ for every $\boldsymbol{y} \in \boldsymbol{P}^n \setminus Q_F$. If $\boldsymbol{x}$ is a regular point of Q_F, then $b(\mathfrak{x}, \mathfrak{y}) = 0$ defines the polar $F(\boldsymbol{x})$ of $\boldsymbol{x}$, and for the points $\boldsymbol{y}\ \iota\ F(\boldsymbol{x})$, $\boldsymbol{y} \neq \boldsymbol{x}$, we now have $\boldsymbol{x} \vee \boldsymbol{y} \subset F(\boldsymbol{x})$. In this case as well as in case a) the line $\boldsymbol{h} = \boldsymbol{x} \vee \boldsymbol{y}$ is called a *tangent* for the quadric Q_F. More generally, we say that a projective subspace $\boldsymbol{B} \in \mathfrak{P}^n$ is *tangent* to the quadric Q_F at a regular point $\boldsymbol{x} \in Q_F$ if $\boldsymbol{x} \in \boldsymbol{B} \subset F(\boldsymbol{x})$. If $\boldsymbol{x} \in Q_F$ is a regular point of the quadric, then $T_{\boldsymbol{x}} Q_F := F(\boldsymbol{x})$ is the *tangent hyperplane* of the quadric Q_F at the point $\boldsymbol{x}$. The totality of tangent subspaces at $\boldsymbol{x}$ forms the bundle $F(\boldsymbol{x})/\boldsymbol{x}$. For the singular points $\boldsymbol{x} \in \operatorname{Ker} F$ tangent subspaces will not be defined.

c) $\operatorname{rk}(F|\boldsymbol{h}) = 2$: In this case $\boldsymbol{h}$ is called a *secant or secant line of* Q_F. In fact, equality holds in (40) if and only if $b(\mathfrak{x}, \mathfrak{y}) \neq 0$. Hence $\boldsymbol{x}$ has to be a regular point of Q_F. Moreover, the rank condition together with (40) implies that $Q_F \cap \boldsymbol{h}$ contains at least one more point $\boldsymbol{z} \neq \boldsymbol{x}$. Since b is σ-Hermitean, the equation

$$\xi = -b(\mathfrak{y}, \mathfrak{x})^{-1} b(\mathfrak{y}, \mathfrak{y})/2 \tag{41}$$

determines a solution for (40). (Because of $\operatorname{char} K \neq 2$, $2 \neq 0$ belongs to the center $Z(K^*)$.) As Examples 6, 7 show, there may well be further solutions; if, however, K is a field, and b is bilinear, then (41) determines the unique point $\boldsymbol{y} \in Q_F \cap \boldsymbol{h}$ different from $\boldsymbol{x}$, cf. Example 5. □

The following proposition contains a geometric interpretation of isotropic subspaces:

Proposition 18. *Let* $\operatorname{char} K \neq 2$, *and let* F *be a polar map of the projective geometry* $\mathfrak{P}^n$ *described by a* σ*-Hermitean biform* b. *Moreover, consider a projective subspace* $\boldsymbol{o} \neq \boldsymbol{A} = \pi(\boldsymbol{W})$ *with associated vector space* $\boldsymbol{W}$. *Then* $\boldsymbol{W}$ *is isotropic if and only if* $\operatorname{Ker}(F|\boldsymbol{A}) \neq \boldsymbol{o}$. *Here,*

$$\operatorname{Ker}(F|\boldsymbol{A}) = (\boldsymbol{A} \wedge \operatorname{Ker} F) \cup \boldsymbol{B}(\boldsymbol{A}), \tag{42}$$

where $\boldsymbol{B}(\boldsymbol{A})$ *denotes the set of those regular points of the quadric* Q_F *at which* $\boldsymbol{A}$ *is tangent to* Q_F.

Proof. The defect subspace of $b|\boldsymbol{W} \times \boldsymbol{W}$ is the vector space associated with $\operatorname{Ker}(F|\boldsymbol{A})$, which immediately implies the first assertion. Moreover, $\boldsymbol{x} \in \operatorname{Ker}(F|\boldsymbol{A})$ if and only if

$$\boldsymbol{x} \in \boldsymbol{A} \subset F(\boldsymbol{x}). \tag{43}$$

For this, there are just two mutually excluding possibilities: Either $F(\boldsymbol{x}) = \boldsymbol{P}^n$, i.e. $\boldsymbol{x} \in \boldsymbol{A} \wedge \operatorname{Ker} F$, or $\operatorname{Dim} F(\boldsymbol{x}) = n - 1$, and then $\boldsymbol{x}$ belongs to $\boldsymbol{B}(\boldsymbol{A})$. □

Corollary 19. *Under the assumptions of Proposition* 18 , *let* F *be a polarity. Then* $\boldsymbol{W} \neq \boldsymbol{o}$ *is isotropic if and only if* $\boldsymbol{A} = \pi(\boldsymbol{W})$ *is tangent to* Q_F *in at least one point* $\boldsymbol{x}$; *the set* $\boldsymbol{B}(\boldsymbol{A})$ *of points of contact is a projective subspace of* $\boldsymbol{A}$, *whose vector space is the defect subspace of* $b|\boldsymbol{W} \times \boldsymbol{W}$. □

Exercise 8. Consider the hyperboloid and show that there are tangent subspaces $\boldsymbol{A} \subset \boldsymbol{P}^n$ for which $\boldsymbol{B}(\boldsymbol{A}) \neq Q_F \cap \boldsymbol{A}$; hence, apart from the points of contact, other points of Q_F may belong to $\boldsymbol{A}$ as well (cf. Figure 2.1).

1.9.9 Dualization: Coquadrics

Although we defined correlations and, more generally, correlative maps as monotonously decreasing maps between projective geometries, we nevertheless preferred to use points as the basic elements, e.g. in the definition of quadrics. The notion which is, in this sense, dual to that of a correlative map now starts from hyperplanes as the basic elements; to each hyperplane $\boldsymbol{U} \in \boldsymbol{P}'^n$ a point $\boldsymbol{x} = F(\boldsymbol{U}) \in \boldsymbol{Q}^m$ or the nopoint $\boldsymbol{o}$ are assigned, where F is determined by a σ-linear map $a : \boldsymbol{V}' \to \boldsymbol{W}$, σ an anti-isomorphism:

$$\boldsymbol{U} = [\mathfrak{u}] \in \boldsymbol{P}'^n_o \longmapsto \boldsymbol{x} = [a(\mathfrak{u})] \in \boldsymbol{Q}^m_o. \tag{44}$$

The pencil of hyperplanes incident with $\boldsymbol{x}$ is determined by $\tau(\boldsymbol{x}) = \boldsymbol{x}^\perp$, and we thus obtain an F-correspondence between hyperplanes $\boldsymbol{U} \in \boldsymbol{P}'^n$, $\boldsymbol{D} = [\mathfrak{d}] \in \boldsymbol{Q}'^m$ via

$$\boldsymbol{D} \ \iota \ F(U] \Longleftrightarrow B(\mathfrak{d}, \mathfrak{u}) := (\mathfrak{d}|a(\mathfrak{u})) = 0. \tag{45}$$

Now, the σ-biform B is linear in $\mathfrak{d} \in \boldsymbol{W}'$ and σ-linear in $\mathfrak{u} \in \boldsymbol{V}'$. Again, let $\operatorname{char} K \neq 2$, $\mathfrak{Q}^m = \mathfrak{P}^n$, and let B be σ-Hermitean over $\boldsymbol{V}' \times \boldsymbol{V}'$, cf. (2). Then

$$B(\mathfrak{u}, \mathfrak{u}) = 0 \tag{46}$$

defines the set $Q'_F := \{\boldsymbol{U} = [\mathfrak{u}] \in \boldsymbol{P}'^n | \boldsymbol{U} \ \iota \ \boldsymbol{F}(\boldsymbol{U})\}$, which will be called the *coquadric* associated with F. The traditional names for Q_F and Q'_F, respectively, are *curve* ($n = 2$), *surface* ($n = 3$), or *hypersurface* (n arbitrary) *of second order* or *second class*. Dual to the notion of the tangent hyperplane $F(\boldsymbol{x})$ at a regular point $\boldsymbol{x} \in Q_F$ is the notion of the *point of contact* $F(\boldsymbol{U}) \neq \boldsymbol{o}$ at the regular hyperplane $\boldsymbol{U} \in Q'_F$. If F is a polarity, then for the pairs $(\boldsymbol{x}, \boldsymbol{X})$ of pole and polar, $\boldsymbol{X} = F(\boldsymbol{x})$, we immediately have

$$\boldsymbol{x} \ \iota \ F(\boldsymbol{x}) \Longleftrightarrow \boldsymbol{X} \ \iota \ F(\boldsymbol{X}), \tag{47}$$

and hence

$$Q'_F = F(Q_F), \ \ Q_F = F(Q'_F) : \tag{48}$$

The coquadric Q'_F (quadric Q_F) is the set of tangent hyperplanes (points of contact) of the quadric Q_F (coquadric Q'_F).

Exercise 9. Let F be a polarity, and let $(u_j) = (\sigma(x^i))(a_{ij})$ be its matrix representation in homogeneous coordinates. Show that, in the same coordinate system, equation (46) for the coquadric Q'_F corresponding to F is described by

$$\sum_{i,j=0}^{n} u_i c^{ij} \sigma(u_j) = 0 \text{ with } (c^{ij}) := (a_{ij})^{-1}. \tag{49}$$

1.10 Restrictions and Extensions of Scalars

In order to further investigate the dependence of the quadric on the scalar domain, particularly obvious in Examples 9.2, 9.3, it will be useful to consider how the projective geometry changes under its variation. In this section, let L always be a skew field, and let $K \subset L$ be a skew subfield. Then L is also a right vector space over K; we will always assume $r := \dim_K L < \infty$. Generalizing Definition II.8.5.4 L is then called a *finite extension of K of degree r*. For the cases we will mainly deal with here the degrees are:

$$\dim_{\mathbf{R}} \mathbf{C} = \dim_{\mathbf{C}} \mathbf{H} = 2, \ \dim_{\mathbf{R}} \mathbf{H} = 4.$$

Fix a basis (a_ρ), $\rho = 1, \ldots, r$, for L such that $a_1 = 1$ is the unit element of L, and let $\boldsymbol{V}$ be a vector space over L. *Restricting the scalars* to K, more precisely, performing scalar multiplication in the vector space $\boldsymbol{V}$ only with elements $\mu \in K$,

$$(\mathfrak{x}, \mu) \in \boldsymbol{V} \times K \longmapsto \mathfrak{x}\mu \in \boldsymbol{V},$$

$\boldsymbol{V}$ becomes a vector space over K to be denoted by $\boldsymbol{V}_{|K}$. Decomposing the scalars $\lambda \in L$ with respect to the basis (a_ρ) of L over K the scalar operations of L on $\boldsymbol{V}$ are transformed into corresponding ones of K on $\boldsymbol{V}_{|K}$. Applying this to the vector coordinates of $\boldsymbol{V}$ over L the following is easily verified (cf. Exercise I.4.4.8)

Lemma 1. *Let $\boldsymbol{V}$ be an n-dimensional right vector space over L, and let $(\mathfrak{b}_l)$ be a basis for $\boldsymbol{V}$. Then $(\mathfrak{b}_l a_\rho)$, $l = 1, \ldots, n$, $\rho = 1, \ldots, r$, is a basis for the vector space $\boldsymbol{V}_{|K}$ over K obtained by restricting the scalars, hence*

$$\dim_K \boldsymbol{V}_{|K} = \dim_L \boldsymbol{V} \cdot \dim_K L. \tag{1}$$

□

1.10.1 Hopf Fibrations

The vector spaces $\boldsymbol{V}^n$ over L, $\boldsymbol{V}^{nr}_{|K}$, and L over K occurring in (1) determine projective geometries. We will study the projective point spaces $\boldsymbol{P}^{n-1}(\boldsymbol{V})$, $\boldsymbol{P}^{nr-1}(\boldsymbol{V}_{|K})$, $\boldsymbol{P}^{r-1}(L_{|K})$, and the canonical maps (cf. (1.2))

$$\begin{aligned} \pi &: \boldsymbol{V} \setminus \{\mathfrak{o}\} \longrightarrow \boldsymbol{P}^{n-1}(\boldsymbol{V}), \\ \pi_K &: \boldsymbol{V} \setminus \{\mathfrak{o}\} \longrightarrow \boldsymbol{P}^{nr-1}(\boldsymbol{V}_{|K}). \end{aligned}$$

Projectively, restricting the scalars results in blowing up each point $\boldsymbol{x} \in \boldsymbol{P}^{n-1}(\boldsymbol{V})$ to a projective subspace $\theta^{-1}(\boldsymbol{x}) \subset \boldsymbol{P}^{nr-1}(\boldsymbol{V}_{|K})$ isomorphic to $\boldsymbol{P}^{r-1}(L_{|K})$. More precisely, we have

Proposition 2. *Let $\boldsymbol{V}$ be an n-dimensional right vector space over the skew field L, let $\boldsymbol{P}^{n-1}(\boldsymbol{V})$ be the corresponding projective space, and let $\boldsymbol{P}^{nr-1}(\boldsymbol{V}_{|K})$ be the projective space associated with the restriction $\boldsymbol{V}_{|K}$ to the skew subfield $K \subset L$, $r = \dim_K L < \infty$. Then there is a uniquely determined surjective map*

$$\theta : \boldsymbol{P}^{nr-1}(\boldsymbol{V}_{|K}) \longrightarrow \boldsymbol{P}^{n-1}(\boldsymbol{V}) \tag{2}$$

satisfying the following equation involving the canonical maps

$$\pi = \theta \circ \pi_K . \tag{3}$$

The preimage $\theta^{-1}(\boldsymbol{x})$ is a projective $(r-1)$-plane in $\boldsymbol{P}^{nr-1}(\boldsymbol{V}_{|K})$. If $(\mathfrak{b}_l)$, $l = 1, \ldots, n$, is a basis of $\boldsymbol{V}$ over L, and U_1 is the domain of definition for the inhomogeneous projective coordinates (ξ_1^j), cf. Lemma 2.6, then

$$\boldsymbol{x} \in U_1 \longmapsto s(\boldsymbol{x}) := \mathfrak{b}_1 + \sum_{j=2}^{n} \mathfrak{b}_j \xi_1^j(\boldsymbol{x}) \tag{4}$$

determines a section s for the canonical map π, i.e. $\pi \circ s = \mathrm{id}_{U_1}$, and the map

$$\Phi : \boldsymbol{y} = [\mathfrak{y}] \in \theta^{-1}(U_1) \longmapsto (\theta(\boldsymbol{y}), [\mathfrak{y}/s(\theta(\boldsymbol{y}))]_K) \in U_1 \times \boldsymbol{P}^{r-1}(L_{|K}) \tag{5}$$

is a bijection (similarly for the other domains U_j).

Proof. For each $\mathfrak{x} \in \boldsymbol{V} \setminus \{\mathfrak{o}\}$ we have $\pi_K(\mathfrak{x}) = \mathfrak{x}K \subset \mathfrak{x}L = \pi(\mathfrak{x})$, and hence $\theta(\mathfrak{x}K) := \mathfrak{x}L$ is surjective and uniquely determined. Obviously, $\mathfrak{x}L = \mathfrak{y}L$ if and only if there is $k \in L^*$ such that $\mathfrak{y} = \mathfrak{x}k$. Decomposing k with respect to the basis (a_ρ) of L we obtain for the complete preimage of $\boldsymbol{x} = \mathfrak{x}L \in \boldsymbol{P}^{n-1}(\boldsymbol{V})$:

$$\theta^{-1}([\mathfrak{x}]_L) = [\mathfrak{x}]_K \vee [\mathfrak{x}a_2]_K \vee \ldots \vee [\mathfrak{x}a_r]_K . \tag{6}$$

Since the vectors $(\mathfrak{x}a_\rho)$, $\rho = 1, \ldots, r$, are linearly independent over K for each $\mathfrak{x} \neq \mathfrak{o}$, the corresponding points span an $(r-1)$-plane in $\boldsymbol{P}^{nr-1}(\boldsymbol{V}_{|K})$. To render (5) meaningful, recall the definition of the ratio of two vectors from I.4.3. By the definition of θ, $\boldsymbol{y} = [\mathfrak{y}]_K \in \theta^{-1}(\boldsymbol{x})$, if $\mathfrak{y} = s(\boldsymbol{x})k$ for some $k \in L^*$; here k is uniquely determined by $\boldsymbol{y}$ up to a right factor $\kappa \in K^*$. Hence

$$[\mathfrak{y}/s(\theta(\boldsymbol{y}))]_K = [k]_K \in \boldsymbol{P}^{r-1}(L_{|K})$$

is uniquely determined by $\boldsymbol{y}$. This readily implies that Φ is a bijection. □

The situation described in Proposition 2 is the projective analog of a topological fibration: $\boldsymbol{P}^{nr-1}(\boldsymbol{V}_{|K})$ is represented as a disjoint union of the "fibres" (6), parametrized by $\boldsymbol{x} = [\mathfrak{x}]_L$ varying in the space $\boldsymbol{P}^{n-1}(\boldsymbol{V})$. The latter is covered by "chart domains" U_j whose preimages $\theta^{-1}(U_j)$ are direct products as described by (5), where Φ maps the fibres $\theta^{-1}(\boldsymbol{x})$ onto the fibres $p_1^{-1}(z)$, $z = (\theta(x), [k]_K)$, of the product on the right-hand side. It is easily verified that the second projection from the direct product onto the fibres,

$$p_2 \circ \Phi|\theta^{-1}(\boldsymbol{x}) : \theta^{-1}(\boldsymbol{x}) \longrightarrow \boldsymbol{P}^{r-1}(L_{|K}).$$

is a projectivity. Using inhomogeneous coordinates to introduce a topology or the structure of a differentiable manifold on the spaces over $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$, all the maps occurring in (3)–(5) become open and continuous or arbitrarily often differentiable, respectively, and θ turns out to be a fibration even in the topological sense.

Example 1. The fibrations θ are closely related to the fibrations of spheres described in 1935 by H. Hopf [55]. To see this we define the covering maps $\lambda : S^N \to \boldsymbol{P}^N_{\mathbf{R}}$. Look at S^N as the unit sphere in $\mathbf{R}^{\mathbf{N+1}}$ and interpret $\boldsymbol{P}^N_{\mathbf{R}}$ as the set of lines through its center, then

$$\lambda : \mathfrak{y} \in S^N \longmapsto \lambda(\mathfrak{y}) := [\mathfrak{y}]_{\mathbf{R}} \in \boldsymbol{P}^N_{\mathbf{R}} \tag{7}$$

is a double covering: In fact, the preimages of any point $\boldsymbol{y} = [\mathfrak{y}]_{\mathbf{R}} \in \boldsymbol{P}^N_{\mathbf{R}}$ are $\pm\mathfrak{y} \in S^N$, and each projective line has the corresponding principal circle in S^N as its preimage. Hence for $K = \mathbf{R} \subset L = \mathbf{C}$ the map

$$\theta_o := \theta \circ \lambda : S^{2n-1} \longrightarrow \boldsymbol{P}^{n-1}_{\mathbf{C}} \tag{8}$$

is a fibration of the sphere S^{2n-1} by circles S^1. In particular, for $n = 2$ the projective line is homeomorphic to the Riemann sphere, $\boldsymbol{P}^1_{\mathbf{C}} \cong \hat{\mathbf{C}} \cong S^2$ (cf. Example 1.3), and so we obtain the *Hopf fibration*: The sphere S^3 is fibred into circles S^1 over the sphere S^2,

$$\theta_o : S^3 \longrightarrow S^2, \qquad \theta_o^{-1}(\boldsymbol{x}) \cong S^1. \tag{9}$$

Analogously, embedding $\mathbf{R} \subset \mathbf{H}$ leads to the maps

$$\theta_o = \theta \circ \lambda, \text{ with } \lambda : S^{4n-1} \to \boldsymbol{P}^{4n-1}_{\mathbf{R}}, \ \theta : \boldsymbol{P}^{4n-1}_{\mathbf{R}} \to \boldsymbol{P}^{n-1}_{\mathbf{H}}, \tag{10}$$

where θ and θ_o are the fibrations by projective spaces $\boldsymbol{P}^3_{\mathbf{R}}$ and spheres S^3, respectively; in fact, by Proposition 2 we have

$$\theta^{-1}(\boldsymbol{x}) \cong \boldsymbol{P}^3_{\mathbf{R}}, \ \theta_o^{-1}(\boldsymbol{x}) = \lambda^{-1} \circ \theta^{-1}(\boldsymbol{x}) \cong S^3. \tag{11}$$

For $n = 2$ we use $\hat{\mathbf{H}} \cong S^4$ to obtain a fibration of the sphere S^7 by three-dimensional spheres S^3 over the sphere S^4,

$$\theta_o : S^7 \longrightarrow S^4, \qquad \theta_o^{-1}(\boldsymbol{x}) \cong S^3. \tag{12}$$

Replacing the quaternions by the *Cayley numbers*or *octonions* as they are also called, one obtains a further *Hopf fibration* (cf. N. Steenrod [100], §II.20)

$$\theta_o : S^{15} \longrightarrow S^8, \qquad \theta_o^{-1}(\boldsymbol{x}) \cong S^7. \tag{13}$$

Finally, the embedding $\mathbf{C} \subset \mathbf{H}$ leads to the series of fibrations

$$\theta : \boldsymbol{P}_{\mathbf{C}}^{2n-1} \longrightarrow \boldsymbol{P}_{\mathbf{H}}^{n-1}. \tag{14}$$

By Exercise II.8.9.9 we have $\mathbf{H}_{|\mathbf{C}} \cong \mathbf{C}^2$, and hence the fibres are

$$\theta^{-1}(\boldsymbol{x}) \cong \boldsymbol{P}_{\mathbf{C}}^{1} \cong S^2. \tag{15}$$

□

Exercise 1. Prove under the assumptions of Proposition 2: a) For every three collinear points $\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2 \in \boldsymbol{P}^{n-1}(\boldsymbol{V})$ there always are three collinear points $\boldsymbol{b}_0, \boldsymbol{b}_1, \boldsymbol{b}_2 \in \boldsymbol{P}^{nr-1}(\boldsymbol{V}_{|K})$ with $\theta(\boldsymbol{b}_i) = \boldsymbol{a}_i$. – b) For every four collinear points $\boldsymbol{a}_i \in \boldsymbol{P}^{n-1}(\boldsymbol{V})$, $i = 0, \ldots, 3$, there are four collinear points $\boldsymbol{b}_i \in \boldsymbol{P}^{nr-1}(\boldsymbol{V}_{|K})$ such that $\theta(\boldsymbol{b}_i) = \boldsymbol{a}_i$ if and only if $K \cap \mathrm{CR}(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3) \neq \emptyset$.

For the subsequent sections the following considerations will be useful: If $\boldsymbol{V}, \boldsymbol{W}$ are vector spaces over the skew field L, and $\boldsymbol{P}(\boldsymbol{V}), \boldsymbol{P}(\boldsymbol{W})$ are the corresponding projective point spaces, then a map $f : \boldsymbol{V} \setminus \{\mathfrak{o}\} \to \boldsymbol{W} \setminus \{\mathfrak{o}\}$ generates a unique map $F : \boldsymbol{P}(\boldsymbol{V}) \to \boldsymbol{P}(\boldsymbol{W})$ with $F(\pi(\mathfrak{x})) = \pi(f(\mathfrak{x}))$ if and only if for all $\mathfrak{x} \in \boldsymbol{V} \setminus \{\mathfrak{o}\}$ and each $\lambda \in L^*$ the vectors $f(\mathfrak{x}\lambda)$ and $f(\mathfrak{x})$ are linearly independent. For this to hold, it is sufficient that there is a function $h : L^* \to L^*$ satisfying $f(\mathfrak{x}\lambda) = f(\mathfrak{x})h(\lambda)$ for all $\mathfrak{x} \in \boldsymbol{V} \setminus \{\mathfrak{o}\}$ and each $\lambda \in L^*$ (or $f(\mathfrak{x}\lambda) = h(\lambda)f(\mathfrak{x})$, in case $\boldsymbol{V}$ is a right and $\boldsymbol{W}$ is a left vector space, etc.).

Exercise 2. Prove under the assumptions and in the notations of Proposition 2: If $h : L^* \to L^*$ is a map with $h(K^*) \subset K^*$, and $f : \boldsymbol{V} \setminus \{\mathfrak{o}\} \to \boldsymbol{V} \setminus \{\mathfrak{o}\}$ is a map with $f(\mathfrak{x}\lambda) = f(\mathfrak{x})h(\lambda)$, then the maps $F : \boldsymbol{P}(\boldsymbol{V}_{|K}) \to \boldsymbol{P}(\boldsymbol{V}_{|K})$, $\hat{F} : \boldsymbol{P}(\boldsymbol{V}) \to \boldsymbol{P}(\boldsymbol{V})$ generated by f are well-defined, and, moreover, $\theta \circ F = \hat{F} \circ \theta$.

1.10.2 Complex Structures

The restriction of a complex vector space $\boldsymbol{V}$ to the real numbers is also called its *realification*. This leads to the notion of a complex structure on a real vector space and to the representation of the algebra of complex linear maps as a subalgebra of the real endomorphism algebra. In particular, we obtain a representation of the complex linear group $\boldsymbol{GL}(n, \mathbf{C})$ as a subgroup of the real linear group $\boldsymbol{GL}(2n, \mathbf{R})$. In the sequel we will describe some consequences for the projective properties of the fibration

$$\theta : \boldsymbol{P}^{2n-1}(\boldsymbol{V}_{|\mathbf{R}}) \longrightarrow \boldsymbol{P}^{n-1}(\boldsymbol{V}).$$

In $\boldsymbol{V}$ scalar multiplication by $\mathrm{i} = \sqrt{-1}$ defines a $\mathbf{C}$-linear and all the more $\mathbf{R}$-linear isomorphism, the *complex structure*:

$$I(\mathfrak{x}) = \mathfrak{x} \cdot \mathrm{i},\ I^2 = -\,\mathrm{id}_{\boldsymbol{V}}\,. \tag{16}$$

Since I has no real eigenvalues, it generates an equally denoted fixed-point free involution $I \in \boldsymbol{PL}(\boldsymbol{P}_{\mathbf{R}}^{2n-1})$, which hence has the normal form (4.42). In fact, if $\mathfrak{a}_l$, $l = 1, \ldots, n$, is a basis of $\boldsymbol{V}$, then the decomposition of the complex coordinates into real and imaginary part,

$$\mathfrak{x} = \sum_{l=1}^{n} \mathfrak{a}_l z^l = \sum_{l=1}^{n} \mathfrak{a}_l (x^l + \mathrm{i}\, x^{n+l}) = \sum_{\nu=1}^{2n} \mathfrak{a}_\nu x^\nu$$

with

$$\mathfrak{a}_{n+l} := \mathfrak{a}_l \,\mathrm{i} = I(\mathfrak{a}_l), \text{ i.e. } I(\mathfrak{a}_{n+l}) = -\mathfrak{a}_l, \tag{17}$$

yields a basis for $\boldsymbol{V}^{2n}_{\mathbf{R}}$, in which I has this normal form. If, conversely, $\boldsymbol{W}^{2n}$ is a $2n$-dimensional real vector space, and $I \in \boldsymbol{GL}(\boldsymbol{W}^{2n}, \mathbf{R})$ is a linear automorphism of $\boldsymbol{W}$ with $I^2 = -\,\mathrm{id}_{\boldsymbol{W}}$, then

$$(\mathfrak{x}, z) \in \boldsymbol{W} \times \mathbf{C} \longmapsto \mathfrak{x} z := \mathfrak{x}\zeta + I(\mathfrak{x})\eta \in \boldsymbol{W} \text{ for } z = \zeta + \mathrm{i}\,\eta \tag{18}$$

defines the structure of an n-dimensional complex vector space on $\boldsymbol{W}$; we will denote it by $\boldsymbol{V}^n$, and so, obviously, $\boldsymbol{W}^{2n} = \boldsymbol{V}_{|\mathbf{R}}$. For this reason, an $\mathbf{R}$-linear map I on $\boldsymbol{W}$ satisfying $I^2 = -\,\mathrm{id}_{\boldsymbol{W}}$ is called a *complex structure* (cf. Definition II.8.9.4) on $\boldsymbol{W}$; according to the result of Exercise 4.15 all complex structures are similar. Of course, any complex linear map $a : \boldsymbol{V} \to \boldsymbol{V}$ is real linear as well. Expressing it in a basis $(\mathfrak{a}_\nu)$, $\nu = 1, \ldots, 2n$, of $\boldsymbol{V}_{|\mathbf{R}}$ satisfying (17) we obtain for the real and imaginary part, using the obvious notation:

$$a(\mathfrak{z}) = (A + \mathrm{i}\, B)(\mathfrak{x} + \mathrm{i}\,\mathfrak{y}) = (A\mathfrak{x} - B\mathfrak{y}) + \mathrm{i}(A\mathfrak{y} + B\mathfrak{x}); \tag{19}$$

here $A, B \in \boldsymbol{M}_n(\mathbf{R})$ are real square matrices, and $\mathfrak{x}, \mathfrak{y} \in \mathbf{R}^{\mathbf{n}}$, cf. Exercise I.5.4.18.

Exercise 3. Show with the notations just introduced: a) Restricting the scalars from $\mathbf{C}$ to $\mathbf{R}$ defines a homomorphic embedding of the real endomorphism algebras

$$c\colon\ (\mathrm{End}_{\mathbf{C}}(\boldsymbol{V}^n))_{|\mathbf{R}} \longrightarrow \mathrm{End}_{\mathbf{R}}(\boldsymbol{W}^{2n}). \tag{20}$$

b) A linear map $b \in \mathrm{End}_{\mathbf{R}}(\boldsymbol{W}^{2n})$ belongs to $(\mathrm{End}_{\mathbf{C}}(\boldsymbol{V}^n))_{|\mathbf{R}}$ if and only if it commutes with the complex structure I, i.e., it satisfies $I \circ b = b \circ I$, and this again holds if and only if its matrix (β^λ_μ) has the following block structure with respect to a basis (17) (cf. Exercise I.5.4.18)

$$C := (\beta^\lambda_\mu) = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \text{ with } A, B \in \boldsymbol{M}_n(\mathbf{R}). \tag{21}$$

c) The determinant of any matrix of the form (21) is non-negative. (Note that in every basis (17) I has the matrix

$$I : \begin{pmatrix} 0 & -1_n \\ 1_n & 0 \end{pmatrix} \text{ with } 1_n = (\delta^i_j) \in \boldsymbol{M}_n(\mathbf{R}). \tag{22}$$

Hence we may suppose that the matrix of the complex endomorphism a has Jordan normal form (cf. Proposition I.5.8.4 and Corollary I.5.8.2); use the special form the matrices A, B from (19) have in this case, insert it into (21) and apply the Laplace expansion theorem to compute the determinant of C recursively.) – d) $\det(C) = |\det(b)|^2$. – e) After realification of $\boldsymbol{V}^n$ the complex linear group $\boldsymbol{GL}(\boldsymbol{V}^n)$ becomes a subgroup of the real linear group $\boldsymbol{GL}(\boldsymbol{W}^{2n})$. Show that the subgroup of projectivities generated by all $g \in \boldsymbol{GL}(\boldsymbol{V}^n)$ acts transitively on the real projective space $\boldsymbol{P}(\boldsymbol{W}^{2n})$.

Exercise 4. Let I be a complex structure on the real vector space $\boldsymbol{W}^{2n}$. Denote by $p_I \in \boldsymbol{PL}(\boldsymbol{P}^{2n-1}_{\mathbf{R}})$ the projective involution generated by I. Again, let $\boldsymbol{V}^n$ be the complex vector space determined by I, and let $\theta : \boldsymbol{P}^{2n-1}_{\mathbf{R}} \to \boldsymbol{P}^{n-1}_{\mathbf{C}}$ be the fibration from Proposition 2. Let $p_a : \boldsymbol{P}^{2n-1}_{\mathbf{R}} \to \boldsymbol{P}^{2n-1}_{\mathbf{R}}$ denote the projective map generated by the linear map $a \in \mathrm{End}_{\mathbf{R}}(\boldsymbol{W}^{2n})$. Prove that there exists a map $\hat{p}_a : \boldsymbol{P}^{n-1}_{\mathbf{C}} \to \boldsymbol{P}^{n-1}_{\mathbf{C}}$ satisfying $\theta \circ p_a = \hat{p}_a \circ \theta$ if and only if $p_a \circ I = I \circ p_a$; here $\hat{p}_a$ is projective or anti-projective.

1.10.3 Quaternionic Structures

Restricting in (20) the homomorphism c to the linear group $\boldsymbol{GL}(\boldsymbol{V}^n, \mathbf{C})$ the matrices in (21) are all invertible, hence we obtain a *real representation of* $\boldsymbol{GL}(\boldsymbol{V}^n, \mathbf{C})$ *as a subgroup of* $\boldsymbol{GL}(2n, \mathbf{R})$. Now we want to apply the restriction of scalars from $\mathbf{H}$ to $\mathbf{C}$ in a similar way to produce a *complex representation of the quaternionic linear group* $\boldsymbol{GL}(n, \mathbf{H})$ *in* $\boldsymbol{GL}(2n, \mathbf{C})$.

Example 2. If $\boldsymbol{U}^n$ is an n-dimensional right vector space over the quaternions $\mathbf{H}$, then by restricting the scalars to $\mathbf{C}$ (cf. § I.2.3 and Exercise II.8.9.9) we obtain a $2n$-dimensional vector space $\boldsymbol{V}^{2n} = \boldsymbol{U}^n_{|\mathbf{C}}$ over $\mathbf{C}$ together with the fibration (14). To see this we write the quaternions in the form

$$q = z + \mathrm{j}\, w \in \mathbf{H}, \text{ with } z,\, w \in \mathbf{C}$$

and apply this representation, e.g., to the coordinates of a point. Note that the dilation $d_{\mathrm{j}} : \mathfrak{x} \mapsto \mathfrak{x}\,\mathrm{j}$ is neither $\mathbf{H}$- nor $\mathbf{C}$-linear; because of $z\,\mathrm{j} = \mathrm{j}\,\bar{z}$ for $z \in \mathbf{C}$, it rather defines a conjugate linear map $J(\mathfrak{x}) := \mathfrak{x}\,\mathrm{j}$ with

$$J(\mathfrak{x}\lambda) = J(\mathfrak{x})\bar{\lambda} \quad (\lambda \in \mathbf{C}),\ J^2 = -\mathrm{id}_{\boldsymbol{V}}\,. \tag{23}$$

Thus J determines a fixed-point free, involutive collineation on $\boldsymbol{P}^{2n-1}_{\mathbf{C}}(\boldsymbol{V})$, which has the normal form (4.43) in a suitable basis of $\boldsymbol{V}^{2n}$. Starting from an arbitrary basis $(\mathfrak{a}_l)$, $l = 1, \ldots, n$, of the quaternionic vector space $\boldsymbol{U}^n$ such a basis is constructed as

$$\mathfrak{a}_{n+l} := \mathfrak{a}_l\,\mathrm{j} = J(\mathfrak{a}_l), \text{ i.e. } J(\mathfrak{a}_{n+l}) = -\mathfrak{a}_l. \tag{24}$$

If, conversely, J is an involutive conjugate linear map on the complex vector space $\boldsymbol{V}^{2n}$ satisfying (23), then

$$(\mathfrak{x}, q) \in \boldsymbol{V} \times \mathbf{H} \longmapsto \mathfrak{x}q := \mathfrak{x}z + J(\mathfrak{x})w \in \boldsymbol{V} \text{ for } q = z + \mathrm{j}\,w \in \mathbf{H} \ \ (z,\, w \in \mathbf{C}) \tag{25}$$

defines the structure of an n-dimensional, right, quaternionic vector space over $\boldsymbol{V}$, that will again be denoted by $\boldsymbol{U}^n$; here $\boldsymbol{U}^n_{|\mathbf{C}} = \boldsymbol{V}^{2n}$. For this reason, J is called a *quaternionic structure* on $\boldsymbol{V}^{2n}$. Exercise 4.15 immediately implies that all quaternionic structures are similar: For two quaternionic structures J_1, J_2 there is always $c \in \boldsymbol{GL}(\boldsymbol{V}^{2n}, \mathbf{C})$ such that $J_2 = c \circ J \circ c^{-1}$. Since each $\mathbf{H}$-linear map a is also $\mathbf{C}$-linear, decomposing the coordinates into their complex components with respect to a basis (24) we obtain (cf. Exercise II.8.9.9):

$$a(\mathfrak{z}) \hat{=} (A + \mathrm{j}\, B)(\mathfrak{x} + \mathrm{j}\, \mathfrak{y}) = (A\mathfrak{x} - \bar{B}\mathfrak{y}) + \mathrm{j}(B\mathfrak{x} + \bar{A}\mathfrak{y}). \tag{26}$$

Here $\bar{A}$ denotes the complex conjugate matrix of A. □

Exercise 5. Show in the notations of Example 2 that the canonical embedding

$$\iota : \mathrm{End}_{\mathrm{H}}(\boldsymbol{U}^n) \longrightarrow (\mathrm{End}_{\mathrm{C}}(\boldsymbol{V}^{2n}))_{|\mathbf{R}} \tag{27}$$

is a homomorphism of real algebras. (Note that according to Proposition II.7.2.5 $\mathrm{End}_{\mathrm{H}}(\boldsymbol{U}^n)$ is only a $Z(\mathbf{H})$-algebra.) A $\mathbf{C}$-linear map $b \in \mathrm{End}_{\mathrm{C}}(\boldsymbol{V}^{2n})$ belongs to $\mathrm{End}_{\mathrm{H}}(\boldsymbol{U}^n)$ if and only if it commutes with the quaternionic structure J, i.e.

$$J \circ b = b \circ J. \tag{28}$$

This again holds if and only if its matrix has the block structure

$$(\beta^\lambda_\mu) = \begin{pmatrix} A & -\bar{B} \\ B & \bar{A} \end{pmatrix} \text{ with } A, B \in \boldsymbol{M}_n(\mathbf{C}). \tag{29}$$

with respect to a basis (24).

Exercise 6. Show that every complex matrix of the form (29) has real determinant. If (29) is the matrix of an $\mathbf{H}$-linear map b, then the *norm* $N(b) := N_{\mathrm{C}}(b) := \det(\beta^\lambda_\mu)$ does not depend on the choice of the basis $(\mathfrak{a}_l)$. (See also Exercise 8.)

Exercise 7. Let $(\mathfrak{a}_l)$ be a basis of the quaternionic vector space $\boldsymbol{U}^n$. Splitting the vector coordinates into their real components:

$$\mathfrak{x} = \mathfrak{a}_l q^l = \mathfrak{a}_l x^l_0 + \mathfrak{a}_l\, \mathrm{i}\, x^l_1 + \mathfrak{a}_l\, \mathrm{j}\, x^l_2 + \mathfrak{a}_l\, \mathrm{k}\, x^l_3 \tag{30}$$

(sum convention, $l = 1, \ldots, n$) we obtain the basis $(\mathfrak{a}_l, \mathfrak{a}_l\,\mathrm{i}, \mathfrak{a}_l\,\mathrm{j}, \mathfrak{a}_l\,\mathrm{k})$ for the real vector space $\boldsymbol{W}^{4n} = \boldsymbol{U}^n_{|\mathbf{R}}$. Now we combine the coordinates of $\mathfrak{x}$ into blocks $\mathfrak{x}_\alpha := (x^l_\alpha) \in \mathbf{R}^n$, $\alpha = 0, \ldots, 3$. Then each $\mathbf{R}$-linear map $a \in \mathrm{End}_{\mathbf{R}}(\boldsymbol{W}^{4n})$ can be represented by a block matrix:

$$(\mathfrak{y}_\alpha) = (A_{\alpha\beta})(\mathfrak{x}_\beta),\ \mathfrak{x}_\beta \in \mathbf{R}^n, A_{\alpha\beta} \in \boldsymbol{M}_n(\mathbf{R}).$$

Show that a is $\mathbf{H}$-linear if and only if its block matrix has the following special form:

$$(A_{\alpha\beta}) = \begin{pmatrix} A_0 & -A_1 & -A_2 & -A_3 \\ A_1 & A_0 & -A_3 & A_2 \\ A_2 & A_3 & A_0 & -A_1 \\ A_3 & -A_2 & A_1 & A_0 \end{pmatrix}. \tag{31}$$

Here the blocks of (31) depend on the quaternionic coordinate matrix of a via

$$(q_{lm}) = A_0 + \mathrm{i}\,A_1 + \mathrm{j}\,A_2 + \mathrm{k}\,A_3 \in \boldsymbol{M}_n(\mathbf{H}). \tag{32}$$

One obtains essentially the same block matrix structure starting from (29) and passing to real coordinates.

Exercise 8. Let $\boldsymbol{V}^n$ be an n-dimensional right vector space over $\mathbf{H}$ and consider $a \in \mathrm{End}_{\mathbf{H}}(\boldsymbol{V}^n)$. Prove: a) There is a basis $(\mathfrak{a}_s)$, $s = 1, \ldots, n$, of $\boldsymbol{V}^n$ such that a has a Jordan matrix (I.5.8.29) with respect to $(\mathfrak{a}_s)$ in which the diagonal entries $\lambda_\kappa = \xi_\kappa + \mathrm{i}\,\eta_\kappa$ are complex numbers with $\eta_\kappa \geq 0$. – b) Two such maps a, b are $\mathbf{H}$-similar, i.e., there is a $c \in GL(\boldsymbol{V}^n, \mathbf{H})$ with $b = c \circ a \circ c^{-1}$ if and only if they have identical Jordan matrices with respect to suitable bases. – c) (cf. Exercise 6)

$$N(a) \geq 0. \tag{33}$$

(Hint. Consider a first as a $\mathbf{C}$-linear map on $\boldsymbol{V}^n_{|\mathbf{C}}$. If $\boldsymbol{W} \subset \boldsymbol{V}^n_{|\mathbf{C}}$ is a complex subspace in which $a|\boldsymbol{W}$ is described by a Jordan box, then examine $J(\boldsymbol{W})$.)

1.10.4 Projective Extensions

Restricting the scalar domain of the vector space $\boldsymbol{V}$ determining a projective geometry from L to $K \subset L$ does not change the structure of the Abelian group $[\boldsymbol{V}, +]$ of the vector space; by Proposition 2 the restricted scalar multiplication generates a new projective geometry with scalars K as well as the surjective map θ from the point space of the K-geometry onto the point space of the L-geometry, cf. (2). Frequently, one wants, however, to embed a projective K-point space $\boldsymbol{P}^n_K$ into a projective L-point space $\boldsymbol{P}^n_L$, i.e., to find an injective map $\iota : \boldsymbol{P}^n_K \to \boldsymbol{P}^n_L$ preserving as much as possible of the projective structure. In this situation, one might make good use of the richer and often simpler possibilities for the solution of algebraic equations over L to draw conclusions concerning the K-geometry. Particularly interesting, and frequently applied in algebraic geometry, is the transition from the real to the complex. The simplifications taking place here are, e.g., evident in the classification of polar maps, cf. Propositions 9.4 and 9.5. Often, extending the scalars proceeds quite directly: Describe the objects of interest in homogeneous or inhomogeneous coordinates with respect to K, and then just allow these coordinates to take values in all of L. To provide a conceptual foundation for this straightforward procedure we again start from the situation described in the beginning of this section.

Definition 1. Let $K \subset L$ be a skew subfield. Restricting the scalars each (e.g. right) vector space $\boldsymbol{V}$ over L becomes a vector space over K. A K-subspace $\boldsymbol{W} \subset \boldsymbol{V}_{|K}$ is called a *K-form of $\boldsymbol{V}$* if $\boldsymbol{W}$ is a minimal K-subspace generating $\boldsymbol{V}$, i.e. if $\mathfrak{L}_L(\boldsymbol{W}) = \boldsymbol{V}$[1]. If this holds, then $\boldsymbol{V}$ is called an *L-extension* of $\boldsymbol{W}$. □

[1] In the case $\mathbf{R} \subset \mathbf{C}$ instead of $\mathbf{R}$-form the term $\mathbf{R}$-structure is frequently used; to avoid confusion with the complex and quaternionic structures defined above, and for the sake of a unified notion we prefer the name $\mathbf{R}$-form here as well.

Lemma 3. *If $\boldsymbol{V}_1$, $\boldsymbol{V}_2$ are right L-extensions of the K-vector space $\boldsymbol{W}$, then $\boldsymbol{V}_1 = \boldsymbol{V}_2$; set $\boldsymbol{V} := \boldsymbol{V}_1$. Each K-basis of the K-structure $\boldsymbol{W}$ is also an L-basis of the extension $\boldsymbol{V}$. In particular,* $\dim_K \boldsymbol{W} = \dim_L \boldsymbol{V}$.

Proof. By definition each of the vector spaces $\boldsymbol{V}_i$, $i = 1, 2$, is the L-linear span of $\boldsymbol{W}$. This implies $\boldsymbol{V}_1 = \boldsymbol{V}_2$. If $(\mathfrak{a}_1, \dots, \mathfrak{a}_r)$ are K-linearly independent vectors of $\boldsymbol{W}$, then they are L-linearly independent in $\boldsymbol{V}$ as well: Assume, e.g. that $\mathfrak{a}_r$ is an L-linear combination of the others,

$$\mathfrak{a}_r = \sum_{\rho=1}^{r-1} \mathfrak{a}_\rho \lambda_\rho, \ \lambda_\rho \in L.$$

Consider the subspace $\boldsymbol{W}_1$ in $\boldsymbol{W} = \boldsymbol{W}_1 \oplus \mathfrak{a}_r K$ complementary to $\mathfrak{a}_r K$, $\boldsymbol{W}_1 = [\mathfrak{a}_1, \dots, \mathfrak{a}_{r-1}]_K$. Then the L-linear span of $\boldsymbol{W}_1$ is obviously equal to $\boldsymbol{V}$, and hence $\boldsymbol{W}$ cannot be a minimal L-subspace generating the space $\boldsymbol{V}$. □

This already justifies the method for finite-dimensional vector spaces mentioned above: If $(\mathfrak{a}_i)$, $i = 1, \dots, n$, is a K-basis for $\boldsymbol{W}$, then each element of the L-extension $\boldsymbol{V}$ has a unique basis representation

$$\mathfrak{x} = \sum_{i=1}^{n} \mathfrak{a}_i \lambda_i, \ \lambda_i \in L.$$

Conversely, if $(\mathfrak{b}_i)$, $i = 1, \dots, n$, is an arbitrary L-basis of $\boldsymbol{V}$, then the K-linear span $\boldsymbol{W} := \mathfrak{L}_K((\mathfrak{b}_i), i = 1, \dots, n)$ is a K-form of $\boldsymbol{V}$. Returning to the bases we immediately obtain:

Corollary 4. *Under the assumptions of Definition 1 let* $\dim \boldsymbol{V} = n < \infty$. *Then there exists a K-form $\boldsymbol{W} \subset \boldsymbol{V}$. The L-linear group $\boldsymbol{GL}(\boldsymbol{V}, L)$ acts transitively on the set of all K-forms on $\boldsymbol{V}$. The latter is thus a homogeneous space isomorphic to*

$$\boldsymbol{GL}(n, L)/\boldsymbol{GL}(n, K).$$

Proof. It suffices to note that the isotropy group of a K-form $\boldsymbol{W}$ consists of all L-linear transformations $h \in \boldsymbol{GL}(\boldsymbol{V}, L)$ with $h(\boldsymbol{W}) = \boldsymbol{W}$. Since the L-linear transformations are K-linear as well, the assertion $h \in \boldsymbol{GL}(\boldsymbol{W}, K)$ follows. Fixing an arbitrary K-basis of $\boldsymbol{W}$ as an L-basis for $\boldsymbol{V}$, the embedding $\boldsymbol{GL}(\boldsymbol{W}, K) \subset \boldsymbol{GL}(\boldsymbol{V}, L)$ is realized by the group of all the L-linear transformations whose matrix belongs to $\boldsymbol{GL}(n, K)$. □

Corollary 5. *Let $\sigma \in$* $\mathrm{Aut}(L)$ *be an automorphism that preserves K, $\sigma(K) = K$. If $a : \boldsymbol{W}_1 \to \boldsymbol{W}_2$ is a K-linear or K-σ-linear map, respectively, then there is a uniquely determined linear or L-σ-linear map $\tilde{a} : \boldsymbol{V}_1 \to \boldsymbol{V}_2$ of the L-extensions $\boldsymbol{V}_1, \boldsymbol{V}_2$ of $\boldsymbol{W}_1$ or $\boldsymbol{W}_2$, respectively, extending a: $\tilde{a}|\boldsymbol{W}_1 = a$. Expressed in terms of the category of linear maps the assignment $a \mapsto \tilde{a}$ forms a covariant functor from the category of right vector spaces over K to the category of right vector spaces over L. Here*

$$\widetilde{(a+b)} = \tilde{a} + \tilde{b}, \; \widetilde{\alpha a} = \alpha \tilde{a} \; \; (\alpha \in Z(L) \cap K).$$

P r o o f. In the finite-dimensional case we consider K-bases $(\mathfrak{a}_i)$, $(\mathfrak{b}_\alpha)$ of $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$, respectively, and also allow for elements $\lambda^i \in L$ in the basis representation of a,

$$a(\mathfrak{x}) = a(\mathfrak{a}_i \lambda^i) = \mathfrak{b}_\alpha a_i^\alpha \sigma(\lambda^i).$$

For infinite-dimensional vector spaces one proceeds similarly making use of the existence of bases (cf. Exercise I.4.4.7). □

We call $\tilde{a}$ the *linear* (or *σ-linear*) *extension* of the map a. Analogous considerations work for left vector spaces. If the parities of the vector spaces are different, then the automorphisms are to be replaced by anti-automorphisms.

Example 3. Consider the inclusion of fields $\mathbf{R} \subset \mathbf{C}$ and the complex plane $\boldsymbol{V} = \mathbf{C}^2$. Then $\boldsymbol{W} := (1,0)\mathbf{R} \oplus (0,1)\mathbf{R}$ is an $\mathbf{R}$-form on $\boldsymbol{V}$. The real subspace $\boldsymbol{U} := (1,0)\mathbf{R} \oplus (\mathrm{i},0)\mathbf{R}$ only spans the complex line $(1,0)\mathbf{C}$, so it is no $\mathbf{R}$-form on $\boldsymbol{V}$. The $\mathbf{R}$-linear map a from $\boldsymbol{W} = \mathbf{R}^2$ to itself defined by

$$a: \; (y^i) = (a^i_j)(x^j), \; (x^j) \in \mathbf{R}^2, \; (a^i_j) \in \boldsymbol{M}_2(\mathbf{R}),$$

has the linear extension $\tilde{a} \in \mathrm{End}_{\mathbf{C}}(\boldsymbol{V})$ determined by simply allowing $(x^j) \in \mathbf{C}^2$ in the definition, and the conjugate linear extension described by

$$\hat{a} : \, (x^i) \in \boldsymbol{V} \longmapsto (\bar{y}^i) := (a^i_j)(\bar{x}^j).$$

□

The extension of vector spaces and σ-linear maps has immediate consequences in projective geometry:

Corollary 6. *Let K be a skew subfield of the skew field L. For every finite-dimensional right projective space $\boldsymbol{P}^n = \boldsymbol{P}(\boldsymbol{W}^{n+1}, K)$ there is a unique* right projective extension $\boldsymbol{P}^n_L = \boldsymbol{P}(\boldsymbol{V}^{n+1}, L)$ *corresponding to the L-extension $\boldsymbol{V}$ of the K-vector space $\boldsymbol{W}$; the point space $\boldsymbol{P}^n$ is canonically embedded into $\boldsymbol{P}^n_L$ by*

$$\iota : \boldsymbol{x} = \mathfrak{x}K \in \boldsymbol{P}^n \longmapsto \iota(\boldsymbol{x}) := \mathfrak{x}L \in \boldsymbol{P}^n_L. \tag{34}$$

The L-extensions of the projective subspace $\boldsymbol{B}^k := \boldsymbol{b}_0 \vee \ldots \vee \boldsymbol{b}_k \subset \boldsymbol{P}^n$ is

$$\boldsymbol{B}^k_L = (\boldsymbol{b}_0 \vee \ldots \vee \boldsymbol{b}_k)_L = \iota(\boldsymbol{b}_0) \vee \ldots \vee \iota(\boldsymbol{b}_k). \tag{35}$$

If $\sigma \in \mathrm{Aut}\, L$ is an automorphism with $\sigma(K) = K$, and f is the collinear map induced by a σ-K-linear map $a : \boldsymbol{W} \to \boldsymbol{W}$, then there exists a unique collinear map $\tilde{f} : \boldsymbol{P}^n_L \to \boldsymbol{P}^n_L$ extending f and induced by the σ-L-linear map $\tilde{a} : \boldsymbol{V} \to \boldsymbol{V}$. Moreover,

$$\widetilde{\mathrm{id}}_{\boldsymbol{P}^n} = \mathrm{id}_{\boldsymbol{P}^n_L}, \;\; \widetilde{f \circ g} = \tilde{f} \circ \tilde{g}. \tag{36}$$

□

We leave it to the reader to formulate corresponding conclusions for maps between different projective spaces, for left projective geometries, and for projective geometries of different parity. Corollary 6 allows to extend the projective geometries over a skew field K functorially to projective geometries over every skew field L containing K. According to Corollary 6 in particular the projectivities $g \in \boldsymbol{PL}(\boldsymbol{W}^{n+1})$ are extended as projectivities $\tilde{g} \in \boldsymbol{PL}(\boldsymbol{V}^{n+1})$, so that $\boldsymbol{PL}(\boldsymbol{W}^{n+1})$ naturally appears as a subgroup of $\boldsymbol{PL}(\boldsymbol{V}^{n+1})$. Obviously, the action of $\boldsymbol{PL}(\boldsymbol{W}^{n+1})$ over $\boldsymbol{P}^n_L$ is not transitive; in fact, it leaves $\boldsymbol{P}^n$ invariant. The properties invariant under this subgroup describe the position of configurations in the projective geometry $\mathfrak{P}(\boldsymbol{V})$ with respect to the embedding defined by (34); in the following Sections 5 and 6 we will discuss this in greater detail.

Furthermore, consider the K-restriction $\boldsymbol{V}_{|K}$ of $\boldsymbol{V}$; the notions referring to it are marked by the label "K". Obviously, this restriction is not equal to the original K-vector space $\boldsymbol{W}$; in fact, $\dim_K \boldsymbol{V}_{|K} = \dim_K \boldsymbol{W} \cdot \dim_K L$. The subspaces of $\boldsymbol{V}_{|K}$ are called *K-subspaces of $\boldsymbol{V}$*; obviously $\boldsymbol{V}$ itself is a K-subspace of $\boldsymbol{V}$, for whose dimension (1) holds. The projective subspaces corresponding to them will also be called *K-subspaces*. E.g., the image of the embedding $\boldsymbol{P}^n_K := \iota(\boldsymbol{P}^n) \subset \boldsymbol{P}^n_L$ defined in (34) is an n-dimensional projective K-subspace that will be called a *projective K-form* on $\boldsymbol{P}^n_L$. In particular, the points $\boldsymbol{x} \in \boldsymbol{P}^n_K$ are called *K-points* in $\boldsymbol{P}^n_L$. A frame $(\boldsymbol{a}_i; \boldsymbol{e})$ for $\boldsymbol{P}^n_K$ is called a *K-frame* of $\boldsymbol{P}^n_L$.

Exercise 9. Let $\boldsymbol{P}^n_K \subset \boldsymbol{P}^n_L$ be a projective K-form on $\boldsymbol{P}^n_L$. Using the notation for projective K-subspaces $\boldsymbol{A}, \boldsymbol{B} \subset \boldsymbol{P}^n_K$ just fixed prove:

$$(\boldsymbol{A} \vee_K \boldsymbol{B})_K = (\boldsymbol{A}_L \vee_L \boldsymbol{B}_L) \cap \boldsymbol{P}^n_K, \tag{37}$$

$$(\boldsymbol{A} \wedge_K \boldsymbol{B})_K = (\boldsymbol{A}_L \wedge_L \boldsymbol{B}_L) \cap \boldsymbol{P}^n_K, \tag{38}$$

$$\boldsymbol{x} \iota_K \boldsymbol{A} \Longleftrightarrow \boldsymbol{x} \in \boldsymbol{P}^n_K \text{ and } \boldsymbol{x} \iota_L \boldsymbol{A}_L. \tag{39}$$

Exercise 10. Each K-basis $(\mathfrak{a}_i)$ for $\boldsymbol{V}^{n+1}$ determines homogeneous coordinates on $\boldsymbol{P}^n_K$ and on $\boldsymbol{P}^n_L$. Prove that a point $\boldsymbol{x} \in \boldsymbol{P}^n_L$ lies in $\boldsymbol{P}^n_K$ if and only if the ratios of its homogeneous coordinates with respect to a K-basis $(\mathfrak{a}_i)$ satisfy $x^i \cdot (x^j)^{-1} \in \hat{K}$, $i, j = 0, \dots, n$. Hence

$$\boldsymbol{P}^n_K = \{\boldsymbol{x} = [\sum_{i=0}^{n} \mathfrak{a}_i x^i]_L | x^i \in K, \ (x^i) \neq \mathfrak{o}\}. \tag{40}$$

Corollary 7. *Under the assumptions of Corollary 6 each projectivity f on $\boldsymbol{P}^n_K$ has an extension $\tilde{f}$ to $\boldsymbol{P}^n_L$, $\tilde{f}|\boldsymbol{P}^n_K = f$. If, conversely, F is a projectivity satisfying*

$$F(\boldsymbol{P}^n_K) = \boldsymbol{P}^n_K \tag{41}$$

on $\boldsymbol{P}^n_L$, then $F|\boldsymbol{P}^n_K$ is a collineation. □

Example 4. The following example shows that (in the notations of Corollary 7) $F|\boldsymbol{P}_K^n$ is not necessarily a projectivity, if it satisfies (41): Consider the usual embeddings $K = \mathbf{C} \subset L = \mathbf{H}$, $\mathbf{C} = \mathfrak{L}_{\mathbf{R}}(\{1, \mathrm{i}\})$. Then the map a with $a((q^s)) := (\mathrm{j}\, q^s)$, $s = 0, \ldots, n$, is a linear map of the right vector space $\mathbf{H}^{n+1}$. Because of

$$[a((z^s))] = [(\mathrm{j}\, z^s)] = [(\bar{z}^s\, \mathrm{j})] = [(\bar{z}^s)],$$

the projectivity on $\boldsymbol{P}_{\mathbf{H}}^n$ defined by $F([\mathfrak{x}]) = [a(\mathfrak{x})]$ determines on the $\mathbf{C}$-form $\boldsymbol{P}_{\mathbf{C}}^n \subset \boldsymbol{P}_{\mathbf{H}}^n$ which corresponds to the standard basis nothing but the anti-projectivity generated by conjugation. □

Nevertheless, we have

Corollary 8. *If L is a field, and F is a projectivity on $\boldsymbol{P}_L^n$ satisfying* (41), *then $F|\boldsymbol{P}_K^n$ again is a projectivity.*

To prove this, it suffices to note that F leaves the cross ratio invariant, which takes only values in $\hat{K}$ for throws from $\boldsymbol{P}_K^n$, and to apply Proposition 4.6. □

1.10.5 The Projective K-Geometry of $\boldsymbol{P}_L^n$

According to Corollary 4 all projective K-forms on $\boldsymbol{P}_L^n$ are projectively equivalent. Now we again fix one of these forms and denote by $\widetilde{\boldsymbol{PL}}_{n,K} \subset \boldsymbol{PL}(\boldsymbol{P}_L^n)$ the subgroup of all projectivities on $\boldsymbol{P}_L^n$ which are extensions of K-projectivities on $\boldsymbol{P}_K^n$; according to Corollary 8, in the case of a field L this group is the isotropy group of $\boldsymbol{P}_K^n$ for the action of $\boldsymbol{PL}(P_L^n)$ described in Corollary 4. Obviously, we have the isomorphy

$$\widetilde{\boldsymbol{PL}}_{n,K} \cong \boldsymbol{PL}(\boldsymbol{P}_K^n). \tag{42}$$

The invariants of the transformation group $[\widetilde{\boldsymbol{PL}}_{n,K}, \boldsymbol{P}_L^n]$ describe the geometric properties of figures in $\boldsymbol{P}_L^n$ with respect to the projective K-form $\boldsymbol{P}_K^n$, in short, the *projective K-geometry* of $\boldsymbol{P}_L^n$.

Example 5. For the applications particularly important is the case $K = \mathbf{R}$, $L = \mathbf{C}$, i.e the real geometry of complex projective space. Example 4.7 immediately shows that the notion of a real structure introduced there essentially amounts to an $\mathbf{R}$-form: *Staudt chains* are the same as $\mathbf{R}$-forms on $\boldsymbol{P}_{\mathbf{C}}^n$ (cf. the following exercise). □

Exercise 11. Let J denote a real structure on the complex projective space $\boldsymbol{P}_{\mathbf{C}}^n$ and at the same time the conjugate linear map of the associated vector $\boldsymbol{V}^{n+1}$ generating it, cf. Example 4.7. Prove: a) The Staudt chain corresponding to J is an $\mathbf{R}$-form on $\boldsymbol{P}_{\mathbf{C}}^n$, and each $\mathbf{R}$-form on $\boldsymbol{P}_{\mathbf{C}}^n$ is a Staudt chain. – b) If $\boldsymbol{P}_{\mathbf{R}}^n \subset \boldsymbol{P}_{\mathbf{C}}^n$ is an $\mathbf{R}$-form, then the basis $(\mathfrak{b}_j)$ of $\boldsymbol{V}^{n+1}$ is an $\mathbf{R}$-basis if and only if J can be represented in the normal form (4.46) with respect to $(\mathfrak{b}_j)$. – c)

$$\widetilde{\boldsymbol{PL}}_{n,\mathbf{R}} = \{f \in \boldsymbol{PL}(\boldsymbol{P}_{\mathbf{C}}^n) | f \circ J = J \circ f\}. \tag{43}$$

Example 6. Consider the extension from $\mathbf{C}$ to $\mathbf{H}$. Let $\boldsymbol{P}^n_{\mathbf{C}} \subset \boldsymbol{P}^n_{\mathbf{H}}$ be a $\mathbf{C}$-form on quaternionic projective space, and let $(\mathfrak{b}_s), s = 0, \ldots, n$, be a $\mathbf{C}$-basis of the associated vector spaces $\boldsymbol{W} = \boldsymbol{W}^{n+1}_{\mathbf{C}} \subset \boldsymbol{V}^{n+1}_{\mathbf{H}} = \boldsymbol{V}$. In the decomposition

$$x^s = z^s + \mathrm{j}\, w^s,\ z^s, w^s \in \mathbf{C},$$

the equations

$$w^s = 0,\ s = 0, \ldots, n, \tag{44}$$

characterize the subset $\boldsymbol{P}^n_{\mathbf{C}} \subset \boldsymbol{P}^n_{\mathbf{H}}$; the (z^s) are just the homogeneous coordinates of the points from $\boldsymbol{P}^n_{\mathbf{C}}$. With respect to $\boldsymbol{P}^n_{\mathbf{C}} \subset \boldsymbol{P}^n_{\mathbf{H}}$ the involutive projectivity J on $\boldsymbol{P}^n_{\mathbf{H}}$ with coordinate representation $J : (x^s) \mapsto (\mathrm{i}\, x^s)$ (cf. Example 4.8, in particular (4.52)) has analogous properties to the real structure J from Exercise 11:

a) We have

$$\boldsymbol{P}^n_{\mathbf{C}} = \{\boldsymbol{x} \in \boldsymbol{P}^n_{\mathbf{H}} | J(\boldsymbol{x}) = \boldsymbol{x}\}; \tag{45}$$

the involutions J with normal form (4.52) and the $\mathbf{C}$-forms $\boldsymbol{P}^n_{\mathbf{C}} \subset \boldsymbol{P}^n_{\mathbf{H}}$ bijectively correspond to one another.

b) For a given $\boldsymbol{P}^n_{\mathbf{C}} \subset \boldsymbol{P}^n_{\mathbf{H}}$, the $(\boldsymbol{b}_s)$ form a $\mathbf{C}$-basis of $\boldsymbol{P}^n_{\mathbf{C}}$ if and only if J has the normal form

$$J(\mathfrak{b}_s) = \mathfrak{b}_s\, \mathrm{i},\ s = 0, \ldots, n, \tag{46}$$

with respect to $(\boldsymbol{b}_s)$. As Example 4 shows, $\widetilde{\boldsymbol{PL}}_{n,\mathbf{C}}$ is no longer the isotropy group G of $\boldsymbol{PL}_{n,\mathbf{C}}$ under the action of $\boldsymbol{PL}(\boldsymbol{P}^n_{\mathbf{H}})$. Making use of the continuity of projectivities on $\boldsymbol{P}^n_{\mathbf{H}}$ the following is easy to show:

c) The isotropy group G of a $\mathbf{C}$-structure $\boldsymbol{P}^n_{\mathbf{C}} \subset \boldsymbol{P}^n_{\mathbf{H}}$ is

$$G = \{f \in \boldsymbol{PL}(\boldsymbol{P}^n_{\mathbf{H}}) | f \circ J = J \circ f\}. \tag{47}$$

The map $f \in G \mapsto f|\boldsymbol{P}^n_{\mathbf{C}} \in \operatorname{Aut} \boldsymbol{P}^n_{\mathbf{C}}$ is an isomorphism from G onto the group of collineations on $\boldsymbol{P}^n_{\mathbf{C}}$ consisting of projectivities and anti-projectivities. □

Exercise 12. Prove the statements in Example 6. Hint. Take the situation described in Example 4 into account.

Example 7. Now we indicate a scheme generalizing Examples 5 and 6. Note first that for each automorphism $\sigma \in \operatorname{Aut} L$ the set K_σ of fixed elements, $K_\sigma := \{\xi \in L | \sigma(\xi) = \xi\}$, is a skew subfield of the skew field L. Suppose that the following condition is satisfied:

B) *Let L be a skew field,* $\operatorname{char} L \neq 2$, *and let* $\sigma \in \operatorname{Aut} L$ *be an involutive automorphism,* $\sigma^2 = \mathrm{id}_L$, $\sigma \neq \mathrm{id}_L$. *Set* $K = K_\sigma$.

In general, K is not commutative. In Example 5, $\sigma(z) = \bar{z}$ is the conjugation, and in Example 6 $\sigma = \sigma_{\mathrm{i}}$ is the inner automorphism

$$\sigma_{\mathrm{i}}(z + \mathrm{j}\, w) = -\mathrm{i}(z + \mathrm{j}\, w)\, \mathrm{i} = z - \mathrm{j}\, w. \tag{48}$$

Let then $\boldsymbol{V}^{n+1}$ be a right vector space over L, and let $(\mathfrak{a}_i)$ be an arbitrary basis for $\boldsymbol{V}$. Setting

$$J(\sum_{i=0}^{n} \mathfrak{a}_i x^i) := \sum_{i=0}^{n} \mathfrak{a}_i \sigma(x^i) \tag{49}$$

defines a σ-linear involution on $\boldsymbol{V}$, and this again determines an equally denoted collineation $J \in \mathrm{Aut}(\boldsymbol{P}^n(\boldsymbol{V}))$. We call J a *σ-conjugation* if, for a suitably chosen basis, J is represented in the form (49). If the related automorphism is clear from the context, J will simply be called the conjugation. Two elements $\boldsymbol{A}, \boldsymbol{B}$ (points, vectors, subspaces, ...) are called *conjugate*, if $J(\boldsymbol{A}) = \boldsymbol{B}$ holds. □

Proposition 9. *Under Assumption* B), $K = K_\sigma$, *each σ-conjugation J determines a K-form on $\boldsymbol{P}^n_L := \boldsymbol{P}^n(\boldsymbol{V})$ by*

$$\boldsymbol{P}^n_K := \{\boldsymbol{x} \in \boldsymbol{P}^n(\boldsymbol{V}) | J(\boldsymbol{x}) = \boldsymbol{x}\}; \tag{50}$$

moreover, for each K-form $\boldsymbol{P}^n_K \subset \boldsymbol{P}^n_L$ there is a unique σ-conjugation J satisfying (50).

Proof. Let $(\mathfrak{a}_s)$ be a basis for $\boldsymbol{V}$ in which (49) holds. Then the K-form (40) determined by this basis obviously belongs to the fixed-point set of J. If, conversely, $J(\boldsymbol{x}) = \boldsymbol{x}$, then for the vector $(x^s) \in L^{n+1}$ given by the homogeneous coordinates of $\boldsymbol{x}$ with respect to the basis $[\mathfrak{a}_s]$ there is $q \in L^*$ with $\sigma(x^s) = x^s q$. This implies $\sigma^2(x^s) = x^s = \sigma(x^s)\sigma(q) = x^s q\sigma(q)$. Since at least one $x^s \neq 0$, we obtain

$$q\sigma(q) = 1, \tag{51}$$

and each coordinate $x^s \neq 0$ of $\boldsymbol{x}$ satisfies $(x^s)^{-1}\sigma(x^s) = q$. We prove

Lemma 10. *Assume that* (51) *holds for $q \in L$. Under Assumption* B) *any solution $\xi_0 \in L$ of*

$$\xi^{-1}\sigma(\xi) = q \tag{52}$$

generates all solutions setting $\xi = \kappa\xi_0$, $\kappa \in K^*$.

Proof. From $\sigma(\kappa) = \kappa$ we immediately conclude that $\xi = \kappa\xi_0$ is a solution. Let now $\xi \in L^*$ be any solution of (52). Because of $\sigma(\xi_0^{-1})\xi_o = \sigma(q)$, we have

$$\sigma(\xi\xi_0^{-1}) = \sigma(\xi)\sigma(\xi_0)^{-1} = \xi q\sigma(q)\xi_0^{-1} = \xi\xi_0^{-1}.$$

Consequently, $\kappa := \xi\xi_0^{-1} \in K^*$, hence $\xi = \kappa\xi_0$. □

Since the non-vanishing coordinates satisfy equation (52), we may write $x^s = \kappa^s\xi_0$ with $\kappa^s \in K$, $s = 0, \ldots, n$, $\xi_0 \in L^*$; for $x^s = 0$ we set $\kappa^s = 0$. By Exercise 10 this means nothing but $\boldsymbol{x} \in \boldsymbol{P}^n_K$. If now, conversely, we start from a K-form, then a σ-conjugation J satisfying (50) is defined by taking (49) with respect to a K-basis $(\mathfrak{a}_s)$. Let, finally, J_1 be another σ-conjugation with fixed-point set $\boldsymbol{P}^n_K$. Then $J_1|\boldsymbol{P}^n_K = J|\boldsymbol{P}^n_K = \mathrm{id}_{\boldsymbol{P}^n_K}$, and $J_1(\mathfrak{a}_s) = J(\mathfrak{a}_s) = \mathfrak{a}_s$ implies $J_1 = J$. □

The result of the following exercise will be needed later.

Exercise 13. Prove assuming that B) is satisfied: a) There exists a $q \in L^*$ with $\sigma(q) = -q$. – b) If $q_0 \in L^*$ is a fixed element such that $\sigma(q_0) = -q_0$, then every $x \in L$ can be uniquely represented in the form $x = \alpha + q_0\beta$ with $\alpha, \beta \in K$, and, moreover, $\dim_K L = 2$. – c) The map $\alpha \in K \mapsto \tau(\alpha) := q_0 \alpha q_0^{-1} \in L$ is an automorphism of K; τ is involutive if and only if q_0^2 belongs to the center $Z(K)$ of K. – d) Prove by means of an example that there not necessarily is a $q \in L^*$ satisfying at simultaneously $\sigma(q) = -q$ and $\sigma(q) = q^{-1}$. Hint. Consider the conjugation in $\mathbf{Q}[\sqrt{-3}]$.

1.10.6 *K*-Classification of Projective *L*-Subspaces

Now we will turn to the projective K-geometry of the L-subspaces of $\boldsymbol{P}_L^n$. Definition 1 easily implies:

Lemma 11. *In the assumptions and notations of Definition 1 the following holds: If $\boldsymbol{A} \subset \boldsymbol{P}_L^n$ is a projective L-subspace, then $\boldsymbol{A} \cap \boldsymbol{P}_K^n$ is a projective K-subspace, and*

$$\operatorname{Dim}_K \boldsymbol{A} \cap \boldsymbol{P}_K^n \le \operatorname{Dim}_L \boldsymbol{A}. \tag{53}$$

In (53) *equality holds if and only if*

$$(\boldsymbol{A} \cap \boldsymbol{P}_K^n)_L = \boldsymbol{A}. \tag{54}$$

□

Proposition 12. *Assuming that* B) *holds consider the projective extension $\boldsymbol{P}_K^n \subset \boldsymbol{P}_L^n$, where $\boldsymbol{P}_K^n$ is defined by the σ-conjugation according to* (50). *Then* (54) *holds for a projective L-subspace $\boldsymbol{A} \subset \boldsymbol{P}_L^n$ if and only if $J(\boldsymbol{A}) = \boldsymbol{A}$.*

P r o o f. Using a K-basis $(\mathfrak{a}_i)$, $i = 0, \ldots, l = \operatorname{Dim}_K \boldsymbol{A} \cap \boldsymbol{P}_K^n$, to describe $\boldsymbol{A} \cap \boldsymbol{P}_K^n$ equation (54) immediately implies $l = \operatorname{Dim}_L \boldsymbol{A}$, i.e. equality in (53), and $J(\boldsymbol{A}) = \boldsymbol{A}$. Conversely, assume that $J(\boldsymbol{A}) = \boldsymbol{A}$ holds. Thus we again find a K-basis $(\mathfrak{a}_i)$, $i = 0, \ldots, l$, for $\boldsymbol{A} \cap \boldsymbol{P}_K^n$. Suppose that $k := \operatorname{Dim}_L \boldsymbol{A} > l$. Then

$$\boldsymbol{B} := (\boldsymbol{A} \cap \boldsymbol{P}_K^n)_L \subset \boldsymbol{A}, \; l = \operatorname{Dim}_L B < k,$$

and there is a point $\boldsymbol{b} \in \boldsymbol{A} \setminus \boldsymbol{B}$. Consequently, $J(\boldsymbol{b}) \in J(\boldsymbol{A} \setminus \boldsymbol{B}) = \boldsymbol{A} \setminus \boldsymbol{B}$; in particular, we have $\boldsymbol{b} \ne J(\boldsymbol{b})$. Choose a vector $\mathfrak{b}$ with $\boldsymbol{b} = [\mathfrak{b}]$. Obviously,

$$J(\mathfrak{b} + J(\mathfrak{b})) = \mathfrak{b} + J(\mathfrak{b}), \; J(\mathfrak{b} - J(\mathfrak{b})) = -(\mathfrak{b} - J(\mathfrak{b})), \tag{55}$$

and hence

$$\boldsymbol{a} := [\mathfrak{b} + J(\mathfrak{b})] \in \boldsymbol{A} \cap \boldsymbol{P}_K^n, \; \boldsymbol{a}' := [\mathfrak{b} - J(\mathfrak{b})] \in \boldsymbol{A} \cap \boldsymbol{P}_K^n.$$

Since $(\mathfrak{a}_i)$ is a K-basis for this K-subspace, Exercise 13 implies

$$\mathfrak{a} := \mathfrak{b} + J(\mathfrak{b}) = \sum_0^l \mathfrak{a}_i \gamma^i, \; \mathfrak{a}' := \mathfrak{b} - J(\mathfrak{b}) = \sum_0^l \mathfrak{a}_i \gamma'^i q, \tag{56}$$

γ^i, $\gamma'^i \in K$ and $q \in L^*$ with

$$\sigma(q) = -q. \tag{57}$$

From this we immediately conclude

$$2\mathfrak{b} = \sum_{0}^{l} \mathfrak{a}_i(\gamma^i + \gamma'^i q), \tag{58}$$

and because of $\operatorname{char} L \neq 2$ we obtain $2\mathfrak{b} \neq \mathfrak{o}$ and $\boldsymbol{b} = [\mathfrak{b}] \in \boldsymbol{B}$. But this contradicts the choice of $\boldsymbol{b}$. Hence $l = k$, and Lemma 11 implies (54). □

From Proposition 12 the classification of L-subspaces in the projective K-geometry on $\boldsymbol{P}_L^n$ can easily be derived:

Proposition 13. *Maintaining the assumptions of Proposition* 12 *consider two projective subspaces $\boldsymbol{A}, \boldsymbol{A}_1 \subset \boldsymbol{P}_L^n$. Then there is a projectivity $f \in \widetilde{\boldsymbol{PL}}_{n,K}$ with $f(\boldsymbol{A}) = \boldsymbol{A}_1$ if and only if*

$$\operatorname{Dim}_L \boldsymbol{A} = \operatorname{Dim}_L \boldsymbol{A}_1 \textit{ and } \operatorname{Dim}_K \boldsymbol{A} \cap \boldsymbol{P}_K^n = \operatorname{Dim}_K \boldsymbol{A}_1 \cap \boldsymbol{P}_K^n. \tag{59}$$

Proof. Because of $f(\boldsymbol{P}_K^n) = \boldsymbol{P}_K^n$, the necessity of condition (59) is clear. To prove the converse, we adapt a K-basis to the space $\boldsymbol{A}$. First we choose a basis $\mathfrak{a}_0, \ldots, \mathfrak{a}_d \in \boldsymbol{W}$ for the space $\boldsymbol{A} \cap \boldsymbol{P}_K^n$; here $\boldsymbol{W} = \boldsymbol{W}^{n+1}$ denotes the K-vector space corresponding to the K-structure $\boldsymbol{P}_K^n$. Now

$$\boldsymbol{A} \cap \boldsymbol{P}_K^n = J(\boldsymbol{A}) \cap \boldsymbol{P}_K^n = (\boldsymbol{A} \wedge J(\boldsymbol{A})) \cap \boldsymbol{P}_K^n,$$

and Proposition 12 implies

$$d = \operatorname{Dim}_K \boldsymbol{A} \cap \boldsymbol{P}_K^n = \operatorname{Dim}_L \boldsymbol{A} \wedge J(\boldsymbol{A}), \tag{60}$$

since $\boldsymbol{A} \wedge J(\boldsymbol{A})$ is J-invariant. Let $k = \operatorname{Dim}_L \boldsymbol{A}$; complete the basis $\mathfrak{a}_0, \ldots, \mathfrak{a}_d$ of $\boldsymbol{A} \wedge J(\boldsymbol{A})$ by the vectors $\mathfrak{b}_{d+1}, \ldots, \mathfrak{b}_k \in \boldsymbol{V}^{n+1}$, $\boldsymbol{P}_L^n = \boldsymbol{P}^n(\boldsymbol{V})$, so that they together form a basis for $\boldsymbol{A}$. Then

$$J(\boldsymbol{A}) = [\mathfrak{a}_o] \vee \ldots \vee [\mathfrak{a}_d] \vee [J(\mathfrak{b}_{d+1})] \vee \ldots \vee [J(\mathfrak{b}_k)],$$

and $\boldsymbol{A} \vee J(\boldsymbol{A})$ is spanned by the $2k - d + 1$ vectors $(\mathfrak{a}_\alpha)$, $\alpha = 0, \ldots, d$, $(\mathfrak{b}_\nu)$, $(J(\mathfrak{b}_\nu))$, $\nu = d+1, \ldots, k$. By the dimension formula we have

$$\operatorname{Dim}_L \boldsymbol{A} \vee J(\boldsymbol{A}) = 2k - \operatorname{Dim}_L \boldsymbol{A} \wedge J(\boldsymbol{A}) = 2k - d, \tag{61}$$

and hence the vectors just introduced are linearly independent. Now we fix an arbitrary $q \in L^*$ satisfying equation (57) (cf. Exercise 13), and define

$$\mathfrak{a}_\nu := \mathfrak{b}_\nu + J(\mathfrak{b}_\nu),\ \hat{\mathfrak{a}}_\nu := (\mathfrak{b}_\nu - J(\mathfrak{b}_\nu))q^{-1}. \tag{62}$$

Because of $J(\mathfrak{a}_\nu) = \mathfrak{a}_\nu$ and $J(\hat{\mathfrak{a}}_\nu) = \hat{\mathfrak{a}}_\nu$ the $\mathfrak{a}_\nu, \hat{\mathfrak{a}}_\nu$ lie in $\boldsymbol{W}^{n+1}$, and, conversely, we can express the $\mathfrak{b}_\nu, J(\mathfrak{b}_\nu)$ linearly by them ($\operatorname{char} L \neq 2$):

$$\mathfrak{b}_\nu = (\mathfrak{a}_\nu + \hat{\mathfrak{a}}_\nu q)/2,\ J(\mathfrak{b}_\nu) = (\mathfrak{a}_\nu - \hat{\mathfrak{a}}_\nu q)/2. \tag{63}$$

Hence the $2k - d + 1$ vectors $\mathfrak{a}_0, \ldots, \mathfrak{a}_k, \hat{\mathfrak{a}}_{d+1}, \ldots, \hat{\mathfrak{a}}_k$ form a K-basis for the J-invariant subspace $\boldsymbol{A} \vee J(\boldsymbol{A})$. We complete them choosing $(\mathfrak{a}_\lambda)$ for $\lambda = 2k - d + 1, \ldots, n$, to obtain a K-basis of $\boldsymbol{V}^{n+1}$. If the subspace $\boldsymbol{A}_1$ determines the same dimensions d, k as $\boldsymbol{A}$, then, in a similar way, we can assign to it a K-basis of $\boldsymbol{V}$ that has the same position properties with respect to $\boldsymbol{A}_1$ as the K-basis $\mathfrak{a}_0, \ldots, \mathfrak{a}_n$ to $\boldsymbol{A}$. These bases determine a projectivity $f \in \widetilde{\boldsymbol{PL}}_{n,K}$ with $f(\boldsymbol{A}) = \boldsymbol{A}_1$. □

Corollary 14. *Under the assumptions of Proposition* 12, *let $\boldsymbol{A} \subset \boldsymbol{P}^n_L$ be an L-subspace. Then the dimensions $k = \operatorname{Dim}_L \boldsymbol{A}$, $d = \operatorname{Dim}_K \boldsymbol{A} \cap \boldsymbol{P}^n_K$ satisfy the inequalities*

$$\max(-1, 2k - n) \le d \le k \le n. \tag{64}$$

All pairs (d, k) satisfying these inequalities occur as dimensions $k = \operatorname{Dim}_L \boldsymbol{A}$, $d = \operatorname{Dim}_K \boldsymbol{A} \cap \boldsymbol{P}^n_K$ for suitable subspaces.

Proof. The necessity of (64) follows from (60),

$$\operatorname{Dim}_L \boldsymbol{A} \vee J(\boldsymbol{A}) \le \min(n, 2k+1) \tag{65}$$

together with the dimension formula in Proposition 1.1. If, conversely, the pair (d, k) satisfies inequality (64), then, because of $d \le k \le n$, we first find $d + 1$ independent points $\boldsymbol{a}_\alpha \in \boldsymbol{P}^n_K$, $\alpha = 0, \ldots, d$. Inequality (64) implies $n - d \ge 2(k - d)$; hence we find $2(k - d)$ independent points $\boldsymbol{a}_\nu = [\mathfrak{a}_\nu]$, $\boldsymbol{a}'_\nu = [\mathfrak{a}'_\nu] \in \boldsymbol{P}^n_K$, $\nu = d + 1, \ldots, k$, such that $\boldsymbol{a}_0, \ldots, \boldsymbol{a}_k, \boldsymbol{a}'_{d+1}, \ldots, \boldsymbol{a}'_k$ are independent. We now again choose $q \in L^*$ with (57) and define $\mathfrak{b}_\nu$ for $\nu = d + 1, \ldots, k$, by (63). Setting $\boldsymbol{b}_\nu := [\mathfrak{b}_\nu]$ we thus conclude that $\boldsymbol{A} = \boldsymbol{a}_0 \vee \ldots \vee \boldsymbol{a}_d \vee \boldsymbol{b}_{d+1} \vee \ldots \vee \boldsymbol{b}_k$ is a k-dimensional subspace of $\boldsymbol{P}^n_L$ with $\operatorname{Dim}_K \boldsymbol{A} \cap \boldsymbol{P}^n_K = d$. □

Exercise 14. Let L be a skew field, and let $K \subset Z(L)$ be a subfield belonging to the center of L; moreover, let $(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$ be a throw in $\boldsymbol{P}^1_L$. Prove: a) There is a K-line $\boldsymbol{P}^1_K \subset \boldsymbol{P}^1_L$ with $\boldsymbol{x}_i \in \boldsymbol{P}^1_K$, $i = 0, \ldots, 3$, if and only if $\mathrm{CR}(\boldsymbol{x}_0, \boldsymbol{x}_1; \boldsymbol{x}_2, \boldsymbol{x}_3)$ belongs to K. – b) For every throw $(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$ in $\boldsymbol{P}^1_{\mathrm{H}}$ there is a complex line $\boldsymbol{P}^1_{\mathrm{C}} \subset \boldsymbol{P}^1_{\mathrm{H}}$ with $\boldsymbol{x}_i \in \boldsymbol{P}^1_{\mathrm{C}}$, $i = 0, \ldots, 3$.

1.10.7 Extensions of Null Systems, Polar Maps, and Quadrics

In this section we want to investigate the relations between fof null systems, polarities, and quadrics in $\boldsymbol{P}^n_K$ to those in the projective extension $\boldsymbol{P}^n_L$. To this end, we need a few simple preparations. To each K-form $\boldsymbol{W}_K$ on the right L-vector space $\boldsymbol{V}^{n+1}$ there corresponds a uniquely determined K-form on the dual vector space $\boldsymbol{V}'$:

$$\boldsymbol{W}'_K := \{\mathfrak{u} \in \boldsymbol{V}' \,|\, \mathfrak{u}(\boldsymbol{W}_K) \subset K\}; \tag{66}$$

$\boldsymbol{W}'_K$ is the $(n+1)$-dimensional K-vector space of $\boldsymbol{V}'$ arising from linearly extending the linear forms on $\boldsymbol{W}_K$ to $\boldsymbol{V}$; as already the notation indicates, it will always be identified with the dual of $\boldsymbol{W}_K$ in the sequel. Obviously, the basis dual to a K-basis $(\mathfrak{a}_i)$ for $\boldsymbol{V}$ is a K-basis of $\boldsymbol{V}'$ in the dual K-form $\boldsymbol{W}'_K$. We again interpret the elements of the corresponding projective spaces $\boldsymbol{P}'^n_L \supset \boldsymbol{P}'^n_K$ as hyperplanes in $\boldsymbol{P}^n_L$ or $\boldsymbol{P}^n_K$, respectively.

Exercise 15. Let $\dim_K L = r < \infty$. Prove that for each hyperplane $\boldsymbol{U} = [\mathfrak{u}] \in \boldsymbol{P}'^n_L$ the inequalities $n - r \leq \operatorname{Dim}_K \boldsymbol{U} \cap \boldsymbol{P}^n_K \leq n-1$ hold; moreover, $\operatorname{Dim}_K \boldsymbol{U} \cap \boldsymbol{P}^n_K = n-1$ holds if and only if $\boldsymbol{U} \in \boldsymbol{P}'^n_K$.

First we will discuss the extension of an auto-correlative map F on $\boldsymbol{P}^n_K$ to an auto-correlative map $\tilde{F}$ on $\boldsymbol{P}^n_L$. According to Definition 7.2 F is generated by an α-linear map $a : \boldsymbol{W}_K \to \boldsymbol{W}'_K$, where α is an anti-automorphism of K. Using τ-linear extension the following is easy to show:

Lemma 15. *Let the auto-correlative map F on $\boldsymbol{P}^n_K$ be induced by the α-linear map $a : \boldsymbol{W}_K \to \boldsymbol{W}'_K$. F can be extended to an auto-correlative map $\tilde{F}$ on $\boldsymbol{P}^n_L$ if and only if the anti-automorphism α of K has an extension to an anti-automorphism τ of L. For each of these extensions τ there is a unique extension $\tilde{F} = \tilde{F}_\tau$ induced by the τ-linear extension $\tilde{a}$ of a; with respect to a K-basis $(\mathfrak{a}_i)$ of $\boldsymbol{V}$ (Note the sum convention!) it is determined by*

$$\tilde{a}(\mathfrak{a}_i x^i) = \tau(x^i)a(\mathfrak{a}_i).$$

□

Example 8. If L is commutative and F is a *null system* over the projective space $\boldsymbol{P}^n_K$ with the subfield K determined by the alternating bilinear form $b(\mathfrak{x}, \mathfrak{y})$, $\mathfrak{x}, \mathfrak{y} \in \boldsymbol{W}^{n+1}_K$, then the linear extension $b_L(\mathfrak{x}, \mathfrak{y})$, $\mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}^{n+1}_L$, defines a null system of the same rank over the projective extension $\boldsymbol{P}^n_L$. In fact, the matrix of b_L with respect to a K-basis of $\boldsymbol{V}$ coincides with the matrix of b with respect to this basis. Compare also Exercise II.7.9.10. □

We will now discuss extensions of polar maps in a number of examples.

Example 9. Take $K = \mathbf{R}$ and $L = \mathbf{C}$. By Exercise I.2.1.3 we have $\alpha = \mathrm{id}_{\mathbf{R}}$, and there are two extensions of this automorphism to $\mathbf{C}$: $\tau = \mathrm{id}_{\mathbf{C}}$, or τ is the conjugation on $\mathbf{C}$. In the first case, each polar map $F_{r,l}$ of rank r and index l leads to a polar map F_r of rank r on $\boldsymbol{P}^n_{\mathbf{C}}$, cf. Proposition 9.5. If, conversely, $(\mathfrak{a}_j)$ is a basis of the complex vector space $\boldsymbol{V}^{n+1}$, with respect to which the auto-polar map F_r is in normal form, then this implies for the $\mathbf{R}$-structure $\boldsymbol{P}^n_{\mathbf{R}} \subset \boldsymbol{P}^n_{\mathbf{C}}$ determined by the basis

$$\mathfrak{b}_\lambda = \mathfrak{a}_\lambda \mathrm{i},\ \lambda = 0, \dots, l-1,\ \mathfrak{b}_\nu = \mathfrak{a}_\nu,\ \nu = l, \dots, n,$$

that $F_r|\boldsymbol{P}^n_{\mathbf{R}}$ has rank r, index l, and extends as F_r to $\boldsymbol{P}^n_{\mathbf{C}}$. If, on the other hand, τ is the conjugation, then extending a map of type $F_{r,l}$ yields a map on $\boldsymbol{P}^n_{\mathbf{C}}$, denoted in the same way, that is determined by a Hermitean form, cf. Proposition 9.7. □

Example 10. Consider $K = \mathbf{R}$, $L = \mathbf{H}$, and extend $\mathrm{id}_{\mathbf{R}}$ as the conjugation τ. Then $F_{r,l}$ again extends to the maps from Proposition 9.7 denoted in the same way. On the other hand, because of (9.19) we can also extend the identity $\mathrm{id}_{\mathbf{R}}$ to $\mathbf{H}$ as the anti-automorphism τ_{i} (cf. (9.17) with $q = \mathrm{i}$). It is easy to show that the polar map $F_{r,l}$ on $\boldsymbol{P}^n_{\mathbf{R}}$ then extends to a map of type J_r on $\boldsymbol{P}^n_{\mathbf{H}}$, cf. Proposition 9.14. □

Example 11. Consider the embedding $\mathbf{C} \subset \mathbf{H}$. By Proposition 9.5, for $\alpha = \mathrm{id}_{\mathbf{C}}$ there only exist the auto-polar maps of type F_r. A straightforward verification shows that the anti-automorphism τ_{j} (by (9.17) with $q = \mathrm{j}$) extends α to $\mathbf{H}$; here the extension of F_r leads to a polar map of type J_r. If, however, α is the conjugation on $\mathbf{C}$, then there exist two extensions to anti-automorphisms, the conjugation τ on $\mathbf{H}$ and the map τ_{i} already discussed in Example 10; the former associates the map on $\boldsymbol{P}^n_{\mathbf{H}}$ denoted in the same way with the polar map $F_{r,l}$ on $\boldsymbol{P}^n_{\mathbf{C}}$, whereas the latter again leads to a map of type J_r. □

The following example shows that, in general, the restriction of an auto-polar map F on $\boldsymbol{P}^n_L$ to a given K-form $\boldsymbol{P}^n_K \subset \boldsymbol{P}^n_L$ cannot be described by a polar map on $\boldsymbol{P}^n_K$.

Example 12. Let L be a field extension of K with rank s; let $c_1, \ldots, c_s$ be a basis of L over K, and assume that $\mathrm{char}\, K \neq 2$. Consider a polar map on $\boldsymbol{P}^n_L$ defined by a symmetric bilinear form $b : \boldsymbol{V} \times \boldsymbol{V} \to L$. If $\boldsymbol{W}^{n+1}_K \subset \boldsymbol{V}^{n+1}$ is the K-subspace determining the K-form, then we decompose $b|\boldsymbol{W}_K \times \boldsymbol{W}_K$ according to

$$b(\mathfrak{x}, \mathfrak{y}) = \sum_{\nu=1}^{s} c_\nu b^\nu(\mathfrak{x}, \mathfrak{y}), \ \mathfrak{x}, \mathfrak{y} \in \boldsymbol{W}^{n+1}_K.$$

The b^ν are symmetric bilinear forms over $\boldsymbol{W}_K$, which, in general, are independent and different from zero; they define s polar maps on $\boldsymbol{P}^n_K$ that have to be considered simultaneously. Hence, in general, investigating the K-geometry of polar maps on $\boldsymbol{P}^n_L$, and consequently also that of quadrics, might be quite extensive. □

Exercise 16. Consider an $\mathbf{R}$-form $\boldsymbol{P}^n_{\mathbf{R}} \subset \boldsymbol{P}^n_{\mathbf{C}}$. Let F be a polar map on $\boldsymbol{P}^n_{\mathbf{C}}$ of type $F_{r,l}$ determined by the Hermitean form b over the complex vector space $\boldsymbol{V}^{n+1}$. Prove the following equation:

$$b(\mathfrak{x}, \mathfrak{y}) = \alpha(\mathfrak{x}, \mathfrak{y}) + \mathrm{i}\, \beta(\mathfrak{x}, \mathfrak{y}), \ \mathfrak{x}, \mathfrak{y} \in \boldsymbol{W}^{n+1}_{\mathbf{R}}, \tag{67}$$

where $\alpha(\mathfrak{x}, \mathfrak{y})$ is a symmetric and $\beta(\mathfrak{x}, \mathfrak{y})$ an alternating real bilinear form on $\boldsymbol{W}^{n+1}_{\mathbf{R}}$. If, conversely, α and β are given as symmetric and alternating bilinear forms on $\boldsymbol{W}^{n+1}_{\mathbf{R}}$, then there is a unique extension of the form b defined by (67) to a Hermitean form over $\boldsymbol{V} \times \boldsymbol{V}$. This again determines a polar map on $\boldsymbol{P}^n_{\mathbf{C}}$ of a certain type $F_{r,l}$ (cf. Proposition 9.7). What are the relations between the invariants of b, α and β?

Exercise 17. Again, let an $\mathbf{R}$-form $\boldsymbol{P}^n_{\mathbf{R}} \subset \boldsymbol{P}^n_{\mathbf{C}}$ be given. Assume that the polar map F on $\boldsymbol{P}^n_{\mathbf{C}}$ is defined by the symmetric bilinear form b that has the decomposition (67) on the $\mathbf{R}$-form. Prove that the complex quadric Q_F meets the $\mathbf{R}$-form in the intersection of two quadrics:

$$Q_F \cap \boldsymbol{P}^n_{\mathbf{R}} = Q_\alpha \cap Q_\beta. \tag{68}$$

2

Cayley-Klein Geometries

According to F. Klein's Erlanger Programm [65], that we already described in § I.6.5, geometry and the theory of transformation groups are closely related. The largest groups, from which we will start in this chapter, are the linear groups $\boldsymbol{GL}(\boldsymbol{V}^{n+1})$ of finite-dimensional vector spaces over a skew field K and the projective groups $\boldsymbol{PL}(\boldsymbol{P}^n)$ assigned to them, as they were defined in Chapter 1. In this chapter we will summarize their properties in the cases of greatest interest for applications, i.e. $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$, and will also add a few facts. Distinguishing an object F in one of the projective geometries determined by these groups as the ***absolute*** the subgroup G_F of all projective transformations leaving F fixed, i.e. the *isotropy group of* F (cf. Appendix A.3), is a finer projective geometry. It studies all the properties remaining invariant when the projective action is restricted to the subgroup G_F. Since this group is a subgroup of the projective group, in general there are more preserved properties than the projective ones; hence they form what is called a *supergeometry* of the projective geometry, cf. O. Giering [44]. A. Cayley [29], [30] and F. Klein [60], [61], see also [62], Bd. 1, studied these geometries in the particular case that the polarity F belongs to a real projective geometry and that invariant metrics exist. The frequently quoted lectures by F. Klein [63], [64] contain detailed discussions of these geometries as applications of his Erlanger Programm. In this connection the terms *Cayley-Klein spaces, geometries, and metrics* are often used. We will not specify these notions here, even the extent of which is not agreed upon in the literature. The book by O. Giering [44] already mentioned presents such an approach. It also contains a detailed bibliography extending up to about 1980, including many comments in the text and more special results than we intend to present here.

2.1 The Classical Groups

The name *"classical groups"* is a rather historic-literary description of a class of linear or matrix groups arising not only in geometry but playing a fundamental part also in many other mathematical branches ranging from algebra, number theory, and analysis to theoretical physics. Although this term appears in the titles of two monographs from different periods of time, first probably in the year 1939 on H. Weyl's ground-breaking book [110], and then again in 1971 on J. Dieudonné's report on results [36], the notion classical group itself is not defined in either of them. The same is true of the very clearly written lecture notes [26] by P. J. Cameron, which emphasize the classical groups and the associated projective geometries over finite fields. The book [47] by L. Grove focusses on the algebraic aspects of the theory of classical groups. As in these texts we will use the term here in the same somewhat vague sense. In Section 2.1 we start by describing these "classical" groups, whose geometries will then be investigated in the subsequent sections.[1]

2.1.1 The Linear and the Projective Groups

We will even refer to the linear groups and the projective ones derived from them as *"classical"*, although, as a rule, they are not simple Lie groups. They were already introduced in the first chapter, since they determine the automorphism groups of projective geometries. The classical groups considered below typically are subgroups of these linear or the projective groups.

As before, the group of linear automorphisms of an n-dimensional vector space $\boldsymbol{V}^n$ over the skew field K is denoted by $\boldsymbol{GL}(\boldsymbol{V}^n)$. It is isomorphic to the group of invertible square matrices of order n:

$$\boldsymbol{GL}(\boldsymbol{V}^n) \cong \boldsymbol{GL}(n, K). \tag{1}$$

For a field K the group $\boldsymbol{GL}(n, K) \subset \boldsymbol{M}_n(K)$ consists of the matrices whose determinant is different from zero, cf. Corollary I.5.4.7. For linear endomorphisms a of the vector space $\boldsymbol{V}^n$ the determinant of the matrix (a^i_j) for a does not depend on the chosen basis; it is called the *norm* $N(a)$ of the endomorphism, cf. Proposition I.5.7.2, Definition I.5.7.2. Based on a generalization of the determinant to matrices over skew fields by J. Dieudonné [35], there is a characterization similar to the one above this case as well, cf. E. Artin [4], § IV.1. Since the only non-commutative skew field we consider here is that of the quaternions, the norm definition contained in Exercise 1.10.6 is sufficient for our purposes. For the norm homomorphism

$$a \in \boldsymbol{GL}(\boldsymbol{V}^n) \longmapsto N(a) \in K^* \ (K \text{ a field or } K = \mathbf{H}) \tag{2}$$

[1] According to the classification of complex and real simple Lie algebras, going back to W. Killing and E. Cartan, there are finitely many infinite series of simple Lie algebras, which are called *classical,* and finitely many so-called *exceptional algebras,* cf. J. Tits [105], [49], [84]. The terminology we use here is in accordance with this scheme. As already mentioned in the introduction, the finitely many simple groups corresponding to the exceptional algebras and their geometries will not be discussed in this book.

the kernel Ker N is just the *special linear group*

$$\boldsymbol{SL}(n,K) \cong \boldsymbol{SL}(\boldsymbol{V}^n) := \{a \in \boldsymbol{GL}(\boldsymbol{V}^n) | N(a) = 1\}, \tag{3}$$

that will also be viewed as a classical one. The same holds true for the *projective groups*

$$\boldsymbol{PL}_n \cong \boldsymbol{GL}(n+1,K)/Z_{n+1},$$

where

$$Z_{n+1} := \{(\delta^i_j)k | k \in Z(K^*)\} \tag{4}$$

denotes the center of $\boldsymbol{GL}(n+1,K)$, cf. Exercise II.7.4.7 and Definition 1.4.3. For the fields $K = \mathbf{R}, \mathbf{C}$ we already described the connection between the projective and the special linear groups in Example 1.4.5; in the notations introduced there we have

$$\boldsymbol{PL}_n(\mathbf{R}) \cong \boldsymbol{SL}(n+1,\mathbf{R}),\ n = 2k, \tag{5}$$

$$\boldsymbol{PL}_n(\mathbf{R}) \cong |\boldsymbol{SL}|(n+1,\mathbf{R})/\{\pm I_{n+1}\},\ n = 2k+1, \tag{6}$$

$$\boldsymbol{PL}_n(\mathbf{C}) \cong \boldsymbol{SL}(n+1,\mathbf{C})/K_{n+1}, \tag{7}$$

where

$$K_{n+1} = \{(\delta^i_j)k | k = e^{2\pi\,\mathrm{i}\,l/n+1},\ l = 0,\ldots,n\}$$

denotes the group of the $(n+1)$-st roots of unity. In the case of the quaternions we obtain a similar representation in a straightforward way: According to Exercise 1.10.6 the norm $N(g)$ is positive for each $g \in \boldsymbol{GL}(n,\mathbf{H})$. Hence g can uniquely be written in the form

$$g = g_o\rho,\ N(g_o) = 1,\ \rho > 0,\ N(g) = \rho^{2n}.$$

This implies

$$\boldsymbol{PL}_n(\mathbf{H}) \cong \boldsymbol{SL}(n+1,\mathbf{H})/\{\pm I_{n+1}\}. \tag{8}$$

Exercise 1. Show that the the normal subgroups occurring in the denominators of (6)–(8) are nothing but the centers of the numerators.

Exercise 2. Let $\boldsymbol{V}_{|\mathbf{C}}$ denote the complex restriction of the right quaternionic vector space $\boldsymbol{V}^n$. Show that the group $\boldsymbol{SL}(n,\mathbf{H})$ is isomorphic to the subgroup of all linear maps $g \in \boldsymbol{SL}(\boldsymbol{V}_{|\mathbf{C}})$ commuting with the quaternionic structure J of $\boldsymbol{V}_{|\mathbf{C}}$ (cf. Exercise 1.10.5). In addition, find natural embeddings of $\boldsymbol{SL}(n,\mathbf{C})$ and $\boldsymbol{SL}(n,\mathbf{R})$ into $\boldsymbol{SL}(n,\mathbf{H})$ that, with corresponding identifications, lead to the following chain of subgroups

$$\boldsymbol{SL}(n,\mathbf{R}) \subset \boldsymbol{SL}(n,\mathbf{C}) \subset \boldsymbol{SL}(n,\mathbf{H}) \subset \boldsymbol{SL}(2n,\mathbf{C}).$$

Remark. The subgroup of $\boldsymbol{SL}(2n,\mathbf{C})$ isomorphic to the group $\boldsymbol{SL}(n,\mathbf{H})$ is often denoted as $\boldsymbol{SU}^*(2n)$, cf. S. Helgason [49], S. 445.

2.1.2 The Projective Isotropy Group of a Correlation

To define the remaining classical groups we will now proceed as follows: In the projective geometry $\mathfrak{P}^n$ we distinguish a symmetric correlation F, i.e. a non-degenerate null system or a polarity, which will be called the *absolute correlation*; its isotropy group for the action of the projective group $\boldsymbol{PL}_n$:

$$\boldsymbol{PG}_n(F) := \{g \in \boldsymbol{PL}_n \mid g \circ F \circ g^{-1} = F\}, \tag{9}$$

cf. Definition 1.8.2, is called the *projective group of the correlation* F. Let now $b = b_F$ be any σ-biform defining the correlation F. According to Proposition 1.7.6 it is uniquely determined up to a proportionality factor $\kappa \in K^*$ and non-degenerate by Corollary 1.8.3. Its isotropy group for the action of the linear group $\boldsymbol{GL}(\boldsymbol{V})$ described by (I.8.12) with $\kappa = 1$ is denoted by $\boldsymbol{UG}_m(b)$, $m = \dim \boldsymbol{V}$; hence

$$g \in \boldsymbol{UG}_m(b) \Longleftrightarrow g \in \boldsymbol{GL}(\boldsymbol{V}^m) \text{ and } b(g\mathfrak{x}, g\mathfrak{y}) = b(\mathfrak{x}, \mathfrak{y}) \text{ for all } \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}^m. \tag{10}$$

The observations in Section 1.8, cf. (1.8.12), imply

Lemma 1. *Let b be a σ-biform defining the absolute correlation F, and let $g \in \boldsymbol{PL}_n$ be a projectivity induced by the linear transformation $a \in \boldsymbol{GL}(\boldsymbol{V}^{n+1})$. Then $g \in \boldsymbol{PG}_n(F)$ if and only if*

$$b(a\mathfrak{x}, a\mathfrak{y}) = \kappa(a) b(\mathfrak{x}, \mathfrak{y}) \text{ for all } \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}^{n+1}; \tag{11}$$

here $\kappa(a) \in Z(K)^*$ is a scalar depending only on a satisfying $\sigma(\kappa(a)) = \kappa(a)$. The canonical map $p : \boldsymbol{GL}(\boldsymbol{V}^{n+1}) \longrightarrow \boldsymbol{PL}_n$ thus yields the inclusion

$$p(\boldsymbol{UG}_{n+1}(b)) \subset \boldsymbol{PG}_n(F). \tag{12}$$

Proof. Relation (1.8.12) immediately implies assertion (11) with $\kappa(a) \in K^*$. Since b is non-degenerate, there are $\mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}^{n+1}$ such that $b(\mathfrak{x}, \mathfrak{y}) = 1$. By (11) we have for arbitrary $\lambda \in K$

$$\begin{aligned} b(a\mathfrak{x}\lambda, a\mathfrak{y}) &= \kappa(a) b(\mathfrak{x}\lambda, \mathfrak{y}) = \kappa(a)\sigma(\lambda) \\ &= \sigma(\lambda) b(a\mathfrak{x}, a\mathfrak{y}) = \sigma(\lambda)\kappa(a) b(\mathfrak{x}, \mathfrak{y}) = \sigma(\lambda)\kappa(a). \end{aligned}$$

Since σ is surjective, $\kappa(a)$ belongs to the center of K. Similarly,

$$\begin{aligned} b(a\mathfrak{y}, a\mathfrak{x}) &= \kappa(a) b(\mathfrak{y}, \mathfrak{x}) = \kappa(a) \\ &= \sigma(b(a\mathfrak{x}, a\mathfrak{y})) = \sigma(\kappa(a)). \end{aligned}$$

□

Obviously, the set of all $a \in \boldsymbol{GL}(\boldsymbol{V}^{n+1})$ satisfying (11) is a subgroup; it is called the *conformal group of the biform* b and is denoted by $\boldsymbol{CUG}_{n+1}(b)$. Inserting $a = a_2 \circ a_1$ into (11) we obtain

Corollary 2. *The map* $\kappa : \boldsymbol{CUG}_{n+1}(b) \to Z(K)^*$ *is a group homomorphism. In particular, for a field* K *it satisfies*

$$\kappa(a)^{n+1} = \sigma(N(a))N(a). \tag{13}$$

Proof. Choose a basis, replace (11) by the corresponding matrix equation,

$$(\sigma(a_i^k))(b_{kl})(a_j^l) = \kappa(a)(b_{ij}), \tag{14}$$

cf. (1.8.13), and compute the determinant for the matrix (a_{ij}) of a, $N(a) = \det(a_{ij})$. This immediately leads to formula (13). □

To describe the closely related groups $\boldsymbol{PG}_n(F)$, $\boldsymbol{UG}_{n+1}(b)$, $\boldsymbol{CUG}_{n+1}(b)$ in detail requires to have a classification of biforms over the skew field K at ones disposal. For $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$ this task was accomplished in Section 1.9 for the polarities and (in the case of an arbitrary K) in Section 1.8 for the null systems. In the sequel we restrict ourselves to these important cases and refer for the general theory to J. Dieudonné [36].

2.1.3 The Symplectic Groups

Let $F : \mathfrak{P}^{2n-1} \to \mathfrak{P}^{2n-1}$ be a non-degenerate null system, and consider a biform $b : \boldsymbol{V} \times \boldsymbol{V} \to K$ determining F. Then, according to Proposition 1.8.5, K has to be a field, and b is bilinear as well as alternating; moreover, V has even dimension, $\dim \boldsymbol{V} = 2n$. For fixed K and n, all these null systems F are equivalent. In a suitably chosen basis the matrix of b has the form (1.8.19) with $r = n$. Frequently, a somewhat different normal form is preferred: Applying the coordinate transformation

$$\hat{x}^i = x^{2i-1},\ \hat{x}^{n+i} = x^{2i},\ i = 1, \ldots, n, \tag{15}$$

in the new coordinates the matrix of b has the form

$$(b_{lk}) = J_n := \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}. \tag{16}$$

A vector space $\boldsymbol{V}^{2n}$ with a distinguished non-degenerate, alternating, bilinear form b is called *symplectic* just like the bases (or coordinates) in which b has normal form (16). If (ω^a) is a *symplectic cobasis*, i.e. a basis dual to a symplectic one, then, in it, the alternating 2-form b is expressed as

$$b = \omega^1 \wedge \omega^{n+1} + \omega^2 \wedge \omega^{n+2} + \ldots + \omega^n \wedge \omega^{2n}, \tag{17}$$

and vice versa (cf. Example II.8.3.2).

The transformations $a \in \boldsymbol{Sp}(\boldsymbol{V}^{2n}) := \boldsymbol{UG}_{2n}(b)$ are also called *symplectic*, and the group $\boldsymbol{Sp}(\boldsymbol{V}^{2n})$ is the *symplectic group*. Since $\kappa(a) = 1$, for $a \in \boldsymbol{UG}_m(b)$ relation (13) implies:

$$\sigma(N(a))N(a) = 1, \qquad a \in \boldsymbol{UG}_m(b),\ K \text{ a field;} \tag{18}$$

thus for bilinear b and $a \in \boldsymbol{UG}_m(b)$ we always have $N(a)^2 = 1$ (cf. Exercise 3). More precisely, symplectic transformations satisfy

$$N(a) = 1 \qquad a \in \boldsymbol{Sp}(\boldsymbol{V}^{2n}). \tag{19}$$

For $n = 1$, i.e. in a two-dimensional vector space, the alternating 2-form (17) coincides with the area form determining the equi-affine geometry, expressed as a scalar multiple of the determinant. So the symplectic group $\boldsymbol{Sp}(\boldsymbol{V}^2)$ is equal to the special linear group $\boldsymbol{SL}(\boldsymbol{V}^2)$, and the symplectic geometry on $\boldsymbol{V}^2$ coincides with its equi-affine geometry. Formulas (6), (7) show that the geometry determined by the symplectic group on the real or complex projective line presents nothing new as compared to projective geometry; it is for the projective space $\boldsymbol{P}^3$ that the symplectic geometry first becomes interesting.

Exercise 3. Prove (19) under the assumption $\operatorname{char} K = 0$ or $\operatorname{char} K > n$. Hint. Consider the alternating $2n$-form $b^n := \bigwedge_1^n b$. (19) also holds without the assumption above; the proof, however, is more difficult, cf. E. Artin [4], Proposition 3.25.

Exercise 4. Let $[\boldsymbol{V}^{2n}, b]$ be a symplectic vector space. Prove: A transformation $g \in \boldsymbol{GL}(\boldsymbol{V}^{2n})$ is symplectic if and only if its matrix with respect to a symplectic basis has the block structure

$$(a^i_j) = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \text{ with } A, B, C, D \in \boldsymbol{M}_n(K), \text{ satisfying}$$

$$A'C = C'A,\ B'D = D'B,\ A'D - C'B = I_n. \tag{20}$$

These matrices are, of course, also called *symplectic*. They form a subgroup of $\boldsymbol{GL}(2n, K)$ isomorphic to $\boldsymbol{Sp}(\boldsymbol{V}^{2n})$ denoted by $\boldsymbol{Sp}(n, K)$ and again called the *symplectic group*.

The projective space $\boldsymbol{P}^{2n-1} = \boldsymbol{P}(\boldsymbol{V}^{2n})$ corresponding to a symplectic vector space where the null system F determined by an alternating biform b is distinguished is called *projective symplectic*; the projectivities preserving F form the *projective symplectic group* $\boldsymbol{PSp}_n := \boldsymbol{PG}_{2n-1}(F)$. Now we consider the cases $K = \mathbf{R}, \mathbf{C}$ separately.

Example 1. Let $\boldsymbol{P}^{2n-1}$ be a real projective symplectic space. By (6), considering only the transformations a with $N(a)^2 = 1$ we already obtain all projectivities, which by (13) implies $\kappa(a)^{2n} = 1$. The map s_o with the coordinate representation

$$s_o :\ y^i = x^{n+i},\ y^{n+i} = x^i,\ i = 1, \ldots, n, \tag{21}$$

satisfies $\kappa(s_o) = -1$. Since, by Corollary 2, $\kappa : |\boldsymbol{SL}|(2n, \mathbf{R}) \longrightarrow \{\pm 1\}$ is a homomorphism as well, we conclude

$$\boldsymbol{PSp}_n(\mathbf{R}) \cong p(\boldsymbol{Sp}(n, \mathbf{R})) \cup p(s_o(\boldsymbol{Sp}(n, \mathbf{R})). \tag{22}$$

Each real, projective symplectic transformation can thus be described by a symplectic transformation $a \in \boldsymbol{Sp}(\boldsymbol{V}^{2n})$ or a transformation of the form $s_0 \circ a$, $a \in \boldsymbol{Sp}(\boldsymbol{V}^{2n})$. As only the dilations belong to $\operatorname{Ker} p$, the homomorphism theorem implies

$$\boldsymbol{PSp}_n(\mathbf{R}) \cong (\boldsymbol{Sp}(\boldsymbol{V}^{2n}) \cup s_o(\boldsymbol{Sp}(\boldsymbol{V}^{2n}))/\{\pm I_{2n}\}.$$

□

Example 2. Consider the complex projective symplectic space $\boldsymbol{P}^{2n-1} = \boldsymbol{P}(\boldsymbol{V}^{2n})$. Since for each dilation d_μ, $\mu \in \mathbf{C}^*$, $\kappa(d_\mu) = \mu^2$ by (11), we may always suppose that the symplectic projectivity $g = p(a) \in \boldsymbol{PSp}_n$ is generated by a symplectic transformation $a \in \boldsymbol{Sp}(\boldsymbol{V}^{2n})$; in fact, $p(a) = p(a \circ d_\mu)$ and

$$\kappa(a \circ d_\mu) = \kappa(a)\mu^2 = 1$$

always has a solution. Hence we have

$$\boldsymbol{PSp}_n(\mathbf{C}) = p(\boldsymbol{Sp}(\boldsymbol{V}^{2n})) \cong \boldsymbol{Sp}(n, \mathbf{C})/\{\pm I_{2n}\},$$

since only the elements $\pm I_{2n}$ belong to $\boldsymbol{Sp}(n, \mathbf{C}) \cap K_{2n}$ (cf. (7)). □

2.1.4 The Orthogonal Groups

Now we want to discuss the polarities F. For $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$ they were classified in Section 1.9. The results of this classification are shown in Table 2.1 at the end of Section 2.1.6. We will discuss the cases one by one.

Example 3. Let $\boldsymbol{V}^n$ be a real vector space, and let b be a distinguished symmetric, non-degenerate, bilinear form of index l on V. Then $[\boldsymbol{V}^n, b]$ is called a *pseudo-Euclidean vector space* and the corresponding isotropy group is called the *pseudo-orthogonal group* $\boldsymbol{O}(l, n-l) := \boldsymbol{UG}_n(b)$. Since $-b$ is invariant if b is, we have

$$\boldsymbol{O}(l, n-l) = \boldsymbol{O}(n-l, l),$$

and we may assume $0 \leq l \leq n/2$. Obviously, for $l = 0$ we just have a *Euclidean vector space* and the *orthogonal group* $\boldsymbol{O}(n) = \boldsymbol{O}(0, n)$ (cf. Corollary I.6.2.4). For the theory of relativity the *Minkowski space* is of fundamental importance, i.e. the four-dimensional pseudo-Euclidean vector space of index 1. The associated group $\boldsymbol{O}(1,3)$ is called the *Lorentz group*. Since $\sigma = \mathrm{id}_{\mathbf{R}}$, we conclude from (18)

$$N(a)^2 = 1, \qquad a \in \boldsymbol{O}(l, n-l). \tag{23}$$

In order to determine the isotropy group $\boldsymbol{PO}(l, n-l) := \boldsymbol{PG}_{n-1}(F_{n,l})$ of the corresponding polarity $F_{n,l}$, again by (5) and (6) we may start from the linear transformations a with $N(a)^2 = 1$; according to (13) these transformations satisfy $\kappa(a)^n = 1$. If $n \neq n - l$, i.e. $l \neq n/2$, then by the Theorem of Inertia, Proposition I.5.9.6, in fact $\kappa(a) = 1$, and, moreover,

$$\boldsymbol{PO}(l, n-l) = p(\boldsymbol{O}(l, n-l)) \cong \boldsymbol{O}(l, n-l)/\{\pm I_n\}, \qquad l \neq n/2. \tag{24}$$

Next we consider the case $\boldsymbol{O}(n,n)$, i.e. $F_{2n,n}$, and, introducing coordinates in which b has normal form, we see that by (21) the map s_o again has the factor $\kappa(s_o) = -1$. As in the case of the symplectic group this implies

$$\begin{aligned} &\boldsymbol{PO}(n,n) = p(\boldsymbol{O}(n,n) \cup s_o(\boldsymbol{O}(n,n))) \text{ and} \\ &\boldsymbol{PO}(n,n) \cong (\boldsymbol{O}(n,n) \cup s_o(\boldsymbol{O}(n,n)))/\{\pm I_{2n}\}. \end{aligned} \tag{25}$$

□

Example 4. The *complex Euclidean vector space* $[\boldsymbol{V}^n, b]$ is defined in analogy to Example 3; this time, let b be a non-degenerate, symmetric, bilinear form. Its isotropy group is denoted by $\boldsymbol{O}(n, \mathbf{C}) := \boldsymbol{UG}_n(b)$ and called the *complex orthogonal group*. As in Examples 2, 3 we conclude for the isotropy group of the polarity $\boldsymbol{PO}(n, \mathbf{C}) := \boldsymbol{PG}_{n-1}(F_n)$:

$$\boldsymbol{PO}(n, \mathbf{C}) \cong \boldsymbol{O}(n, \mathbf{C})/\{\pm I_n\}.$$

□

Proposition 1.9.3, see also Proposition I.5.9.6 and Corollary I.5.9.3, implies that for any non-degenerate symmetric bilinear form b and $K = \mathbf{R}, \mathbf{C}$ we can always find bases $(\mathfrak{e}_i)$ of $\boldsymbol{V}^n$ in which the *orthogonality relations*

$$b(\mathfrak{e}_i, \mathfrak{e}_j) = \delta_{ij}\epsilon_i, \qquad i, j = 1, \ldots, n,$$

hold; here $\epsilon_i = -1$ for $K = \mathbf{R}$, $i = 1, \ldots, l$, and $\epsilon_i = 1$ else. Such bases are called *pseudo-orthonormal* or *orthonormal*, respectively, if $l = 0$ or $K = \mathbf{C}$. A linear transformation a belongs to the isotropy group $\boldsymbol{UG}_n(b)$ if its matrix (a_i^k) is *pseudo-orthogonal* with respect to a pseudo-orthonormal basis, i.e., if the *orthogonality relations*

$$\sum_{k=1}^{n} \epsilon_k a_i^k a_j^k = \epsilon_i \delta_{ij}, \qquad i, j = 1, \ldots, n, \tag{26}$$

hold. The groups of pseudo-orthogonal matrices isomorphic to the isotropy groups are denoted like these. Again, relation (13) implies that for pseudo-orthogonal matrices $\det(a_i^k)^2 = 1$. The subgroups of matrices (or linear transformations, respectively) whose determinant (norm) is equal to 1 are called *special pseudo-orthogonal groups* and denoted by $\boldsymbol{SO}(l, n-l)$. For $l = 0$ and $K = \mathbf{R}$ it is called the *special orthogonal group* $\boldsymbol{SO}(n)$ or the *special complex orthogonal group* $\boldsymbol{SO}(n, \mathbf{C})$.

2.1.5 The Unitary Groups

Let now $K = \mathbf{C}$ or $K = \mathbf{H}$, and consider a non-degenerate Hermitean form b of index l, cf. Lemma 1.9.8. The related isotropy groups are called *pseudo-unitary*, or *unitary* in the case $l = 0$. For $K = \mathbf{C}$ they are denoted by $\boldsymbol{U}(l, n-l)$, $\boldsymbol{U}(n) := \boldsymbol{U}(0, n)$, and for $K = \mathbf{H}$ by $\boldsymbol{Sp}(l, n-l)$, $\boldsymbol{Sp}(n) := \boldsymbol{Sp}(0, n)$. As in Example 3 we may suppose $0 \leq l \leq n/2$. The notion of a pseudo-orthonormal basis defined above is also applied to pseudo-unitary groups. The matrices of pseudo-unitary transformations with respect to these bases are again called pseudo-unitary. They are characterized by *orthogonality relations* analogous to (26):

$$\sum_{k=1}^{n} \epsilon_k \bar{a}_i^k a_j^k = \epsilon_i \delta_{ij}, \qquad i, j = 1, \ldots, n. \tag{27}$$

For $K = \mathbf{C}$ we compute the determinant or use (13) to obtain

$$|N(a)| = |\det(a_j^i)| = 1. \tag{28}$$

Note that, e.g., $a = \mathrm{id}_{\boldsymbol{V}} \cdot \mathrm{e}^{\mathrm{i}\varphi}$ is always pseudo-unitary, whereas, in general, we have $N(a) = \mathrm{e}^{\mathrm{i}\,n\varphi} \neq 1$. The kernel of the norm homomorphism ($K = \mathbf{C}$),

$$N : a \in \boldsymbol{U}(l, n-l) \longmapsto N(a) \in S^1 = \{z \in \mathbf{C} |\; |z| = 1\},$$

is called the *special pseudo-unitary group*

$$\boldsymbol{SU}(l, n-l) := \{a \in \boldsymbol{U}(l, n-l) |\; N(a) = 1\}, \tag{29}$$

the *special unitary group* is defined as $\boldsymbol{SU}(n) := \boldsymbol{SU}(0, n)$ (cf. (I.6.2.33)).

For the *quaternionic pseudo-unitary groups* $\boldsymbol{Sp}(l, n-l)$ we now prove the following result describing complex realizations of these groups:

Proposition 3. *Let $\boldsymbol{V} = \boldsymbol{V}^n$ be an n-dimensional, right vector space over $\mathbf{H}$, let b be a non-degenerate, τ-Hermitean biform of index l over $\boldsymbol{V}$, and let $\tau(q) = \bar{q}$ be the conjugation in $\mathbf{H}$ (cf. Proposition 1.9.7). Moreover, denote by $\boldsymbol{V}_{|\mathbf{C}}$ the complex restriction of $\boldsymbol{V}$. Decompose $b(\mathfrak{x}, \mathfrak{y})$ into its two complex components according to*

$$b(\mathfrak{x}, \mathfrak{y}) = \alpha(\mathfrak{x}, \mathfrak{y}) + \mathrm{j}\,\beta(\mathfrak{x}, \mathfrak{y}) \qquad (\mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}), \tag{30}$$

$\alpha(\mathfrak{x}, \mathfrak{y}), \beta(\mathfrak{x}, \mathfrak{y}) \in \mathbf{C}$. Then:
1) α is non-degenerate and Hermitean of index $2l$ over $\boldsymbol{V}_{|\mathbf{C}}$;
2) β is a non-degenerate, alternating, bilinear form over $\boldsymbol{V}_{|\mathbf{C}}$;
3) the group $\boldsymbol{Sp}(l, n-l)$ is isomorphic to the group of complex linear transformations

$$\boldsymbol{Sp}(l, n-l) \cong \hat{\boldsymbol{U}}(2l, 2(n-l)) \cap \boldsymbol{Sp}(n, \mathbf{C}), \tag{31}$$

where $\hat{\boldsymbol{U}}(2l, 2(n-l))$ is the group isomorphic to $\boldsymbol{U}(2l, 2(n-l))$, consisting of the transformations leaving the Hermitean form α invariant whose matrix

(α_{ij}) is the following block matrix in a symplectic basis (with respect to β) $(\mathfrak{a}_j, \mathfrak{a}_{n+j})$, $j = 1, \dots, n$:

$$(\alpha_{ij}) = K_{l,n-l} := \begin{pmatrix} -I_l & 0 & 0 & 0 \\ 0 & I_{n-l} & 0 & 0 \\ 0 & 0 & -I_l & 0 \\ 0 & 0 & 0 & I_{n-l} \end{pmatrix}. \tag{32}$$

P r o o f. Since b is a τ-biform, we have

$$b(\mathfrak{y}, \mathfrak{x}) = \alpha(\mathfrak{y}, \mathfrak{x}) + \mathrm{j}\,\beta(\mathfrak{y}, \mathfrak{x}) = \overline{b(\mathfrak{x}, \mathfrak{y})} = \overline{\alpha(\mathfrak{x}, \mathfrak{y})} + \overline{\mathrm{j}\,\beta(\mathfrak{x}, \mathfrak{y})}.$$

Now

$$\overline{\mathrm{j} \cdot \beta(\mathfrak{x}, \mathfrak{y})} = \overline{\beta(\mathfrak{x}, \mathfrak{y})} \cdot \bar{\mathrm{j}} = -\mathrm{j} \cdot \beta(\mathfrak{x}, \mathfrak{y}).$$

Comparing the components shows that α is Hermitean, and β is bilinear as well as alternating. So let $(\mathfrak{a}_i)$, $i = 1, \dots, n$, be a pseudo-orthogonal basis with respect to b, i.e., suppose that

$$b(\mathfrak{a}_k, \mathfrak{a}_l) = \epsilon_k \delta_{kl}, \qquad k, l = 1, \dots, n. \tag{33}$$

Then $(\mathfrak{a}_k, \mathfrak{a}_{n+k})$, $k = 1, \dots, n$, with $\mathfrak{a}_{n+k} := \mathfrak{a}_k \epsilon_k \mathrm{j}$ is a basis for $\boldsymbol{V}_{\mathbf{C}}$. A straightforward computation shows that this basis is symplectic with respect to β, and that the biform α has the matrix (32) in it. This again implies that each $\mathbf{H}$-linear transformation a leaving b invariant preserves α as well as β as a $\mathbf{C}$-linear transformation, hence $a \in \hat{\boldsymbol{U}}(2l, 2(n-l)) \cap \boldsymbol{Sp}(n, \mathbf{C})$. The following exercise shows that each transformation $a \in \hat{\boldsymbol{U}}(2l, 2(n-l)) \cap \boldsymbol{Sp}(n, \mathbf{C})$ von $\boldsymbol{V}_{|\mathbf{C}}$ is $\mathbf{H}$-linear. This implies (31). □

Exercise 5. Complete the computations left out in the proof of Proposition 3. Hint. Represent the matrix of the $\mathbf{C}$-linear map a with respect to the basis $(\mathfrak{a}_k)$, $k = 1 \dots, 2n$, defined in the proof as a block matrix of the form

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}, \qquad A, B, C, D \in \boldsymbol{M}_n(\mathbf{C}).$$

Then derive the conditions to be satisfied by $a \in \hat{\boldsymbol{U}}(2l, 2(n-l))$ as well as $a \in \boldsymbol{Sp}(n, \mathbf{C})$, and conclude from them two expressions for a^{-1} as block matrices of the same type. Their comparison shows $a(\mathfrak{x}\,\mathrm{j}) = a(\mathfrak{x})\,\mathrm{j}$. This implies that a is $\mathbf{H}$-linear.

Proposition 3 shows the close relation of $\boldsymbol{Sp}(l, n-l)$ with the symplectic group $\boldsymbol{Sp}(n, \mathbf{C})$; moreover, it justifies the notation. In particular,

$$\boldsymbol{Sp}(n) \cong \boldsymbol{Sp}(n, \mathbf{C}) \cap \boldsymbol{U}(2n), \tag{34}$$

where the basis $(\mathfrak{a}_k)$, $k = 1, \dots, 2n$, defined above is symplectic and, at the same time, orthonormal. The following exercise formulates a result similar to Proposition 3 for the real restriction of a complex pseudo-unitary vector space.

Exercise 6. Let $\boldsymbol{V} = \boldsymbol{V}^n$ be an n-dimensional complex vector space, and let $\boldsymbol{V}_{|\mathbf{R}}$ be its real restriction; moreover, let b be a Hermitean biform over $\boldsymbol{V}$ of index l. Prove: a) The decomposition of $b(\mathfrak{x}, \mathfrak{y})$ into its real and imaginary parts,

$$b(\mathfrak{x}, \mathfrak{y}) = \alpha(\mathfrak{x}, \mathfrak{y}) + \mathrm{i}\,\beta(\mathfrak{x}, \mathfrak{y}) \qquad \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V},$$

defines a non-degenerate symmetric bilinear form α of index $2l$ as well as a non-degenerate alternating bilinear form β over $\boldsymbol{V}_{|\mathbf{R}}$. – b) The pseudo-unitary group $\boldsymbol{U}(l, n-l)$ is isomorphic to the group of real linear transformations,

$$\boldsymbol{U}(l, n-l) \cong \hat{\boldsymbol{O}}(2l, 2(n-l)) \cap \boldsymbol{Sp}(n, \mathbf{R}),$$

where $\hat{\boldsymbol{O}}(2l, 2(n-l))$ denotes the group isomorphic to the pseudo-orthogonal group $\boldsymbol{O}(2l, 2(n-l))$ of real linear transformations which leave the bilinear form α invariant and whose matrix has the form (32) with respect to a symplectic basis $(\mathfrak{a}_j, \mathfrak{a}_{n+j})$, $j = 1, \ldots, n$, of $\boldsymbol{V}_{|\mathbf{R}}$. Hint. If $(\mathfrak{a}_j)$, $j = 1, \ldots, n$, is a pseudo-orthonormal basis of $\boldsymbol{V}$, then define $\mathfrak{a}_{n+j} := \mathfrak{a}_j\,\mathrm{i}\,\epsilon_j$ and proceed as in the proofs of Proposition 3, Exercise 5.

Exercise 7. Consider the action of the unitary group $\boldsymbol{U}(n)$ on the Hermitean vector space $\boldsymbol{V}^n$. Let $G_{2n,k}(\mathbf{R})$, $0 < k < 2n$, denote the *Graßmann manifold of k-dimensional real subspaces of the real restriction $\boldsymbol{V}_{|\mathbf{R}}$ of $\boldsymbol{V}$*. Prove: a) $G_{2n,k}(\mathbf{R})$ is invariant under the action of $\boldsymbol{U}(n)$. – b) For $k = 1$ and $k = 2n-1$ the action of $\boldsymbol{U}(n)$ on $G_{2n,k}(\mathbf{R})$ is transitive, and for $1 < k < 2n-1$ it is not transitive.

Example 5. Now consider the Hermitean polarities $H_{n,l}$ of $\boldsymbol{P}^{n-1}(K)$ for $K = \mathbf{C}, \mathbf{H}$. Because of (7) and (8), we may again start from $\boldsymbol{SL}(n, K)$ in order to determine their isotropy groups:

$$\boldsymbol{PU}(l, n-l) := \boldsymbol{PG}_{n-1}(H_{n,l}) \text{ for } K = \mathbf{C},$$
$$\boldsymbol{PSp}(l, n-l) := \boldsymbol{PG}_{n-1}(H_{n,l}) \text{ for } K = \mathbf{H};$$

by (13) we obtain $\kappa^n(a) = 1$, and, applying once again the Theorem of Inertia (Lemma 1.9.8), for $l \neq n/2$ we even arrive at $\kappa(a) = 1$. This leads to

$$\boldsymbol{PU}(l, n-l) \cong \boldsymbol{SU}(l, n-l)/K_n \text{ for } K = \mathbf{C}, l \neq n/2,$$
$$\boldsymbol{PSp}(l, n-l) := \boldsymbol{Sp}(l, n-l)/\{\pm I_n\} \text{ for } K = \mathbf{H}, l \neq n/2.$$

For $l = n/2$ use (21) and construct from s_o a transformation with $\kappa(s_o) = -1$, this yields a result similar to (25). □

2.1.6 The Quaternionic Skew Hermitean Polarities

Next we want to study the polarities $H_n := J_n$ of the quaternionic projective geometry $\mathfrak{P}^{n-1}(\mathbf{H})$, cf. Proposition 1.9.14; each of them is described by a skew Hermitean biform (1.9.22) with $r = n$. Once more we decompose this biform b according to (30) into both its complex components α, β and consider the complex restriction $\boldsymbol{V}^n_{|\mathbf{C}}$. Since b is skew Hermitean with respect to $\mathbf{H}$

(cf. 1.9.16), the biform α is complex skew Hermitean, and β is symmetric as well as $\mathbf{C}$-bilinear. Let $(\mathfrak{b}_j)$, $j = 1, \ldots, n$, be a basis of $\boldsymbol{V}^n$ with respect to which the matrix of b has the normal form (1.9.22). Then $(\mathfrak{b}_l, \mathfrak{b}_{n+l})$, $l = 1, \ldots, n$, with $\mathfrak{b}_{n+l} := \mathfrak{b}_l\,\mathrm{j}$ is a basis of $\boldsymbol{V}^n_{|\mathbf{C}}$; denoting by

$$x^l = \xi^l + \mathrm{j}\,\xi^{n+l},\ l = 1, \ldots, n,\ \xi^l, \xi^{n+l} \in \mathbf{C}, \tag{35}$$

the quaternionic coordinates of the vector $\mathfrak{x}$ with respect to $(\mathfrak{b}_l)$, the ξ^k, $k = 1, \ldots, 2n$, are its complex coordinates. In these, the biforms α, β are expressed as follows:

$$\alpha(\mathfrak{x}, \mathfrak{y}) = \mathrm{i}\sum_{l=1}^{n}(\bar{\xi}^l \eta^l - \bar{\xi}^{n+l}\eta^{n+l}), \tag{36}$$

$$\beta(\mathfrak{x}, \mathfrak{y}) = -\,\mathrm{i}\sum_{l=1}^{n}(\xi^l \eta^{n+l} + \xi^{n+l}\eta^l). \tag{37}$$

This shows that both biforms are non-degenerate. Hence β is complex Euclidean. Changing coordinates in $\boldsymbol{V}_{|\mathbf{C}}$ according to

$$\xi^l = (\xi'^l + \mathrm{i}\,\xi'^{n+l})/\sqrt{2},\ \xi^{n+l} = (\xi'^l - \mathrm{i}\,\xi'^{n+l})\,\mathrm{i}\,/\sqrt{2}$$

we arrive at normal forms for both biforms:

$$\alpha(\mathfrak{x}, \mathfrak{y}) = \sum_{l=1}^{n}(\bar{\xi}'^{n+l}\eta'^l - \bar{\xi}'^l\eta'^{n+l}), \tag{38}$$

$$\beta(\mathfrak{x}, \mathfrak{y}) = \sum_{l=1}^{n}(\xi'^l\eta'^l + \xi'^{n+l}\eta'^{n+l}). \tag{39}$$

The subgroup of complex orthogonal transformations leaving α invariant is denoted by $\boldsymbol{O}^*(2n)$; set, moreover,

$$\boldsymbol{SO}^*(2n) := \boldsymbol{SO}(2n, \mathbf{C}) \cap \boldsymbol{O}^*(2n). \tag{40}$$

Exercise 8. Prove that $\boldsymbol{SO}^*(2n)$ is a subgroup of $\boldsymbol{SU}^*(2n)$ (cf. the remark following Exercise 2).

Exercise 9. Prove that the subgroup of $\boldsymbol{GL}(\boldsymbol{V}^n_{|\mathbf{C}})$ leaving α invariant is isomorphic to the subgroup $\boldsymbol{U}(n, n)$.

Let now $a \in \boldsymbol{GL}(\boldsymbol{V}^n)$ be an $\mathbf{H}$-linear transformation satisfying (11). Lemma 1 then implies $\kappa(a) \in \mathbf{R}^*$, and we may always suppose $\kappa = \pm 1$. Left multiplication by the imaginary unit $\mathrm{k} \in \mathbf{H}^*$, i.e. applying the map

$$s_o : \mathfrak{x} = \sum_{l=1}^{n} \mathfrak{b}_l x^l \longmapsto s_o(\mathfrak{x}) = \sum_{l=1}^{n} \mathfrak{b}_l \,\mathrm{k}\cdot x^l, \tag{41}$$

yields an element $s_o \in \boldsymbol{SL}(\boldsymbol{V}^n)$ satisfying (14) with $\kappa(s_o) = -1$. Since we only have to consider elements $a \in \boldsymbol{SL}(\boldsymbol{V}^n)$ by (8), for the isotropy group of H_n a result similar to (25) holds:

$$\boldsymbol{PG}_{n-1}(H_n) \cong (\boldsymbol{SO}^*(2n) \cup s_o(\boldsymbol{SO}^*(2n)))/\{\pm I_{2n}\}. \tag{42}$$

Exercise 10. Determine the matrix for the map s_o in the complex restriction $\boldsymbol{V}^n_{|\mathbb{C}}$ defined by (41), show $s_o \in \boldsymbol{SL}(\boldsymbol{V}^n)$, and $\kappa(s_o) = -1$.

Exercise 11. Let $\boldsymbol{V}^n$ be a vector space over the skew field K, and let b be a non-degenerate σ-Hermitean or skew σ-Hermitean biform over $\boldsymbol{V}$. Moreover, let $a \in \boldsymbol{GL}(\boldsymbol{V}^n)$ be a linear transformation preserving orthogonality:

$$b(\mathfrak{x}, \mathfrak{y}) = 0 \text{ always implies } b(a(\mathfrak{x}), a(\mathfrak{y})) = 0.$$

Prove that $a \in \boldsymbol{CU}_n(b)$.

2.1.7 $\boldsymbol{SL(n, K)}$ is Generated by Transvections.

To prove the invariance of a geometric property under all the transformations from a certain group G it is often useful to know a system of generators for the group consisting of particularly simple transformations. In fact, it then suffices to verify the invariance for these generating transformations. In this section we will prove that the special linear group over a field K is generated by transvections. A *transvection* is understood to be a projectivity $f \in \boldsymbol{PL}_n$ for which there is a hyperplane $\boldsymbol{H} \subset \boldsymbol{P}^n$ that f leaves point-wise fixed, i.e. $f|\boldsymbol{H} = \mathrm{id}_{\boldsymbol{H}}$, and which has no additional fixed points. Obviously, this definition also makes sense for projective spaces over skew fields. Considering, as in Exercise 1.5.3, affine space as a subset of projective space the translations are just the restrictions of transvections, whose fixed-point set is the improper hyperplane, to affine space. Note that now, in projective geometry, different transvections have different hyperplanes as their fixed-point sets, in contrast to the translations of the projectively embedded affine geometry, whose fixed-point set is the distinguished improper hyperplane we keep fixed. The basis representations of transvections have particularly simple matrices, if the points $\boldsymbol{a}_j = [\mathfrak{a}_j] \in \boldsymbol{H}$, $j = 1, \ldots, n$, of the frame are fixed points themselves. Let B denote the matrix of the transvection b in a corresponding basis with $\boldsymbol{a}_0 = [\mathfrak{a}_0] \notin \boldsymbol{H}$, then, just like for the translations, we have

$$B = \begin{pmatrix} 1 & \mathfrak{o}' \\ \mathfrak{a} & I_n \end{pmatrix} \text{ with } \mathfrak{a} \neq \mathfrak{o} \in K^n. \tag{43}$$

At first, the conditions in the definition only lead to the form

K	polarity	biform b	$\boldsymbol{UG}_n(b)$
$\mathbf{R}$	$F_{n,l}$	bilinear, symmetric, of index l	$\boldsymbol{O}(l,n-l)$, $\boldsymbol{O}(n) := \boldsymbol{O}(0,n)$
$\mathbf{C}$	F_n	bilinear, symmetric	$\boldsymbol{O}(n,\mathbf{C})$
$\mathbf{C}$	$H_{n,l}$	Hermitean of index l	$\boldsymbol{U}(l,n-l)$, $\boldsymbol{U}(n) := \boldsymbol{U}(0,n)$
$\mathbf{H}$	$H_{n,l}$	Hermitean of index l	$\boldsymbol{Sp}(l,n-l)$, $\boldsymbol{Sp}(n) := \boldsymbol{Sp}(0,n)$
$\mathbf{H}$	H_n	quaternionic skew Hermitean	$\boldsymbol{O}^*(2n)$

The normal forms for the non-degenerate biforms above are

$$\begin{aligned} F_{n,l}: &\quad -\sum_{\alpha=1}^{l} x^\alpha y^\alpha + \sum_{\kappa=l+1}^{n} x^\kappa y^\kappa,\ 0 \le l \le n/2, \\ F_n: &\quad \sum_{j=1}^{n} x^j y^j, \\ H_{n,l}: &\quad -\sum_{\alpha=1}^{l} \bar{x}^\alpha y^\alpha + \sum_{\kappa=l+1}^{n} \bar{x}^\kappa y^\kappa,\ 0 \le l \le n/2, \\ H_n: &\quad \sum_{j=1}^{n} \bar{x}^j\, \mathrm{i}\, y^j. \end{aligned}$$

Table 2.1. Polarities, biforms, and their isotropy groups

$$\begin{pmatrix} a & \mathfrak{o}' \\ \mathfrak{a} & I_n\lambda \end{pmatrix} \text{ with } a \in K^*,\ \mathfrak{a} \in K^n,\ \lambda \in Z^*.$$

Since, by Definition 1.4.3 for the projective groups (cf. (1.4.22)), the kernel of the canonical map $\boldsymbol{GL}(n+1,K) \to \boldsymbol{PL}_n$ is the center of $\boldsymbol{GL}(n+1,K)$, i.e.

$$Z(n+1) = \{I_{n+1}\lambda | \lambda \in Z^*\},$$

we may multiply the matrix by λ^{-1}, or suppose without loss of generality that $\lambda = 1$. Now it is easy to check that the corresponding projectivity has no fixed point outside of $\boldsymbol{H}$ if and only if $a = 1$ and $\mathfrak{a} \neq \mathfrak{o}$ both hold. The following exercise characterizes the linear transvections.

Exercise 12. A linear transformation $b \in \boldsymbol{GL}(\boldsymbol{V}^{n+1})$ is a transvection if and only if it has the form

$$b(\mathfrak{x}) = \mathrm{id}_{\boldsymbol{V}}(\mathfrak{x}) + \mathfrak{a}\omega(\mathfrak{x}),$$

where $\omega \in \boldsymbol{V}'$ is a linear form, $\omega \neq 0$, satisfying $\omega(\mathfrak{a}) = 0$. This again is the case if and only if $\mathrm{rk}(b - \mathrm{id}_{\boldsymbol{V}}) = 1$ and $(b - \mathrm{id}_{\boldsymbol{V}})^2 = 0$.

The isomorphies (5)–(8) show that for the scalar domains $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$ the projective groups essentially are quotients of special linear groups. Hence the following proposition is of particular interest for projective geometries over these scalar domains:

Proposition 4. *Let K be a field. Then the special linear group $\boldsymbol{SL}(n+1, K)$ and the group of projectivities $\boldsymbol{PSL}(n+1, K) \subset \boldsymbol{PL}_n$ corresponding to it are generated by transvections.*

Proof. We prove this by induction, starting with $n = 1$. Let G denote the subgroup of projectivities generated by the transvections. According to (43) each transvection can be generated by a linear transvection with determinant one. Hence we have $G \subset \boldsymbol{PSL}(n+1, K)$. The transvections leaving, say, the point ∞ of a projective scale on the line invariant are the translations represented in this scale as $x \mapsto x + a$, and the group they generate acts transitively on the complement of ∞. Thus G acts *doubly transitively* on the projective line $\boldsymbol{P}^1$, i.e., we can always find an element $h \in G$ transforming the points x, y with $x \neq y$ into the points $h(x) = 0$, $h(y) = 1$. (We identify the points with the corresponding elements of a fixed projective scale.) Now take an arbitrary $g \in \boldsymbol{PSL}(2, K)$, $g \neq e$. We then find $x_1 \in \boldsymbol{P}^1$ such that $g(x_1) = \infty$. If $x_1 \neq \infty$, then there is a transvection t_1 with the " hyperplane " $\boldsymbol{H} = \{0\}$ and $t_1(\infty) = x_1$, so that for $g_1 = g \circ t_1$ the point $g_1(\infty) = \infty$ is fixed. Hence if g_1 can be represented as the product of transvections, this is also true of g; so we may suppose $g(\infty) = \infty$. Moreover, we find a transvection t_0 with fixed point ∞ mapping $y_0 = g(0)$ to 0. Thus $g_0 = t_0 \circ g$ satisfies $g_0(0) = 0$. What we now still have to prove amounts to showing that a projectivity $g \in \boldsymbol{PSL}(2, K)$ with $g(0) = 0$ and $g(\infty) = \infty$ can be represented as a product of transvections. Such a transformation is generated by a linear transformation with a diagonal matrix, and the equation below displays a decomposition into transvections:

$$\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a-1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -a^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a-a^2 & 1 \end{pmatrix}. \tag{44}$$

This proves the statement for $n = 1$. Suppose that the assertion is already proved for $n - 1$, and take $g \in \boldsymbol{PSL}(n+1) \subset \boldsymbol{PL}_n$. Since the group G generated by the transvections acts transitively on the projective space $\boldsymbol{P}^n$, we may suppose that g has a fixed point, which we choose as the vertex $\boldsymbol{a}_{n+1} = [\mathfrak{a}_{n+1}]$ of a coordinate simplex. In a corresponding basis the matrix B of g then has the form

$$B = \begin{pmatrix} C & \mathfrak{o} \\ \mathfrak{c} & a^{-1} \end{pmatrix},$$

where $\mathfrak{c} = \{\gamma_{n+1}^1, \ldots, \gamma_{n+1}^n\}$ and $\det(C) = a$. It is easy to show that each matrix of the form

$$T(\mathfrak{b}) = \begin{pmatrix} I_n & \mathfrak{o} \\ \mathfrak{b} & 1 \end{pmatrix}$$

generates a transvection in $\boldsymbol{P}^n$. Hence it suffices to show that

$$B_1 = BT(-a\mathfrak{c}) = \begin{pmatrix} C & \mathfrak{o} \\ \mathfrak{o}' & a^{-1} \end{pmatrix}$$

is a product of transvections. Considering a decomposition analogous to (44) one sees that the transformation with the matrix

$$A = \begin{pmatrix} a^{-1} & \mathfrak{o}' & 0 \\ \mathfrak{o} & I_{n-1} & \mathfrak{o} \\ 0 & \mathfrak{o}' & a \end{pmatrix}$$

is a product of four transvections. Multiplying B_1 by A from the right leads to the matrix

$$B_1 A = \begin{pmatrix} C_1 & \mathfrak{o} \\ \mathfrak{o}' & 1 \end{pmatrix}$$

with $\det(C_1) = 1$. By the induction hypothesis we can represent C_1 as the product of finitely many transvection matrices T_α of order n. Extending these as

$$\hat{T}_\alpha = \begin{pmatrix} T_\alpha & \mathfrak{o} \\ \mathfrak{o}' & 1 \end{pmatrix}$$

to matrices of order $n+1$ we obtain a representation for $B_1 A$ as a finite product of transvection matrices, which completes the proof. □

>From Proposition 4 we easily obtain the following statement concerning generators for the projective groups in the cases of most interest for us, $K = \mathbf{R}, \mathbf{C}$:

Corollary 5. *The projective groups $\boldsymbol{PL}_{2k}(\mathbf{R}), \boldsymbol{PL}_n(\mathbf{C})$ are generated by their transvections. To obtain a system of generators for the groups $\boldsymbol{PL}_{2k+1}(\mathbf{R})$ of real projectivities in odd dimensions it suffices to take all transvections and add a single transformation reversing the orientation with determinant -1.*

P r o o f. For $K = \mathbf{R}, \mathbf{C}$ the assertion immediately follows from Formulas (5)–(7) together with Proposition 4. □

2.2 Vector Spaces with Scalar Product

In the preceding section we determined the isotropy groups of biforms for the natural action of the linear groups and at the same time introduced the so-called classical groups. The geometries of these groups have some properties in common which we want to study in this section; below we will discuss some of the geometries important for later applications in detail.

2.2.1 Vector, Projective, and Affine Geometries

A *linear group* is understood to be a subgroup G of the general linear group $\boldsymbol{GL}(\boldsymbol{V})$ of a finite-dimensional vector space $\boldsymbol{V}$ over a skew field K, or, having fixed a basis, the corresponding matrix group isomorphic to it. Its *vector geometry* is determined by the transformation group $[G, \boldsymbol{V}]$, where, as a subgroup of $\boldsymbol{GL}(\boldsymbol{V})$, G acts linearly on $\boldsymbol{V}$. The contents of this geometry is formed by the G-invariant properties of the space $\boldsymbol{V}$ itself and of all the spaces functorially related to it: set products, tensor spaces etc. In particular, this includes the *projective geometry* $[G, \mathfrak{P}(\boldsymbol{V})]$, where the action of G is defined by the projectivities corresponding to the elements of G. To study the *affine geometry* of the group G we start from Formula (1.5.8) describing the general affine group of the affine space; taking not any element $B \in \boldsymbol{GL}(n, K)$ but only those from the linear group $G \subset \boldsymbol{GL}(n, K)$ uniquely determines a subgroup $GA \subset \mathfrak{A}(n)$ of the affine group. The transformation group $[GA, \boldsymbol{A}^n]$ for this affine action of GA over the affine space $\boldsymbol{A}^n$ then determines the affine geometry corresponding to $[G, \boldsymbol{V}]$. The group GA may also be defined as the subgroup of the affine group $\mathfrak{A}(n)$ generated by the union of the group of translations $\boldsymbol{T}(\boldsymbol{A}^n)$ with the group $G_{|\boldsymbol{A}}$; in the case of a field K it is the semi-direct product of $\boldsymbol{T}(\boldsymbol{A}^n)$ with the group $G_{|\boldsymbol{A}}$, cf. Exercise 1.5.1.

$>$From these rather general observations together with what we discussed in Sections I.4–6 as well as in the preceding chapter it is obvious that the vectorial point of view, at least methodically, dominates both the other two; thus we will emphasize it in this section and fix a corresponding terminology.

Definition 1. A *vector space with scalar product* is a pair $[\boldsymbol{V}, \langle , \rangle]$ with the following properties:

1. $\boldsymbol{V}$ is a finite-dimensional right vector space over the skew field K.
2. The *scalar product* $\langle , \rangle$ is a σ-biform over $\boldsymbol{V}$, $\langle \mathfrak{x}, \mathfrak{y} \rangle = b(\mathfrak{x}, \mathfrak{y})$, where σ is an involutive anti-automorphism of K.
3. The scalar product is non-degenerate, i.e., for each $\mathfrak{x} \in \boldsymbol{V}$, $\mathfrak{x} \neq \mathfrak{o}$, the linear form $\langle \mathfrak{x}, \cdot \rangle$ over $\boldsymbol{V}$ is different from 0.
4. The scalar product may be an alternating or a symmetric bilinear form, or σ-Hermitean.
5. If $\langle , \rangle$ is not alternating, then suppose that $\operatorname{char} K \neq 2$.

□

Let $G \subset \boldsymbol{GL}(\boldsymbol{V})$ denote the *group of automorphisms of the vector space with scalar product*, i.e. the linear automorphisms g satisfying $\langle g\mathfrak{x}, g\mathfrak{y} \rangle = \langle \mathfrak{x}, \mathfrak{y} \rangle$ for all $\mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}$. As described in the introductory paragraph, this group G then also determines the associated projective and affine geometries, respectively. According to this definition we will speak of a *projective* or *affine space with scalar product* whenever a scalar product is distinguished in the underlying vector space. Both, the *geometry* and the underlying *space* are called *symplectic* if the scalar product is alternating, otherwise they are called *Hermitean.*

If K is a field and the scalar product is bilinear and symmetric, then the geometry is called *orthogonal*; hence this is a special case of a Hermitean geometry. Talking about the *index* of a Hermitean form always refers to the cases $K = \mathbf{R}, \mathbf{C}$ or $\mathbf{H}$; this notion is not to be confused with that of the index of a space with scalar product (see Definition 3 below). For $K = \mathbf{R}$ and a scalar product with positive index l, the corresponding geometry is called *pseudo-Euclidean*, for $K = \mathbf{C}, \mathbf{H}$ and $l > 0$ these spaces are called *pseudo-unitary*. The *Euclidean* and *unitary spaces* discussed in Chapter I.6 are included as the case $l = 0$.

2.2.2 Subspaces

Throughout this section we will suppose that $\boldsymbol{V}$ is a vector space with scalar product $\langle , \rangle$. The notions *isotropic* and *totally isotropic* introduced for subspaces $\boldsymbol{W} \subset \boldsymbol{V}$ in Definition 1.9.3 will always refer to the scalar product $\langle , \rangle$. In particular, a vector $\mathfrak{x} \in \boldsymbol{V}$ is called *isotropic* if $\mathfrak{x} \neq \mathfrak{o}$ and $\langle \mathfrak{x}, \mathfrak{x} \rangle = 0$. Likewise, *orthogonality* is always taken with respect to this scalar product: The vectors $\mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}$ are called *orthogonal* if $\langle \mathfrak{x}, \mathfrak{y} \rangle = 0$; subspaces $\boldsymbol{U}, \boldsymbol{W} \subset \boldsymbol{V}$ are *orthogonal* if their vectors $\mathfrak{x} \in \boldsymbol{U}, \mathfrak{y} \in \boldsymbol{W}$ are orthogonal (cf. Definition I.6.1.1). Definition I.6.2.16 of the subspace $\boldsymbol{W}^\perp \subset \boldsymbol{V}$ orthogonal to a subspace $\boldsymbol{W} \subset \boldsymbol{V}$ applies:

$$\boldsymbol{W}^\perp := \{\mathfrak{y} \in \boldsymbol{V} | \langle \mathfrak{x}, \mathfrak{y} \rangle = 0 \text{ for all } \mathfrak{x} \in \boldsymbol{W}\}.$$

We will not, however, call it the orthogonal complement, since, in general, the condition $\boldsymbol{W} \cap \boldsymbol{W}^\perp = \{\mathfrak{o}\}$ is not satisfied. Instead, the following is an immediate consequence of the definitions

Proposition 1. *A subspace $\boldsymbol{W} \subset \boldsymbol{V}$ is isotropic if and only if the intersection $\boldsymbol{W} \cap \boldsymbol{W}^\perp$ is non-trivial; together with $\boldsymbol{W}$ its orthogonal subspace $\boldsymbol{W}^\perp$ is also isotropic.* □

Corollary 2. *If $\boldsymbol{W} \subset \boldsymbol{V}$ is not isotropic, then $[\boldsymbol{W}, \langle , \rangle | \boldsymbol{W} \times \boldsymbol{W}]$ is a vector space of the same type as $\boldsymbol{V}$, i.e. alternating or σ-Hermitean. If $\langle , \rangle$ has index l on $\boldsymbol{V}$, then $\langle , \rangle | \boldsymbol{W} \times \boldsymbol{W}$ has index $l_{\boldsymbol{W}} \leq l$.* □

Exercise 1. Let $\dim \boldsymbol{V} = n < \infty$. Prove Proposition 1. Show, moreover, that $\boldsymbol{W}$ is isotropic if and only if $\boldsymbol{W} + \boldsymbol{W}^\perp \neq \boldsymbol{V}$, and that the following properties, familiar from the theory of unitary vector spaces (Chapter I.6), also hold for vector spaces with scalar product:

$$\dim \boldsymbol{W} + \dim \boldsymbol{W}^\perp = \dim \boldsymbol{V}, \tag{1}$$

$$(\boldsymbol{W}^\perp)^\perp = \boldsymbol{W}, \tag{2}$$

$$(\boldsymbol{U} + \boldsymbol{W})^\perp = \boldsymbol{U}^\perp \cap \boldsymbol{W}^\perp, \tag{3}$$

$$(\boldsymbol{U} \cap \boldsymbol{W})^\perp = \boldsymbol{U}^\perp + \boldsymbol{W}^\perp, \tag{4}$$

$$\boldsymbol{U} \subset \boldsymbol{W} \text{ implies } \boldsymbol{W}^\perp \subset \boldsymbol{U}^\perp. \tag{5}$$

2.2.3 E. Witt's Theorem

Now we want to formulate a fundamental proposition by E. Witt concerning the extension of isomorphisms from subspaces to isomorphisms of vector spaces with scalar product. To this end we first define the notion of a morphism for the class of pairs $[\boldsymbol{W}, b]$ consisting of a vector space over a fixed skew field K together with a σ-biform b, where σ is a fixed anti-automorphism of K. Note that any subspace $\boldsymbol{W} \subset \boldsymbol{V}$ of a space with scalar product b determines such a pair by restriction $[\boldsymbol{W}, b|\boldsymbol{W} \times \boldsymbol{W}]$. Since $b|\boldsymbol{W} \times \boldsymbol{W}$ may be degenerate, $\boldsymbol{W}$ itself not necessarily has to be a vector space with scalar product in the sense of Definition 1.

Definition 2. Let $[\boldsymbol{W}, b]$ and $[\tilde{\boldsymbol{W}}, \tilde{b}]$ be vector spaces with σ-biforms over the same skew field K. A linear map $\varphi : \boldsymbol{W} \to \tilde{\boldsymbol{W}}$ satisfying

$$b(\mathfrak{x}, \mathfrak{y}) = \tilde{b}(\varphi(\mathfrak{x}), \varphi(\mathfrak{y})), \qquad (\mathfrak{x}, \mathfrak{y} \in \boldsymbol{W}) \tag{6}$$

is called a *morphism* of these pairs and an *isomorphism* if φ is bijective. If $\boldsymbol{W}, \tilde{\boldsymbol{W}}$ are subspaces of vector spaces with scalar product, then we will call them *isomorphic* if there exists an isomorphism

$$\varphi : [\boldsymbol{W}, \langle , \rangle | \boldsymbol{W} \times \boldsymbol{W}] \to [\tilde{\boldsymbol{W}}, \langle , \rangle | \tilde{\boldsymbol{W}} \times \tilde{\boldsymbol{W}}].$$

□

Obviously, for fixed σ and K the class of pairs $[\boldsymbol{W}, b]$, b alternating or σ-Hermitean, respectively, with the morphisms defined above form a category.

Exercise 2. In the notations and assumptions of Definition 2, assume that the map $\varphi : \boldsymbol{W} \to \tilde{\boldsymbol{W}}$ is a morphism, and that b is non-degenerate. Prove that φ is injective, and $[\boldsymbol{W}, b]$ is isomorphic to $[\varphi(\boldsymbol{W}), \tilde{b}|\varphi(\boldsymbol{W}) \times \varphi(\boldsymbol{W})]$.

In the following subsections we will prove *E. Witt's Theorem* :

Proposition 3. (E. Witt). *Let $\boldsymbol{V}, \tilde{\boldsymbol{V}}$ be finite-dimensional vector spaces with scalar products over a skew field K, which are isomorphic. Let, moreover, $\varphi : \boldsymbol{W} \to \tilde{\boldsymbol{W}}$ be an isomorphism from the subspace $\boldsymbol{W} \subset \boldsymbol{V}$ onto the subspace $\tilde{\boldsymbol{W}} \subset \tilde{\boldsymbol{V}}$. Then there is an isomorphism $\psi : \boldsymbol{V} \to \tilde{\boldsymbol{V}}$ of the vector spaces with scalar product such that $\varphi = \psi|\boldsymbol{W}$.*

2.2.4 Properties of Isotropic Subspaces

To prove this proposition we will need a few facts concerning isotropic subspaces that will also frequently be applied elsewhere.

Lemma 4. *Let $\mathfrak{v} \in \boldsymbol{V} \setminus \{\mathfrak{o}\}$ be an isotropic vector in the vector space $\boldsymbol{V}$ with scalar product $\langle , \rangle$. Then there always is an isotropic vector $\hat{\mathfrak{v}} \in \boldsymbol{V}$ such that $\langle \mathfrak{v}, \hat{\mathfrak{v}} \rangle = 1$; $\mathfrak{v}, \hat{\mathfrak{v}}$ are linearly independent.*

P r o o f. Since $\mathfrak{v} \neq \mathfrak{o}$ and the scalar product is non-degenerate, we find a vector $\mathfrak{w} \in \boldsymbol{V}$ with $\langle \mathfrak{v}, \mathfrak{w} \rangle \neq 0$. As $\langle \mathfrak{v}, \mathfrak{w} \rangle$ is linear in $\mathfrak{w}$, we may normalize $\mathfrak{w}$ to have $\langle \mathfrak{v}, \mathfrak{w} \rangle = 1$. If the scalar product is σ-Hermitean, we try to find $\hat{\mathfrak{v}}$ in the form $\hat{\mathfrak{v}} = \mathfrak{v}\xi + \mathfrak{w}$, $\xi \in K$. Now for σ-Hermitean scalar products we have $\sigma(\langle \mathfrak{w}, \mathfrak{w} \rangle) = \langle \mathfrak{w}, \mathfrak{w} \rangle$, so the vector $\hat{\mathfrak{v}}$ with $\xi = -\langle \mathfrak{w}, \mathfrak{w} \rangle / 2$ is a solution; condition 5 in Definition 1 guarantees its existence. In a symplectic space each vector $\mathfrak{w} \neq \mathfrak{o}$ is isotropic; so we may set $\hat{\mathfrak{v}} := \mathfrak{w}$. If $\mathfrak{v}$ and $\hat{\mathfrak{v}}$ were linearly dependent, then $\hat{\mathfrak{v}} = \mathfrak{v}\lambda$, hence $\langle \mathfrak{v}, \hat{\mathfrak{v}} \rangle = \langle \mathfrak{v}, \mathfrak{v} \rangle \lambda = 0$ contradicting $\langle \mathfrak{v}, \hat{\mathfrak{v}} \rangle = \langle \mathfrak{v}, \mathfrak{w} \rangle = 1$. □

Exercise 3. Let $\boldsymbol{V}^n$, $n = 2m$, be an n-dimensional symplectic vector space. Show that for every $k \leq m$ there is a totally isotropic subspace of dimension k.

Proposition 5. *Let $\boldsymbol{W}^m \subset \boldsymbol{V}^n$ be a totally isotropic subspace of the vector space $\boldsymbol{V}$ with scalar product, and let $(\mathfrak{b}_\alpha)$, $\alpha = 1, \ldots, m$, be a basis of $\boldsymbol{W}^m$. Then there is a sequence $(\hat{\mathfrak{b}}_\alpha)$, $\alpha = 1, \ldots, m$, of isotropic vectors in the space $\boldsymbol{V}$ such that*

$$\langle \mathfrak{b}_\beta, \hat{\mathfrak{b}}_\alpha \rangle = \delta_{\alpha\beta}, \ \langle \hat{\mathfrak{b}}_\beta, \hat{\mathfrak{b}}_\alpha \rangle = 0, \qquad (\alpha, \beta = 1, \ldots, m),$$

and, moreover, the sequence

$$\mathfrak{b}_1, \hat{\mathfrak{b}}_1, \ldots, \mathfrak{b}_m, \hat{\mathfrak{b}}_m \tag{7}$$

is linearly independent.

P r o o f. For $m = 1$ the assertion follows from Lemma 4. Suppose that Proposition 5 is already proved for all totally isotropic subspaces $\boldsymbol{W}_o$ with $\dim \boldsymbol{W}_o < m$ and consider the linear span $\boldsymbol{W}_o := \mathfrak{L}(\mathfrak{b}_1, \ldots, \mathfrak{b}_{m-1})$. By the induction hypothesis we find corresponding vectors $\tilde{\mathfrak{b}}_1, \ldots, \tilde{\mathfrak{b}}_{m-1}$ with the properties stated in Proposition 5 for $m-1$ instead of m as well as $\tilde{\mathfrak{b}}_\alpha$ instead of $\hat{\mathfrak{b}}_\alpha$. Set

$$\boldsymbol{U}^{2m-2} := \mathfrak{L}(\mathfrak{b}_1, \ldots, \mathfrak{b}_{m-1}, \tilde{\mathfrak{b}}_1, \ldots, \tilde{\mathfrak{b}}_{m-1}).$$

Looking at the scalar products of these basis vectors it is obvious that $\boldsymbol{U}$ is not isotropic. Hence $\boldsymbol{U}^\perp$ is also not isotropic, and, moreover,

$$\boldsymbol{V} = \boldsymbol{U} \oplus \boldsymbol{U}^\perp. \tag{8}$$

>From (8) we obtain the decomposition $\mathfrak{b}_m = \mathfrak{u}_m + \mathfrak{v}_m$, where $\mathfrak{u}_m \in \boldsymbol{U}$ has the representation

$$\mathfrak{u}_m = \sum_{\alpha=1}^{m-1} \mathfrak{b}_\alpha \lambda_\alpha + \sum_{\alpha=1}^{m-1} \tilde{\mathfrak{b}}_\alpha \tilde{\lambda}_\alpha.$$

Forming the scalar product of this equation from the left with $\mathfrak{b}_\beta$, the relation $\langle \mathfrak{b}_\beta, \mathfrak{b}_m \rangle = \langle \mathfrak{b}_\beta, \mathfrak{u}_m \rangle = 0$ leads to the equality $\tilde{\lambda}_\beta = 0$, hence

$$\mathfrak{u}_m = \sum_{\alpha=1}^{m-1} \mathfrak{b}_\alpha \lambda_\alpha. \tag{9}$$

Thus we have $\mathfrak{v}_m \neq \mathfrak{o}$ and $\langle \mathfrak{b}_m, \mathfrak{b}_m \rangle = \langle \mathfrak{v}_m, \mathfrak{v}_m \rangle = 0$; in fact, $\mathfrak{b}_m$ and $\mathfrak{u}_m$ are isotropic as elements of $\boldsymbol{W}$. Relation (9) together with the linear independence of $(\mathfrak{b}_\alpha)$, $\alpha = 1, \ldots, m$, implies that $\mathfrak{v}_m$ is different from zero and hence also isotropic. According to Lemma 4 there is an isotropic vector $\hat{\mathfrak{b}}_m \in \boldsymbol{U}^\perp$ with $\langle \mathfrak{v}_m, \hat{\mathfrak{b}}_m \rangle = 1$, since $\boldsymbol{U}^\perp$ it not isotropic. Because of $\langle \mathfrak{b}_m, \hat{\mathfrak{b}}_m \rangle = \langle \mathfrak{v}_m, \hat{\mathfrak{b}}_m \rangle = 1$, the space $\boldsymbol{A} := \mathfrak{L}(\mathfrak{b}_m, \hat{\mathfrak{b}}_m)$ is not isotropic, and hence the same holds for $\boldsymbol{A}^\perp$. As $\mathfrak{b}_1, \ldots, \mathfrak{b}_{m-1}$ lie in $\boldsymbol{A}^\perp$, by the induction hypothesis there are vectors $\hat{\mathfrak{b}}_1, \ldots, \hat{\mathfrak{b}}_{m-1} \in \boldsymbol{A}^\perp$ with the properties stated in Proposition 5 (for $m-1$ instead of m). It is easily verified that the sequence (7) formed by these vectors together with $\mathfrak{b}_m, \hat{\mathfrak{b}}_m$ meets all the requirements. □

Corollary 6. *Let $\boldsymbol{V}^n$ be a vector space with scalar product, and let $\boldsymbol{W}^k \subset \boldsymbol{V}^n$ be a totally isotropic subspace. Then $k \leq n/2$.* □

Exercise 4. Let $\boldsymbol{V}^n$ be a real pseudo-Euclidean or a complex pseudo-unitary vector space with scalar product of index $l \leq n/2$. Prove that for each $k \leq l$ there is a totally isotropic subspace of dimension k in $\boldsymbol{V}$, and that there is no totally isotropic subspace of any dimension $k > l$.

Definition 3. Let $\boldsymbol{V}^n$ be a vector space with scalar product. The maximum of the dimensions of totally isotropic subspaces $\boldsymbol{W} \subset \boldsymbol{V}^n$ is called *the index of $\boldsymbol{V}^n$*. □

This definition is not in accordance with that of the index for a symmetric bilinear form (in the case $K = \mathbf{R}$) or a Hermitean form ($K = \mathbf{C}, \mathbf{H}$), cf. Definition I.5.9.7 and Section 1.9. If the form is non-degenerate, then its index l_o is related to the index l of a vector space with scalar product just defined via

$$l = \min(l_o, n - l_o) \tag{10}$$

(cf. Exercise 4). The index introduced in Definition 3 makes sense for an arbitrary skew field K and any scalar product. Exercise 3, e.g., implies that every symplectic vector space $\boldsymbol{V}^{2m}$ has index m.

Let $\boldsymbol{U} \subset \boldsymbol{V}$ be an arbitrary subspace. Then by Definition 1.9.2 the subspace $\boldsymbol{U} \cap \boldsymbol{U}^\perp$ is the defect subspace of $\boldsymbol{U}$. Hence for the *defect*

$$\operatorname{def} \boldsymbol{U} := \operatorname{def}(\langle , \rangle | \boldsymbol{U} \times \boldsymbol{U})$$

we have

$$\operatorname{def} \boldsymbol{U} = \operatorname{def} \boldsymbol{U}^\perp = \dim(\boldsymbol{U} \cap \boldsymbol{U}^\perp). \tag{11}$$

Corollary 7. *Let $\boldsymbol{U} \subset \boldsymbol{V}$ be an arbitrary subspace. Then the dimension of every minimal non-isotropic subspace $\boldsymbol{W}$ containing $\boldsymbol{U}$ satisfies*

$$\dim \boldsymbol{W} = \dim \boldsymbol{U} + \operatorname{def} \boldsymbol{U}. \tag{12}$$

Proof. First, the following is straightforward.

Lemma 8. *If $\boldsymbol{U}_1$ is a subspace of $\boldsymbol{U}$ complementary to the defect subspace $\boldsymbol{U}_o$, then $\boldsymbol{U} = \boldsymbol{U}_o \oplus \boldsymbol{U}_1$ is a decomposition of $\boldsymbol{U}$ into orthogonal subspaces, and $\boldsymbol{U}_1$ it not isotropic.* □

Therefore, the subspace $\boldsymbol{U}_1^\perp \subset \boldsymbol{V}$ is not isotropic as well, and $\boldsymbol{U}_o \subset \boldsymbol{U}_1^\perp$. According to Proposition 5 there is a non-isotropic subspace $\boldsymbol{W}_1$ with $\boldsymbol{U}_o \subset \boldsymbol{W}_1 \subset \boldsymbol{U}_1^\perp$, namely the subspace spanned by the sequence (7) with $m = \operatorname{def} \boldsymbol{U}$, $\mathfrak{b}_1, \ldots, \mathfrak{b}_m$ a basis of $\boldsymbol{U}_o$. Hence $\boldsymbol{W} := \boldsymbol{W}_1 \oplus \boldsymbol{U}_1 \supset \boldsymbol{U}$ is a non-isotropic subspace satisfying (12). By Proposition 5 we may perform this construction within each non-isotropic subspace containing $\boldsymbol{U}$. In the case of a minimal subspace $\boldsymbol{W}$ containing U the result has to be $\boldsymbol{W}$ itself, proving the assertion. □

2.2.5 The Proof of E. Witt's Theorem

Now we turn to the proof of E. Witt's Theorem and start by proving

Lemma 9. *Let $\boldsymbol{V}, \tilde{\boldsymbol{V}}$ be vector spaces with scalar products over K, both symplectic or both σ-Hermitean. Moreover, let $\boldsymbol{U}$ be an arbitrary subspace, and let $\boldsymbol{W} \supset \boldsymbol{U}$ be a minimal, non-isotropic subspace containing $\boldsymbol{U}$. Then every injective morphism $\varphi : \boldsymbol{U} \to \tilde{\boldsymbol{V}}$ can be extended to an injective morphism $\psi : \boldsymbol{W} \to \tilde{\boldsymbol{V}}$.*

P r o o f. We start from the decomposition constructed in the proof of Corollary 7:

$$\boldsymbol{W} = \boldsymbol{U}_o \oplus \hat{\boldsymbol{U}}_o \oplus \boldsymbol{U}_1,$$

where we have to set

$$\boldsymbol{W}_1 = \boldsymbol{U}_o \oplus \hat{\boldsymbol{U}}_o \text{ and } \hat{\boldsymbol{U}}_o = \mathfrak{L}(\hat{\mathfrak{b}}_1, \ldots, \hat{\mathfrak{b}}_m),\ m = \operatorname{def} \boldsymbol{U}.$$

The image

$$\tilde{\boldsymbol{U}} := \varphi(\boldsymbol{U}) = \varphi(\boldsymbol{U}_o) \oplus \varphi(\boldsymbol{U}_1)$$

is isomorphic to $\boldsymbol{U}$, since φ is injective. Consequently,

$$(\mathfrak{c}_\alpha) := (\varphi(\mathfrak{b}_\alpha)),\ \alpha = 1, \ldots, m,$$

is a basis for the defect subspace $\tilde{\boldsymbol{U}}_o := \varphi(\boldsymbol{U}_o)$ of $\tilde{\boldsymbol{U}}$, and $\tilde{\boldsymbol{U}}_1 := \varphi(\boldsymbol{U}_1)$ is a non-isotropic subspace of $\tilde{\boldsymbol{U}}$ complementary and orthogonal to $\tilde{\boldsymbol{U}}_o$. Hence $\tilde{\boldsymbol{U}}_1^\perp \subset \tilde{\boldsymbol{V}}$ is not isotropic, and $\tilde{\boldsymbol{U}}_o$ is a totally isotropic subspace of $\tilde{\boldsymbol{U}}_1^\perp$. By Proposition 5 we find a vector sequence $\mathfrak{c}_1, \hat{\mathfrak{c}}_1, \ldots, \mathfrak{c}_m, \hat{\mathfrak{c}}_m$ in $\tilde{\boldsymbol{V}}$ analogous to (7) with corresponding properties. It is easy to see that φ is extended to an injective morphism $\psi : \boldsymbol{W} \to \tilde{\boldsymbol{V}}$ by setting

$$\psi | \boldsymbol{U} := \varphi,\ \psi(\hat{\mathfrak{b}}_\alpha) := \hat{\mathfrak{c}}_\alpha,\ \alpha = 1, \ldots, m.$$

If both spaces are symplectic, then

$$\langle \hat{\mathfrak{c}}_\alpha, \mathfrak{c}_\alpha \rangle = -\langle \mathfrak{c}_\alpha, \hat{\mathfrak{c}}_\alpha \rangle = -1 = \langle \hat{\mathfrak{b}}_\alpha, \mathfrak{b}_\alpha \rangle,$$

whereas, because of $\sigma(1) = 1$, in the σ-Hermitean case

$$\langle \hat{\mathfrak{c}}_\alpha, \mathfrak{c}_\alpha \rangle = \sigma(\langle \mathfrak{c}_\alpha, \hat{\mathfrak{c}}_\alpha \rangle) = 1 = \langle \hat{\mathfrak{b}}_\alpha, \mathfrak{b}_\alpha \rangle.$$

□

Note that, in Lemma 9, we did not suppose isomorphy of $\boldsymbol{V}$ and $\tilde{\boldsymbol{V}}$. Towards the end of the proof we implicitly used the simple construction described in the following Exercise:

Exercise 5. Let $\boldsymbol{U} = \bigoplus_{i=1}^r \boldsymbol{U}_i$ be a direct sum of pairwise orthogonal subspaces $\boldsymbol{U}_i \subset \boldsymbol{V}$. Let, moreover, $\varphi_i : \boldsymbol{U}_i \to \tilde{\boldsymbol{V}}$ be injective morphisms, for which $\tilde{\boldsymbol{U}} := \bigoplus_{i=1}^r \varphi_i(\boldsymbol{U}_i)$ is a direct sum of pairwise orthogonal subspaces in $\tilde{\boldsymbol{V}}$. Show that, under these conditions,

$$\varphi(\sum_{i=1}^r \mathfrak{x}_i) := \sum_{i=1}^r \varphi_i(\mathfrak{x}_i), \qquad \mathfrak{x}_i \in \boldsymbol{U}_i,$$

defines an isomorphism from $\boldsymbol{U}$ onto $\tilde{\boldsymbol{U}}$.

First we will now prove E. Witt's Theorem for symplectic vector spaces. By Lemma 9 it suffices to prove the statement for non-isotropic subspaces; note that $\boldsymbol{W}$ and $\psi(\boldsymbol{W})$ are isomorphic. Since $\boldsymbol{V}$ and $\tilde{\boldsymbol{V}}$ may be supposed to be isomorphic, without loss of generality we can assume $\boldsymbol{V} = \tilde{\boldsymbol{V}}$. In fact, $\rho : \boldsymbol{V} \to \tilde{\boldsymbol{V}}$ is an isomorphism, and hence it suffices to find an automorphism α of $\boldsymbol{V}$ mapping $\boldsymbol{W}$ into $\rho^{-1}(\tilde{\boldsymbol{W}})$; then $\psi = \rho \circ \alpha$ is the isomorphism we need. With these observations the symplectic case is straightforward: For any non-isotropic subspace $\boldsymbol{W} \subset \boldsymbol{V}$ of the symplectic vector space $\boldsymbol{V}$, $\tilde{\boldsymbol{W}} = \varphi(\boldsymbol{W})$ is also not isotropic. According to Proposition 1 the same holds for the orthogonal subspaces $\boldsymbol{W}^\perp$ and $\tilde{\boldsymbol{W}}^\perp$. Since the dimensions of both are equal, there is an isomorphism $\varphi^\perp : \boldsymbol{W}^\perp \to \tilde{\boldsymbol{W}}^\perp$ (cf. Lemma 1.8.4 and Proposition 1.8.5). Then by Exercise 5 the map $\psi = \varphi \oplus \varphi^\perp$ composed of φ and $\varphi^\perp$ provides the automorphism $\psi \in \boldsymbol{Sp}(\boldsymbol{V})$. Obviously, one can prove E. Witt's Theorem using the same method for all the scalar products classified in Sections 2.1.4 to 2.1.6 ($K = \mathbf{R}, \mathbf{C}, \mathbf{H}$), looking at the normal forms and, if necessary, applying the Theorem of Inertia for the associated σ-biforms.

With this done we next prove E. Witt's Theorem in the general case. As above we suppose $\boldsymbol{V} = \tilde{\boldsymbol{V}}$ and proceed by induction on $m := \dim \boldsymbol{W}$. For $m = 0$ the assertion is trivial. So assume that the proposition holds for all dimensions less than m. In $\boldsymbol{W}$ we choose any subspace $\boldsymbol{U} \subset \boldsymbol{W}$ of dimension $\dim \boldsymbol{U} = m - 1$. Then $\varphi_1 := \varphi|\boldsymbol{U}$ is an isomorphism onto $\tilde{\boldsymbol{U}} := \varphi(\boldsymbol{U})$, which we can extend to an automorphism $\psi_1 : \boldsymbol{V} \to \boldsymbol{V}$ by the induction hypothesis. Set $\boldsymbol{W}_1 := \psi_1^{-1}(\tilde{\boldsymbol{W}})$. Then we have $\boldsymbol{U} = \psi_1^{-1}(\tilde{\boldsymbol{U}})$, and $\varphi_o := \psi_1^{-1} \circ \varphi : \boldsymbol{W} \to \boldsymbol{W}_1$ is an isomorphism leaving every element of the subspace $\boldsymbol{U} \subset \boldsymbol{W}$ fixed. If we succeed in finding an extension of φ_o to an automorphism ψ_o of

$\boldsymbol{V}$ under the additional assumption that there is a subspace $\boldsymbol{U}^{m-1} \subset \boldsymbol{W}^m$ satisfying $\varphi_o|\boldsymbol{U} = \mathrm{id}_{\boldsymbol{U}}$, then the proof is complete. In fact, we can then extend $\varphi_o = \psi_1^{-1} \circ \varphi$ to an automorphism ψ_o of $\boldsymbol{V}$, for which $\psi_o(\boldsymbol{W}) = \boldsymbol{W}_1$ holds. So $\psi = \psi_1 \circ \psi_o$ is an automorphism of $\boldsymbol{V}$ satisfying $\psi|\boldsymbol{W} = \varphi$. The next step is to prove an auxiliary statement generalizing Exercise 5:

Lemma 10. *Let $\boldsymbol{V}$ be a vector space with scalar product, let $\boldsymbol{W}_i \subset \boldsymbol{V}$, $i = 1, 2$, be subspaces, and let $u_i : \boldsymbol{W}_i \to \boldsymbol{V}$ be injective morphisms. Suppose, moreover, that*

$$\boldsymbol{W}_1 \cap \boldsymbol{W}_2 = \{\boldsymbol{o}\} \text{ and } u_1(\boldsymbol{W}_1) \cap u_2(\boldsymbol{W}_2) = \{\boldsymbol{o}\}.$$

Finally, assume that for all $\mathfrak{x}_i \in \boldsymbol{W}_i$

$$\langle u_1(\mathfrak{x}_1), u_2(\mathfrak{x}_2)\rangle = \langle \mathfrak{x}_1, \mathfrak{x}_2\rangle. \tag{13}$$

Then

$$u(\mathfrak{x}_1 + \mathfrak{x}_2) := u_1(\mathfrak{x}_1) + u_2(\mathfrak{x}_2), \qquad \mathfrak{x}_i \in \boldsymbol{W}_i, \tag{14}$$

is an injective morphism from $\boldsymbol{W}_1 \oplus \boldsymbol{W}_2$ into $\boldsymbol{V}$.

Proof. Because of $\boldsymbol{W}_1 \cap \boldsymbol{W}_2 = \{\boldsymbol{o}\}$, Definition (14) is correct. It is easy to verify that for all $\mathfrak{x}, \mathfrak{y} \in \boldsymbol{W}_1 \oplus \boldsymbol{W}_2$ the condition $\langle u(\mathfrak{x}), u(\mathfrak{y})\rangle = \langle \mathfrak{x}, \mathfrak{y}\rangle$ is satisfied. Since $u : \boldsymbol{W}_1 \oplus \boldsymbol{W}_2 \to u_1(\boldsymbol{W}_1) \oplus u_2(\boldsymbol{W}_2)$ is surjective, and as the dimensions of both spaces coincide by $u_1(\boldsymbol{W}_1) \cap u_2(\boldsymbol{W}_2) = \{\boldsymbol{o}\}$, u is injective as well. $\square$

Let now $\boldsymbol{U}^{m-1} \subset \boldsymbol{W}^m$ be a subspace with $\varphi|\boldsymbol{U} = \mathrm{id}_{\boldsymbol{U}}$, and let $\mathfrak{a} \in \boldsymbol{W} \setminus \boldsymbol{U}$ be a complementary vector. Consider the map $\alpha := \varphi - \iota_{\boldsymbol{W}} : \boldsymbol{W} \to \boldsymbol{V}$, where $\iota_{\boldsymbol{W}}$ denotes the embedding of $\boldsymbol{W}$ into $\boldsymbol{V}$. Since we can represent any vector $\mathfrak{w} \in \boldsymbol{W}$ as $\mathfrak{w} = \mathfrak{u} + \mathfrak{a}\xi$ with $\mathfrak{u} \in \boldsymbol{U}$, $\xi \in K$, the image $\boldsymbol{H} := \alpha(\boldsymbol{W})$ is equal to $\mathfrak{L}(\varphi(\mathfrak{a}) - \mathfrak{a})$, i.e. $\dim \boldsymbol{H} \leq 1$. If $\dim \boldsymbol{H} = 0$, then $\varphi(\mathfrak{a}) = \mathfrak{a}$, hence $\varphi = \mathrm{id}_{\boldsymbol{W}}$, and $\psi = \mathrm{id}_{\boldsymbol{V}}$ is a suitable extension. Let then $\dim \boldsymbol{H} = 1$. We consider $\boldsymbol{H}^\perp$ and show that $\boldsymbol{U} \subset \boldsymbol{H}^\perp$. In fact,

$$\langle \varphi(\mathfrak{x}) - \mathfrak{x}, \mathfrak{y}\rangle = \langle \varphi(\mathfrak{x}), \mathfrak{y}\rangle - \langle \mathfrak{x}, \mathfrak{y}\rangle = \langle \varphi(\mathfrak{x}), \mathfrak{y} - \varphi(\mathfrak{y})\rangle \tag{15}$$

for $\mathfrak{x}, \mathfrak{y} \in \boldsymbol{W}$. Inserting here $\mathfrak{x} = \mathfrak{a}$ and $\mathfrak{y} \in \boldsymbol{U}$ we conclude $\langle \varphi(\mathfrak{a}) - \mathfrak{a}, \mathfrak{y}\rangle = 0$. Let now $\boldsymbol{W}_2 \subset \boldsymbol{H}^\perp$ be a subspace satisfying $\boldsymbol{W} \cap \boldsymbol{W}_2 = \{\mathfrak{o}\}$ and $\tilde{\boldsymbol{W}} \cap \boldsymbol{W}_2 = \{\mathfrak{o}\}$. We show that taking $u_1 = \varphi$, $\boldsymbol{W}_1 = \boldsymbol{W}$, $u_2 = \iota_{\boldsymbol{W}_2}$ all assumptions of Lemma 10 are satisfied: Indeed, from $u_1(\mathfrak{x}_1) - \mathfrak{x}_1 \in \boldsymbol{H}$ and $\mathfrak{x}_2 \in \boldsymbol{H}^\perp$ we obtain for the scalar product $\langle u_1(\mathfrak{x}_1) - \mathfrak{x}_1, \mathfrak{x}_2\rangle = 0$; this implies (13). Hence we can extend φ to an injective morphism $\varphi_1 : \boldsymbol{W} \oplus \boldsymbol{W}_2 \to \boldsymbol{V}$. In order to apply this construction we will distinguish two cases: 1. $\boldsymbol{W}$ is not contained in $\boldsymbol{H}^\perp$. Then by (15) $\varphi(\boldsymbol{W}) = \tilde{\boldsymbol{W}}$ does also not lie in $\boldsymbol{H}^\perp$. Since $\boldsymbol{U} \subset \boldsymbol{H}^\perp$ has dimension $m - 1$, for every subspace $\boldsymbol{W}_2 \subset \boldsymbol{H}^\perp$ complementary to $\boldsymbol{U}$ we have

$$\dim \boldsymbol{W}_2 = n - m, \qquad n = \dim \boldsymbol{V}.$$

Similarly, $\tilde{\boldsymbol{W}} \cap \boldsymbol{H}^{\perp} = \boldsymbol{U}$, i.e. $\tilde{\boldsymbol{W}} \cap \boldsymbol{W}_2 = \{\mathfrak{o}\}$. Hence $\boldsymbol{W} \oplus \boldsymbol{W}_2 = \boldsymbol{V}$, and by Lemma 10 we obtain an extension of φ to an automorphism ψ of $\boldsymbol{V}$. – 2. Consider now the case $\boldsymbol{W} \subset \boldsymbol{H}^{\perp}$. Then from (15) we conclude $\tilde{\boldsymbol{W}} \subset \boldsymbol{H}^{\perp}$. First we will show that φ can be extended to an injective morphism $\hat{\varphi} : \boldsymbol{H}^{\perp} \to \boldsymbol{V}$. If $\boldsymbol{W} = \tilde{\boldsymbol{W}} \subset \boldsymbol{H}^{\perp}$, then we find a subspace $\boldsymbol{W}_2 \subset \boldsymbol{H}^{\perp}$ complementary to $\boldsymbol{W}$. Otherwise, $\boldsymbol{W} \cap \tilde{\boldsymbol{W}} = \boldsymbol{U}$, and hence we could find a subspace $\boldsymbol{W}_2$ complementary to $\boldsymbol{W}$ in $\boldsymbol{H}^{\perp}$ with $\boldsymbol{W}_2 \cap \tilde{\boldsymbol{W}} = \{\mathfrak{o}\}$: To see this, it is sufficient to note that $\boldsymbol{W} + \tilde{\boldsymbol{W}}$ has dimension $\dim \boldsymbol{U} + 2$; if $\mathfrak{b}$ and $\tilde{\mathfrak{b}}$ are vectors spanning with $\boldsymbol{U}$ the subspace $\boldsymbol{W}$ and $\tilde{\boldsymbol{W}}$, respectively, and if $\boldsymbol{U}_1$ is a subspace of $\boldsymbol{H}^{\perp}$ complementary to $\boldsymbol{W} + \tilde{\boldsymbol{W}}$, then $\boldsymbol{W}_2 = \mathfrak{L}(\mathfrak{b} + \tilde{\mathfrak{b}}) + \boldsymbol{U}_1$ is such a complementary subspace. By Lemma 10 and using the argument above we obtain the extension $\hat{\varphi}$ of φ to $\boldsymbol{H}^{\perp}$ we wanted to construct. It remains to extend $\hat{\varphi}$ to an automorphism ψ of $\boldsymbol{V}$. In the second case $\boldsymbol{W} + \tilde{\boldsymbol{W}} \subset \boldsymbol{H}^{\perp}$ implies $\varphi(\mathfrak{a}) - \mathfrak{a} \in \boldsymbol{H} \cap \boldsymbol{H}^{\perp}$; hence $\boldsymbol{H}^{\perp}$ is an isotropic hyperplane. Thus, $\boldsymbol{V}$ is the smallest non-isotropic subspace containing $\boldsymbol{H}^{\perp}$, and by Lemma 9 we can extend $\hat{\varphi}$ to an automorphism ψ of $\boldsymbol{V}$. □

2.2.6 Transitivity Results

The importance of E. Witt's Theorem lies in the wealth of geometric applications to actions of the classical groups on subspaces; the classification of subspaces with respect to this action is reduced to the classification of biforms on them, which is known in the essential cases. In particular, this leads to interesting transitivity results:

Corollary 11. *The automorphism group $G \subset \boldsymbol{GL}(\boldsymbol{V})$ of a vector space with scalar product $\langle , \rangle$ acts transitively on the isomorphy classes of subspaces $\boldsymbol{W} \subset \boldsymbol{V}$ with the induced biform $b = \langle , \rangle | \boldsymbol{W} \times \boldsymbol{W}$* (cf. Definition 2). □

As a consequence this almost allows to completely describe the orbit decomposition of the Grassmann manifold $G_{n,k}$ of k-dimensional subspaces of an n-dimensional vector space with scalar product, see the examples below.

Example 1. Let $\boldsymbol{V}^n$ be an n-dimensional pseudo-Euclidean vector space with a scalar product of index l. By the Theorem of Inertia, Proposition I.5.9.6, the isomorphy class of the bilinear form $b_{\boldsymbol{W}} := \langle , \rangle | \boldsymbol{W} \times \boldsymbol{W}$ is determined by its rank $r_{\boldsymbol{W}}$ together with its index $l_{\boldsymbol{W}}$; we may, however, replace the rank by the *defect* $d_{\boldsymbol{W}} = \dim \boldsymbol{W} - r_{\boldsymbol{W}}$. By Corollary 11 the pseudo-orthogonal group $\boldsymbol{O}(l, n-l)$ acts transitively on the set $G_{n,k,s,d}$ of k-dimensional subspaces with index s and defect d. In the case of a Euclidean vector space, i.e. if $\langle , \rangle$ is positive definite, we always have $l_{\boldsymbol{W}} = d_{\boldsymbol{W}} = 0$; hence the orthogonal group $\boldsymbol{O}(n)$ acts transitively on the Grassmann manifold $G_{n,k}$ of k-dimensional subspaces in $\boldsymbol{V}^n$. □

Example 2. Let $[\boldsymbol{A}^n, \boldsymbol{V}^n, K, \langle , \rangle]$ be an affine geometry over a vector space with scalar product. The *isomorphy type of a k-plane $\boldsymbol{H}^k \subset \boldsymbol{A}^n$* is understood to be the isomorphy type of the vector space $\boldsymbol{W}$ corresponding to $\boldsymbol{H}$ with the induced biform $b_{\boldsymbol{W}}$. Since the group of translations $\boldsymbol{T}(\boldsymbol{A})$ acts transitively over the k-planes parallel to a fixed k-plane, Corollary 11 implies: The subgroup $\boldsymbol{GA} \subset \mathfrak{A}(n)$ of the affine group generated by the automorphism group $\boldsymbol{G} \subset \boldsymbol{GL}(\boldsymbol{V})$ together with the group of translations acts transitively on each of the sets of k-planes of the same isomorphy type. $\square$

Corollary 12. *The automorphism group G of a vector space with scalar product acts transitively over the set of k-dimensional, totally isotropic subspaces. If $\boldsymbol{W}^k, \tilde{\boldsymbol{W}}^k$ are two such subspaces, and if $(\mathfrak{b}_\alpha)$, $(\mathfrak{c}_\alpha)$, $\alpha = 1, \ldots, k$, are arbitrary bases for $\boldsymbol{W}$ and $\tilde{\boldsymbol{W}}$, respectively, then there is a transformation $g \in G$ with $g(\mathfrak{b}_\alpha) = \mathfrak{c}_\alpha$, $\alpha = 1, \ldots, k$.*

Proof. Obviously, the linear map $\varphi : \boldsymbol{W} \to \boldsymbol{V}$ determined by $\varphi(\mathfrak{b}_\alpha) = \mathfrak{c}_\alpha$ is a linear morphism, which can thus be extended to an automorphism $g \in G$ by E. Witt's Theorem. $\square$

Example 3. The symplectic group $\boldsymbol{Sp}(\boldsymbol{V}^{2n})$, and hence the projective symplectic group $\boldsymbol{PSp}_n$ as well, acts transitively on the projective symplectic space $\boldsymbol{P}^{2n-1}$. $\square$

Example 4. Let $Q \subset \boldsymbol{P}^n$ be any quadric of the projective space $\boldsymbol{P}^n$ determined by the non-degenerate σ-biform b. By $\langle , \rangle := b$ we define a scalar product on the associated vector space $\boldsymbol{V}$. Note that b, and hence also the scalar product, are determined by the quadric only up to a factor $k \in K^*$; choosing a fixed value for this factor is called a *gauge*; the automorphism group G of $[\boldsymbol{V}, \langle , \rangle]$ does not depend on the choice of this gauge factor. The results of this section imply: 1. The group $\boldsymbol{PG}$ of projectivities on $\boldsymbol{P}^n$ preserving Q acts transitively on Q. – 2. The dimension of a projective k-plane $\boldsymbol{A} \subset Q$ satisfies $\operatorname{Dim} \boldsymbol{A} \leq (n-1)/2$. – 3. Let m be the maximal dimension of a k-plane in Q. Then $m+1$ coincides with the index l of the scalar product; for this reason, l is also called the *index of the quadric*. For every $k < l$ the group $\boldsymbol{PG}$ acts transitively on the set Q_k of projective k-planes contained in Q. $\square$

Example 5. As discussed above, *Euclidean geometry* is a specialization of affine geometry, cf. I.6.5. Taking $K = \mathbf{R}$ and choosing the scalar product positive definite it is a special case of Example 2. From the projective point of view it looks as follows: Distinguish a hyperplane $\boldsymbol{H}^{n-1}$ as the absolute in real projective space (often also called the hyperplane at infinity). Then the resulting geometry is that of the affine space $\boldsymbol{A}^n = \boldsymbol{P}^n \setminus \boldsymbol{H}^{n-1}$, cf. Section 1.5. There we already know that the points $\boldsymbol{x} \in \boldsymbol{H}^{n-1}$ correspond to the directions in $\boldsymbol{A}^n$. By Exercise 1.5.4 the vectors in affine geometry may be interpreted as the elements of the vector space $\boldsymbol{W}^n$ corresponding to the absolute hyperplane $\boldsymbol{H}^{n-1}$. Prescribing now in $\boldsymbol{H}^{n-1}$ an *absolute polarity*, the correspondence $\langle , \rangle := b$ associated with it (Definition 1.7.4) defines orthogonality: two directions $\boldsymbol{x} = [\mathfrak{x}]$, $\boldsymbol{y} = [\mathfrak{y}]$ are orthogonal if and only if $\langle \mathfrak{x}, \mathfrak{y} \rangle = 0$. If the scalar product has index 0, then it defines the empty quadric $Q = \emptyset$ in $\boldsymbol{H}^{n-1}$; fixing a gauge, e.g. distinguishing a vector $\mathfrak{x} \in \boldsymbol{W}^n$, $\mathfrak{x} \neq \mathfrak{o}$, as a unit vector completes the projective view on Euclidean geometry, cf. also Exercise 1.9.6. Choosing instead of the positive definite scalar product one of index l leads, in a similar way, to the whole series of *pseudo-Euclidean geometries*, and in the case $K = \mathbf{C}$ one obtains the *complex-Euclidean geometry* this way. Just as in Example 2, a variety of transitivity statements follow. □

2.2.7 Neutral Vector Spaces

We now again refer to Proposition 5 and call a vector space with scalar product *neutral* if it may be represented as the direct sum of two totally isotropic subspaces (cf. N. Bourbaki [19], §IX.4.2; E. Artin [4] calls such spaces hyperbolic, a terminology in conflict with the frequently used notion of hyperbolic space in the geometry of constant negative sectional curvature, cf. Section 6 below). The term neutral will also be applied to the associated scalar product. Obviously, every symplectic vector space is neutral. In the following exercises the reader can find some applications of this notion.

Exercise 6. Let $\boldsymbol{V}^n$ be an n-dimensional neutral vector space, and consider a decomposition of $\boldsymbol{V}$ into the direct sum of two totally isotropic subspaces, $\boldsymbol{V} = \boldsymbol{W}_1 \oplus \boldsymbol{W}_2$. Prove: a) The map

$$\varphi : \mathfrak{x} \in \boldsymbol{W}_2 \longmapsto \varphi(\mathfrak{x}) := \langle \mathfrak{x}, . \rangle | \boldsymbol{W}_1 \in \boldsymbol{W}_1' \tag{16}$$

is σ-linear and injective. – b) Both subspaces have equal dimension,

$$\dim \boldsymbol{W}_1 = \dim \boldsymbol{W}_2, \tag{17}$$

and φ is a σ-linear bijection. – c) If $(\mathfrak{x}_\alpha)$, $\alpha = 1, \ldots, k$, is a basis for $\boldsymbol{W}_1$, then there exists a uniquely determined basis $(\mathfrak{y}_\beta)$, $\beta = 1, \ldots, k$, for $\boldsymbol{W}_2$ such that

$$\langle \mathfrak{x}_\alpha, \mathfrak{y}_\beta \rangle = \delta_{\alpha\beta},\ \alpha, \beta = 1, \ldots, k. \tag{18}$$

Exercise 7. If $\boldsymbol{V}_1, \boldsymbol{V}_2$ are two neutral, σ-Hermitean vector spaces of equal finite dimension over the skew field K, then they are isomorphic (the analogue of Proposition 1.8.5). Each pseudo-Euclidean vector space $\boldsymbol{V}^{2m}$ of index m is neutral; the same holds for vector spaces with the Hermitean scalar products $H_{2m,m}$ in the cases $K = \mathbf{C}, \mathbf{H}$.

Exercise 8. The vector space $\boldsymbol{V}^n$ with scalar product has index k if and only if it contains a maximal neutral subspace $\boldsymbol{W} \subset \boldsymbol{V}$ of dimension $2k$, and vice versa. A complex vector space $\boldsymbol{V}^n$ with the scalar product F_n or a quaternionic vector space $\boldsymbol{V}^n$ with the skew Hermitean scalar product H_n is neutral if and only if n is even.

Exercise 9. Let V^n be an n-dimensional vector space with scalar product. Prove:
a) There is a decomposition

$$V^n = W_1 \oplus W_2 \oplus U, \ \dim W_1 = \dim W_2 = \text{ index } V^n, \tag{19}$$

such that W_1, W_2 are totally isotropic, and $U = (W_1 \oplus W_2)^\perp$ does not contain any isotropic vector. Such a decomposition is called a *Witt decomposition* . – b) If $V^n = \tilde{W}_1 \oplus \tilde{W}_2 \oplus \tilde{U}$ is a second Witt decomposition, and if, moreover, $(\mathfrak{x}_\alpha)$, $(\tilde{\mathfrak{x}}_\alpha)$, $\alpha = 1, \ldots, k$, are bases of W_1 and $\tilde{W}_1$, respectively, then there is an automorphism g in the automorphism group G such that

$$g(\mathfrak{x}_\alpha) = \tilde{\mathfrak{x}}_\alpha, \ \alpha = 1, \ldots, k, \ g(W_2) = \tilde{W}_2, g(U) = \tilde{U}. \tag{20}$$

– c) If $W_i, \tilde{W}_i$, $i = 1, 2$, are two pairs of totally isotropic subspaces of maximal dimension with $W_1 \cap W_2 = \tilde{W}_1 \cap \tilde{W}_2 = \{\mathfrak{o}\}$, then there always is a $g \in G$ $[V^n, \langle, \rangle]$ with $g(W_i) = \tilde{W}_i$, $i = 1, 2$.

Exercise 10. Let V^n be a vector space with scalar product $\langle, \rangle$. Prove that for two isomorphic subspaces $U, \tilde{U} \subset V^n$ the corresponding orthogonal subspaces $U^\perp, \tilde{U}^\perp$ are also isomorphic.

Exercise 11. Let $[V^n, \langle, \rangle]$ be a neutral space, and let $\mathfrak{x}_\alpha, \mathfrak{y}_\beta$, $\alpha, \beta = 1, \ldots, k = n/2$, be a basis adapted to a Witt decomposition $V^n = W_1 \oplus W_2$ having property (18). Prove:
a) If g is an automorphism of V^n satisfying $g|W_1 = \mathrm{id}_{W_1}$, then, with respect to this basis, g is represented by a block matrix of the form

$$\begin{pmatrix} I_k & A \\ 0 & I_k \end{pmatrix} \in GL(2k, K), \tag{21}$$

where, for symplectic spaces, the matrix $A = A'$ is symmetric, whereas, for σ-Hermitean spaces, the matrix A is skew σ-Hermitean, i.e.

$$A' = -\sigma(A). \tag{22}$$

– b) Conversely, each matrix with the above properties defines an automorphism g of V^n with $g|W_1 = \mathrm{id}_{W_1}$.

According to Exercise 7, neutral σ-Hermitean vector spaces of equal dimension over the same skew field are isomorphic. This leaves the task to determine the automorphism groups O_n of these vector spaces with scalar product. We only want to derive a simple result in that direction for later use.

Proposition 13. *Let V^2 be a two-dimensional neutral vector space with symmetric bilinear scalar product over a field K. Then the* special orthogonal group

$$SO_2 := \{g \in O_2 |\, \det g = 1\} \tag{23}$$

is isomorphic to the multiplicative group K^* of K, and the orthogonal group O_2 is the semi-direct product of a subgroup isomorphic to $\mathbf{Z}_2$ with the normal subgroup SO_2.

Proof. Because of char $K \neq 2$ we may choose an isotropic basis $\mathfrak{z}_1, \mathfrak{z}_2$ for $\boldsymbol{V}^2$ such that $\langle \mathfrak{z}_1, \mathfrak{z}_2 \rangle = 1/2$. For $g \in \boldsymbol{O}_2$ we start from the ansatz

$$g\mathfrak{z}_1 = \mathfrak{z}_1\alpha + \mathfrak{z}_2\beta, \; g\mathfrak{z}_2 = \mathfrak{z}_1\gamma + \mathfrak{z}_2\delta$$

for the basis representation, and make use of the invariance of the scalar product to derive the conditions

$$\alpha\beta = \gamma\delta = 0, \; \alpha\delta + \beta\gamma = 1. \tag{24}$$

Hence, if $\beta = 0$, then $\delta \neq 0$ and $\gamma = 0$; this implies the isomorphy

$$g: \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} \in \boldsymbol{SO}_2 \longmapsto \alpha \in K^*. \tag{25}$$

For $\beta \neq 0$ we similarly obtain $\alpha = 0$, $\gamma = \beta^{-1}$, i.e.

$$g: \begin{pmatrix} 0 & \gamma \\ \gamma^{-1} & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \gamma^{-1} & 0 \\ 0 & \gamma \end{pmatrix}. \tag{26}$$

Thus $\boldsymbol{O}_2 = \boldsymbol{SO}_2 \cup s(\boldsymbol{SO}_2)$, where s is the transformation interchanging $\mathfrak{z}_1$ with $\mathfrak{z}_2$; it is easy to check that $\boldsymbol{O}_2$ is the semi-direct product of $\boldsymbol{SO}_2 = \operatorname{Ker}\det$ with $\{\mathrm{id}_{\boldsymbol{V}}, s\} \cong \mathbf{Z}_2$. $\square$

Exercise 12. Prove under the assumptions of Proposition 13: $\boldsymbol{O}_2$ is Abelian if and only if K is isomorphic to the field $\mathbf{Z}_3$ with 3 elements.

2.2.8 Tensors and Volume Functions

In this section we want to extend the biforms to tensor spaces; moreover, we will discuss volume functions on vector spaces with scalar product. To this end, suppose that the scalar domain is a field.

Exercise 13. Let K be a field, and let b be an alternating or σ-Hermitean biform over the vector space $\boldsymbol{V}^n$. Prove: a) The biform b is non-degenerate on $\boldsymbol{V}^n$, if for each basis $(\mathfrak{a}_i)$ of $\boldsymbol{V}^n$

$$\det(b(\mathfrak{a}_i, \mathfrak{a}_j)) \neq 0. \tag{27}$$

– b) The bilinear or σ-linear and linear extensions

$$b^{\otimes p}(\mathfrak{x}_1 \otimes \ldots \otimes \mathfrak{x}_p, \mathfrak{y}_1 \otimes \ldots \otimes \mathfrak{y}_p) := \prod_1^p b(\mathfrak{x}_\alpha, \mathfrak{y}_\alpha), \qquad p \in \mathbf{N}, \tag{28}$$

or

$$b^{(p)}(\mathfrak{x}_1 \wedge \ldots \wedge \mathfrak{x}_p, \mathfrak{y}_1 \wedge \ldots \wedge \mathfrak{y}_p) := \det(b(\mathfrak{x}_\alpha, \mathfrak{y}_\beta)), \qquad 0 < p \leq n, \tag{29}$$

respectively, define bilinear or σ-Hermitean biforms on the tensor spaces $\bigotimes^p \boldsymbol{V}^n$, $\bigwedge^p \boldsymbol{V}^n$ (cf. also Example II.8.3.4); the biforms $b^{\otimes p}$, $b^{(p)}$ are non-degenerate if and only if b is non-degenerate. – c) If b is σ-Hermitean, then this also holds for $b^{\otimes p}$ and $b^{(p)}$. – d) For an alternating biform b, $b^{\otimes p}$ and $b^{(p)}$ are alternating, if p is odd, and they are symmetric, if p is even.

As in Exercise 13, let $\boldsymbol{V}^n$ be a vector space with scalar product over the field K, and denote by $\boldsymbol{UG}_n$ its automorphism group. In general, according to (1.13) we only know that $\sigma(N(a))N(a) = 1$ for $a \in \boldsymbol{UG}_n$. Hence, in order to achieve invariance for a volume we have to confine the action to the *special unitary group*

$$\boldsymbol{SUG}_n := \{a \in \boldsymbol{UG}_n | N(a) = 1\}. \tag{30}$$

Let $[., \ldots, .]$ be a volume function, and let $(\mathfrak{a}_j)$ be a unit basis, i.e., suppose that

$$[\mathfrak{a}_1, \ldots, \mathfrak{a}_n] = 1 \text{ and } [\mathfrak{x}_1, \ldots, \mathfrak{x}_n] = \det(\xi_k^j),$$

if $\mathfrak{x}_k = \mathfrak{a}_j \xi_k^j$ is the basis representation of $\mathfrak{x}_k$, $k = 1, \ldots, n$. We prove

Lemma 14. *Let $[\boldsymbol{V}^n, \langle , \rangle]$ be a vector space with scalar product over a field K. For $k \in \mathbf{N}$ and $\mathfrak{b}_\kappa \in \boldsymbol{V}^n$, $\kappa = 1, \ldots, k$,*

$$\det(\langle \mathfrak{b}_\kappa, \mathfrak{b}_\lambda \rangle) = 0 \tag{31}$$

if and only if the sequence $(\mathfrak{b}_\kappa)$, $\kappa = 1, \ldots, k$, is linearly dependent, or if the subspace spanned by it is isotropic.

Proof. Linear dependence of the $\mathfrak{b}_\lambda$ implies linear dependence of the columns in the matrix $(\langle \mathfrak{b}_\kappa, \mathfrak{b}_\lambda \rangle)$ leading to (31). Let thus the $(\mathfrak{b}_\kappa)$ be linearly independent; then the assertion is a consequence of the assertion in Exercise 13 (with $n = k$, $\boldsymbol{V}^k = \mathfrak{L}(\mathfrak{b}_1, \ldots, \mathfrak{b}_k)$, $b = \langle , \rangle$). □

Lemma 15. *With the same assumptions as in Lemma* 14 *let $(\mathfrak{a}_\kappa)$, $\kappa = 1, \ldots, k$, be a basis of $\boldsymbol{W}^k \subset \boldsymbol{V}^n$; let $[., \ldots, .]$ be the volume function on $\boldsymbol{W}^k$ satisfying $[\mathfrak{a}_1, \ldots, \mathfrak{a}_k] = 1$* (cf. Proposition I.4.7.5). *Suppose, moreover,*

$$\mathfrak{x}_\lambda = \mathfrak{a}_\kappa \xi_\lambda^\kappa \in \boldsymbol{W}^k, \ \kappa, \lambda = 1, \ldots, k. \tag{32}$$

Then

$$\det(\langle \mathfrak{x}_\kappa, \mathfrak{x}_\lambda \rangle) = a \cdot [\mathfrak{x}_1, \ldots, \mathfrak{x}_k] \cdot \sigma([\mathfrak{x}_1, \ldots, \mathfrak{x}_k]) \tag{33}$$

with

$$a := \det(\langle \mathfrak{a}_\kappa, \mathfrak{a}_\lambda \rangle) = \sigma(a). \tag{34}$$

□

Remarks. 1. If $\boldsymbol{W}^k$ is isotropic, then $a = 0$, and (33) reduces to $\det(\langle \mathfrak{x}_\kappa, \mathfrak{x}_\lambda \rangle) = 0$. – 2. The matrix $(\langle \mathfrak{a}_\kappa, \mathfrak{a}_\lambda \rangle)$ formed by the scalar products of the basis vectors is called the *Gram matrix.* For $\sigma = \mathrm{id}_K$ and $a = 1$, a condition always satisfied by orthonormal and symplectic bases, this is nothing but Formula I.(6.3.8) for the *Gram determinant.* – 3. Consider $K = \mathbf{R}$ and a pseudo-orthonormal basis $(\mathfrak{a}_\alpha)$. Then $a = (-1)^l$, where l is the index of $\boldsymbol{W}^k$, and hence

$$[\mathfrak{x}_1, \ldots, \mathfrak{x}_k]^2 = |\det(\langle \mathfrak{x}_\kappa, \mathfrak{x}_\lambda \rangle)|,$$

cf. II(8.3.47), where the absolute value on the right-hand side is missing.

Corollary 16. *If $k = n$ in Lemma (15) , i.e., if $[.,\ldots,.]$ is a volume function on $\boldsymbol{V}^n$, then the constant a does not depend on the choice of the unit basis $(\mathfrak{a}_i)$, $[\mathfrak{a}_1,\ldots,\mathfrak{a}_n] = 1$.*

Proof. For a second unit basis $(\mathfrak{b}_i)$ of $\boldsymbol{V}^n$ we conclude from (33) and (34) that

$$b := \det(\langle \mathfrak{b}_i, \mathfrak{b}_j \rangle) = a.$$

□

2.2.9 The General Vector Product

It turns out that the vector product and its properties as described in Proposition I.6.3.2 (cf. also Exercise II.8.3.14) can be generalized to a large extent:

Proposition 17. *Let $\boldsymbol{V}^n$ be a vector space over a field K with the σ-biform $\langle,\rangle$ as scalar product ($\sigma = \mathrm{id}_K$ for $\langle,\rangle$ bilinear). Let, moreover, $[.,\ldots,.]$ be a volume function on $\boldsymbol{V}^n$. Then for every $(n-1)$-tuple $(\mathfrak{c}_\alpha)$, $\alpha = 1,\ldots,n-1$, of vectors $\mathfrak{c}_\alpha \in \boldsymbol{V}^n$ there is a uniquely determined vector $\mathfrak{c} \in \boldsymbol{V}^n$ satisfying*

$$\langle \mathfrak{c}, \mathfrak{x} \rangle = [\mathfrak{c}_1,\ldots,\mathfrak{c}_{n-1},\mathfrak{x}] \text{ for all } \mathfrak{x} \in \boldsymbol{V}^n. \tag{35}$$

Proof. Since the scalar product can be viewed as a σ-linear bijection via $\mathfrak{y} \in \boldsymbol{V} \mapsto \langle \mathfrak{y}, . \rangle \in \boldsymbol{V}'$, and as the volume function is linear in each argument, we immediately conclude the existence of a vector $\mathfrak{c}$ which is uniquely determined by the $(n-1)$-tuple $(\mathfrak{c}_\alpha)$ (cf. Definition 1.7.2). This vector $\mathfrak{c}$ is called the *vector product* of the vectors $\mathfrak{c}_\alpha$; it is denoted by $\mathfrak{c}_1 \times \ldots \times \mathfrak{c}_{n-1}$. □

Proposition 18. *Under the assumptions of Proposition 17 the vector product has the following properties:*

1. *The map*

$$(\mathfrak{c}_\alpha) \in \prod^{n-1} \boldsymbol{V}^n \mapsto \mathfrak{c}_1 \times \ldots \times \mathfrak{c}_{n-1} \in \boldsymbol{V}^n$$

   is σ-linear in each argument $\mathfrak{c}_\alpha$.
2. *The map is alternating.*
3. *$\mathfrak{c}_1 \times \ldots \times \mathfrak{c}_{n-1} = \mathfrak{o}$ if and only if the sequence $(\mathfrak{c}_\alpha)$, $\alpha = 1,\ldots,n-1$, is linearly dependent.*
4. *The vector product is orthogonal to each of its factors:*

$$\langle \mathfrak{c}_1 \times \ldots \times \mathfrak{c}_{n-1}, \mathfrak{c}_\alpha \rangle = 0,\ \alpha = 1,\ldots,n-1. \tag{36}$$

5. *The scalar product of the vector product with itself is computed by*

$$\det(\langle \mathfrak{c}_\alpha, \mathfrak{c}_\beta \rangle) = a \langle \mathfrak{c}_1 \times \ldots \times \mathfrak{c}_{n-1}, \mathfrak{c}_1 \times \ldots \times \mathfrak{c}_{n-1} \rangle, \tag{37}$$

   where a is the constant associated with the volume function according to (16) .

6. *The vector product $\mathfrak{c}_1 \times \ldots \times \mathfrak{c}_{n-1}$ is isotropic if and only if the subspace spanned by the factors, $\mathfrak{L}(\mathfrak{c}_1, \ldots, \mathfrak{c}_{n-1})$, is isotropic.*
7. *For any unitary transformation $g \in \boldsymbol{UG}_n$ the vector product transforms according to*

$$g(\mathfrak{c}_1) \times \ldots \times g(\mathfrak{c}_{n-1}) = g(\mathfrak{c}_1 \times \ldots \times \mathfrak{c}_{n-1}) N(g^{-1}). \tag{38}$$

Proof. Properties 1 to 4 immediately follow from the definition. To prove (37) we apply (33) to $(\mathfrak{c}_1, \ldots, \mathfrak{c}_{n-1}, \mathfrak{c})$ with $k = n$, $\mathfrak{c} = \mathfrak{c}_1 \times \ldots \times \mathfrak{c}_{n-1}$. From (36) we obtain

$$\det(\langle \mathfrak{c}_\alpha, \mathfrak{c}_\beta \rangle)\langle \mathfrak{c}, \mathfrak{c} \rangle = a\langle \mathfrak{c}, \mathfrak{c} \rangle^2;$$

here we made use of (35) with $\mathfrak{x} = \mathfrak{c}$ and $\langle \mathfrak{c}, \mathfrak{c} \rangle = \sigma(\langle \mathfrak{c}, \mathfrak{c} \rangle)$. If $\mathfrak{c} = \mathfrak{o}$, then (37) is trivial by 3 and Lemma 14. According to 4, the subspace $\boldsymbol{W}^{n-1} = \mathfrak{L}(\mathfrak{c}_1, \ldots, \mathfrak{c}_{n-1})$ is isotropic for $\mathfrak{c} \neq \mathfrak{o}$ if and only if $\mathfrak{c}$ is isotropic. In fact, $\mathfrak{c}$ is orthogonal to $\boldsymbol{W}^{n-1}$. This implies 6, and by Lemma 14 property (37) follows from the equation just proved. To prove statement 7 we set $\hat{\mathfrak{c}} := g(\mathfrak{c}_1) \times \ldots \times g(\mathfrak{c}_{n-1})$. Then I.(5.7.32) implies

$$\begin{aligned}
\langle \hat{\mathfrak{c}}, \mathfrak{x} \rangle &= [g(\mathfrak{c}_1), \ldots, g(\mathfrak{c}_{n-1}), g(g^{-1}(\mathfrak{x}))],\\
&= [\mathfrak{c}_1, \ldots, \mathfrak{c}_{n-1}, g^{-1}(\mathfrak{x})] N(g),\\
&= \langle \mathfrak{c}, g^{-1}(\mathfrak{x}) \rangle N(g),\\
&= \langle g(\mathfrak{c}) N(g^{-1}), \mathfrak{x} \rangle;
\end{aligned}$$

here we made use of (1.18):

$$N(g^{-1}) = N(g)^{-1} = \sigma(N(g)) \text{ for } g \in \boldsymbol{UG}_n. \tag{39}$$

□

Remark. Since there exists an order in $\mathbf{R}$, and, moreover, each positive number has a positive root, Assertion 5 in Proposition I.6.3.2 is a specialization of Assertion 5 in Proposition 18 that is impossible in general; for the same reason, the former Property 6 cannot be formulated in the general case, whereas the new Property 6 makes no sense for Euclidean vector spaces.

Exercise 14. In the notations and assumptions of Proposition 17, let the $\mathfrak{c}_i$, $i = 1, \ldots, n$, be orthogonal. Prove: a) If $(\mathfrak{c}_i)$ is a unit basis, $[\mathfrak{c}_1, \ldots, \mathfrak{c}_n] = 1$, then

$$a = \prod_{i=1}^{n} \langle \mathfrak{c}_i, \mathfrak{c}_i \rangle.$$

b) Show by means of an example that the converse of a) does not hold in general. c) If $(\mathfrak{b}_\alpha)$, $\alpha = 1, \ldots, n-1$, is an orthogonal sequence of vectors with $\langle \mathfrak{b}_\alpha, \mathfrak{b}_\alpha \rangle \neq 0$, then

$$\mathfrak{b}_n := \mathfrak{b}/\langle \mathfrak{b}, \mathfrak{b} \rangle \text{ with } \mathfrak{b} := \mathfrak{b}_1 \times \ldots \times \mathfrak{b}_{n-1} \tag{40}$$

completes the sequence $(\mathfrak{b}_\alpha)$ to a unit basis of $\boldsymbol{V}^n$. – d) Generalizing (37) we have for arbitrary vectors $\mathfrak{x}_\alpha, \mathfrak{y}_\beta \in \boldsymbol{V}^n$

$$\langle \mathfrak{x}_1 \times \ldots \times \mathfrak{x}_{n-1}, \mathfrak{y}_1 \times \ldots \times \mathfrak{y}_{n-1} \rangle = a \det(\langle \mathfrak{x}_\alpha, \mathfrak{y}_\beta \rangle). \tag{41}$$

(Hint. Recall the linearity or σ-linearity of the functions occurring in (41) in all variables and use basis representations.)

2.2.10 Adjoint Linear Maps

Let $[\boldsymbol{V}, \langle , \rangle]$, $[\boldsymbol{W}, \langle , \rangle]$ be two finite-dimensional vector spaces with scalar product (not necessarily of equal type) over the same skew field K. Then with each linear map $a : \boldsymbol{V} \to \boldsymbol{W}$ an *adjoint linear map*[1] $a^* : \boldsymbol{W} \to \boldsymbol{V}$ is associated which is uniquely defined by the equation

$$\langle a^* \mathfrak{y}, \mathfrak{x} \rangle = \langle \mathfrak{y}, a\mathfrak{x} \rangle \text{ for all } \mathfrak{x} \in \boldsymbol{V}, \mathfrak{y} \in \boldsymbol{W}. \tag{42}$$

This notion generalizes the one for Euclidean or unitary vector spaces, denoted by a' in § I.6.4. In these geometries the *self-adjoint endomorphisms* are of particular interest; they are defined by the condition $a^* = a$ for $a \in \mathrm{End}(\boldsymbol{V})$ and turn out to be diagonalizable, cf. Proposition I.6.4.2. Moreover, they are closely related to the symmetric and Hermitean bilinear forms, respectively. To a large extent, the classification of quadrics in Euclidean geometry can be derived from the theory of self-adjoint linear maps. Quite similarly, we will be able to reduce classification problems for some special geometries with scalar product to the classification of special linear endomorphisms in the following. For simplicity, we restrict the considerations to the case $[\boldsymbol{V}, \langle , \rangle] = [\boldsymbol{W}, \langle , \rangle]$, i.e. $a \in \mathrm{End}(\boldsymbol{V})$. First we note

Proposition 19. *Let $\boldsymbol{V}$ be a vector space with scalar product, and let σ be the corresponding anti-automorphism of the scalar domain K. Then, for each linear map $a \in \mathrm{End}(\boldsymbol{V})$ there is a uniquely determined linear map $a^* \in \mathrm{End}(\boldsymbol{V})$ satisfying* (42). *The map $\varphi : a \mapsto a^*$ is an involutive σ-anti-automorphism of the endomorphism algebra* $\mathrm{End}(\boldsymbol{V})$*; moreover,*

$$(a + b)^* = a^* + b^*, \tag{43}$$

$$(a\gamma)^* = a^* \sigma(\gamma) \textit{ for all } \gamma \in Z(K), \tag{44}$$

$$(a \circ b)^* = b^* \circ a^*, \tag{45}$$

$$(a^*)^* = a. \tag{46}$$

Proof. The proof of this proposition is a simple verification based on the related definitions. For fixed $\mathfrak{y}$, the left-hand side of (42) is a linear function in $\mathfrak{x}$. Since the scalar product is non-degenerate, it defines a bijection between $\boldsymbol{V}$ and its dual space $\boldsymbol{V}'$; hence, there is precisely one vector $a^*\mathfrak{y}$ satisfying (42). Thus the map a^* is uniquely defined. The linearity of a^* is easy to check. In fact, for all $\eta \in K$ we have

[1] Do not confuse this notion of the adjoint linear map with that of the contragredient isomorphism, which is frequently also denoted by a^*, cf. Example 1.7.4.

$$\langle a^*(\mathfrak{y}\eta), \mathfrak{x}\rangle = \langle \mathfrak{y}\eta, a\mathfrak{x}\rangle = \sigma(\eta)\langle \mathfrak{y}, a\mathfrak{x}\rangle = \sigma(\eta)\langle a^*\mathfrak{y}, \mathfrak{x}\rangle = \langle (a^*\mathfrak{y})\eta, \mathfrak{x}\rangle.$$

Recall that the endomorphism algebra is an algebra only over the center of the scalar domain, which, obviously, is mapped into itself by σ. Let us conclude the proof by showing (44). For $\gamma \in Z(K)$ we have

$$\langle (a\gamma)^*\mathfrak{y}, \mathfrak{x}\rangle = \langle \mathfrak{y}, (a\gamma)\mathfrak{x}\rangle = \langle \mathfrak{y}, a\mathfrak{x}\rangle\gamma = \gamma\langle a^*\mathfrak{y}, \mathfrak{x}\rangle = \langle a^*\mathfrak{y}\sigma(\gamma), \mathfrak{x}\rangle = \langle (a^*\sigma(\gamma))\mathfrak{y}, \mathfrak{x}\rangle.$$

The proof of the remaining statements is left to the reader. □

The scalar product determines a σ-linear bijection from the vector space $\boldsymbol{V}$ onto its dual $\boldsymbol{V}'$, assigning to each $\mathfrak{y} \in \boldsymbol{V}$ a linear form $\omega_{\mathfrak{y}}(\mathfrak{x}) := \langle \mathfrak{y}, \mathfrak{x}\rangle$. Comparing the definition of the adjoint map a^* to Definition 1.6.1, (16), of the dual map shows that, essentially, both describe the same (apart from the application of σ^{-1}). a^* expresses, however, the dual map in the vector space itself, which is possible due to the available scalar product. Having any at our disposal, we can completely do without the introduction of dual vectors, as it is frequently done in elementary presentations of Euclidean geometry.

Exercise 15. Prove that, in analogy to the situation in unitary vector spaces, the following holds: A linear map $a \in \mathrm{End}(\boldsymbol{V})$ is an automorphism of the vector space with scalar product if and only if a is invertible, and $a^{-1} = a^*$. Show, moreover, that Formulas (43)–(46) also hold for arbitrary linear maps a, b between possibly different vector spaces with scalar products over the same skew field K.

Now we want to establish a relation between biforms and linear maps, that is defined by the scalar product. In fact, projectively interpreted, the scalar product is a distinguished correlation F; more generally, each σ-biform is a correlative map H, cf. Proposition 1.7.6. Thus, $F^{-1} \circ H$ is a collinear map that, except for trivial cases, is induced by a semi-linear map a. Expressed differently, for all $\mathfrak{x} \in \boldsymbol{V}$ we have $F([a(\mathfrak{x})]) = H([\mathfrak{x}])$; projectively, this amounts to saying that *the H-polar of $\boldsymbol{x} = [\mathfrak{x}]$ coincides with the F-polar of* $[a(\mathfrak{x})]$. This corresponds to the relation we now want to describe. Here we will consider the general case of a σ-Hermitean scalar product; for bilinear scalar products assume that K is a field and $\sigma = \mathrm{id}_K$.

Lemma 20. *Let $a \in \mathrm{End}(\boldsymbol{V})$ be a linear map. Then*

$$b_a(\mathfrak{x}, \mathfrak{y}) := \langle a\mathfrak{x}, \mathfrak{y}\rangle,\ \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}, \tag{47}$$

is a σ-biform. For each σ-biform b over $\boldsymbol{V} \times \boldsymbol{V}$ there is a uniquely determined endomorphism $a \in \mathrm{End}(\boldsymbol{V})$ such that (47) holds with $b = b_a$. With respect to the standard actions of the linear group over the endomorphism algebra, $l_g(a) = g \cdot a \cdot g^{-1}$ for $a \in \mathrm{End}(\boldsymbol{V}), g \in \boldsymbol{GL}(\boldsymbol{V})$, and on the σ-biforms $b \mapsto gb$, defined by

$$gb(\mathfrak{x}, \mathfrak{y}) := b(g^{-1}\mathfrak{x}, g^{-1}\mathfrak{y}),\ \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V},$$

assigning $a \in \mathrm{End}(\boldsymbol{V}) \mapsto b_a$ as well as taking the inverse are G-maps for the automorphism group G of the vector space with scalar product $[\boldsymbol{V}, \langle,\rangle]$.

Proof. The first assertion is easily verified. So let, conversely, b be a σ-biform; σ denotes the involutive anti-automorphism of K fixed by the scalar product. For fixed $\mathfrak{x}$, $b(\mathfrak{x}, \mathfrak{y})$ is a linear function of $\mathfrak{y}$. Hence, there is precisely one vector $a(\mathfrak{x})$ for which (47) holds with b_a replaced by b and all $\mathfrak{y} \in \boldsymbol{V}$. The linearity of a is proved as above for a^*. The final assertion is immediate. The linear map a and the σ-biform $b = b_a$ bijectively assigned to it via (47) are called mutually *associated.* The next exercise describes the relation between the matrices for b_a and a. □

Exercise 16. Let $(\mathfrak{a}_i)$ be a basis of the vector space $\boldsymbol{V}$ with scalar product, and let

$$(\epsilon_{ij}) := (\langle \mathfrak{a}_i, \mathfrak{a}_j \rangle), (\epsilon^{ij}) := (\epsilon_{ij})^{-1}. \tag{48}$$

be the matrix of the scalar product and its inverse. a) Denoting by (α^i_j) the matrix of $a \in End(\boldsymbol{V})$ and by (β_{ij}) the matrix of the biform defined by (47), show that

$$(\alpha^k_i) = (\epsilon^{kj})(\sigma(\beta_{ij}))'. \tag{49}$$

Here $(\gamma_{ij})'$ denotes the transposed matrix of (γ_{ij}). – b) Denote by $({\alpha^*}^i_k)$ the matrix of the adjoint map for $a \in End(\boldsymbol{V})$ in the above basis. Prove:

$$({\alpha^*}^h_i) = (\epsilon^{hk})(\sigma(\alpha^j_k))'(\epsilon_{ji}). \tag{50}$$

Following the terminology used by A. I. Maltzev [75] for real and complex vector spaces with scalar product we define:

Definition 4. Let $\boldsymbol{V}$ be a vector space with scalar product. A linear map $a \in \mathrm{End}(\boldsymbol{V})$ is called *symmetric* if $a = a^*$; it is called *skew-symmetric* if it satisfies $a = -a^*$. Symmetric linear endomorphisms are often also called *self-adjoint operators*, cf. Definition I.6.4.2. □

Exercise 17. Prove: a) If the scalar product is bilinear and symmetric, then the map a is symmetric or skew-symmetric if and only if the biform b_a associated with a has the same property. – b) If the scalar product is σ-Hermitean, then the map a is symmetric or skew-symmetric if and only if the biform b_a associated with a is σ-Hermitean or *skew σ-Hermitean*,respectively, (cf. Definition 1.9.1); the last property means that

$$b_a(\mathfrak{y}, \mathfrak{x}) = -\sigma(b_a(\mathfrak{x}, \mathfrak{y})) \text{ for all } \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}.$$

– c) If the scalar product is alternating, then the biform b_a associated with a is symmetric or skew-symmetric, respectively, if and only if a is skew-symmetric or symmetric. – d) The map $a \mapsto b_a$ is σ-linear over the center $Z(K)$ of the scalar domain in the following sense:

$$b_{a+c} = b_a + b_c,\ b_{a_\mu} = \sigma(\mu) b_a \text{ for all } \mu \in Z(K).$$

The following result is frequently used:

Lemma 21. *Let a be a symmetric or skew-symmetric endomorphism of the vector space $\boldsymbol{V}$ with scalar product, and let $\boldsymbol{W} \subset \boldsymbol{V}$ be an invariant subspace for a, i.e. $a(\boldsymbol{W}) \subset \boldsymbol{W}$. Then the orthogonal subspace $\boldsymbol{W}^\perp$ is also invariant for a: $a(\boldsymbol{W}^\perp) \subset \boldsymbol{W}^\perp$.*

Proof. For all $\mathfrak{x} \in \boldsymbol{W}^{\perp}$ and any $\mathfrak{w} \in \boldsymbol{W}$ we have

$$\langle a(\mathfrak{w}), \mathfrak{x} \rangle = \pm \langle \mathfrak{w}, a(\mathfrak{x}) \rangle = 0,$$

since $a(\mathfrak{w}) \in \boldsymbol{W}$. Hence $a(\mathfrak{x}) \in \boldsymbol{W}^{\perp}$. □

The following is easy to prove:

Lemma 22. *Let $\boldsymbol{V} = \bigoplus_1^k \boldsymbol{W}_\kappa$ be a decomposition of the vector space with scalar product into the direct sum of mutually orthogonal subspaces. Then the restriction of the scalar product from $\boldsymbol{V}$ to each of the subspaces $\boldsymbol{W}_\kappa$ defines the structure of a vector space with scalar product of the same type as $\boldsymbol{V}$, i.e. bilinearly symmetric, bilinearly alternating, or σ-Hermitean. If, moreover, $a \in \mathrm{End}(\boldsymbol{V})$ is a linear map leaving each of the subspaces invariant, $a(\boldsymbol{U}_\kappa) \subset \boldsymbol{U}_\kappa$, then a is an automorphism, a symmetric or a skew-symmetric map if and only if each restriction $a|\boldsymbol{U}_\kappa$ has this property.* □

>From the theory of unitary vector spaces we know that the symmetric maps considered there have a particularly simple structure: There is an orthonormal basis consisting of eigenvectors of the map, and all the eigenvalues are real. So, the matrix of the map in this basis is diagonal with real numbers on the principal diagonal. The classification of quadrics in Euclidean geometry as well as the classification of real quadratic or Hermitean forms under the orthogonal or unitary group, respectively, all are based on this very fact, cf. Chapter I.6. For symmetric maps in general vector spaces with scalar product the situation is considerably more complicated. Since classification results for symmetric and skew-symmetric maps have interesting geometric applications, we now want to derive some structural properties of these maps. Reading the already cited books by A. I. Maltzev [75] is very stimulating, further results in this direction can be found there. The theory probably goes back to K. Weierstrass [109] and L. Kronecker [70], cf. F. R. Gantmacher [41], [43], [42].

The fact that all the eigenvalues of a symmetric endomorphism in a unitary space are real (cf. Proposition I.6,4.2) arises as a special case of the following

Lemma 23. *Let $\boldsymbol{V}$ be a vector space with a σ-Hermitean scalar product over a field K or over the skew field $\mathbf{H}$ of quaternions, where in the latter case $\sigma = \tau$ is the conjugation. If α is an eigenvalue of the symmetric endomorphism a of $\boldsymbol{V}$, and if there exists a non-isotropic eigenvector $\mathfrak{x}$ for the eigenvalue α, then $\alpha = \sigma(\alpha)$. If a is skew-symmetric, then, under the same assumptions, $\alpha = -\sigma(\alpha)$.*

Proof. We have $\langle a\mathfrak{x}, \mathfrak{x} \rangle = \langle \mathfrak{x}, a\mathfrak{x} \rangle$, since a is symmetric. The eigenvector property implies

$$\sigma(\alpha)\langle \mathfrak{x}, \mathfrak{x} \rangle = \langle \mathfrak{x}, \mathfrak{x} \rangle \alpha.$$

If K is a field and $\mathfrak{x}$ is not isotropic, then this yields the assertion. For a τ-Hermitean scalar product over the quaternions note that $\langle \mathfrak{x}, \mathfrak{x} \rangle \in \mathbf{R}$, i.e., this number belongs to the center of $\mathbf{H}$; hence, $\alpha \in \mathbf{R}$. For a skew-symmetric endomorphism, the only thing to be changed in the equation for the scalar product is the sign. □

The following simple example shows that, for indefinite pseudo-unitary scalar products there are symmetric endomorphisms having only isotropic eigenvectors; for the eigenvalues we then have $\alpha \neq \sigma(\alpha)$.

Example 6. Consider the two-dimensional pseudo-unitary vector space $\boldsymbol{V}$ of index 1 over the field $\mathbf{C}$ of complex numbers. If $(\xi^l), (\eta^l)$ are the coordinates of the vectors $\mathfrak{x}, \mathfrak{y}$ in the orthonormal basis $(\mathfrak{e}_1, \mathfrak{e}_2)$, then their scalar product is

$$\langle \mathfrak{x}, \mathfrak{y} \rangle = \bar{\xi}^1 \eta^1 - \bar{\xi}^2 \eta^2 .$$

Hence, $\mathfrak{a}_1 := \mathfrak{e}_1 + \mathfrak{e}_2, \mathfrak{a}_2 := \mathfrak{e}_1 - \mathfrak{e}_2$ are isotropic vectors also forming a basis. A straightforward calculation shows that the endomorphism defined by

$$a(\mathfrak{a}_1) = \mathfrak{a}_1 \, \mathrm{i}, \; a(\mathfrak{a}_2) = -\mathfrak{a}_2 \, \mathrm{i}$$

is symmetric; here i denotes the imaginary unit. Obviously, this endomorphism has the required properties. Note, in addition, that the eigensubspaces are not orthogonal; in fact, $\langle \mathfrak{a}_1, \mathfrak{a}_2 \rangle = 2$. □

Exercise 18. Under the assumptions of the preceding example show the following: If the symmetric linear map $b \in \operatorname{End}(\boldsymbol{V})$ has the eigenvectors $\mathfrak{a}_1, \mathfrak{a}_2$, then the corresponding eigenvalues are conjugate.

The invariants of an endomorphism a for the action of the linear group $\boldsymbol{GL}(\boldsymbol{V})$ are determined by its Jordan normal form if K contains all eigenvalues of a, i.e., if it is a splitting field for the characteristic polynomial of a, cf. Proposition I.5.8.4. In this connection we define the notion of a *root subspace* $\boldsymbol{U}_\alpha$ of an endomorphism $a \in \operatorname{End}(\boldsymbol{V})$ for the scalar $\alpha \in K$:

$$\boldsymbol{U}_\alpha := \{\mathfrak{x} \in \boldsymbol{V} | \text{ there is an } s \in \mathbf{N}_0 \text{ with } (a - e\alpha)^s(\mathfrak{x}) = \mathfrak{o}\}.$$

Here and in the sequel we set $e := \mathrm{id}_{\boldsymbol{V}}$. In § I.5.8 these spaces are called *nil-subspaces*. It is easy to see that each root subspace is invariant for a. Obviously, the root subspace satisfies $\boldsymbol{U}_\alpha \neq \mathfrak{o}$ if and only if α is an eigenvalue of a. We show

Proposition 24. *Let $\boldsymbol{V}$ be a vector space with scalar product over a field K, and let $a \in \operatorname{End}(\boldsymbol{V})$ be a symmetric endomorphism. If $\sigma(\alpha) \neq \beta$, then the root subspaces $\boldsymbol{U}_\alpha, \boldsymbol{U}_\beta$ are mutually orthogonal.* [2]

---

[2] This proposition corrects Proposition 5 in Section 101, p. 250, in A. I. Maltzev [75]. There, due to a mistake in the proof, $\alpha \neq \beta$ is assumed instead of $\alpha \neq \sigma(\beta)$. Example 6 shows that the proposition as formulated there is wrong.

P r o o f. If α or β are no eigenvalues of a, then the assertion is trivial. Take now $\mathfrak{x} \in \boldsymbol{U}_\alpha$ and $\mathfrak{y} \in \boldsymbol{U}_\beta$. So there are natural numbers $s, t \in \mathbf{N}_0$ for which

$$(a - e\alpha)^s(\mathfrak{x}) = \mathfrak{o},\ (a - e\beta)^t(\mathfrak{y}) = \mathfrak{o}.$$

We proceed by induction on the number $s + t$. For $s + t = 1$ we have $s = 0$ or $t = 0$, and the assertion is trivially satisfied. Assume now that the claim is true for all s, t with $s + t \leq k$, and take $s + t = k + 1$. Set

$$\mathfrak{x}_1 := (a - e\alpha)(\mathfrak{x}), \mathfrak{y}_1 := (a - e\beta)(\mathfrak{y}),$$

Then

$$(a - e\alpha)^{s-1}(\mathfrak{x}_1) = \mathfrak{o},\ (a - e\beta)^{t-1}(\mathfrak{y}_1) = \mathfrak{o},$$

i.e., the induction hypothesis is satisfied for the vectors $\mathfrak{x}_1, \mathfrak{y}$ and $\mathfrak{x}, \mathfrak{y}_1$. Hence,

$$\langle \mathfrak{x}_1, \mathfrak{y} \rangle = \langle (a - e\alpha)(\mathfrak{x}), \mathfrak{y} \rangle = \langle a\mathfrak{x}, \mathfrak{y} \rangle - \sigma(\alpha)\langle \mathfrak{x}, \mathfrak{y} \rangle = 0, \tag{51}$$
$$\langle \mathfrak{x}, \mathfrak{y}_1 \rangle = \langle \mathfrak{x}, (a - e\beta)(\mathfrak{y}) \rangle = \langle \mathfrak{x}, a\mathfrak{y} \rangle - \beta\langle \mathfrak{x}, \mathfrak{y} \rangle = 0. \tag{52}$$

Because of the symmetry of a, taking the difference yields $(\sigma(\alpha) - \beta)\langle \mathfrak{x}, \mathfrak{y} \rangle = 0$. Since by assumption $\sigma(\alpha) \neq \beta$, we arrive at the assertion $\langle \mathfrak{x}, \mathfrak{y} \rangle = 0$. □

Corollary 25. *If under the assumptions of Proposition* 24 α *is an eigenvalue of* a *with* $\alpha \neq \sigma(\alpha)$, *then the associated root subspace* $\boldsymbol{U}_\alpha$ *is totally isotropic.*

P r o o f. With $\beta = \alpha$ the assumptions of Proposition 24 are satisfied. Hence $\boldsymbol{U}_\alpha$ is orthogonal to itself. □

Almost identical proofs lead to corresponding results for skew-symmetric operators:

Proposition 26. *Let* $\boldsymbol{V}$ *be a vector space with scalar product over a field* K, *and let* $a \in \mathrm{End}(\boldsymbol{V})$ *be a skew-symmetric endomorphism. If* $\sigma(\alpha) \neq -\beta$, *then the root subspaces* $\boldsymbol{U}_\alpha, \boldsymbol{U}_\beta$ *are mutually orthogonal.*

P r o o f. Because of the skew-symmetry of the endomorphism, we just have to add equations (51) and (52) instead of taking the difference; everything else remains unchanged. □

Corollary 27. *Under the assumptions of Proposition* 26 *let* α *be an eigenvalue of* a *with the property* $\alpha \neq -\sigma(\alpha)$. *Then the associated root subspace* $\boldsymbol{U}_\alpha$ *is totally isotropic.* □

Let us again consider symmetric endomorphisms, and define

$$\boldsymbol{W}_0 := \bigoplus_{\alpha \neq \sigma(\alpha)} \boldsymbol{U}_\alpha,\ \boldsymbol{W}_1 := \bigoplus_{\alpha = \sigma(\alpha)} \boldsymbol{U}_\alpha. \tag{53}$$

The fact that the sums occurring in (53) are direct immediately follows from Proposition I.5.8.4 concerning the Jordan normal form applied to the split endomorphism $a|\Sigma \boldsymbol{U}_\alpha$, where the sum runs over all eigenvalues of a belonging to K. Recall that an endomorphism a is said to be *split* if the scalar domain K is a splitting field for the characteristic polynomial of a. Proposition 24 then immediately implies

Corollary 28. *Under the assumptions of Proposition* 24 *the subspace* $\boldsymbol{W}_0$ *is orthogonal to* $\boldsymbol{W}_1$. *If* K *is a splitting field for the characteristic polynomial of* a, *then*

$$\boldsymbol{V} = \boldsymbol{W}_0 \oplus \boldsymbol{W}_1$$

is an orthogonal decomposition into subspaces invariant for a; *the restrictions of* a *to these subspaces are symmetric endomorphisms.*

Proof. The first assertion is immediate, since each summand occurring in $\boldsymbol{W}_0$ is orthogonal to every summand in $\boldsymbol{W}_1$. If K is a splitting field, then the decomposition of $\boldsymbol{V}$ is just a regrouping of the direct decomposition into root subspaces determined by the Jordan normal form, hence itself a direct sum of mutually orthogonal subspaces, on which the scalar product consequently is non-degenerate, cf. Lemma 22. From here the proof is easily completed. □

Let, moreover, a be a symmetric, split endomorphism. By Proposition 24, the root subspaces occurring in the direct sum representation of $\boldsymbol{W}_1$ are mutually orthogonal. As Example 6 shows, this does not hold for the direct summands in $\boldsymbol{W}_0$. For the decomposition of $\boldsymbol{W}_0$ we have

Proposition 29. *Let* a *be a symmetric, split endomorphism of the vector space* $\boldsymbol{V}$ *with scalar product over the field* K. *If* $\lambda \neq \sigma(\lambda)$ *is an eigenvalue of* a, *then* $\sigma(\lambda)$ *is an eigenvalue, too.* $\boldsymbol{U}_\lambda \oplus \boldsymbol{U}_{\sigma(\lambda)}$ *is a neutral subspace; hence*

$$\dim \boldsymbol{U}_\lambda = \dim \boldsymbol{U}_{\sigma(\lambda)}.$$

Choose one entry from each pair $(\lambda, \sigma(\lambda))$ *and denote the resulting set of representatives by* Λ_o. *Then*

$$\boldsymbol{W}_0 = \bigoplus_{\lambda \in \Lambda_0} (\boldsymbol{U}_\lambda \oplus \boldsymbol{U}_{\sigma(\lambda)}) \tag{54}$$

is a direct orthogonal decomposition into neutral subspaces.

Proof. According to Proposition 24 the summands occurring in the decomposition (54) are mutually orthogonal. Suppose that $\sigma(\lambda)$ is not an eigenvalue of a, then $\boldsymbol{U}_\lambda$ is a totally isotropic subspace orthogonal to $\boldsymbol{V}$ contradicting the assumption that the scalar product is non-degenerate. By Lemma 22 no direct summand of (54) is isotropic, hence each one is a neutral subspace. Exercise 6 implies the assertion concerning the dimensions. □

Similar to (53) we define for skew-symmetric endomorphisms

$$\boldsymbol{M}_0 := \bigoplus_{\alpha \neq -\sigma(\alpha)} \boldsymbol{U}_\alpha, \; \boldsymbol{M}_1 := \bigoplus_{\alpha = -\sigma(\alpha)} \boldsymbol{U}_\alpha. \tag{55}$$

Proposition 26 immediately leads to

Corollary 30. *Under the assumptions of Proposition 26 $\boldsymbol{M}_0$ is orthogonal to $\boldsymbol{M}_1$. If K is a splitting field for the characteristic polynomial of a, then*

$$\boldsymbol{V} = \boldsymbol{M}_0 \oplus \boldsymbol{M}_1$$

is an orthogonal decomposition into subspaces with scalar product invariant for a; the restriction of a to any of these subspaces is a skew-symmetric endomorphism. □

The analogue of Proposition 29 is proved almost literally repeating the preceding argument:

Proposition 31. *Let a be a skew-symmetric, split endomorphism of the vector space $\boldsymbol{V}$ with scalar product over the field K. If $\lambda \neq -\sigma(\lambda)$ is an eigenvalue of a, then $-\sigma(\lambda)$ is an eigenvalue, too. $\boldsymbol{U}_\lambda \oplus \boldsymbol{U}_{-\sigma(\lambda)}$ is a neutral subspace; hence $\dim \boldsymbol{U}_\lambda = \dim \boldsymbol{U}_{-\sigma(\lambda)}$. Choose one entry from each pair $(\lambda, -\sigma(\lambda))$ and denote the resulting set of representatives by Λ_o. Then*

$$\boldsymbol{M}_0 = \bigoplus_{\lambda \in \Lambda_0} (\boldsymbol{U}_\lambda \oplus \boldsymbol{U}_{-\sigma(\lambda)}) \tag{56}$$

is a direct orthogonal decomposition into neutral subspaces. □

2.2.11 Properties of the Root Subspaces

In this section we will derive some special properties of the root subspaces for symmetric or skew-symmetric endomorphisms. We will frequently refer to the following assumption:

Assumption A. Let $\boldsymbol{V}$ be a vector space with scalar product over a field K, $\operatorname{char} K \neq 2$, and let $a \in \operatorname{End}(\boldsymbol{V})$ be a split endomorphism.

Recall that for a split endomorphism the root subspaces $\boldsymbol{U}_\lambda$ can be represented as the direct sum of *Jordan cells* Δ_λ^k:

$$\boldsymbol{U}_\lambda \cong \Delta_\lambda^{k_1} \oplus \Delta_\lambda^{k_2} \oplus \ldots \oplus \Delta_\lambda^{k_r}, \; k_1 \geq k_2 \geq \ldots \geq k_r. \tag{57}$$

Here Δ_λ^k is a subspace invariant for a with a basis $(\mathfrak{a}_i)$, $i = 1, \ldots, k$, in which $a|\Delta_\lambda^k$ is represented by a matrix of the form

$$\begin{pmatrix} \lambda & 0 & \ldots & 0 & 0 \\ 1 & \lambda & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda & 0 \\ 0 & 0 & \ldots & 1 & \lambda \end{pmatrix}. \tag{58}$$

Finding the Jordan normal form for an endomorphism amounts to representing the vector space $\boldsymbol{V}$ as the direct sum of Jordan cells. The matrix of a in a basis adapted to these Jordan cells is quasi-diagonal; along the diagonal it contains *Jordan boxes* , i.e. matrices of the form (58). The type of a Jordan normal form, i.e. the sequences of the eigenvalues and the dimensions of the Jordan cells corresponding to the eigenvalues, are uniquely determined up to the order of the summands by the endomorphism a. Jordan cells may occur repeatedly; the number of cells of the form Δ_λ^k is called the *multiplicity* of this cell in the Jordan decomposition for a. In general, the Jordan cells themselves are not uniquely determined; only the type of the Jordan normal form, i.e. the eigenvalues, the dimensions of the corresponding Jordan cells, and their multiplicity, are invariants of the endomorphism. Two split linear endomorphisms are similar if and only if their Jordan normal forms coincide up to the order of the summands (cf. Section I.5.8 in [82]). Let G denote the group of automorphisms of the vector space with scalar product $[\boldsymbol{V}, \langle,\rangle]$; then G also acts on $\mathrm{End}(\boldsymbol{V})$ by

$$(g, a) \in G \times \mathrm{End}(\boldsymbol{V}) \longmapsto g \circ a \circ g^{-1} \in \mathrm{End}(\boldsymbol{V}),$$

i.e. by restricting the action of the linear automorphisms to $G \subset \boldsymbol{GL}(\boldsymbol{V})$. Two linear maps transformed into one another under this action are called *G-congruent.* Hence the equality of the Jordan normal forms is a necessary, but in general not sufficient condition for the G-congruence of two linear maps.

The specific properties of the Jordan normal forms for symmetric or skew-symmetric endomorphisms will be studied in greater detail now.

Proposition 32. *Under Assumption A, let $\lambda \neq \sigma(\lambda)$ be an eigenvalue of the symmetric endomorphism a. For each Jordan cell Δ_λ^k in the Jordan normal form of a there is a Jordan cell $\Delta_{\sigma(\lambda)}^k$ of the same dimension k so that $\Delta_\lambda^k \oplus \Delta_{\sigma(\lambda)}^k$ is a neutral subspace of $\boldsymbol{V}$. The multiplicities of Δ_λ^k and $\Delta_{\sigma(\lambda)}^k$ in the Jordan decomposition of a coincide.*

Proof. By Proposition 29 it suffices to prove the statement separately for each of the summands in (54); thus we may assume

$$\boldsymbol{V} = \boldsymbol{U}_\lambda \oplus \boldsymbol{U}_{\sigma(\lambda)}.$$

The proof will proceed by induction on the dimension $m = \dim \boldsymbol{U}_\lambda$. For $m = 1$ the assertion immediately follows from Proposition 29. Suppose that the proposition is already proved for all dimensions less than m, and let k be the minimum of all the dimensions of Jordan cells occurring in the Jordan decomposition of a. Let the cell Δ_λ^k belong to $\boldsymbol{U}_\lambda$; if necessary, interchange λ and $\sigma(\lambda)$. Then there is a decomposition (57) with $k = k_r$. Set

$$\hat{\boldsymbol{U}}_\lambda := \Delta_\lambda^{k_1} \oplus \Delta_\lambda^{k_2} \oplus \ldots \oplus \Delta_\lambda^{k_{r-1}}.$$

In Exercise 6 we noted that in a decomposition of a neutral vector space into the direct sum of totally isotropic subspaces each of these subspaces determines a canonical σ-linear bijection onto the dual of the other subspace. Consider the decomposition $\boldsymbol{U}_\lambda = \hat{\boldsymbol{U}}_\lambda \oplus \Delta_\lambda^k$ and define

$$\boldsymbol{B} := \{\mathfrak{x} \in \boldsymbol{U}_{\sigma(\lambda)} | \langle \mathfrak{x}, \hat{\boldsymbol{U}}_\lambda \rangle = 0\} = \hat{\boldsymbol{U}}_\lambda^\perp \cap \boldsymbol{U}_{\sigma(\lambda)},$$
$$\boldsymbol{M} := \{\mathfrak{x} \in \boldsymbol{U}_{\sigma(\lambda)} | \langle \mathfrak{x}, \Delta_\lambda^k \rangle = 0\} = (\Delta_\lambda^k)^\perp \cap \boldsymbol{U}_{\sigma(\lambda)}.$$

>From the result of Exercise 6 mentioned above we immediately obtain

$$\dim \boldsymbol{B} = k,\ \dim \boldsymbol{M} = m - k,\ \boldsymbol{U}_{\sigma(\lambda)} = \boldsymbol{M} \oplus \boldsymbol{B}.$$

The invariance of the subspaces $\hat{\boldsymbol{U}}_\lambda, \Delta_\lambda^k$ under the action of a implies the invariance of the subspaces $\boldsymbol{B}, \boldsymbol{M}$ under the map a. A Jordan decomposition for the restrictions of a to these subspaces leads to a Jordan decomposition for the restriction of a to $\boldsymbol{U}_{\sigma(\lambda)}$. Since k is the minimum of the dimensions of all Jordan cells for a, $\boldsymbol{B}$ itself has to be a Jordan cell of the form $\Delta_{\sigma(\lambda)}^k$; we identify $\boldsymbol{B} = \Delta_{\sigma(\lambda)}^k$. It remains to show that $\boldsymbol{H} := \Delta_\lambda^k \oplus \Delta_{\sigma(\lambda)}^k$ is a neutral subspace of $\boldsymbol{V}$. Let $\mathfrak{x}$ be a vector from the defect subspace $\boldsymbol{D} = \boldsymbol{H} \cap \boldsymbol{H}^\perp$. Then

$$\mathfrak{x} \in (\Delta_\lambda^k \oplus \Delta_{\sigma(\lambda)}^k)^\perp = (\Delta_\lambda^k)^\perp \cap (\Delta_{\sigma(\lambda)}^k)^\perp.$$

Since $\mathfrak{x} \in \boldsymbol{H}$, we have the representation

$$\mathfrak{x} = \mathfrak{x}_0 + \mathfrak{x}_1 \text{ with } \mathfrak{x}_0 \in \Delta_\lambda^k \subset \boldsymbol{U}_\lambda, \mathfrak{x}_1 \in \Delta_{\sigma(\lambda)}^k \subset \boldsymbol{U}_{\sigma(\lambda)},$$

Now $\mathfrak{x}$ also belongs to

$$(\Delta_{\sigma(\lambda)}^k)^\perp = \boldsymbol{B}^\perp = \hat{\boldsymbol{U}}_\lambda \oplus \boldsymbol{U}_{\sigma(\lambda)},$$

since for the neutral vector space $\boldsymbol{V}$ the following relations hold

$$\boldsymbol{U}_\lambda^\perp = \boldsymbol{U}_\lambda,\ \boldsymbol{U}_{\sigma(\lambda)}^\perp = \boldsymbol{U}_{\sigma(\lambda)}.$$

Hence we obtain for the first component of $\mathfrak{x}$:

$$\mathfrak{x}_0 \in \Delta_\lambda^k \cap \hat{\boldsymbol{U}}_\lambda = \{\mathfrak{o}\}.$$

Thus, because of

$$(\Delta_\lambda^k)^\perp = \boldsymbol{U}_\lambda \oplus ((\Delta_\lambda^k)^\perp \cap \boldsymbol{U}_{\sigma(\lambda]}) = \boldsymbol{U}_\lambda \oplus \boldsymbol{M},$$

we have

$$\mathfrak{x} = \mathfrak{x}_1 \in \Delta_{\sigma(\lambda)}^k \cap \boldsymbol{M} = \boldsymbol{B} \cap \boldsymbol{M} = \{\mathfrak{o}\}.$$

Therefore $\mathfrak{x} = \mathfrak{o}$, and $\boldsymbol{H}$ is an a-invariant neutral subspace. Hence $\boldsymbol{H}^\perp$ is also a-invariant. Since the dimension of the root subspace of $\boldsymbol{H}^\perp$ corresponding to λ is less than m, the assertion follows from the induction hypothesis. Note that, once the Jordan normal form of a is known, the proof provides an efficient procedure to determine the decomposition of $\boldsymbol{U}_\lambda \oplus \boldsymbol{U}_\sigma(\lambda)$ into the direct sum of mutually orthogonal neutral subspaces of the form $\Delta_\lambda^k \oplus \Delta_{\sigma(\lambda)}^k$. Here, k runs through the dimensions of the Jordan cells occurring in $\boldsymbol{U}_\lambda$. □

The corresponding proposition for skew-symmetric endomorphisms, proved along the same lines, is

Proposition 33. *Under Assumption A let $\lambda \neq -\sigma(\lambda)$ be an eigenvalue of the skew-symmetric endomorphism a. For each Jordan cell Δ^k_λ in a Jordan normal form of a there is a Jordan cell $\Delta^k_{-\sigma(\lambda)}$ of equal dimension k so that $\Delta^k_\lambda \oplus \Delta^k_{-\sigma(\lambda)}$ is a neutral subspace of $\boldsymbol{V}$. The multiplicities of Δ^k_λ and $\Delta^k_{-\sigma(\lambda)}$ in the Jordan decomposition for a coincide.* □

The following two exercises again imply that the unitary spaces with their positive definite scalar product are particularly fortunate exceptions; since there are no isotropic vectors, their symmetric and skew-symmetric endomorphisms have very simple structures.

Exercise 19. Prove supposing that Assumption A holds: If Δ^k_λ is a Jordan cell of a symmetric or skew-symmetric endomorphism with $k > 1$, then the eigenvector of this cell, i.e. the vector $\mathfrak{a}_k$ in a corresponding Jordan basis, is isotropic.

Exercise 20. Under Assumption A let a be a symmetric (or skew-symmetric) endomorphism, let $\lambda \neq \sigma(\lambda)$ (or $\lambda \neq -\sigma(\lambda)$) and μ be eigenvalues of a. Prove: If Δ^k_λ, Δ^l_μ are Jordan cells for a, and if $\mu \neq \sigma(\lambda)$ (or $\mu \neq -\sigma(\lambda)$) or $k \neq l$, then the subspace $\Delta^k_\lambda \oplus \Delta^l_\mu$ is isotropic.

Exercise 21. Under Assumption A let a be a symmetric (or skew-symmetric) endomorphism, and let $\lambda \neq \sigma(\lambda)$ (or $\lambda \neq -\sigma(\lambda)$) be an eigenvalue of a. Consider one of the neutral subspaces occurring in Proposition 32 (or 33), $\boldsymbol{B} = \Delta^k_\lambda \oplus \Delta^k_{\sigma(\lambda)}$ (or $\boldsymbol{B} = \Delta^k_\lambda \oplus \Delta^k_{-\sigma(\lambda)}$). Prove: The restriction $a|\Delta^k_\lambda$ uniquely determines the restriction of a to the other direct summand of $\boldsymbol{B}$. (Hint. Apply Exercise 6c).)

The following proposition shows that, for sufficiently large index of the scalar product, arbitrary endomorphisms, and hence also arbitrary Jordan normal forms, may occur as components of symmetric or skew-symmetric endomorphisms. Therefore, enumerative classifications will rapidly become complicated with increasing index and dimension.

Proposition 34. *Let $\boldsymbol{V} = \boldsymbol{B}_0 \oplus \boldsymbol{B}_1$ be a neutral vector space represented as the direct sum of two totally isotropic subspaces. Then each linear endomorphism a of $\boldsymbol{B}_0$ can be uniquely extended to a symmetric endomorphism b of $\boldsymbol{V}$ leaving $\boldsymbol{B}_1$ invariant. Similarly, there exists a uniquely determined skew-symmetric extension with this property.*

Proof. Let $(\mathfrak{a}_i), i = 1, \ldots, m$, be an arbitrary basis of $\boldsymbol{B}_0$, and let $(\mathfrak{b}_i)$ be the basis of $\boldsymbol{B}_1$ satisfying

$$\langle \mathfrak{a}_i, \mathfrak{b}_j \rangle = \delta_{ij}, \ i, j = 1, \ldots, m,$$

that is uniquely determined according to Exercise 6c). If (α^i_j) is the matrix of a, then this relation immediately implies that the restriction of the symmetric extension b we are looking for has the transposed σ-conjugate matrix to (α^i_j) in the basis $(\mathfrak{b}_i)$ of $\boldsymbol{B}_1$. Using the sum convention and setting

$$b(\mathfrak{b}_j) = \mathfrak{b}_k \beta_j^k,$$

the extension property implies

$$\beta_j^k = \langle \mathfrak{a}_k, b(\mathfrak{b}_j)\rangle = \langle a(\mathfrak{a}_k), \mathfrak{b}_j\rangle = \sigma(\alpha_k^j).$$

For the matrix of the skew-symmetric extension we similarly obtain $(\gamma_j^i) = -(\sigma(\alpha_j^i))'$. □

2.3 The Projective Geometry of a Polarity

In this section we consider σ-Hermitean and, especially, symmetric bilinear scalar products from the projective point of view. So we only discuss polarities; in order to investigate degenerate cases one could proceed as in Exercise 1.9.7, i.e. consider polarities on a cone section. Here we will always work in a right vector space $\boldsymbol{V}^{n+1}$ over a skew field K, char $K \neq 2$, with a σ-Hermitean scalar product $\langle , \rangle$. $\boldsymbol{P}^n$ denotes the projective space associated with $\boldsymbol{V}$, and F is the polarity defined by $\langle , \rangle$ determining the, possibly empty, quadric $Q = Q_F$. Finally, $\boldsymbol{PG}_n = \boldsymbol{PG}_n(F)$ is the projective group of the polarity F.

Obviously, these assumptions will later have to be specialized; e.g., in the second part of this section we will consider non-degenerate, symmetric bilinear forms.

2.3.1 The Quadric of a Polarity

For the projective interpretation of the notions related to the scalar product we recall that the subspace $\boldsymbol{W}^\perp$ orthogonal to any subspace $\boldsymbol{W}$ with respect to $\langle , \rangle$ is the one assigned to it by means of the polarity (cf. (1.7.22)):

$$F: \ \boldsymbol{W} \in \mathfrak{P} \longmapsto F(\boldsymbol{W}) := \boldsymbol{W}^\perp \in \mathfrak{P}.$$

The totally isotropic subspaces are those contained in the quadric, $\boldsymbol{E} \subset Q$ (Corollary 1.9.15); by Corollary 2.6 their projective dimension satisfies the inequality

$$\operatorname{Dim} \boldsymbol{E} \leq (n-1)/2. \tag{1}$$

In Section 1.9 we pointed out that an isotropic subspace intersect the quadric Q in its defect subspace $\boldsymbol{W}_o = \boldsymbol{W} \cap \boldsymbol{W}^\perp$ (Corollary 1.9.19). E. Witt's Theorem immediately implies

Corollary 1. *Two points $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{P}^n$ are $\boldsymbol{PG}_n$-equivalent if they can be represented by vectors $\mathfrak{a}, \mathfrak{b} \in \boldsymbol{V}$ satisfying*

$$\boldsymbol{a} = [\mathfrak{a}], \ \boldsymbol{b} = [\mathfrak{b}], \ \langle \mathfrak{a}, \mathfrak{a}\rangle = \langle \mathfrak{b}, \mathfrak{b}\rangle. \tag{2}$$

Dually, two hyperplanes $\boldsymbol{A}, \boldsymbol{B} \subset \boldsymbol{P}^n$ are $\boldsymbol{PG}_n$-equivalent if they can be represented by equations

$$\boldsymbol{A}: \ \langle \mathfrak{a}, \mathfrak{x} \rangle = 0, \qquad \boldsymbol{B}: \ \langle \mathfrak{b}, \mathfrak{x} \rangle = 0$$

with $\boldsymbol{PG}_n$-equivalent poles $\boldsymbol{A}^\perp = [\mathfrak{a}]$, $\boldsymbol{B}^\perp = [\mathfrak{b}]$ whose representing vectors satisfy (2). In particular, $\boldsymbol{PG}_n$ acts transitively on the points as well as on the tangent hyperplanes of Q. □

In the next Section the orbits of points, or, more generally, of point sequences, will be examined more closely. For now we only note the following

Example 1. Let $K = \mathbf{R}$, and let $F = F_{n+1,l}$ be a polarity of index l, where $0 \leq l \leq (n+1)/2$. For $0 < l < (n+1)/2$ there are three orbits of $\boldsymbol{PO}(l, n+1-l)$ in $\boldsymbol{P}^n$:

$$\begin{aligned} Q &= \{\boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{P}^n | \langle \mathfrak{x}, \mathfrak{x} \rangle = 0\}, \\ I(Q) &:= \{\boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{P}^n | \langle \mathfrak{x}, \mathfrak{x} \rangle < 0\}, \\ A(Q) &:= \{\boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{P}^n | \langle \mathfrak{x}, \mathfrak{x} \rangle > 0\}. \end{aligned}$$

The set $I(Q)$ is called the *inner region*, and $A(Q)$ the *outer region* of the quadric. Because of $\langle \mathfrak{x}\lambda, \mathfrak{x}\lambda \rangle = \langle \mathfrak{x}, \mathfrak{x} \rangle \lambda^2$, with a suitable normalization we can always represent the point $\boldsymbol{x}$ by a vector $\mathfrak{x}$ satisfying

$$\langle \mathfrak{x}, \mathfrak{x} \rangle = 0, \ \langle \mathfrak{x}, \mathfrak{x} \rangle = -1 \ \text{ or } \langle \mathfrak{x}, \mathfrak{x} \rangle = 1. \tag{3}$$

In the case $l = 0$ we have $Q = I(Q) = \emptyset$, and $A(Q) = \boldsymbol{P}^n$; the group $\boldsymbol{PO}(n+1)$ acts transitively on $\boldsymbol{P}^n$. The geometry determined by this action on the projective space is called *elliptic geometry*; we will discuss it in greater detail below. If $l = 1$, then the point $\boldsymbol{x}$ belongs to $I(Q)$ if and only if its polar $F(\boldsymbol{x})$ does not intersect the quadric Q. For all l such that $0 < l < (n+1)/2$ a point $\boldsymbol{x}$ belongs to $A(Q)$ if and only if its polar $F(\boldsymbol{x})$ intersects the quadric Q in a quadric $F(\boldsymbol{x}) \cap Q$ of the same index l, whereas the polar $F(\boldsymbol{x})$ of a point $\boldsymbol{x} \in I(Q)$ meets the quadric Q in a quadric of smaller index. To see this, it suffices to adapt any pseudo-orthonormal basis to the subspaces $\boldsymbol{x}$ and $\boldsymbol{x}^\perp = F(\boldsymbol{x})$ and then apply the Theorem of Inertia. The case $l = (n+1)/2$ holds a kind of special position, cf. Example 1.3 and Exercise 1. M. Stary [99] takes the properties of polars in real projective geometry mentioned here as the starting point for a definition of the inner and the outer region for a quadric in the case of an arbitrary field K with $\operatorname{char} K \neq 2$. □

Exercise 1. Let $Q \subset \boldsymbol{P}^n(\mathbf{R})$ be a non-degenerate quadric. Prove that there is a projectivity $a \in \boldsymbol{PO}(l, n+1-l)$ transforming the outer region $A(Q)$ into the inner region $I(Q)$ and conversely if and only if the index l of the polarity determining Q satisfies $l = (n+1)/2$.

Exercise 2. Consider the pseudo-Euclidean vector space $\boldsymbol{V} = \boldsymbol{V}^{n+1}$ of index l, $0 < l \leq (n+1)/2$. Prove that the associated quadric $Q_l \subset \boldsymbol{P}^n(\mathbf{R})$ can be represented as the orbit space

$$Q_l \approx (S^{l-1} \times S^{n-l})/\mathbf{Z}_2,$$

where $\mathbf{Z}_2 = \{\pm \mathrm{id}_{\boldsymbol{V}}\}$ and S^{l-1}, S^{n-l} are hyperspheres in two complementary subspaces of $\boldsymbol{V}$. (Hint. Write the equation of the quadric in the form

$$\sum_{a=1}^{l} |x^a|^2 = \sum_{k=l+1}^{n+1} |x^k|^2, \tag{4}$$

where the x^i, $i = 1, \ldots, n+1$, are the homogeneous coordinates of the point $\boldsymbol{x} \in Q$.) In particular, $Q_1 \approx S^{n-1}$ is a hypersphere of the projective space $\boldsymbol{P}^n(\mathbf{R})$. Simple differential-geometric arguments imply that the non-empty, non-degenerate quadrics Q_l are compact, connected hypersurfaces in $\boldsymbol{P}^n(\mathbf{R})$.

Exercise 3. Prove that under the assumptions of Exercise 2 the inner and the outer region of a quadric Q_l are open, connected subsets of $\boldsymbol{P}^n(\mathbf{R})$.

Exercise 4. Let $\boldsymbol{V}^n$ be the n-dimensional pseudo-Euclidean vector space of index l. By $E_{n,k,s,d}$ we denote the set of k-dimensional subspaces $\boldsymbol{H}^k \subset \boldsymbol{V}^n$ on which the scalar product has index s and defect d, $0 \leq s$, $0 \leq d$, $s + d \leq k$. Set $p := k - s - d$. Prove that $E_{n,k,s,d}$ is non-empty if and only if $s + d \leq l$ and $p + d \leq n - l$.

Example 2. Let $\boldsymbol{P}^n(\mathbf{C})$ be the complex orthogonal projective space with the quadric Q determined by a bilinear scalar product $\langle , \rangle$. Since each point $\boldsymbol{x} \in \boldsymbol{P}^n \setminus Q$ can be represented as $\boldsymbol{x} = [\mathfrak{x}]$, $\mathfrak{x} \in \boldsymbol{V}$ with $\langle \mathfrak{x}, \mathfrak{x} \rangle = 1$, the group $G := \boldsymbol{PO}(n+1, \mathbf{C})$ acts transitively both on Q and on $\boldsymbol{P}^n \setminus Q$. Under the action of G the Graßmann manifold $\boldsymbol{P}_{n,k}$ of k-planes decomposes into finitely many orbits, which are characterized by the rank r of $\langle , \rangle | \boldsymbol{W} \times \boldsymbol{W}$ on the corresponding $(k+1)$-dimensional subspaces $\boldsymbol{W} \subset \boldsymbol{V}$ (or, equivalently, by their defect $d = k - r + 1$). □

Exercise 5. Prove under the assumptions of Example 2: For every $d \in \mathbf{N}$ such that $0 \leq d \leq (n+1)/2$ and $d \leq k + 1$, there are subspaces $\boldsymbol{W}^{k+1} \subset \boldsymbol{V}^{n+1}$ with defect d. Discuss all the possible positions of a k-plane relative to the quadric Q for dimensions $n \leq 4$.

Exercise 6. Find a homeomorphism from the Graßmann manifold $\tilde{\boldsymbol{G}}_2(\mathbf{R}^{n+1})$ of oriented two-dimensional subspaces in the $(n+1)$-dimensional real vector space onto the quadric $Q \subset \boldsymbol{P}^n(\mathbf{C})$ from Example 2. Hint. Consider the complex vector space $\mathbf{C}^{n+1}$ of $\boldsymbol{P}^n(\mathbf{C})$ as the complexification of the real Euclidean vector space $\mathbf{R}^{n+1}$ for the bilinear extension of the scalar product, and take a positively oriented, orthonormal basis $(\mathfrak{a}_1, \mathfrak{a}_2)$ for $\boldsymbol{U}^2 \subset \mathbf{R}^{n+1}$. Justify the definition of the homeomorphism $\varphi(\boldsymbol{U}^2) := [\mathfrak{a}_1 + \mathfrak{a}_2\, \mathrm{i}]_{\mathbf{C}}$ and prove the claim.

Example 3. The Hermitean scalar products of index l in the projective geometries over $\mathbf{C}$ and $\mathbf{H}$ form a picture that is , in many ways, similar to the pseudo-orthogonal geometry. Again, the equation of the quadric Q_l can be written in the form (4), where now, of course, one has to take the norm square with respect to $\mathbf{C}$ or $\mathbf{H}$, respectively. For $l = 0$ the quadric Q_0 is again empty, and the projective unitary group $\boldsymbol{PU}(n+1)$ or $\boldsymbol{PSp}(n+1)$ acts transitively on the projective space $\boldsymbol{P}^n$. Otherwise, inner and outer regions can be defined for Q as in Example 1. Passing to the realifications (4) becomes an equation in the real coordinates, which already has the normal form (4), of course with $K = \mathbf{R}$; for every norm square one has to write two or four real squares, respectively. Hence, in the realification this leads to special cases of the quadrics considered in Example 1. In the complex case the inverse image $\theta^{-1}(Q_l)$ for the Hopf fibration $\theta : \boldsymbol{P}^{2n+1}(\mathbf{R}) \to \boldsymbol{P}^n(\mathbf{C})$, cf. (1.10.3), is a hyperquadric $Q_{2l} \subset \boldsymbol{P}^{2n+1}(\mathbf{R})$, whereas in the quaternionic case the quadric Q_l has a real hyperquadric Q_{4l} as its inverse image under the Hopf fibration $\theta : \boldsymbol{P}^{4n+3}(\mathbf{R}) \to \boldsymbol{P}^n(\mathbf{H})$. Restricting the Hopf fibrations to these quadrics leads to corresponding fibrations of Q_{2l} over $Q_l(\mathbf{C})$ into one-dimensional and of Q_{4l} over $Q_l(\mathbf{H})$ into three-dimensional real projective spaces, respectively. In the simplest case $(l, n) = (1, 1)$ the base $Q_1(\mathbf{C})$ is the circle

$$|z^0|^2 = |z^1|^2, \text{ i.e. } |\zeta| = |z^1/z^0| = 1$$

in the Riemann sphere $S^2 = \boldsymbol{P}^1(\mathbf{C})$, and $\theta^{-1}(Q_1(\mathbf{C}))$ is the quadric $Q_2 \subset \boldsymbol{P}^3(\mathbf{R})$ described by

$$(x^0)^2 + (y^0)^2 = (x^1)^2 + (y^1)^2.$$

Note that by restricting the scalars the polarity induces more involved structures on the inverse images than just the polarities corresponding to the mentioned (real) quadrics (cf. Example 1.10.12 and Exercise 1.10.16). $\square$

Example 4. Little appears to be known concerning the quadrics associated with the quaternionic skew Hermitean polarities H_n in $\boldsymbol{P}^{n-1}(\mathbf{H})$. Let b denote the skew Hermitean form determining H_n. Because of $b(\mathfrak{x}, \mathfrak{x}) \in \mathbf{R}^\perp$, we obtain from $\langle \mathfrak{x}, \mathfrak{x} \rangle = -\mathrm{i}\, b(\mathfrak{x}, \mathfrak{x}) = 0$ three quadratic equations in real coordinates; hence the quadric they define is a $(4n-7)$-dimensional real submanifold of the $(4n-4)$-dimensional space $\boldsymbol{P}^{n-1}(\mathbf{H})$. Passing from Formulas (1.36) and (1.37) or their equivalent versions (1.38), (1.39) to the realification of $\boldsymbol{V}_{|\mathbf{C}}$, i.e. to the real coordinates, we see that this real submanifold is the intersection of three real hyperquadrics of index $2n$. For $n = 2$ it is a circle in $S^4 = \boldsymbol{P}^1(\mathbf{H})$, cf. Example 1.9.7. $\square$

2.3.2 Efficiency

We now again consider a general polarity F and will prove certain efficiency results for the actions of its projective group $\boldsymbol{PG}_n(F)$.

Proposition 2. *Let F be a polarity of $\boldsymbol{P}^n$ defined by the σ-biform $\langle , \rangle$, $n \geq 2$; suppose that the corresponding quadric Q_F is not empty. If for some $g \in \boldsymbol{PG}_n(F)$*

$$g|Q_F = \mathrm{id}_{Q_F},$$

then $g = \mathrm{id}_{\boldsymbol{P}^n}$.

Proof. We pass to the right vector space $\boldsymbol{V}^m$, $m = n+1$, with the σ-Hermitean scalar product $\langle , \rangle$ and prove the following statement equivalent to Proposition 2:

Proposition 3. *If $\boldsymbol{V}^m$, $m > 2$, contains isotropic vectors, and if for some $g \in \boldsymbol{CUG}_m$ and any isotropic vector $\mathfrak{v} \in \boldsymbol{V}$ the following relation holds*

$$g\mathfrak{v} = \mathfrak{v}c_{\mathfrak{v}} \textit{ with } c_{\mathfrak{v}} \in K^*,$$

then $g = \mathrm{id}_{\boldsymbol{V}} \cdot c$ for some $c \in Z(K)^*$.

Proof. By Lemma 2.4, for each isotropic vector $\mathfrak{v}$ there is an isotropic vector $\hat{\mathfrak{v}} \in \boldsymbol{V}$ linearly independent of it and satisfying $\langle \mathfrak{v}, \hat{\mathfrak{v}} \rangle = 1$. Then the linear span $\boldsymbol{W}^2 := \mathfrak{L}(\mathfrak{v}, \hat{\mathfrak{v}})$ is a non-isotropic subspace, and, moreover, $\boldsymbol{V} = \boldsymbol{W} \oplus \boldsymbol{W}^\perp$. Choose a basis $(\mathfrak{b}_\alpha)$, $\alpha = 1, \dots, m-2$, in the likewise non-isotropic subspace $\boldsymbol{W}^\perp$ containing at least one non-isotropic vector, say $\mathfrak{b}_1$. Set $\lambda_\alpha := \langle \mathfrak{b}_\alpha, \mathfrak{b}_\alpha \rangle$ and

$$\mathfrak{v}_\alpha := \mathfrak{b}_\alpha - \hat{\mathfrak{v}}\lambda_\alpha/2 + \mathfrak{v}. \tag{5}$$

It is then easy to verify that the vectors $\mathfrak{v}_\alpha$ are isotropic. By assumption there are elements $c, \hat{c}, c_\alpha \in K^*$ such that

$$g\mathfrak{v} = \mathfrak{v}c,\ g\hat{\mathfrak{v}} = \hat{\mathfrak{v}}\hat{c},\ g\mathfrak{v}_\alpha = \mathfrak{v}_\alpha c_\alpha,\ \alpha = 1, \dots, m-2.$$

Applying these relations to Definition (5) for the $\mathfrak{v}_\alpha$ we obtain

$$\begin{aligned} g\mathfrak{v}_\alpha &= g\mathfrak{b}_\alpha - g\hat{\mathfrak{v}}\lambda_\alpha/2 + g\mathfrak{v} = (\mathfrak{b}_\alpha - \hat{\mathfrak{v}}\lambda_\alpha/2 + \mathfrak{v})c_\alpha \\ &= g\mathfrak{b}_\alpha - \hat{\mathfrak{v}}\hat{c}\lambda_\alpha/2 + \mathfrak{v}c,\ \alpha = 1, \dots, m-2. \end{aligned}$$

Because of $m > 2$, there is at least one such relation. Since g preserves orthogonality and leaves the subspace $\boldsymbol{W}$, hence $\boldsymbol{W}^\perp$ as well, invariant, and as the vectors $\mathfrak{v}, \hat{\mathfrak{v}}, \mathfrak{b}_1, \dots, \mathfrak{b}_{m-2}$ form a basis of $\boldsymbol{V}$, by scalar multiplication with $\hat{\mathfrak{v}}$ we obtain the equations $c = c_\alpha$. The invariance of $\boldsymbol{W}^\perp$ implies $g\mathfrak{b}_\alpha = \mathfrak{b}_\alpha c$, i.e. $g|\boldsymbol{W}^\perp = \mathrm{id}_{\boldsymbol{W}^\perp}\, c$. The linearity of g implies $c \in Z(K)^*$. Comparing the coefficients of the vector $\hat{\mathfrak{v}}$ finally leads to

$$\hat{c}\lambda_\alpha/2 = (\lambda_\alpha/2)c = c(\lambda_\alpha/2).$$

Because of $\lambda_1 \neq 0$ we obtain for $\alpha = 1$ the assertion $\hat{c} = c$ and $g = \mathrm{id}_{\boldsymbol{V}}\, c$. □

Exercise 7. Prove that the assertion of Proposition 2 does, in general, not hold for $n = 1$.

Proposition 4. *Let K be a skew field of characteristic $\operatorname{char} K \neq 2, 3$, or suppose $n > 1$. Consider the polarity F corresponding to a σ-Hermitean scalar product. If all points $\boldsymbol{x} \in \boldsymbol{P}^n \setminus Q_F$ not belonging to the quadric of F are fixed points for the transformation $g \in \boldsymbol{PG}_n(F)$, then $g = \mathrm{id}_{\boldsymbol{P}^n}$.*

Proof. For $Q = Q_F = \emptyset$ the assertion is trivial. Suppose now $n > 1$, and let $\boldsymbol{v} = [\mathfrak{v}] \in Q$. By Lemma 2.4 there is a second isotropic vector $\hat{\mathfrak{v}}$ with $\langle \mathfrak{v}, \hat{\mathfrak{v}} \rangle = 1$. As a neutral subspace, $\boldsymbol{W}^2 := \mathfrak{L}(\mathfrak{v}, \hat{\mathfrak{v}})$ is not isotropic. Because of $\dim \boldsymbol{V} > 2$, there is a non-isotropic vector $\mathfrak{a} \neq \mathfrak{o}$ within the likewise non-isotropic subspace $\boldsymbol{W}^\perp$. Then $\mathfrak{b} = \mathfrak{v} + \mathfrak{a}$ is not isotropic as well, and by assumption there are $a, b \in K^*$ such that

$$g(\mathfrak{a}) = \mathfrak{a}\, a,\ g(\mathfrak{b}) = g(\mathfrak{v}) + g(\mathfrak{a}) = \mathfrak{b}\, b.$$

Hence also

$$g(\mathfrak{v}) = g(\mathfrak{b}) - g(\mathfrak{a}) = (\mathfrak{v} + \mathfrak{a})b - \mathfrak{a}a = \mathfrak{v}b + \mathfrak{a}(b - a).$$

Since $\mathfrak{v}$ is isotropic, this is also true of $g(\mathfrak{v})$, and hence $a = b$ as well as $g(\mathfrak{v}) = \mathfrak{v}b$. So the arbitrarily chosen point $\boldsymbol{v} = [\mathfrak{v}] \in Q$ is also fixed implying the assertion.

Suppose now $n = 1$, $Q \neq \emptyset$, and let $\mathfrak{v}$, $\hat{\mathfrak{v}}$ be defined as above. Then

$$\mathfrak{v}_+ := \mathfrak{v} + \hat{\mathfrak{v}},\ \mathfrak{v}_- := \hat{\mathfrak{v}} - \mathfrak{v} \tag{6}$$

is an orthogonal basis of $\boldsymbol{V}^2$ with

$$\langle \mathfrak{v}_+, \mathfrak{v}_+ \rangle = -\langle \mathfrak{v}_-, \mathfrak{v}_- \rangle = 2. \tag{7}$$

Consider the points

$$\boldsymbol{c}(t) := [\mathfrak{v}_+ + \mathfrak{v}_- t] \in \boldsymbol{P},\ t \in K.$$

For their vectors $\mathfrak{c}(t)$ we have

$$\langle \mathfrak{c}(t), \mathfrak{c}(t) \rangle = 2(1 - \sigma(t)t);$$

in fact, $2 \in Z(K)$: 2 belongs to the ring $\mathbf{Z}1 \subset Z(K)$. Because of $\operatorname{char} K \neq 2, 3$, there is a $t_o \in \mathbf{Z}1$ with $t_o \neq 0, -1, 1$. Since for all $t \in \mathbf{Z}1$ obviously $\sigma(t) = t$, we have $\boldsymbol{c}_o = [\mathfrak{c}(t_o)] \notin Q$, $\mathfrak{c}_o := \mathfrak{c}(t_o) \neq \mathfrak{v}_+$, $\mathfrak{c}_o \neq \mathfrak{v}_-$. By assumption there are scalars $b_+, b_-, c \in K^*$ for which

$$g(\mathfrak{v}_+) = \mathfrak{v}_+ b_+,\ g(\mathfrak{v}_-) = \mathfrak{v}_- b_-, g(\mathfrak{c}_o) = \mathfrak{c}_o c.$$

On the other hand,

$$g(\mathfrak{c}_o) = g(\mathfrak{v}_+) + g(\mathfrak{v}_-)t_o = \mathfrak{v}_+ b_+ + \mathfrak{v}_- b_- t_o = (\mathfrak{v}_+ + \mathfrak{v}_- t_o)c.$$

Comparing coefficients we obtain $c = b_+$, $b_- t_o = t_o c$. Now $t_o \neq 0$ and $t_o \in \mathbf{Z}1 \subset Z(K)$ imply $b_+ = b_- = c$. Since $(\mathfrak{v}_+, \mathfrak{v}_-)$ is a basis of $\boldsymbol{V}^2$, we again arrive at $g = \mathrm{id}_{\boldsymbol{V}}\, c$ for the linear map, i.e., projectively, $g = \mathrm{id}_{\boldsymbol{P}}$. □

Remark. If K is a field, then it suffices to show the existence of $n+2$ fixed points of g in general position, cf. Proposition 1.3.14. Formula (1.3.35) with $f = g$, $\boldsymbol{a}'_i = \boldsymbol{a}_i$, $\boldsymbol{e}' = \boldsymbol{e}$ leads to $g = \mathrm{id}_{\boldsymbol{P}^n}$.

2.3.3 Orthogonal Geometry. Reflections

In this section we will mainly be concerned with the geometries called *orthogonal* by E. Artin [4]. Let thus F be a polarity defined by a bilinear scalar product $\langle,\rangle$, where the scalar domain K is necessarily a field. The isotropy group of the scalar product defined on an n-dimensional vector space $\boldsymbol{V}^n$ is denoted by $\boldsymbol{O}_n = \boldsymbol{O}_n(\langle,\rangle)$ and called the *orthogonal group*. The most prominent special cases are the real groups $\boldsymbol{O}(n)$, $\boldsymbol{O}(l, n-l)$ and the complex orthogonal group $\boldsymbol{O}(n, \mathbf{C})$.

For the purposes of this section we want to specialize the notion of a reflection already introduced in Example 1.4.6. Let $\boldsymbol{W}^{n-1} \subset \boldsymbol{V}^n$ be a non-isotropic $(n-1)$-dimensional subspace; then its orthogonal complement $\boldsymbol{W}^\perp$ is a one-dimensional, non-isotropic subspace called the *normal* of $\boldsymbol{W}$; each vector $\mathfrak{n} \in \boldsymbol{W}^\perp$, $\mathfrak{n} \neq \mathfrak{o}$, is a *normal vector* of $\boldsymbol{W}$. Then, according to Example 1.4.6, the direct decomposition $\boldsymbol{V} = \boldsymbol{W} \oplus \boldsymbol{W}^\perp$ determines a reflection

$$s_W : \mathfrak{x}_0 + \mathfrak{x}_n \in \boldsymbol{V} \longmapsto \mathfrak{x}_0 - \mathfrak{x}_n \in \boldsymbol{V},\ \mathfrak{x}_0 \in \boldsymbol{W}, \mathfrak{x}_n \in \boldsymbol{W}^\perp,$$

which we call the *orthogonal reflection* in the non-isotropic $(n-1)$-dimensional subspace $\boldsymbol{W}^{n-1}$. In this section, a reflection will always understood to be such an orthogonal reflection. For a normal vector $\mathfrak{n}$ of $\boldsymbol{W}$, $\langle \mathfrak{n}, \mathfrak{n}\rangle = \rho \neq 0$, we have

$$s_W : \mathfrak{x} = \mathfrak{x} \in \boldsymbol{V} \longmapsto s_W(\mathfrak{x}) := \mathfrak{x} - \mathfrak{n} \cdot 2\langle \mathfrak{n}, \mathfrak{x}\rangle/\rho. \tag{8}$$

A straightforward calculation shows that $s_W \in \boldsymbol{O}_n$. The corresponding projective map will be called an *orthogonal reflection* as well. Obviously, the norm (determinant; cf. I(5.7.30)) of a reflection is equal to -1, $N(s_W) = -1$. We will now prove a proposition refining the statement of Lemma II.8.10.6 (cf. E. Artin [4], Theorem 3.20):

Proposition 5. *Let $\boldsymbol{V}^n$ be a vector space over a field K,* char $K \neq 2$, *and let $\langle,\rangle$ be a symmetric bilinear scalar product on $\boldsymbol{V}^n$. Then every orthogonal transformation $g \in \boldsymbol{O}_n$ can be represented as a product of $r \leq n$ reflections.*

Proof. For $n = 0, 1$ or $g = \mathrm{id}_V$ the claim is trivial. So we suppose that the proposition is already proved for all natural numbers $m < n$, and proceed by induction in four steps.

1. Suppose that there exists a non-isotropic vector $\mathfrak{a} \in \boldsymbol{V} \setminus \{\mathfrak{o}\}$ satisfying $g(\mathfrak{a}) = \mathfrak{a}$, and let $\boldsymbol{W} := (\mathfrak{a}K)^\perp$ be the subspace orthogonal to $\mathfrak{a}$. Since $\mathfrak{a}$ is not isotropic, we obtain the orthogonal decomposition $\boldsymbol{V} = \boldsymbol{W} \oplus \mathfrak{a}K$ into g-invariant, non-isotropic subspaces. By the induction hypothesis we find $r \leq n-1$ reflections $\tilde{s}_\nu$ in hyperplanes $\boldsymbol{U}_\nu$ of $\boldsymbol{W}$ such that $g|\boldsymbol{W} = \tilde{s}_r \circ \ldots \circ \tilde{s}_1$. Then, by Exercise 2.5, we have $s_\nu := \tilde{s}_\nu + \mathrm{id}_{\mathfrak{a}K} \in \boldsymbol{O}_n$; this map is a reflection in the subspace $\boldsymbol{W}_\nu := \boldsymbol{U}_\nu \oplus \mathfrak{a}K$, and, moreover, $g = s_r \circ \ldots \circ s_1$ with $r \leq n-1$.

2. Now suppose that there exists a non-isotropic vector $\mathfrak{a} \in \boldsymbol{V} \setminus \{\mathfrak{o}\}$ for which $\mathfrak{b} := g(\mathfrak{a}) - \mathfrak{a} \neq \mathfrak{o}$ is not isotropic. Let s be the reflection in the likewise non-isotropic subspace $\boldsymbol{W} := (\mathfrak{b}K)^\perp$. Because of

$$\langle g(\mathfrak{a}) + \mathfrak{a}, g(\mathfrak{a}) - \mathfrak{a}\rangle = \langle g(\mathfrak{a}), g(\mathfrak{a})\rangle - \langle \mathfrak{a}, \mathfrak{a}\rangle = 0,$$

we have $g(\mathfrak{a}) + \mathfrak{a} \in \boldsymbol{W}$. Hence

$$s(g(\mathfrak{a}) - \mathfrak{a}) = -g(\mathfrak{a}) + \mathfrak{a}, \; s(g(\mathfrak{a}) + \mathfrak{a}) = g(\mathfrak{a}) + \mathfrak{a},$$

and adding both equations we obtain $s \circ g(\mathfrak{a}) = \mathfrak{a}$ for the non-isotropic vector $\mathfrak{a} \neq \mathfrak{o}$. By Step 1 we can represent $s \circ g$ as the product of $r < n$ reflections s_ν, and thus $g = s \circ s_r \circ \ldots \circ s_1$ is the product of $r + 1 \leq n$ reflections.

3. Next we consider the case $\dim \boldsymbol{V} = 2$. By Steps 1 and 2 the assertion holds true, if $\boldsymbol{V}$ does not contain any isotropic vectors. So let $\mathfrak{v} \in \boldsymbol{V}$ be an isotropic vector. By Lemma 2.4 we find an isotropic basis $(\mathfrak{v}, \hat{\mathfrak{v}})$ for $\boldsymbol{V}$ with $\langle \mathfrak{v}, \hat{\mathfrak{v}}\rangle = 1$. Since each isotropic vector in $\boldsymbol{V}$ has to be a multiple of $\mathfrak{v}$ or of $\hat{\mathfrak{v}}$, for $g \in \boldsymbol{O}_2$ there only remain the following possibilities:

a) $\quad g\mathfrak{v} = \hat{\mathfrak{v}}\beta, \; g\hat{\mathfrak{v}} = \mathfrak{v}\beta^{-1},$

b) $\quad g\mathfrak{v} = \mathfrak{v}\beta, \; g\hat{\mathfrak{v}} = \hat{\mathfrak{v}}\beta^{-1}.$

In Case a), the vector $g(\mathfrak{v} + \hat{\mathfrak{v}}\beta) = \mathfrak{v} + \hat{\mathfrak{v}}\beta$ is not isotropic and fixed under g, and using Step 1 we obtain the assertion. In Case b), we may suppose $\beta \neq 1$, since otherwise $g = \mathrm{id}_V$. Then the vectors $\mathfrak{a} := \mathfrak{v} + \hat{\mathfrak{v}}$ and

$$g(\mathfrak{a}) - \mathfrak{a} = \mathfrak{v}(\beta - 1) + \hat{\mathfrak{v}}(\beta^{-1} - 1)$$

are not isotropic, and the assertion follows from 2.

4. We may now suppose $\dim \boldsymbol{V} \geq 3$, that every non-isotropic vector $\mathfrak{a} \in \boldsymbol{V}$ is not invariant, $g(\mathfrak{a}) \neq \mathfrak{a}$, and that $g(\mathfrak{a}) - \mathfrak{a}$ is isotropic. From this we can even conclude that $g(\mathfrak{a}) - \mathfrak{a}$ is isotropic for all $\mathfrak{a} \in \boldsymbol{V}$. In fact, if $\mathfrak{v} \in \boldsymbol{V}$ is isotropic, then the $(n-1)$-dimensional subspace $(\mathfrak{v}K)^\perp$ is also isotropic, and has defect 1. Hence there is a non-isotropic vector $\mathfrak{a} \in (\mathfrak{v}K)^\perp$. So we have

$$\langle \mathfrak{a}, \mathfrak{a}\rangle \neq 0 \text{ und } \langle \mathfrak{a} + \mathfrak{v}c, \mathfrak{a} + \mathfrak{v}c\rangle = \langle \mathfrak{a}, \mathfrak{a}\rangle \neq 0$$

for all $c \in K$. According to our assumption the vectors $g(\mathfrak{a}) - \mathfrak{a}$ and

$$g(\mathfrak{a} + \mathfrak{v}c) - (\mathfrak{a} + \mathfrak{v}c) = g(\mathfrak{a}) - \mathfrak{a} + (g(\mathfrak{v}) - \mathfrak{v})c$$

are isotropic. For the scalar square of the latter vector we thus have

$$0 = 2\langle g(\mathfrak{a}) - \mathfrak{a}, g(\mathfrak{v}) - \mathfrak{v}\rangle c + \langle g(\mathfrak{v}) - \mathfrak{v}, g(\mathfrak{v}) - \mathfrak{v}\rangle c^2.$$

Inserting here $c = \pm 1$ and adding both the resulting equations we obtain $\langle g(\mathfrak{v}) - \mathfrak{v}, g(\mathfrak{v}) - \mathfrak{v}\rangle = 0$, and $g(\mathfrak{v}) - \mathfrak{v}$ is isotropic.

The just proved implies that the image $\boldsymbol{W} := (g - \mathrm{id}_V)(\boldsymbol{V}) \neq \{\mathfrak{o}\}$ is a totally isotropic subspace of $\boldsymbol{V}$. Consider now $\mathfrak{a} \in \boldsymbol{V}$ and $\mathfrak{b} \in \boldsymbol{W}^\perp$. Then we have

$$0 = \langle g(\mathfrak{a}) - \mathfrak{a}, g(\mathfrak{b}) - \mathfrak{b}\rangle = \langle g(\mathfrak{a}), g(\mathfrak{b})\rangle - \langle \mathfrak{a}, g(\mathfrak{b})\rangle - \langle g(\mathfrak{a}) - \mathfrak{a}, \mathfrak{b}\rangle.$$

From $\langle g(\mathfrak{a}), g(\mathfrak{b})\rangle = \langle \mathfrak{a}, \mathfrak{b}\rangle$ and $\mathfrak{b} \in \boldsymbol{W}^{\perp}$ we conclude

$$\langle g(\mathfrak{a}) - \mathfrak{a}, \mathfrak{b}\rangle = 0,\ \langle \mathfrak{a}, \mathfrak{b} - g(\mathfrak{b})\rangle = 0.$$

Since the last equation holds for all $\mathfrak{a} \in \boldsymbol{V}$, we obtain $g(\mathfrak{b}) = \mathfrak{b}$, i.e. $g|\boldsymbol{W}^{\perp} = \mathrm{id}_{W^{\perp}}$. According to our assumption there is no non-isotropic fixed vector. Hence $\boldsymbol{W}^{\perp}$ has to be a totally isotropic subspace as well.

By Corollary 2.6,

$$\dim \boldsymbol{W} \leq n/2,\ \dim \boldsymbol{W}^{\perp} \leq n/2,$$

and because of $\dim \boldsymbol{W} + \dim \boldsymbol{W}^{\perp} = n$ and $\boldsymbol{W} \subset \boldsymbol{W}^{\perp}$, we have

$$\dim \boldsymbol{W} = \dim \boldsymbol{W}^{\perp} = n/2 \text{ und } \boldsymbol{W} = \boldsymbol{W}^{\perp}.$$

Thus by Proposition 2.5, $\boldsymbol{V}$ has to be a neutral space, and, in particular, $n = \dim \boldsymbol{V} = 2r$ is even. The automorphism g leaves each element of $\boldsymbol{W} = \boldsymbol{W}^{\perp}$ fixed and so has the form (2.21). Consequently, $N(g) = 1$. This implies that for a neutral space $\boldsymbol{V}^{2r}$ and every g with $N(g) \neq 1$ the assertion holds. So let now s be an arbitrary reflection and let $g \in \boldsymbol{O}_n$ be an element with $N(g) = 1$. Then we have $N(s \circ g) = N(s) \cdot N(g) = -1$; hence there is a representation $s \circ g = s_k \circ \ldots \circ s_1$ of $s \circ g$ as the product of $k \leq 2r$ reflections. Since $N(s \circ g) = (-1)^k = -1$, k has to be odd and thus must be less than $2r$. This implies: $g = s \circ s_k \circ \ldots \circ s_1$ is represented as the product of at most $n = 2r$ reflections. □

Corollary 6. *For every $g \in \boldsymbol{O}_n$ the norm of its square satisfies $N(g^2) = 1$. The special orthogonal group*

$$\boldsymbol{SO}_n := \{g \in \boldsymbol{O}_n | N(g) = 1\}$$

consists of all orthogonal transformations, which can be represented as the product of an even number of reflections. □

The elements $g \in \boldsymbol{SO}_n$ are also called *proper*, and those from $\boldsymbol{O}_n \setminus \boldsymbol{SO}_n$ *improper motions*, or *rotations* and *rotoinversions*, respectively.

Corollary 7. *If $g \in \boldsymbol{O}_n$ can be represented as the product of $r \leq n$ reflections, then the dimension of the space of fixed vectors, $\boldsymbol{U} := \{\mathfrak{x} \in \boldsymbol{V}^n | g\mathfrak{x} = \mathfrak{x}\}$, is at least* $\dim U \geq n - r$. *Hence, if n is odd (or even), then each proper (improper) motion g has at least one fixed vector $\mathfrak{x} \neq \mathfrak{o}$, i.e., g has a fixed point in $\boldsymbol{P}^{n-1}(\boldsymbol{V})$.*

Proof. Let $g = s_r \circ \ldots \circ s_1$, s_ρ be the reflection in $\boldsymbol{W}_\rho^{n-1}$. Then for all

$$\mathfrak{x} \in \boldsymbol{U}_r := \bigcap_{\rho=1}^{r} \boldsymbol{W}_\rho$$

we have $g(\mathfrak{x}) = \mathfrak{x}$. A straightforward induction proves $\dim \boldsymbol{U}_k \geq n - k$, and $\boldsymbol{U} \supset \boldsymbol{U}_r$ implies $\dim U \geq n - r$. Hence in case n is odd, for every proper motion we have $\dim U \geq 1$, and for even n this holds for each improper motion. □

Exercise 8. Let $g \in \boldsymbol{O}_n$ be an element that cannot be represented as the product of less than n reflections. Prove that then the first (or the last) element in any representation $g = s_n \circ \ldots \circ s_1$ may be chosen to be an arbitrarily given reflection.

Exercise 9. Let $g \in \boldsymbol{O}_n$ be an involution. Prove: a) The decomposition $\boldsymbol{V}^n = \boldsymbol{W}_1 \oplus \boldsymbol{W}_{-1}$ into the eigensubspaces of the only possible eigenvalues ± 1 is orthogonal. – b) The eigensubspaces $\boldsymbol{W}_1, \boldsymbol{W}_{-1}$ are not isotropic. – c) There is no orthogonal reflection in an isotropic hyperplane. (Cf. Example 1.4.6.)

Exercise 10. Prove: a) Any $g \in \boldsymbol{O}_n$ satisfying $g|\boldsymbol{W}^{n-1} = \mathrm{id}_W$ for some isotropic hyperplane $\boldsymbol{W} \subset \boldsymbol{V}^n$ is itself the identity, $g = \mathrm{id}_V$. (Hint. Apply Lemma 2.4 to a vector $\mathfrak{v} \neq \mathfrak{o}$ from the defect subspace of $\boldsymbol{W}$.) – b) If for $g, h \in \boldsymbol{O}_n$ the restrictions to a hyperplane $\boldsymbol{U}^{n-1} \subset \boldsymbol{V}^n$ coincide, $g|\boldsymbol{U} = h|\boldsymbol{U}$, then $g = h$ or $g = s_U \circ h$; if $\boldsymbol{U}$ is isotropic, then necessarily $g = h$.

By Exercise 9, there are no orthogonal reflections in isotropic hyperplanes. This fact is generalized in the following proposition:

Proposition 8. *Let $\boldsymbol{V}^n$ be a vector space with a symmetric, bilinear scalar product $\langle , \rangle$, and let $\varphi : \boldsymbol{U} \to \tilde{\boldsymbol{U}}$ be an isomorphism from the subspace $\boldsymbol{U}$ onto the subspace $\tilde{\boldsymbol{U}}$ (cf. Definition 2.2). Then there is an extension $g \in \boldsymbol{O}_n$ of φ (its existence is guaranteed by E. Witt's Theorem) with prescribed norm $N(g) = \pm 1$ if and only if*

$$\dim \boldsymbol{U} + \operatorname{def} \boldsymbol{U} < \dim \boldsymbol{V}. \tag{9}$$

Proof. If $g_1, g_2 \in \boldsymbol{O}_n$ both are extensions with different norms, then

$$g := g_2^{-1} \circ g_1 \in \boldsymbol{O}_n$$

is a map with norm $N(g) = -1$ satisfying $g|\boldsymbol{U} = \mathrm{id}_U$. Hence it has to be shown that id_U can be extended to an improper motion if (9) holds. We decompose $\boldsymbol{U} = \boldsymbol{U}_o \oplus \boldsymbol{U}_1$ into the orthogonal sum of a defect subspace $\boldsymbol{U}_o$ and a complement $\boldsymbol{U}_1$. Then $\boldsymbol{U}_1$ is non-isotropic; in $\boldsymbol{U}_1^\perp$ there is a totally isotropic subspace $\hat{\boldsymbol{U}}_o$, $\dim \hat{\boldsymbol{U}}_o = \dim \boldsymbol{U}_o$, and $\boldsymbol{U}_o \oplus \hat{\boldsymbol{U}}_o$ is not isotropic (by Proposition 2.5). Obviously, $\dim \boldsymbol{U} \oplus \hat{\boldsymbol{U}}_o = \dim \boldsymbol{U} + \operatorname{def} \boldsymbol{U}$, and $\boldsymbol{U} \oplus \hat{\boldsymbol{U}}_o$ is not isotropic. If (9) holds, then there is a non-isotropic hyperplane $\boldsymbol{W}^{n-1} \supset \boldsymbol{U} \oplus \hat{\boldsymbol{U}}_o$, and the reflection s_W in this hyperplane satisfies $N(s_W) = -1$ as well as $s_W|\boldsymbol{U} = \mathrm{id}_U$. Conversely, suppose now that $\dim \boldsymbol{U} + \operatorname{def} \boldsymbol{U} = n$. Then $\boldsymbol{V}^n = \hat{\boldsymbol{U}}_o \oplus \boldsymbol{U} = \hat{\boldsymbol{U}}_o \oplus \boldsymbol{U}_o \oplus \boldsymbol{U}_1$. Since $\hat{\boldsymbol{U}}_o \oplus \boldsymbol{U}_o = \boldsymbol{U}_1^\perp$ is a neutral space with $\boldsymbol{U}_o \subset \boldsymbol{U}$, $g|\boldsymbol{U}_1^\perp$ is a map satisfying the conditions from Exercise 2.11. From $g|\boldsymbol{U}_1 = \mathrm{id}_{U_1}$ we obtain

$$N(g) = N(g|\boldsymbol{U}_1) \cdot N(g|\boldsymbol{U}_1^\perp) = 1.$$

□

Exercise 11. a) Determine the equations of the quadrics Q_F associated with the real auto-polar maps classified in Proposition 1.9.4. – b) Compute the maximal dimension of a projective subspace contained in Q_F. – c) Prove that each type of these quadrics $Q_F \subset \boldsymbol{P}^n$ can be obtained as the intersection $Q_F = \boldsymbol{P}^n \cap \tilde{Q}$ of an n-plane $\boldsymbol{P}^n \subset \tilde{\boldsymbol{P}}^N$ with a suitable quadric $\tilde{Q}$ in a larger projective space $\tilde{\boldsymbol{P}}^N$. Determine, in addition, the smallest value for N (depending on n and the type of F) allowing such a representation.

Exercise 12. Let $(\boldsymbol{a}_0, \ldots, \boldsymbol{a}_n)$ be a polar simplex of the polarity F in the n-dimensional projective space $\boldsymbol{P}^n$. Denoting by s_i the reflection in the i-th face $F(\boldsymbol{a}_i)$ of the polar simplex prove

$$s_0 \circ s_1 \circ \ldots \circ s_n = \mathrm{id}_{\boldsymbol{P}^n} .$$

2.4 Invariants of Finite Configurations

As in Section 2 of this chapter we start from a vector space with scalar product $[\boldsymbol{V}^{n+1}, \langle , \rangle]$ determining the absolute symmetric auto-correlation F of the associated projective geometry. Its projective isotropy group will again be denoted by $\boldsymbol{PG}_n = \boldsymbol{PG}_n(F)$. According to the Erlanger Programm by F. Klein [65], any action of $\boldsymbol{PG}_n$ as a transformation group on a set M poses the task to describe the orbits of this action by invariants. In this section we will consider some special cases of this kind of problem in the case of transformation groups arising in a natural way from the action of $\boldsymbol{PG}_n$ over $\boldsymbol{P}^n$ as the isotropy group of F: The elements of the set M are to be described by finitely many projective subspaces of $\boldsymbol{P}^n$. For this reason we call them *finite configurations*. The action of any element $g \in \boldsymbol{PG}_n$ on a configuration results from the simultaneous action of g on all the *components* of the configuration, i.e. the subspaces defining the configuration. We start with the discussion of point sequences. Dualizing the results by means of the polarity F immediately leads to invariants for hyperplanes, which thus do not have to be considered separately in this section. In the sections to follow, presenting the elliptic and the hyperbolic geometries, it will become clear that the distance of point pairs and the angle between two hyperplanes are mutually dual notions. In these metric geometries and in the Möbius geometry we will describe the invariants of pairs of subspaces of arbitrary dimension.

2.4.1 $\boldsymbol{PG}_n$-Congruence of Finite Point Sequences

The action of $\boldsymbol{PG}_n$ on point sequences is defined for $g \in \boldsymbol{PG}_n$ by

$$g : (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) \in (\boldsymbol{P}^n)^k \longmapsto g(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) := (g\boldsymbol{x}_1, \ldots, g\boldsymbol{x}_k) \in (\boldsymbol{P}^n)^k. \quad (1)$$

Here $(B)^k$ denotes the k-fold direct product of the set B. Point sequences transformed into one another by such a transformation are called *congruent* or, more precisely, $\boldsymbol{PG}_n$*-equivalent.* We will use a similar terminology also for other transformation groups. In order to not always have to exclude particular cases, we restrict the action to the obviously $\boldsymbol{PG}_n$-invariant subset consisting of sequences containing only mutually different points; so, for an arbitrary set B, we consider

$$M_k(B) := \{(x_1, \ldots, x_k) \in (B)^k | x_\mu \neq x_\nu \text{ for } \mu \neq \nu\}. \tag{2}$$

Invariance of dimension and Lemma 1.1 immediately imply the following necessary condition

Proposition 1. *The following conditions have to be satisfied by two k-tuples $(\boldsymbol{x}_\mu), (\boldsymbol{y}_\mu) \in M_k(\boldsymbol{P}^n)$ in order to be congruent:*

a) For all ρ_λ with $1 \leq \rho_1 < \ldots < \rho_l \leq k$, $1 \leq l \leq k$ the dimensions of corresponding joins have to coincide,

$$\operatorname{Dim} \boldsymbol{x}_{\rho_1} \vee \ldots \vee \boldsymbol{x}_{\rho_l} = \operatorname{Dim} \boldsymbol{y}_{\rho_1} \vee \ldots \vee \boldsymbol{y}_{\rho_l}. \tag{3}$$

b) There exist a $\kappa \in Z(K)^*$ with $\sigma(\kappa) = \kappa$ as well as vectors $\mathfrak{x}_\mu, \mathfrak{y}_\mu \in \boldsymbol{V}^{n+1}$ such that $\boldsymbol{x}_\mu = [\mathfrak{x}_\mu], \boldsymbol{y}_\mu = [\mathfrak{y}_\mu]$ and

$$\langle \mathfrak{y}_\mu, \mathfrak{y}_\nu \rangle = \kappa \langle \mathfrak{x}_\mu, \mathfrak{x}_\nu \rangle \text{ for } 1 \leq \mu \leq \nu \leq k.$$

□

The following shows that these conditions are not sufficient.

Example 1. Take $K = \mathbf{R}$, and let $F_{4,2}$ be the polarity of index 2 in $\boldsymbol{P}^3$. Then there is a totally isotropic line $\boldsymbol{H}^1 \subset Q$. Moreover, it is possible to find two quadruples $(\boldsymbol{x}_\mu), (\boldsymbol{y}_\mu) \in M_4(\boldsymbol{H}^1)$ with differing cross ratios; hence they cannot be transformed into one another by any projectivity. On the other hand, for an arbitrary pair of such quadruples conditions a) and b) from Proposition 1 are satisfied. There is a similar example formed by two isotropic lines in the three-dimensional projective symplectic space. □

Requiring, in addition that the projective span of the point sequences is not isotropic leads to the following sufficient condition for congruence:

Proposition 2. *Let $(\boldsymbol{x}_\mu), (\boldsymbol{y}_\mu) \in M_k(\boldsymbol{P}^n)$ be two point sequences satisfying conditions a), b) from Proposition* 1. *Suppose, in addition:*

c) The projective span $\boldsymbol{x}_1 \vee \ldots \vee \boldsymbol{x}_k$ is not isotropic.

Then there is a projective isomorphism

$$\varphi : \boldsymbol{x}_1 \vee \ldots \vee \boldsymbol{x}_k \longrightarrow \boldsymbol{y}_1 \vee \ldots \vee \boldsymbol{y}_k \tag{4}$$

with the properties

$$\varphi(\boldsymbol{x}_\mu) = \boldsymbol{y}_\mu, \; \mu = 1, \ldots, k, \tag{5}$$

$$\langle\varphi(\mathfrak{x}),\varphi(\mathfrak{y})\rangle=\kappa\langle\mathfrak{x},\mathfrak{y}\rangle \textit{ for all } \mathfrak{x},\mathfrak{y}\in \boldsymbol{x}_1\vee\ldots\vee\boldsymbol{x}_k. \tag{6}$$

If $\operatorname{Dim}\boldsymbol{x}_1\vee\ldots\vee\boldsymbol{x}_k=n$*, or if*

d) There is $\rho\in Z(K)^*$ *such that* $\sigma(\rho)\rho=\kappa$*,*

then the point sequences are congruent.

Proof. Set $\boldsymbol{X}:=\boldsymbol{x}_1\vee\ldots\vee\boldsymbol{x}_k$ and $m:=\operatorname{Dim}\boldsymbol{X}$. Then, in the sequence $(\boldsymbol{x}_\mu)$ there are $m+1$ points generating $\boldsymbol{X}$. The corresponding vectors, say

$$\mathfrak{a}_0:=\mathfrak{x}_{\mu_0},\ldots,\mathfrak{a}_m:=\mathfrak{x}_{\mu_m},$$

then form a basis for $\boldsymbol{X}$ (according to b)). Condition a) implies that the vectors

$$\mathfrak{b}_0:=\mathfrak{y}_{\mu_0},\ldots,\mathfrak{b}_m:=\mathfrak{y}_{\mu_m}$$

form a basis for the space $\boldsymbol{Y}:=\boldsymbol{y}_1\vee\ldots\vee\boldsymbol{y}_k$. Consider the linear isomorphism $\varphi:\boldsymbol{X}\to\boldsymbol{Y}$ determined by $\varphi(\mathfrak{a}_\mu):=\mathfrak{b}_\mu$. From Condition b) we immediately conclude that equation (6) holds. We now prove (5). Let $\mathfrak{x}=\mathfrak{x}_\nu$, $\mathfrak{y}=\mathfrak{y}_\nu$ be two of the mutually corresponding vectors determining the point sequences according to b). Let their basis representations in $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively, be

$$\mathfrak{x}=\sum_{\mu=0}^m\mathfrak{a}_\mu x^\mu,\ \ \mathfrak{y}=\sum_{\mu=0}^m\mathfrak{b}_\mu y^\mu.$$

Forming the scalar product of these basis representations with $\mathfrak{a}_\nu$ and $\mathfrak{b}_\nu$ from the left, respectively, we obtain

$$\langle\mathfrak{a}_\nu,\mathfrak{x}\rangle=\sum_{\mu=0}^m\langle\mathfrak{a}_\nu,\mathfrak{a}_\mu\rangle x^\mu,\ \nu=0,\ldots,m,$$
$$\langle\mathfrak{b}_\nu,\mathfrak{y}\rangle=\sum_{\mu=0}^m\langle\mathfrak{b}_\nu,\mathfrak{b}_\mu\rangle y^\mu,\ \nu=0,\ldots,m.$$

Since all vectors occurring in the second system of equations correspond to those in the point sequences above them, we can apply Condition b) to the second system. Cancelling the scalar κ we obtain a system of equations with the same coefficients as the first one. Since, by Condition c), the space $\boldsymbol{X}$ is not isotropic, the ranks of these systems both are equal to $m+1$; hence their solutions have to coincide: $(y^\mu)=(x^\mu)$. From the definition of φ we obtain

$$\varphi(\mathfrak{x})=\sum_{\mu=0}^m\varphi(\mathfrak{a}_\mu)x^\mu=\sum_{\mu=0}^m\mathfrak{b}_\mu x^\mu=\sum_{\mu=0}^m\mathfrak{b}_\mu y^\mu=\mathfrak{y}.$$

Hence the projective map generated by φ (denoted by the same symbol) satisfies the asserted equations (5) and (6). In the case $m=n$ Lemma 1.1 implies that φ belongs to $\boldsymbol{PG}_n(F)$ completing the argument. For $m<n$ we want to apply E. Witt's Theorem, Proposition 2.3. Since, in general, φ is not an

isomorphism of the subspaces with induced scalar product — this only holds for $\kappa = 1$, cf. (6), we now need to make use of Condition d) and define $\psi(\mathfrak{x}) := \varphi(\mathfrak{x})(\rho)^{-1}$. Then, because of $\rho \in Z(K)^*$, (6) implies:

$$\langle \psi(\mathfrak{x}), \psi(\mathfrak{y}) \rangle = \sigma(\rho^{-1})\rho^{-1}\langle \varphi(\mathfrak{x}), \varphi(\mathfrak{y}) \rangle = (\kappa)^{-1}\kappa\langle \mathfrak{x}, \mathfrak{y} \rangle = \langle \mathfrak{x}, \mathfrak{y} \rangle$$

for all $\mathfrak{x}, \mathfrak{y} \in \boldsymbol{x}_1 \vee \ldots \vee \boldsymbol{x}_k$. Hence ψ is an isomorphism of the subspaces, which, by E. Witt's Theorem, can be extended to an automorphism $g \in \boldsymbol{PG}_n$. As the projective maps generated by the linear maps ψ and φ coincide, g realizes the congruence of the point sequences. □

Example 2. To see that Condition d) is not superfluous in the case $m < n$, take $K = \mathbf{R}$, $n = 4$, let $\boldsymbol{V}^5$ be the pseudo-orthogonal vector space of index 2, and let $(\mathfrak{e}_i)$, $i = 0, \ldots, 4$, be a pseudo-orthonormal basis. Consider the point pairs $\boldsymbol{x}_1 = [\mathfrak{e}_0], \boldsymbol{x}_2 = [\mathfrak{e}_1]$ and $\boldsymbol{y}_1 = [\mathfrak{e}_3], \boldsymbol{y}_2 = [\mathfrak{e}_4]$. Then Conditions a), b), c) with $\kappa = -1$ are satisfied, but d) is not. There is a projective isomorphism $\varphi : \boldsymbol{x}_1 \vee \boldsymbol{x}_2 \to \boldsymbol{y}_1 \vee \boldsymbol{y}_2$; according to the Theorem of Inertia, however, it is impossible to extend it to an isomorphism $g \in \boldsymbol{PG}_4$. This is also expressed in Corollary 3 below: □

Corollary 3. *For $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$ let F be one of the polarities from Table 2.1. Then, in the necessary Condition b) from Proposition 1, κ can be taken to be ± 1; for $F = F_{n+1}$, $F = F_{n+1,l}$, or $F = H_{n+1,l}$, and $l < (n+1)/2$, κ can be taken to be 1; for real null systems κ can be taken to be ± 1, and for complex ones κ can be taken to be 1. In all the cases where $\kappa = 1$, Conditions a), b) (with $\kappa = 1$) and c) together are sufficient for to ensure congruence of the point sequences.*

Proof. The number κ only depends on the transformation to be found. In the Examples 1.3–5 and in Section 2.1.6 we proved that only the values listed in Corollary 3 have to be considered. The statement concerning the null systems follows from the considerations in Section 2.1.3. If, finally, $\kappa = 1$, then d) is trivially satisfied with $\rho = 1$. □

2.4.2 Orbits of Points. Normalized Representatives

Corollary 3.1 already contains a statement concerning the $\boldsymbol{PG}_n$-equivalence of two points included in the more general criteria proved above. By Example 2.3, the projective-symplectic group $\boldsymbol{PSp}_n$ acts transitively on the projective space $\boldsymbol{P}^{2n-1}$. Similarly, the isotropy group $\boldsymbol{PG}_n(F)$ of a polarity acts transitively on the quadric corresponding to F. Hence it will suffice to consider polarities, and here again only the points not belonging to Q. These are determined by non-isotropic vectors: $\boldsymbol{x} = [\mathfrak{x}]$ such that $\langle \mathfrak{x}, \mathfrak{x} \rangle \neq 0$. Obviously, the scalar square depends on the chosen representing vector. Substituting $\mathfrak{x} \mapsto \mathfrak{x}\xi$, $\xi \in K^*$, leads to

$$\langle \mathfrak{x}\xi, \mathfrak{x}\xi \rangle = \sigma(\xi)\langle \mathfrak{x}, \mathfrak{x} \rangle\xi = \bar{\xi}\langle \mathfrak{x}, \mathfrak{x} \rangle\xi. \tag{7}$$

Here and in what follows we will frequently use the notation $\sigma(\xi) = \overline{\xi}$, familiar from the conjugation in $\mathbf{H}$, if it is clear which involutive anti-automorphism σ is meant. It satisfies the usual rules,

$$\overline{\xi + \eta} = \overline{\xi} + \overline{\eta},\ \overline{\xi\eta} = \overline{\eta}\overline{\xi},\ \overline{0} = 0,\ \overline{1} = 1, \tag{8}$$

$$\overline{(-\xi)} = -\bar{\xi},\ \overline{(\xi^{-1})} = \bar{\xi}^{-1}\ (\xi \in K^*),\ \overline{\overline{\xi}} = \xi. \tag{9}$$

Let $K_\sigma := \{\xi \in K | \overline{\xi} = \xi\}$ be the set of fixed elements of σ. K_σ is a subgroup of the additive group of K, but, in general, it is no subring. Nevertheless we have: If $\xi \neq 0$ belongs to K_σ, then ξ^{-1} is also in K_σ; if $K_\sigma \subset Z(K)$, σ is called *central*; in this case, K_σ is a subfield of $Z(K)$, hence also of K. Denoting by K_o the prime field of K, the equation $\overline{1} = 1$ immediately implies $K_o \subset K_\sigma$.

Formula (7) leads to the right action

$$\lambda \in K^*,\ \xi \in K \longmapsto \overline{\lambda}\xi\lambda \in K \tag{10}$$

of K^* over K, for which K_σ is an invariant subset. As is well-known, the scalar squares $\langle \mathfrak{x}, \mathfrak{x} \rangle$ belong to K_σ. We consider the action of K^* over K_σ and fix a representative in each orbit. Let $P \subset K_\sigma$ denote the set of these representatives. Obviously, we have $0 \in P$; take 1 as the representative of the orbit $\{\overline{\lambda}\lambda\}$, $\lambda \in K^*$ and, if $\mu \in P$ and $-\mu$ is not equivalent to μ, take $-\mu$ as the representative for $\{-\overline{\lambda}\mu\lambda\}$; if $\mu \neq 0$ and μ^{-1} is not equivalent to μ, then we choose μ^{-1} as the representative of $\{\overline{\lambda}\mu^{-1}\lambda\}$. Hence, either $-1 = \overline{\lambda}\lambda$ for a certain $\lambda \in K^*$, or $\{0, 1, -1\} \subset P$. Note that Condition d) from Proposition 2 just means that κ is equivalent to 1 under the action (10).

Example 3. Take $\sigma = \mathrm{id}_K$. Then $K = K_\sigma$ is a field; two elements are equivalent under the action (10), if they differ by a square factor, and the orbit space is the union of zero and the quotient group $K^*/(K^*)^2$. For $K = \mathbf{R}$ we thus have $P = \{0, 1, -1\}$; this property is the basis for the definition of the inner and the outer region of a real quadric in Example 3.1. The same example also shows that not every element of P has to occur as the representative for an orbits of a scalar square; e.g., for $l = 0$ the number -1 does not occur. For $K = \mathbf{C}$, as for any algebraically closed field, we have $P = \{0, 1\}$, cf. Example 3.2. If, however, $K = \mathbf{Q}$ is the field of rational numbers, then $\mathbf{Q}^*/(\mathbf{Q}^*)^{\mathbf{2}}$ is infinite; two numbers $\xi, \eta \in \mathbf{Q}^*$ are equivalent if and only if $\xi\eta^{-1}$ is a square, i.e., if all its prime factors have even multiplicity, cf. Proposition I.2.6.5. The difficulties we meet here are similar to those we had to face exploiting Proposition 1.9.3 for the classification of polarities. Without any particular assumptions concerning the scalar domain K satisfactory answers to the initial question of this section will be hard to obtain. □

Now we return to the general case. To each point $\boldsymbol{x} \in \boldsymbol{P}^n$ we assign the element $\xi(\boldsymbol{x}) \in P$ representing $\langle \mathfrak{x}, \mathfrak{x} \rangle$ for an arbitrary representative $\mathfrak{x} \in \boldsymbol{V}^{n+1}$ with $\boldsymbol{x} = [\mathfrak{x}]$. A vector $\mathfrak{x} \in \boldsymbol{V}^{n+1}$ is called *normalized* if $\langle \mathfrak{x}, \mathfrak{x} \rangle \in P$. Note that a normalized representative of a point $\boldsymbol{x}$ is never uniquely determined; in fact,

K	F	σ	K_σ	P	K_1^*
$\mathbf{R}$	$F_{n,l}$	$\mathrm{id}_{\mathbf{R}}$	$\mathbf{R}$	$\{0,1,-1\}$	$\{1,-1\}$
$\mathbf{C}$	F_n	$\mathrm{id}_{\mathbf{C}}$	$\mathbf{C}$	$\{0,1\}$	$\{1,-1\}$
$\mathbf{C}$	$H_{n,l}$	τ	$\mathbf{R}$	$\{0,1,-1\}$	S^1
$\mathbf{H}$	$H_{n,l}$	τ	$\mathbf{R}$	$\{0,1,-1\}$	S^3
$\mathbf{H}$	H_n	τ_{i}	$\mathfrak{L}(1,\mathrm{j},\mathrm{k})$	$\{0,1\}$	S^1

Table 2.2. Normalization for the classical polarities. Here τ denotes the conjugation in $\mathbf{C}$ or $\mathbf{H}$, respectively, $\tau_{\mathrm{i}} = -\mathrm{i}\,\tau\,\mathrm{i}$, $S^1 \subset \mathbf{C}$ denotes the unit circle, and $S^3 \subset \mathbf{H}$ is the unit hypersphere.

$\langle -\mathfrak{x}, -\mathfrak{x}\rangle = \langle \mathfrak{x}, \mathfrak{x}\rangle$. In general, together with $\mathfrak{x}$ the vector $\mathfrak{x}\lambda$ is also normalized if λ belongs to the isotropy group of $\xi = \langle \mathfrak{x}, \mathfrak{x}\rangle \in P$ under the action (10). In Table 2.2 the corresponding data for the classical polarities are displayed; here $K_1^* \subset K^*$ denotes the isotropy group of the element 1 coinciding with that of -1. From Corollary 3.1 we conclude that $\xi(\boldsymbol{x}) = \xi(\boldsymbol{y})$ implies the congruence of $\boldsymbol{x}$ and $\boldsymbol{y}$.

Exercise 1. Verify the data in Table 2.2. Hint. Recall Section 1.9.5, in particular Formulas (19), (20).

Exercise 2. Show that there is an orthogonal basis $(\mathfrak{a}_1, \mathfrak{a}_2)$ for which $\langle \mathfrak{a}_1, \mathfrak{a}_2\rangle = \lambda\delta_{ij}$, and λ is no square using the standard scalar product in the rational vector space $\mathbf{Q}^2$, for which the standard basis $(\mathfrak{e}_1, \mathfrak{e}_2)$ is orthonormal.

Interpreting the Theorem of Inertia by saying that the number $l(\xi)$ of elements in an orthogonal basis belonging to a given representative $\xi \in P$ does not depend on the chosen basis, Exercise 2 shows that such a statement does in general not hold.

Exercise 3. a) Taking F as one of the classical polarities from Table 2.2, and $l < n/2$, prove that two points $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{P}^{n-1}$ are $\boldsymbol{PG}_{n-1}(F)$-equivalent if and only if $\xi(\boldsymbol{x}) = \xi(\boldsymbol{y})$. – b) Show by means of an example that this statement is wrong for $l = n/2$; in this case all points not belonging to the quadric are $\boldsymbol{PG}_{n-1}(F)$-equivalent.

2.4.3 Invariants of Point Pairs

We now want to find the invariants of point pairs $(\boldsymbol{x}_1, \boldsymbol{x}_2) \in M_2(\boldsymbol{P}^n)$, which are important for the foundations of metric geometries. Since two of these points always determine a line, we can pick up the thread of Section 1.9.7 where we studied the position relations of line and quadric. A little more special than there we will here make the following assumption:

Assumption A. *Let $\boldsymbol{V}^{n+1}$ be a vector space with a σ-Hermitean or symmetric scalar product over a skew field of characteristic* $\operatorname{char} K \neq 2$, *and let F be the associated polarity of the n-dimensional projective geometry; the case of an empty quadric $Q = Q_F$ is not excluded.*

Compare Definition 2.1; null systems are not allowed now. Denote by $\boldsymbol{U}_{n+1}$ the isotropy group of the scalar product, i.e. the set of all linear transformations g of $\boldsymbol{V}^{n+1}$ satisfying

$$\langle g\mathfrak{x}, g\mathfrak{y}\rangle = \langle \mathfrak{x}, \mathfrak{y}\rangle \text{ for all } \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}^{n+1}$$

(cf. (1.10)); $\boldsymbol{PU}_{n+1}$ denotes the corresponding group of projectivities. If one of the points belongs to the quadric, then from the results of Section 1.9.7 we immediately obtain

Corollary 4. *Consider, under Assumption A, the pair $(\boldsymbol{x}_1, \boldsymbol{x}_2) \in M_2(\boldsymbol{P}^n)$ and suppose that $\boldsymbol{x}_1 \in Q$. Let, moreover, $\xi = \xi(\boldsymbol{x}_2) \in P$. Then there only are the following, mutually excluding possibilities:*

a) $\xi = 0$ *and* $\langle \mathfrak{x}_1, \mathfrak{x}_2\rangle = 0 \Longleftrightarrow \boldsymbol{x}_1 \vee \boldsymbol{x}_2 \subset Q$.
b) $\xi \neq 0$ *and* $\langle \mathfrak{x}_1, \mathfrak{x}_2\rangle = 0 \Longleftrightarrow \boldsymbol{x}_1 \vee \boldsymbol{x}_2$ *is a tangent of* Q.
c) $\langle \mathfrak{x}_1, \mathfrak{x}_2\rangle \neq 0 \Longleftrightarrow \boldsymbol{x}_1 \vee \boldsymbol{x}_2$ *is a chord of* Q.

The point pairs $(\boldsymbol{x}_1, \boldsymbol{x}_2), (\boldsymbol{y}_1, \boldsymbol{y}_2) \in M_2(\boldsymbol{P}^n)$, $\boldsymbol{x}_1, \boldsymbol{y}_1 \in Q$, are congruent if they share Property a), b) or c) and, in addition, satisfy $\xi(\boldsymbol{x}_2) = \xi(\boldsymbol{y}_2)$. The group $\boldsymbol{PU}_{n+1}$ acts transitively on each of these orbits. If $\boldsymbol{PG}_n(F) = \boldsymbol{PU}_{n+1}$, then the conditions are also necessary for point pairs to be congruent.

Proof. The first claim immediately follows from Example 1.9.8. We will show the asserted congruence of the pairs. In a), this is a special case of Corollary 2.12. In both the other cases we normalize the vectors $\mathfrak{x}_2, \mathfrak{y}_2$ so that $\xi = \langle \mathfrak{x}_2, \mathfrak{x}_2\rangle = \langle \mathfrak{y}_2, \mathfrak{y}_2\rangle$. In c), multiplying $\mathfrak{x}_1$ by a suitable factor we can always arrange $\langle \mathfrak{x}_1, \mathfrak{x}_2\rangle = 1$. Thus, the scalar products of vectors representing both pairs of the same type a), b) or c) can arranged to be equal; assigning these pairs of bases to one another defines an isomorphism of subspaces, which, by, E. Witt's Theorem, Proposition 2.3, can be extended to an automorphism of $\boldsymbol{V}^{n+1}$. We have $\kappa = 1$, hence the corresponding projectivity belongs to $\boldsymbol{PU}_{n+1}$; it realizes the congruence we were looking for. The last claim follows from the invariance of Properties a), b), c) together with the fact that it suffices to consider transformations with $\kappa = 1$. □

Remark. The example of the quadric $Q(F_{4,2})$ in the real three-dimensional projective space $\boldsymbol{P}^3$ shows that the condition $\xi(\boldsymbol{x}_2) = \xi(\boldsymbol{y}_2)$ in Corollary 4 is not necessary, cf. Exercise 3 and Example 6 below.

Example 4. Let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in Q$ be points of the quadric, $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$. Then there only are the following possibilities: Case a): $\boldsymbol{x}_1 \vee \boldsymbol{x}_2 \subset Q$, or Case c): $\boldsymbol{x}_1 \vee \boldsymbol{x}_2$ is a chord of Q. By Corollary 4, the group $\boldsymbol{PU}_{n+1}$ acts transitively, hence $\boldsymbol{PG}_n$ as well, on each of these sets of point pairs. If, in particular, Q is a quadric of index 1 (cf. Example 2.4), then Q contains no line, and this implies: *For quadrics Q of index 1 the groups $\boldsymbol{PU}_{n+1}$ acts transitively on the set of point pairs $(\boldsymbol{x}_1, \boldsymbol{x}_2) \in Q \times Q$, $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$.* In particular, this holds for the polarities $F_{n+1,1}$, $K = \mathbf{R}$, and $H_{n+1,1}$, $K = \mathbf{C}, \mathbf{H}$, from Table 2.1. According to Example 2.4 the group $\boldsymbol{PU}_{n+1}$ acts transitively on the set of lines contained in Q. Then Corollary 4, Case c) with $\xi = 0$, leads to

Corollary 5. *Under Assumption A the group $\boldsymbol{PU}_{n+1}$ acts transitively on the set of chords for the quadric Q.*

□

From now on, we will consider only point pairs $(\boldsymbol{x}_1, \boldsymbol{x}_2) \in M_2(\boldsymbol{P}^n)$ with $\boldsymbol{x}_j \notin Q$, $j = 1, 2$.

Example 5. Let $(\boldsymbol{x}_1, \boldsymbol{x}_2), (\boldsymbol{y}_1, \boldsymbol{y}_2) \in M_2(\boldsymbol{P}^n)$, $\boldsymbol{x}_j, \boldsymbol{y}_j \notin Q$, $j = 1, 2$, be point pairs for which $\boldsymbol{x}_1 \vee \boldsymbol{x}_2$, $\boldsymbol{y}_1 \vee \boldsymbol{y}_2$ are tangent to the quadric Q, but not contained in Q. Then the corresponding subspaces are isotropic, but not totally isotropic. Let $\boldsymbol{z} = Q \cap (\boldsymbol{x}_1 \vee \boldsymbol{x}_2)$ be the unique point of contact of the tangent; it corresponds to the defect subspace of the associated vector subspace. We will show: *The point pairs $(\boldsymbol{x}_1, \boldsymbol{x}_2), (\boldsymbol{y}_1, \boldsymbol{y}_2)$ are $\boldsymbol{PU}_{n+1}$-equivalent if and only if*

$$\xi(\boldsymbol{x}_j) = \xi(\boldsymbol{y}_j), \qquad j = 1, 2. \tag{11}$$

Obviously, condition (11) is necessary. Let us show the converse. Consider $\boldsymbol{z} = [\mathfrak{z}]$, $\boldsymbol{x}_j = [\mathfrak{x}_j]$. Because of $\boldsymbol{x}_1 \notin Q$, the $(\mathfrak{x}_1, \mathfrak{z})$ form a basis for $\mathfrak{L}(\mathfrak{x}_1, \mathfrak{x}_2)$. Let $\xi(\boldsymbol{x}_1) = \langle \mathfrak{x}_1, \mathfrak{x}_1 \rangle$, and let

$$\mathfrak{x}_2 = \mathfrak{x}_1 \alpha + \mathfrak{z} \beta$$

be the basis representation of $\mathfrak{x}_2$. Multiplying by α^{-1}, renormalizing the isotropic vectors $\mathfrak{z}$ as well as $\mathfrak{x}_2$ we obtain

$$\mathfrak{x}_2 = \mathfrak{x}_1 + \mathfrak{z}; \tag{12}$$

indeed, because of $\langle \mathfrak{x}_2, \mathfrak{x}_2 \rangle \neq 0$ and $\mathfrak{x}_1 \neq \mathfrak{x}_2$, α and β both have to be different from zero. Note that (12) even implies

$$\xi = \langle \mathfrak{x}_1, \mathfrak{x}_1 \rangle = \langle \mathfrak{x}_2, \mathfrak{x}_2 \rangle = \langle \mathfrak{x}_1, \mathfrak{x}_2 \rangle, \tag{13}$$

from which we conclude that *all pairs of points different from the point of contact of a tangent are $\boldsymbol{PU}_{n+1}$-equivalent.* According to Assumption (11), we may find a basis $(\mathfrak{y}_1, \hat{\mathfrak{z}})$ for the tangent $\boldsymbol{y}_1 \vee \boldsymbol{y}_2$ with correspondingly equal scalar squares. By E. Witt's Theorem, Proposition 2.3, this implies the assertion. □

Example 6. Let $K = \mathbf{R}$. We consider the polarities $F_{n+1,l}$ of $\boldsymbol{P}^n$. Since here $P = \{0, 1, -1\}$, for $l > 1$ there are *inner* ($\xi = -1$) and *outer* ($\xi = 1$) *tangents.* Figure 2.1 shows part of the hyperboloid

$$Q: \quad -(x^0)^2 - (x^1)^2 + (x^2)^2 + (x^3)^2 = 0 \tag{14}$$

in $\boldsymbol{P}^3$ with one inner and outer tangent, respectively, through a point $\boldsymbol{z} \in Q$. The intersection of Q with the tangent plane $F_{4,2}(\boldsymbol{z})$ at the point $\boldsymbol{z}$ consists of two lines, the *generators of Q through $\boldsymbol{z}$*, separating the inner from the outer tangent. Note that here we have $l = (n+1)/2$; hence there is a projectivity interchanging the inner region $I(Q)$ with the outer region $A(Q)$. So, if $(\boldsymbol{x}_1, \boldsymbol{x}_2), (\boldsymbol{y}_1, \boldsymbol{y}_2) \in M_2(\boldsymbol{P}^3)$ are points, different from the points of contact of two tangents, $\boldsymbol{x}_1 \vee \boldsymbol{x}_2, \boldsymbol{y}_1 \vee \boldsymbol{y}_2$, of Q, then there always is $g \in \boldsymbol{PG}_3(F_{4,2})$ such that $g(\boldsymbol{x}_j) = \boldsymbol{y}_j$, $j = 1, 2$. If $l < (n+1)/2$, then distinguishing the inner from the outer region becomes projectively relevant: Outer tangents cannot be transformed into inner tangents by any $g \in \boldsymbol{PG}_n(F_{n+1,l})$; this group acts transitively on the set of outer and on the set of inner tangents, separately. □

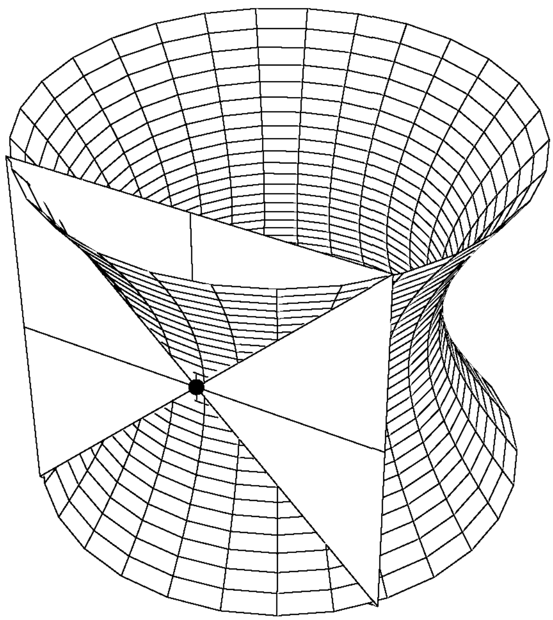


Fig. 2.1. Hyperboloid, tangent plane, inner and outer tangents.

Now we turn to the most interesting case that the joining lines of each point pair are not tangent, i.e.

$$(\boldsymbol{x}_1 \vee \boldsymbol{x}_2) \wedge F(\boldsymbol{x}_1 \vee \boldsymbol{x}_2) = \boldsymbol{o}.$$

The associated vector space $\boldsymbol{W}^2 := \mathfrak{L}(\mathfrak{x}_1, \mathfrak{x}_2)$ is not isotropic. Let $\mathfrak{x}_j$ be normalized representatives:

$$\boldsymbol{x}_j = [\mathfrak{x}_j],\ \xi_j = \xi(\boldsymbol{x}_j) = \langle \mathfrak{x}_j, \mathfrak{x}_j \rangle \in P,\ j = 1, 2.$$

Since the ξ_j are only determined up to renormalization by $\lambda_j \in K^*_{\xi_j}$, i.e. the isotropy group of ξ_j, for the scalar product we obtain with $\hat{\mathfrak{x}}_j = \mathfrak{x}_j \lambda_j$

$$\langle \hat{\mathfrak{x}}_j, \hat{\mathfrak{x}}_j \rangle = \langle \mathfrak{x}_j, \mathfrak{x}_j \rangle,\ \langle \hat{\mathfrak{x}}_1, \hat{\mathfrak{x}}_2 \rangle = \bar{\lambda}_1 \langle \mathfrak{x}_1, \mathfrak{x}_2 \rangle \lambda_2,\ \lambda_j \in K^*_{\xi_j},\ j = 1, 2. \tag{15}$$

Choosing in each class of $(K^*_{\xi_1} \times K^*_{\xi_2})$-equivalent elements of K a representative η, and normalizing this representative by

$$\xi_1 = \xi(\boldsymbol{x}_1),\ \xi_2 = \xi(\boldsymbol{x}_2),\ \eta = \eta(\langle \mathfrak{x}_1, \mathfrak{x}_2 \rangle) \tag{16}$$

we obtain a complete system of invariants, capable of deciding upon $\boldsymbol{PU}_{n+1}$-congruence for point pairs. The examples below display a closer look at some special spaces.

Example 7. Under Assumption A, let $K, \operatorname{char} K \neq 2$, be a field, and suppose that the scalar product is bilinear, i.e. $\sigma = \mathrm{id}_K$. Then we have $K^*_\xi = \{1, -1\}$ for all $\xi \neq 0$; as always $K^*_0 = K^*$. We define

$$\operatorname{Sq}(\boldsymbol{x}_1, \boldsymbol{x}_2) := \frac{\langle \mathfrak{x}_1, \mathfrak{x}_2 \rangle^2}{\langle \mathfrak{x}_1, \mathfrak{x}_1 \rangle \langle \mathfrak{x}_2, \mathfrak{x}_2 \rangle} \qquad (\boldsymbol{x}_1, \boldsymbol{x}_2 \notin Q). \tag{17}$$

□

Obviously, $\operatorname{Sq}(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is independent of the chosen representatives $\mathfrak{x}_j$ for $\boldsymbol{x}_j$, and invariant with respect to transformations from $\boldsymbol{PG}_n(F)$, cf. Lemma 1.1. If $\boldsymbol{x}_1, \boldsymbol{x}_2$ are points on a tangent not belonging to Q, then $\operatorname{Sq}(\boldsymbol{x}_1, \boldsymbol{x}_2) = 1$, cf. (13). We prove

Proposition 6. *Under the assumptions of Example 7, let $\boldsymbol{x}_1, \boldsymbol{x}_2 \notin Q$ be points on a non-tangent line $\boldsymbol{B}$. Then $\operatorname{Sq}(\boldsymbol{x}_1, \boldsymbol{x}_2) = 1$ if and only if $\boldsymbol{x}_1 = \boldsymbol{x}_2$. For points $\boldsymbol{x}_1, \boldsymbol{x}_2$ different from one another, $Q \cap \boldsymbol{B} \neq \emptyset$ if and only if*

$$-\det(\langle \mathfrak{x}_j, \mathfrak{x}_k \rangle) = \langle \mathfrak{x}_1, \mathfrak{x}_2 \rangle^2 - \langle \mathfrak{x}_1, \mathfrak{x}_1 \rangle \langle \mathfrak{x}_2, \mathfrak{x}_2 \rangle$$

is a square in K.

Proof. As the basis for the vector space associated with $\boldsymbol{B}$ we choose an orthogonal basis $(\mathfrak{a}_1, \mathfrak{a}_2)$ such that $[\mathfrak{a}_1] = \boldsymbol{x}_1$. Take $\lambda_j := \langle \mathfrak{a}_j, \mathfrak{a}_j \rangle$, $j = 1, 2$. For a vector $\mathfrak{x}_2$ with $\boldsymbol{x}_2 = [\mathfrak{x}_2]$, let $\mathfrak{x}_2 = \mathfrak{a}_1\alpha + \mathfrak{a}_2\beta$ be its basis representation. Then $\mathrm{Sq}(\boldsymbol{x}_1, \boldsymbol{x}_2) = 1$ if and only if $\alpha^2\lambda_1^2 = \lambda_1(\lambda_1\alpha^2 + \lambda_2\beta^2)$, and this in turn holds if and only if $\lambda_1\lambda_2\beta^2 = 0$. As $\boldsymbol{B}$ is not tangent, i.e., the associated vector space is not isotropic, we have $\lambda_1\lambda_2 \neq 0$, and this implies the first assertion. The second is an immediate consequence of the discussion concerning the condition for $\boldsymbol{y} := [\mathfrak{x}_1 t + \mathfrak{x}_2] \in Q$ with $t \in K$:

$$\langle \mathfrak{y}, \mathfrak{y} \rangle = \langle \mathfrak{x}_1, \mathfrak{x}_1 \rangle t^2 + 2t\langle \mathfrak{x}_1, \mathfrak{x}_2 \rangle + \langle \mathfrak{x}_2, \mathfrak{x}_2 \rangle = 0.$$

□

Exercise 4. Show under the assumptions of Example 7: If the pairs $(\boldsymbol{x}_1, \boldsymbol{x}_2), (\boldsymbol{y}_1, \boldsymbol{y}_2) \in M_2(\boldsymbol{P}^n)$ satisfy $\mathrm{Sq}(\boldsymbol{x}_1, \boldsymbol{x}_2) \neq 0$, then there is $g \in \boldsymbol{PU}_{n+1}$ such that $g\boldsymbol{x}_j = \boldsymbol{y}_j$, $j = 1, 2$, if and only if $\xi(\boldsymbol{x}_1) = \xi(\boldsymbol{y}_1)$ and

$$\mathrm{Sq}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \mathrm{Sq}(\boldsymbol{y}_1, \boldsymbol{y}_2). \tag{18}$$

Exercise 5. In addition to the assumptions of Example 7, suppose that K is algebraically closed. Prove for $\boldsymbol{x}_j, \boldsymbol{y}_j \notin Q$, $j = 1, 2$, that $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is congruent to $(\boldsymbol{y}_1, \boldsymbol{y}_2)$ if and only if (18) holds.

Exercise 5 settles the special case $K = \mathbf{C}$: The pairs of points not belonging to Q are classified by the $\boldsymbol{PG}_n(F)$-invariant $\mathrm{Sq}(\boldsymbol{x}_1, \boldsymbol{x}_2)$.

2.4.4 Real Orthogonal Geometries

Example 8. In this section we will deal with the real polarities $F = F_{n+1,l}$ of the projective space $\boldsymbol{P}^n$ over the field $K = \mathbf{R}$, and start with the case $l = 0$. According to Example 3.1 the group $\boldsymbol{PO}(n+1) = \boldsymbol{PG}_n(F)$ acts transitively on $\boldsymbol{P}^n$. Since the scalar product is positive definite, the associated vector space is Euclidean, and, moreover, $\xi(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in \boldsymbol{P}^n$. Proposition 2 immediately implies that the point pairs are classified by the invariant (17). In order to have a geometric interpretation of (17), we represent the points by unit vectors $\mathfrak{x}_1, \mathfrak{x}_2$ such that $\langle \mathfrak{x}_1, \mathfrak{x}_2 \rangle \geq 0$. Then by Formula I.(6.1.17),

$$\cos\varphi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \langle \mathfrak{x}_1, \mathfrak{x}_2 \rangle = \sqrt{\mathrm{Sq}(\boldsymbol{x}_1, \boldsymbol{x}_2)},$$

the angle $\varphi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is uniquely determined as the angle between the corresponding unit vectors; it satisfies

$$0 \leq \varphi(\boldsymbol{x}_1, \boldsymbol{x}_2) \leq \pi/2, \tag{19}$$

where, according to Proposition 6, $\varphi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is equal to zero if and only if $\boldsymbol{x}_1 = \boldsymbol{x}_2$. The real projective space $\boldsymbol{P}^n(\mathbf{R})$, in which an absolute polarity $F = F_{n+1,0}$ is distinguished, is called the n-dimensional *elliptic space* (cf. Example 3.1); the *elliptic geometry* will be discussed in Section 5 below in greater detail. □

Example 9. Again take $K = \mathbf{R}$, and let $F = F_{n+1,l}$ be a polarity of index $l > 0$. Then for the line $\boldsymbol{B} = \boldsymbol{x}_1 \vee \boldsymbol{x}_2$ determined by two points $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$ the following positions relative to the quadric Q, which is non-empty, are possible (cf. Example 1.9.8):

a) $\boldsymbol{B} \subset Q$; this can only occur for $l > 1$.
b) $\boldsymbol{B}$ is tangent to Q: $\boldsymbol{B} \cap Q$ consists of a single point.
c) $\boldsymbol{B} \cap Q = \emptyset$; for this to be the case, $n > 1$ is necessary.
d) $\boldsymbol{B} \cap Q = \{\boldsymbol{z}_+, \boldsymbol{z}_-\}$ consists of two points $\boldsymbol{z}_+ \neq \boldsymbol{z}_-$.

For $n > 1$, all the cases actually occur, where in Case a), because of $l \leq (n+1)/2$, the inequality $n \geq 3$ has to be satisfied. Cases a) and b) were discussed in Corollary 4 and Examples 5, 6, in general. Case c) can be treated as in the preceding example: this case occurs if and only if the restriction of the scalar product to the subspace $\boldsymbol{W}^2 \subset \boldsymbol{V}^{n+1}$ corresponding to $\boldsymbol{B}$ is positive or negative definite. We now discuss Case d). Since there exist two different points of the quadric on the line $\boldsymbol{B}$, the associated vector subspace $\boldsymbol{W}^2$ has to contain two linearly independent, isotropic vectors. Hence $\boldsymbol{W}^2$ is pseudo-Euclidean of index 1. By Corollary 2.11, all these subspaces are isomorphic, and, on the other hand, each automorphism of such a subspace can be extended to an automorphism of $\boldsymbol{V}^{n+1}$. So, in order to determine the invariants of point pairs it suffices to consider transformations from $\boldsymbol{W}^2$ to itself. Let $(\mathfrak{a}_1, \mathfrak{a}_2)$ be a pseudo-orthonormal basis of $\boldsymbol{W}^2$:

$$\langle \mathfrak{a}_1, \mathfrak{a}_1 \rangle = -1,\ \langle \mathfrak{a}_2, \mathfrak{a}_2 \rangle = 1,\ \langle \mathfrak{a}_1, \mathfrak{a}_2 \rangle = 0.$$

For the points $\boldsymbol{x} \in \boldsymbol{B}$ not belonging to Q we have $\xi(\boldsymbol{x}) = \pm 1$. The correspondingly normalized vectors, with basis representation $\mathfrak{x} = \mathfrak{a}_1 \zeta_1 + \mathfrak{a}_2 \zeta_2$, determine a pair of hyperbolas in $\boldsymbol{W}^2$,

$$\langle \mathfrak{x}, \mathfrak{x} \rangle = -\zeta_1^2 + \zeta_2^2 = \pm 1,$$

with the common asymptotes

$$\langle \mathfrak{x}, \mathfrak{x} \rangle = (-\zeta_1 + \zeta_2)(\zeta_1 + \zeta_2) = 0.$$

The points of intersection with the quadric are determined by the isotropic vectors $\mathfrak{a}_2 \pm \mathfrak{a}_1$. We may always normalize the representing vectors $\mathfrak{x}_j$ of the points $\boldsymbol{x}_j$, $j = 1, 2$, so that

$$\langle \mathfrak{x}_j, \mathfrak{x}_j \rangle = \xi(\boldsymbol{x}_j) = \pm 1 \text{ and } \langle \mathfrak{x}_1, \mathfrak{x}_2 \rangle \geq 0. \tag{20}$$

Hence by Proposition 2 the numbers

$$\xi_1 = \xi(\boldsymbol{x}_1),\ \xi_2 = \xi(\boldsymbol{x}_2),\ \mathrm{Sq}(\boldsymbol{x}_1, \boldsymbol{x}_2), \tag{21}$$

uniquely determining the scalar products (20), form a complete system of invariants for the points pairs considered here, $(\boldsymbol{x}_1, \boldsymbol{x}_2) \in M_2(\boldsymbol{P}^n)$, $\boldsymbol{x}_1, \boldsymbol{x}_2 \notin Q$,

$\boldsymbol{x}_1 \vee \boldsymbol{x}_2$ a chord on Q, with respect to the group $\boldsymbol{PO}(l, n+1-l)$. Note that, for $l = (n+1)/2$ there again exist transformations s_o with $\kappa = -1$. In this case, the ξ-values of corresponding points $\boldsymbol{y} = s_o(\boldsymbol{x})$ satisfy:

$$\xi(\boldsymbol{y}) = \langle \mathfrak{y}, \mathfrak{y} \rangle = -\langle \mathfrak{x}, \mathfrak{x} \rangle = -\xi(\boldsymbol{x}).$$

In general, the congruence condition reads as follows: *Two point pairs, $(\boldsymbol{x}_1, \boldsymbol{x}_2), (\boldsymbol{y}_1, \boldsymbol{y}_2)$, in the set considered here are $\boldsymbol{PG}_n(F)$-congruent if and only if there is an ϵ such that*

$$\xi(\boldsymbol{y}_1) = \epsilon\xi(\boldsymbol{x}_1),\ \xi(\boldsymbol{y}_2) = \epsilon\xi(\boldsymbol{x}_2) \text{ and } \langle \mathfrak{y}_1, \mathfrak{y}_2 \rangle = \epsilon \langle \mathfrak{x}_1, \mathfrak{x}_2 \rangle$$

for suitably normalized representatives. In case $l < (n+1)/2$ necessarily $\epsilon = 1$, and for $l = (n+1)/2$ either $\epsilon = 1$ or $\epsilon = -1$. □

It is common and practical to establish the relation between the function Sq and the familiar trigonometric functions from calculus. According to Proposition 6 and because of $Q \cap (\boldsymbol{x}_1 \vee \boldsymbol{x}_2) \neq \emptyset$, we have

$$\langle \boldsymbol{x}_1, \boldsymbol{x}_2 \rangle^2 \geq \langle \boldsymbol{x}_1, \boldsymbol{x}_1 \rangle \langle \boldsymbol{x}_2, \boldsymbol{x}_2 \rangle. \tag{22}$$

If both points lie in the inner region, or if both points belong to the outer region, then the right-hand side of (22) is positive, and hence

$$\mathrm{Sq}(\boldsymbol{x}_1, \boldsymbol{x}_2) \geq 1; \tag{23}$$

if, however, the points $\boldsymbol{x}_1, \boldsymbol{x}_2$ lie in different regions, then (22) is trivial, and, by (17),

$$\mathrm{Sq}(\boldsymbol{x}_1, \boldsymbol{x}_2) \leq 0. \tag{24}$$

In the former case we define the *hyperbolic angle* $\varphi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ via

$$\cosh \varphi(\boldsymbol{x}_1, \boldsymbol{x}_2) = +\sqrt{\mathrm{Sq}(\boldsymbol{x}_1, \boldsymbol{x}_2)},\ \varphi \geq 0. \tag{25}$$

The monotony of cosh implies that (25) uniquely determines the number $\varphi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ for each point pair $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ with $\xi(\boldsymbol{x}_1) = \xi(\boldsymbol{x}_2)$; according to Proposition 6 $\varphi(\boldsymbol{x}_1, \boldsymbol{x}_2) = 0$ if and only if $\boldsymbol{x}_1 = \boldsymbol{x}_2$. Hence $\varphi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ will also have to be interpreted as a distance; in fact, for $l = 1$, $\varphi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is the starting point for the definition of the distance in hyperbolic geometry, cf. Section 6.

2.4.5 Projective Orthogonal Geometries for Arbitrary Fields

Once again we return to the assumptions of Example 7. In this section we want to show that, under rather general assumptions on the field K, we have the complete elementary trigonometry at our disposal. The trigonometric functions appear as matrix functions of the simplest representations for the orthogonal groups.

Consider a projective line $\boldsymbol{P}^1(K)$ with a non-empty quadric $Q = \{\boldsymbol{z}_1, \boldsymbol{z}_2\}$, cf. Example 1.9.5. The associated vector space $\boldsymbol{V}^2$ thus carries a neutral scalar product, whose orthogonal group $\boldsymbol{O}_2$ was described in Proposition 2.13. Passing from an isotropic basis $(\mathfrak{z}_1, \mathfrak{z}_2)$ of $\boldsymbol{V}^2$ in the normalization

$$\langle \mathfrak{z}_1, \mathfrak{z}_2 \rangle = 1/2, \ \langle \mathfrak{z}_1, \mathfrak{z}_1 \rangle = \langle \mathfrak{z}_2, \mathfrak{z}_2 \rangle = 0$$

to the pseudo-orthonormal basis

$$\mathfrak{a}_1 = \mathfrak{z}_1 - \mathfrak{z}_2, \ \mathfrak{a}_2 = \mathfrak{z}_1 + \mathfrak{z}_2, \tag{26}$$

$$-\langle \mathfrak{a}_1, \mathfrak{a}_1 \rangle = \langle \mathfrak{a}_2, \mathfrak{a}_2 \rangle = 1, \ \langle \mathfrak{a}_1, \mathfrak{a}_2 \rangle = 0,$$

we obtain the matrix representation for $g(\alpha) \in \boldsymbol{SO}_2$ by a straightforward computation:

$$\begin{pmatrix} c(\alpha) & s(\alpha) \\ s(\alpha) & c(\alpha) \end{pmatrix} \in \boldsymbol{SO}_2 \text{ with } c(\alpha) := (\alpha + \alpha^{-1})/2, \ s(\alpha) := (\alpha - \alpha^{-1})/2, \ \alpha \in K^*. \tag{27}$$

For the functions $c(\alpha), s(\alpha)$ we have

$$c^2(\alpha) - s^2(\alpha) = 1, \tag{28}$$

$$c(\alpha_1 \cdot \alpha_2) = c(\alpha_1)c(\alpha_2) + s(\alpha_1)s(\alpha_2), \tag{29}$$

$$s(\alpha_1 \cdot \alpha_2) = s(\alpha_1)c(\alpha_2) + s(\alpha_2)c(\alpha_1). \tag{30}$$

These relations may also be concluded from $N(g(\alpha)) = 1$ together with the fact that $\alpha \in K^* \mapsto g(\alpha) \in \boldsymbol{SO}_2$ is an isomorphism. In this basis, the elements from $\boldsymbol{O}_2 \setminus \boldsymbol{SO}_2$ have the matrices

$$\begin{pmatrix} -c(\alpha) & s(\alpha) \\ -s(\alpha) & c(\alpha) \end{pmatrix}. \tag{31}$$

If there exists a homomorphism exp from the additive to the multiplicative group of K,

$$\exp : t \in K \longmapsto \mathrm{e}^t = \exp t \in K^*, \tag{32}$$

the *hyperbolic trigonometric functions* of the field K are defined by

$$\cosh t := c(\mathrm{e}^t), \ \sinh t := s(\mathrm{e}^t), \ t \in K.$$

With these definitions, equations (27)–(30) become the well-known fundamental formulas of real hyperbolic trigonometry, from which the other formulas may be deduced. Inserting $\alpha = e^t$ one attains all the possible values of the functions $c(\alpha), s(\alpha)$ only if the *exponential* exp is surjective. As is well-known, for $K = \mathbf{R}$ the image $\exp \mathbf{R}$ is the multiplicative group $\mathbf{R}_+^*$ of positive real numbers; hence, all of $\boldsymbol{SO}(1,1)$ can only be attained if $\alpha = -\exp t$ is also permitted:

$$\boldsymbol{SO}(1,1) = \{\pm \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix} \mid t \in \mathbf{R}\}. \tag{33}$$

Obviously, for $\boldsymbol{PSO}(1,1)$ this is irrelevant. The whole group $\boldsymbol{O}(1,1)$ consists of four parts, its *connected components*, each homeomorphic to $\mathbf{R}$; they correspond to the four branches of the pair of hyperbolas

$$x^2 - y^2 = \pm 1.$$

For $K = \mathbf{C}$ the exponential is surjective. By Proposition 2.13, the whole group $\boldsymbol{SO}(2,\mathbf{C})$ is homeomorphic to the punctured Gauß plane $\mathbf{C}^* = \mathbf{C} \setminus \{0\}$ and connected; $\boldsymbol{O}(2,\mathbf{C})$ consists of two such domains defined by $N(g) = \pm 1$, respectively. The topological notions used here are based on the transfer of topologies from $\mathbf{R}$ and $\mathbf{C}$, respectively, using the bijective maps above, which thereby become natural homeomorphisms.

Example 10. Let $\boldsymbol{P}^n$ be a projective space over a field K with the polarity F determined by a scalar product, for which there is an orthonormal basis $(\mathfrak{e}_j)$; this situation will be called a *standard scalar product* on the associated vector space $\boldsymbol{V}^{n+1}$:

$$\langle \mathfrak{e}_i, \mathfrak{e}_j \rangle = \delta_{ij}, \qquad i, j = 0, \dots, n. \tag{34}$$

For this case we now want to determine the orthogonal group $\boldsymbol{O}(2,K)$. Let

$$S^1_K := \{\mathfrak{a} \in \boldsymbol{V}^2 | \; \langle \mathfrak{a}, \mathfrak{a} \rangle = 1\} \tag{35}$$

denote the *unit circle* in $\boldsymbol{V}^2$. Furthermore, let $[\mathfrak{x}, \mathfrak{y}]$ be the volume function on $\boldsymbol{V}^2$ satisfying $[\mathfrak{e}_1, \mathfrak{e}_2] = 1$ (Proposition I.4.7.5). We claim that the map

$$g \in \boldsymbol{SO}(2,K) \longmapsto \mathfrak{a} := g\mathfrak{e}_1 = \mathfrak{e}_1 \alpha + \mathfrak{e}_2 \beta \in S^1_K \tag{36}$$

is bijective; in fact, if $\mathfrak{a} = \mathfrak{e}_1 \alpha + \mathfrak{e}_2 \beta \in S^1_K$, then there is a unique transformation $g \in \boldsymbol{SO}(2,K)$ such that $\mathfrak{a} = g\mathfrak{e}_1$. With respect to the basis $(\mathfrak{e}_1, \mathfrak{e}_2)$ it has the matrix

$$g \mathrel{\hat{=}} \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}, \quad \alpha^2 + \beta^2 = 1. \tag{37}$$

The group of these matrices, which is isomorphic to $\boldsymbol{SO}(2,K)$, will also be denoted by $\boldsymbol{SO}(2,K)$; it is Abelian. The transformations belonging to $\boldsymbol{O}(2,K) \setminus \boldsymbol{SO}(2,K)$ have matrices

$$g \mathrel{\hat{=}} \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}, \quad \alpha^2 + \beta^2 = 1.$$

As the *oriented angle* $\varphi(\mathfrak{a}, \mathfrak{b})$, $\mathfrak{a}, \mathfrak{b} \in S^1_K$, we take the unique group element $\varphi(\mathfrak{a}, \mathfrak{b}) = g \in \boldsymbol{SO}(2,K)$ for which $g\mathfrak{a} = \mathfrak{b}$. Note that this is the "correct" definition for the angle even in the case $K = \mathbf{R}$, since there the number φ is only determined modulo 2π.

Having fixed the unit vector $\mathfrak{e}_1$ (and thereby the orthonormal basis $(\mathfrak{e}_1, \mathfrak{e}_2)$ with $v(\mathfrak{e}_1, \mathfrak{e}_2) = 1$), we may use (36) for the identification:

$$\varphi = \varphi(g) \Longleftrightarrow \varphi = \varphi(\mathfrak{e}_1, g\mathfrak{e}_1) \Longleftrightarrow g\mathfrak{e}_1 = \mathfrak{e}_1 \cos\varphi + \mathfrak{e}_2 \sin\varphi$$

$$\text{with} \quad \cos\varphi := \langle \mathfrak{e}_1, g\mathfrak{e}_1 \rangle, \ \sin\varphi := v(\mathfrak{e}_1, g\mathfrak{e}_1) \quad (g \in \boldsymbol{SO}(2, K)). \tag{38}$$

Hence the trigonometric functions $\cos\varphi, \sin\varphi$ are defined as matrix elements on $\boldsymbol{SO}(2, K)$. By (37), we obtain the matrix representation that is well-known in the case $K = \mathbf{R}$

$$g \mathrel{\hat{=}} \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}, \qquad \alpha = \cos\varphi, \beta = \sin\varphi.$$

In connection with the notion of angle it is common to write the group operation for the angles occurring in identification (38) additively:

$$g = g(\varphi), \ g(\varphi_1 + \varphi_2) = g(\varphi_1) \cdot g(\varphi_2),$$
$$\varphi = \varphi(g), \ \varphi(g_1 \cdot g_2) = \varphi(g_1) + \varphi(g_2).$$

Having agreed on this, $\sin\varphi, \cos\varphi$ become functions on the group $\boldsymbol{SO}(2, K)$ with values in K satisfying the familiar *trigonometric addition formulas*:

$$\cos^2\varphi + \sin^2\varphi = 1, \tag{39}$$
$$\cos(\varphi_1 + \varphi_2) = \cos\varphi_1 \cos\varphi_2 - \sin\varphi_1 \sin\varphi_2, \tag{40}$$
$$\sin(\varphi_1 + \varphi_2) = \sin\varphi_1 \cos\varphi_2 + \sin\varphi_2 \cos\varphi_1. \tag{41}$$

These relations are derived simply by applying the rules of matrix multiplication. Obviously, we also have $\cos 0 = 1, \sin 0 = 0$ and

$$\cos(-\varphi) = \cos\varphi, \ \sin(-\varphi) = -\sin\varphi. \tag{42}$$

Finally, for $\mathfrak{a}, \mathfrak{b} \in S^1_K$ we have

$$\cos\varphi(\mathfrak{a}, \mathfrak{b}) = \langle \mathfrak{a}, \mathfrak{b} \rangle, \ \sin\varphi(\mathfrak{a}, \mathfrak{b}) = [\mathfrak{a}, \mathfrak{b}]. \tag{43}$$

The group element $\varphi(\mathfrak{e}_1, \mathfrak{e}_2)$ will be denoted by $\pi/2$, or, expressed differently,

$$\cos(\pi/2) = 0, \ \sin(\pi/2) = 1. \tag{44}$$

The addition formulas (40),(41) imply the familiar values

$$\cos m\pi = (-1)^m, \ \sin m\pi = 0, \quad m \in \mathbf{Z}. \tag{45}$$

□

Example 11. Now assume that -1 is not a square in K, and consider the algebraic extension $\tilde{K} := K(\mathrm{i})$, $\mathrm{i} := \sqrt{-1}$, of the field K. By $\tilde{\boldsymbol{V}}^2$ we denote the $\tilde{K}$-extension of the vector space $\boldsymbol{V}^2$, cf. Section 1.10.4 or Example II.7.9.9. Extending the standard scalar product bilinearly to $\tilde{\boldsymbol{V}}^2$ the space $[\tilde{\boldsymbol{V}}^2, \tilde{K}, \langle , \rangle]$

becomes a vector space with standard scalar product over $\tilde{K}$, which, moreover, is neutral. The isotropic subspaces are spanned by the vectors

$$\mathfrak{z}_1 := (\mathfrak{e}_1 \,\mathrm{i} + \mathfrak{e}_2)/2, \ \mathfrak{z}_2 := (\mathfrak{e}_2 - \mathfrak{e}_1 \,\mathrm{i})/2. \tag{46}$$

By (2.25), with respect to this isotropic basis each transformation $g \in \boldsymbol{O}(2, \tilde{K})$ has the representation

$$g \in \boldsymbol{SO}(2, \tilde{K}) \longmapsto \gamma \in \tilde{K}^* \text{ with } g\mathfrak{z}_1 = \mathfrak{z}_1\gamma, \, g\mathfrak{z}_2 = \mathfrak{z}_2\gamma^{-1}, \tag{47}$$

where $g \mapsto \gamma$ is a group isomorphism. Extending the transformations $g \in \boldsymbol{O}(2, K)$ linearly we obtain equally denoted orthogonal transformations $g \in \boldsymbol{O}(2, \tilde{K})$. This leads to the canonical embeddings $\boldsymbol{O}(2, K) \subset \boldsymbol{O}(2, \tilde{K})$ and $\boldsymbol{SO}(2, K) \subset \boldsymbol{SO}(2, \tilde{K})$. From (38) and (46) we obtain for $g \in \boldsymbol{SO}(2, K)$

$$g\mathfrak{z}_1 = \mathfrak{z}_1 \,\mathrm{e}^{\mathrm{i}\,\varphi}, \ g\mathfrak{z}_2 = \mathfrak{z}_2 \,\mathrm{e}^{-\,\mathrm{i}\,\varphi}, \quad g \in \boldsymbol{SO}(2, K), \tag{48}$$

this time using a formula, well-known for $K = \mathbf{R}$, as the definition:

$$\mathrm{e}^{\mathrm{i}\,\varphi} := \cos\varphi + \mathrm{i}\sin\varphi \in \tilde{K}^*. \tag{49}$$

This is an isomorphism from $\boldsymbol{SO}(2, K)$ onto the subgroup

$$S^1_K = \{z \in \tilde{K} | \; z\bar{z} = \alpha^2 + \beta^2 = 1\} \subset \tilde{K}^*, \tag{50}$$

which has to be identified with S^1_K. Here $z = \alpha + \mathrm{i}\,\beta \mapsto \bar{z} := \alpha - \mathrm{i}\,\beta$ denotes the canonical isomorphism of the extension $\tilde{K} = K(\mathrm{i})$. Lastly, with the notations (27), we obtain

$$c(\mathrm{e}^{\mathrm{i}\,\varphi}) = \cos(\varphi), \ s(\mathrm{e}^{\mathrm{i}\,\varphi}) = \mathrm{i}\sin(\varphi). \tag{51}$$

Now we consider the pseudo-orthonormal basis (26), which is related to the orthonormal one $(\mathfrak{e}_1, \mathfrak{e}_2)$ via

$$\mathfrak{a}_1 = \mathfrak{e}_1 \,\mathrm{i}, \ \mathfrak{a}_2 = \mathfrak{e}_2.$$

Since (38) makes sense also for $g \in \boldsymbol{SO}(2, \tilde{K})$, we obtain an extension of the trigonometric functions from $\boldsymbol{SO}(2, K)$ to $\boldsymbol{SO}(2, \tilde{K})$; thus, for all $g \in \boldsymbol{SO}(2, \tilde{K})$:

$$\begin{aligned} g\mathfrak{e}_1 &= \mathfrak{e}_1 \cos\psi + \mathfrak{e}_2 \sin\psi, \\ g\mathfrak{e}_2 &= -\mathfrak{e}_1 \sin\psi + \mathfrak{e}_2 \cos\psi. \end{aligned}$$

Comparing the matrix of g resulting from this representation in the basis $(\mathfrak{a}_1, \mathfrak{a}_2)$ with that from (27) leads to a generalization of (51):

$$\cos\psi(\gamma) = \frac{\gamma + \gamma^{-1}}{2}, \ \sin\psi(\gamma) = -\,\mathrm{i}\,\frac{\gamma - \gamma^{-1}}{2}, \ \gamma \in \tilde{K}^*. \tag{52}$$

Using the common definitions in the case $K = \mathbf{R}$, the relations (48) lead to equations for $g \in \boldsymbol{SO}(2, K)$, well-known from hyperbolic trigonometry:

$$\cos\varphi = \frac{\mathrm{e}^{\mathrm{i}\,\varphi} + \mathrm{e}^{-\,\mathrm{i}\,\varphi}}{2} =: \cosh(\mathrm{i}\,\varphi), \ \mathrm{i}\sin\varphi = \frac{\mathrm{e}^{\mathrm{i}\,\varphi} - \mathrm{e}^{-\,\mathrm{i}\,\varphi}}{2} =: \sinh(\mathrm{i}\,\varphi). \tag{53}$$

□

Exercise 6. Let $Q \subset \boldsymbol{P}^n(\boldsymbol{V})$ be the quadric corresponding to the standard scalar product $\langle , \rangle$ of the vector space $\boldsymbol{V}^{n+1}$ over a field K. a) Prove that $Q \neq \emptyset$ if and only if -1 can be represented as a sum of n squares. – b) Find an example with $Q \neq \emptyset$, such that -1 is not a square in K.

Exercise 7. Find an example for a vector space $\boldsymbol{V}$ with standard scalar product over a field K with $\sqrt{-1} \notin K$, which has a neutral subspace $\boldsymbol{W} \subset \boldsymbol{V}$. (Hint. Try vector spaces over finite fields.)

Exercise 8. Prove that there exist positive definite scalar products in the rational vector space $\boldsymbol{V}^2 \cong \mathbf{Q}^2$ for which there are no orthonormal bases.

2.4.6 Plane Cone Sections

In this section we consider *ellipses* in a projective plane $\boldsymbol{P}^2$ over a field K with $\operatorname{char} K \neq 2$. This is meant to be a quadric whose equation can be brought into the following normal form in homogeneous coordinates:

$$-(x^0)^2 + (x^1)^2 + (x^2)^2 = 0. \tag{54}$$

Let $\langle , \rangle$ denote the associated symmetric, bilinear scalar product in the vector space $\boldsymbol{V}^3$ corresponding to $\boldsymbol{P}^2 = \boldsymbol{P}(\boldsymbol{V}^3)$. It is easy to verify that $[\boldsymbol{V}^3, \langle , \rangle]$ is a vector space with scalar product of index 1. For $K = \mathbf{R}$ the associated quadric Q can be identified with the unit circle S^1_K: Since each point $\boldsymbol{x} \in Q$ lies in the complement of the line $x^0 = 0$, chosen as the absolute, i.e., in the Euclidean plane complementary to this line, we may suppose $x^0 = 1$. Thus (54) becomes the equation of the unit circle. In general, however, the intersection of Q with the absolute line is not empty; e.g., in the complex projective plane this intersection consists of the points with coordinates $(0, 1, \mathrm{i})$, $(0, 1, -\mathrm{i})$. The next proposition describes the *central projection* of the ellipse from one of its points onto the tangent at another point. This projection allows to introduce the projective structure of a line on the ellipse.

Proposition 7. *Let Q be the ellipse with normal form* (54) *in the projective plane $\boldsymbol{P}^2 = \boldsymbol{P}(\boldsymbol{V}^3)$ over a field K with* $\operatorname{char} K \neq 2$, *consider $\boldsymbol{n}, \boldsymbol{s} \in Q$, $\boldsymbol{n} \neq \boldsymbol{s}$, and denote by $\boldsymbol{T_n}, \boldsymbol{T_s}$ the tangents at $\boldsymbol{n}, \boldsymbol{s}$. Then the map*

$$\begin{gathered} f : \boldsymbol{x} \in Q \longmapsto \boldsymbol{y} := (\boldsymbol{n} \vee \boldsymbol{x}) \wedge \boldsymbol{T_s} \in \boldsymbol{T_s},\ (\boldsymbol{x} \neq \boldsymbol{n}), \\ f(\boldsymbol{n}) := \boldsymbol{T_n} \wedge \boldsymbol{T_s}, \end{gathered} \tag{55}$$

is a bijection from Q onto $\boldsymbol{T_s}$.

Proof. Because of $\boldsymbol{T_s} \cap Q = \{s\}$ and $\boldsymbol{n} \neq \boldsymbol{s}$, the line $\boldsymbol{n} \vee \boldsymbol{y}$ is defined for all $\boldsymbol{y} \in \boldsymbol{T_s}$. Each line $\boldsymbol{A} \neq \boldsymbol{T_n}$ through $\boldsymbol{n}$ meets the quadric Q in precisely one further point $\boldsymbol{x} = \boldsymbol{A} \cap Q \neq \boldsymbol{n}$, and we have $\boldsymbol{T_n} \cap Q = \{\boldsymbol{n}\}$, cf. Example 1.9.8. Hence f is bijective. Here we made use of the fact that the scalar product has index 1, and that, for this reason, Q cannot contain any line (cf. Corollary 2.2). □

By means of the bijection (55) we now transfer a projective scale from $\boldsymbol{T_s}$ into the ellipse Q. Because of $\boldsymbol{n}, \boldsymbol{s} \in Q$, $\boldsymbol{n} \neq \boldsymbol{s}$, we may choose the representing isotropic vectors $\mathfrak{s}, \mathfrak{n}$, $\boldsymbol{s} = [\mathfrak{s}], \boldsymbol{n} = [\mathfrak{n}]$, such that $\langle \mathfrak{s}, \mathfrak{n} \rangle = -1/2$. Then

$$\mathfrak{a}_0 := \mathfrak{s} + \mathfrak{n},\ \mathfrak{a}_1 := \mathfrak{s} - \mathfrak{n},\ \mathfrak{a}_2 = \mathfrak{a}_0 \times \mathfrak{a}_1 \tag{56}$$

is a pseudo-orthonormal basis for $\boldsymbol{V}^3$. Moreover, because of

$$\langle \mathfrak{s}, \mathfrak{a}_2 \rangle = \langle \mathfrak{n}, \mathfrak{a}_2 \rangle = 0,$$

$\boldsymbol{b} := [\mathfrak{a}_2]$ is the point of intersection of the tangents $\boldsymbol{b} = \boldsymbol{T}_n \wedge \boldsymbol{T_s}$. The vectors

$$\mathfrak{c}_0 := \mathfrak{s},\ \mathfrak{c}_1 := \mathfrak{a}_2$$

form a basis of the vector space corresponding to $\boldsymbol{T_s}$, which, by (1.1.4), determines a projective scale ξ on $\boldsymbol{T_s}$ satisfying

$$\xi(\boldsymbol{s}) = 0,\ \xi(\boldsymbol{b}) = \infty.$$

An elementary calculation of the inverse function f^{-1} leads to a parameter representation for the ellipse with the projective scale ξ as parameter. This provides a proof of Proposition 7 by direct computation:

$$\xi \in \hat{K} \longmapsto \boldsymbol{x}(\xi) = [\mathfrak{s} + \mathfrak{a}_2\xi + \mathfrak{n}\xi^2] \in Q. \tag{57}$$

Exercise 9. Prove under the assumptions of Proposition 7: a) With respect to the basis (56),

$$\varphi \in \boldsymbol{SO}(2, K) \longmapsto [\mathfrak{a}_0 + \mathfrak{a}_1 \cos\varphi + \mathfrak{a}_2 \sin\varphi] \in Q$$

is a bijection that allows to identify Q with $S^1_K \subset \mathfrak{L}(\mathfrak{a}_1, \mathfrak{a}_2)$ (cf. Example 10). – b) Using a) and the basis $\mathfrak{c}_0, \mathfrak{c}_1$ of the vector space associated with $\boldsymbol{T_s}$ defined above, the map (55) has the coordinate representation

$$\boldsymbol{x} = [\mathfrak{a}_0 + \mathfrak{a}_1 \cos\varphi + \mathfrak{a}_2 \sin\varphi] \in Q \longmapsto [\mathfrak{c}_0 + \mathfrak{c}_1\xi(\varphi)] \in \boldsymbol{T_s} \text{ with}$$

$$\xi(\varphi) = \frac{\sin\varphi}{1 + \cos\varphi} \text{ for } \varphi \neq \pi,\ \xi(\pi) = \infty. \tag{58}$$

c) Compute the inverse of (58):

$$\cos\varphi = \frac{1 - \xi^2}{1 + \xi^2},\ \sin\varphi = \frac{2\xi}{1 + \xi^2}.$$

Exercise 10. For $K = \mathbf{R}$ relation (58) can also be written as $\xi(\varphi) = \tan(\varphi/2)$. Consider the field $\mathbf{Z}_3$ and show that, in general, not every angle $\varphi \in \boldsymbol{SO}(2, K)$ can be bisected, i.e., there does not always exist a $\psi \in \boldsymbol{SO}(2, K)$ such that $\varphi = 2\psi$. (Recall that the operation in $\boldsymbol{SO}(2, K)$ is written additively, cf. formulas (38)–(41).)

Exercise 11. Prove that $\boldsymbol{O}(2, K)$ is never Abelian for a field K with char $K \neq 2$. If, in addition, K is a finite field with m elements, then $\boldsymbol{O}(2, K)$ has precisely $2m + 2$ elements.

Transfer the projective scale from the tangent $\boldsymbol{T_s}$ into the ellipse Q, which thereby receives the projective structure of a line. Here we may replace the tangent by an arbitrary line not containing the point $\boldsymbol{n}$; in fact, any two such lines are related to one another by the central projection with center $\boldsymbol{n}$. Denoting by $\hat{f}$ the bijection generated by means of another point $\hat{\boldsymbol{n}} \in Q$, the composition $h := \hat{f} \circ f^{-1}$ preserves cross ratios by construction. Hence it is a projectivity of the tangent $\boldsymbol{T_s}$. For the one-dimensional projective geometry of line pencils, cf. Section 1.6.3, this observation results in a proposition going back to J. Steiner , cf. W. Blaschke [14], S. 48:

Corollary 8. *Under the assumptions of Proposition 7, let $\boldsymbol{n}_1, \boldsymbol{n}_2 \in Q$ be two different points of the ellipse and denote by τ_i, $i = 1, 2$, the line pencils with centers $\boldsymbol{n}_i$. Then the maps*

$$\begin{aligned} &F_1 : \boldsymbol{x} \in Q, \boldsymbol{x} \neq \boldsymbol{n}_1, \longmapsto \boldsymbol{n}_1 \vee \boldsymbol{x} \in \tau_1,\ F_1(\boldsymbol{n}_1) = \boldsymbol{T}_{\boldsymbol{n}_1}, \\ &F_2 : \boldsymbol{x} \in Q, \boldsymbol{x} \neq \boldsymbol{n}_2, \longmapsto \boldsymbol{n}_2 \vee \boldsymbol{x} \in \tau_2,\ F_2(\boldsymbol{n}_2) = \boldsymbol{T}_{\boldsymbol{n}_2}, \\ &H = F_2 \circ F_1^{-1} : \tau_1 \to \tau_2, \end{aligned}$$

are projectivities of the projective geometries defined in the line pencils or on Q, respectively. □

Obviously, for the projectivity H we have

$$H(\boldsymbol{T}_{\boldsymbol{n}_1}) = \boldsymbol{n}_1 \vee \boldsymbol{n}_2,\ H(\boldsymbol{n}_1 \vee \boldsymbol{n}_2) = \boldsymbol{T}_{\boldsymbol{n}_2}. \tag{59}$$

Thus, according to Corollary 8, two line pencils whose centers lie on an ellipse are projectively related via the points of the ellipse. The converse of this fact is the *definition of the ellipse according to J. Steiner*:

Proposition 9. *Let $\boldsymbol{P}^2$ be a projective plane over a field K, char $K \neq 2$. In $\boldsymbol{P}^2$ consider two line pencils τ_1, τ_2 with centers $\boldsymbol{n}_1, \boldsymbol{n}_2 \in \boldsymbol{P}^2$, $\boldsymbol{n}_1 \neq \boldsymbol{n}_2$, and a projectivity $H : \tau_1 \to \tau_2$ of the line pencils such that $H(\boldsymbol{n}_1 \vee \boldsymbol{n}_2) \neq \boldsymbol{n}_1 \vee \boldsymbol{n}_2$. Then the point set*

$$Q := \{\boldsymbol{g} \wedge H(\boldsymbol{g}) | \boldsymbol{g} \in \tau_1\}$$

is an ellipse.

Proof. Choose a suitable basis $(\mathfrak{a}_0, \mathfrak{a}_1, \mathfrak{a}_2)$ for the vector space $\boldsymbol{V}^3$ of $\boldsymbol{P}^2$ that is adapted to the above configuration. Moreover, let

$$\boldsymbol{n}_1 = [\mathfrak{a}_1],\ \boldsymbol{n}_2 = [\mathfrak{a}_2].$$

Then the connecting line $\boldsymbol{g}_0 := \boldsymbol{n}_1 \vee \boldsymbol{n}_2$ is described by the covector $\mathfrak{u}^0$ in the dual basis:

$$\boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{g}_0 \Longleftrightarrow \mathfrak{u}^0(\mathfrak{x}) = x^0 = 0.$$

Consider the line $\boldsymbol{g}_2 \in \tau_1$ with $H(\boldsymbol{g}_2) = \boldsymbol{g}_0$ and set $\boldsymbol{g}_1 := H(\boldsymbol{g}_0) \in \tau_2$. Then we have $\boldsymbol{g}_1 \neq \boldsymbol{g}_2$: Indeed, $\boldsymbol{g}_1 = \boldsymbol{g}_2$ would imply $\boldsymbol{g}_1 = \boldsymbol{n}_1 \vee \boldsymbol{n}_2 = \boldsymbol{g}_2 = \boldsymbol{g}_0$ and hence $H(\boldsymbol{g}_0) = \boldsymbol{g}_0$, contradicting the assumptions. Thus the point $\boldsymbol{p} := \boldsymbol{g}_1 \wedge \boldsymbol{g}_2$ is independent of $\boldsymbol{n}_1, \boldsymbol{n}_2$; choose $\mathfrak{a}_0$ as its representative, $\boldsymbol{p} = [\mathfrak{a}_0]$. So, the pencil τ_1 is determined by the covectors from $[\mathfrak{a}_1]^\perp = [\mathfrak{u}^0, \mathfrak{u}^2]$, and τ_2 by $[\mathfrak{a}_2]^\perp = [\mathfrak{u}^0, \mathfrak{u}^1]$, where $(\mathfrak{u}^0, \mathfrak{u}^1, \mathfrak{u}^2)$ denotes the basis dual to $(\mathfrak{a}_0, \mathfrak{a}_1, \mathfrak{a}_2)$. Let H be the map induced by the linear map $a : [\mathfrak{u}^0, \mathfrak{u}^2] \to [\mathfrak{u}^0, \mathfrak{u}^1]$. Since

$$H(\boldsymbol{g}_0) = \boldsymbol{g}_1 = \boldsymbol{n}_2 \vee \boldsymbol{p} : \; x^1 = 0,$$

we have $a(\mathfrak{u}^0) = \mathfrak{u}^1\alpha$. The condition

$$H(\boldsymbol{g}_2) = \boldsymbol{g}_0, \; \boldsymbol{g}_2 = \boldsymbol{n}_1 \vee \boldsymbol{p} : x^2 = 0,$$

yields $a(\mathfrak{u}^2) = \mathfrak{u}_0\beta$. Letting in both pencils the unit lines $\mathfrak{u}^0 + \mathfrak{u}^2, \mathfrak{u}^0 + \mathfrak{u}^1$ correspond to one another, we obtain a in the form of an assignment by equal coordinates:

$$a(\mathfrak{u}^0\alpha + \mathfrak{u}^2\beta) = \mathfrak{u}^0\beta + \mathfrak{u}^1\alpha.$$

For the point of intersection, $\boldsymbol{x} = \boldsymbol{g} \wedge H(\boldsymbol{g}) = [\mathfrak{x}]$, $\boldsymbol{g} \in \tau_1$, the following equations hold

$$\begin{aligned}(\mathfrak{u}^0\alpha + \mathfrak{u}^2\beta)(\mathfrak{x}) &= x^0\alpha + x^2\beta = 0,\\ (\mathfrak{u}^0\beta + \mathfrak{u}^1\alpha)(\mathfrak{x}) &= x^0\beta + x^1\alpha = 0.\end{aligned}$$

For $\beta = 0$ this yields $\boldsymbol{x} = \boldsymbol{g}_0 \wedge \boldsymbol{g}_1 = \boldsymbol{n}_2$. In case $\beta \neq 0$ we set $\beta = 1$ and, having eliminated α, we obtain the equation of an ellipse in the form

$$(x^0)^2 - x^1x^2 = 0,$$

which is brought into its normal form by the substitution

$$x^1 = y^0 - y^1, \; x^2 = y^0 + y^1, \; x^0 = y^2.$$

By Corollary 8, we thus reach every point of the ellipse. □

Exercise 12. Using Steiner's definition of the ellipse prove that there is a unique ellipse passing through any five points of the plane $\boldsymbol{P}_K^2$ in general position, if $\operatorname{char} K \neq 2$. Hint. Consider the line pencil τ_1, τ_2 through two of the points and use the three remaining ones to construct a suitable projectivity $H : \tau_1 \to \tau_2$.

2.5 Spherical and Elliptic Geometry

In this part we will investigate two classical real geometries, spherical and elliptic, from the projective point of view; so let the base field always be $K = \mathbf{R}$. Because of the practical applications of spherical geometry ($n = 2$, geodesy, astronomy, see e.g., the book, [13] by H.-G. Bigalke), and due to their distinguished position as geometries of constant positive curvature (cf. [113] by J. A. Wolf, or [108] by E. B. Vinberg (ed.)) both geometries occupy a central position in the foundations of geometry. Locally, both geometries do not differ; globally, however, the sphere is a twofold covering of the projective space of the same dimension. This fact leads to close connections between the two geometries.

2.5.1 Spherical as a Covering of Elliptic Geometry

Elliptic geometry was already defined in Example 3.1 as the real projective geometry of a non-degenerate polarity F of index 0. So let $\boldsymbol{V}^{n+1}$ be an $(n+1)$-dimensional Euclidean vector space, and let $\pi : \boldsymbol{V}^{n+1} \to \boldsymbol{P}^n_o$ be the canonical map (1.1.2) onto the associated projective space. The Euclidean, i.e. positive definite, scalar product then defines the polarity F:

$$F : \boldsymbol{A} \in \mathfrak{P}^n \longmapsto \boldsymbol{A}^\perp \in \mathfrak{P}^n.$$

By (1.24) we know that $\boldsymbol{PO}(n+1) = \boldsymbol{O}(n+1)/\{\pm I_{n+1}\}$ is the isotropy group of the polarity F, which, in the sense of F. Klein's Erlanger Programm, determines the elliptic geometry of $\mathfrak{P}^n$. If n is even, $n = 2m$, then $\det(-I_{n+1}) = -1$, and this implies the isomorphy

$$\boldsymbol{PO}(2m+1) \cong \boldsymbol{SO}(2m+1). \tag{1}$$

For what follows, however, it will be advisable to consider the non-effective actions of $\boldsymbol{O}(n+1)$ on the spaces and manifolds occurring in projective geometry.

The orbits of the group $\boldsymbol{O}(n+1)$ for its linear action on the Euclidean vector space $\boldsymbol{V}^{n+1}$ are the origin $\mathfrak{o}$ and the *spheres of radius* r:

$$S^n(r) := \{\mathfrak{x} \in \boldsymbol{V}^{n+1} \mid |\mathfrak{x}| = r\}. \tag{2}$$

In fact, not only $\boldsymbol{O}(n+1)$, but also the special orthogonal group $\boldsymbol{SO}(n+1)$ both act transitively on $S^n(r)$, which is a consequence of the special case $k = 1$ for the following exercise:

Exercise 1. Prove the following simple refinement of E. Witt's Theorem for Euclidean vector spaces: If $\boldsymbol{A}^k, \boldsymbol{B}^k$ are oriented subspaces of the oriented Euclidean vector space $\boldsymbol{V}^n$, $0 \le k < n$, then there is a map $g \in \boldsymbol{SO}(n)$ with $g(\boldsymbol{A}^k) = \boldsymbol{B}^k$ that transforms an arbitrarily chosen, positively oriented, orthonormal basis of $\boldsymbol{A}^k$ into an arbitrarily chosen, positively oriented, orthonormal basis of $\boldsymbol{B}^k$ (cf. Corollary I.6.2.1).

In the sequel, let $(\mathfrak{e}_j)$, $j = 0, \dots, n$, denote a fixed orthonormal basis of the Euclidean vector space $\boldsymbol{V}^{n+1}$. We identify the points of the n-sphere with their position vectors with respect to the chosen fixed origin of the Euclidean point space. As the origin of $S^n(r)$ we choose the point $\mathfrak{e}_0 r$; then, with respect to the basis $(\mathfrak{e}_i)$ and in the usual representation by orthogonal matrices (a_{ij}) (cf. I.6.2, (27)–(29)) its isotropy group $\boldsymbol{O}(n)$ is described by

$$a_{00} = 1, \; a_{0j} = a_{i0} = 0 \text{ for } i, j = 1, \dots, n. \tag{3}$$

The isotropy groups of all the spheres $S^n(r)$ in the Euclidean group $\boldsymbol{E}(n+1)$ coincide, they are all equal to the orthogonal group $\boldsymbol{O}(n+1)$. Within the transitive transformation group $[\boldsymbol{O}(n+1), S^n(r)]$, the isotropy group of the point $\mathfrak{e}_0 r \in S^n(r)$ is equal to $\boldsymbol{O}(n)$, so that all these spheres are even isomorphic as homogeneous spaces (cf. Example I.1.5.3):

$$S^n(r) \cong \boldsymbol{O}(n+1)/\boldsymbol{O}(n).$$

Here, $\boldsymbol{O}(n+1)/\boldsymbol{O}(n)$ denotes the quotient of the group $\boldsymbol{O}(n+1)$ by its subgroup $\boldsymbol{O}(n)$, cf. Appendix A.3 or Definition I.3.1.1; in geometry, it is common to call it a *quotient space*. The radius $r > 0$ is just a parameter describing the relative size of the sphere with respect to a fixed scale. *Spherical geometry* is now understood to be the geometry of the homogeneous space $[\boldsymbol{O}(n+1), S^n(r)]$ in the sense of F. Klein's Erlanger Programm. Until the end of this section we will use the abbreviations

$$G_n := \boldsymbol{O}(n+1), \; SG_n := \boldsymbol{SO}(n+1).$$

To have a clear picture, we consider an n-dimensional sphere as embedded into the Euclidean $(n+1)$-dimensional vector space by its very definition (2) (or, via $\boldsymbol{x} = \boldsymbol{o} + \mathfrak{x}$, into the associated point space). Disregarding the actual value of the radius r, e.g. setting $r = 1$, all the following considerations remain valid for the transformation group $[G_n, G_n/G_{n-1}]$ as well. In the following, we will write $S^n := S^n(1)$.

We now restrict the canonical map $\pi : \boldsymbol{V}^{n+1} \to \boldsymbol{P}^n$ to a particular n-sphere:

$$p := \pi_{|S^n(r)} : \mathfrak{x} \in S^n(r) \longmapsto \boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{P}^n. \tag{4}$$

Obviously, p is surjective; each point $\boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{P}^n$ has precisely two pre-images

$$p^{-1}(\boldsymbol{x}) = \{\mathfrak{x}, -\mathfrak{x}\}.$$

In general, the pre-image $p^{-1}(\boldsymbol{A}) = \boldsymbol{A} \cap S^n(r)$ of a k-plane $\boldsymbol{A} \subset \boldsymbol{P}^n$ is the intersection of the corresponding, equally denoted $(k+1)$-dimensional subspace $\boldsymbol{A} \subset \boldsymbol{V}^{n+1}$ with the n-sphere $S^n(r)$; these intersections are called *great k-spheres*, $k = -1, \dots, n$. The *spherical polarity*, also denoted by F, is defined by orthogonality:

$$F : \boldsymbol{A} \cap S^n(r) \longmapsto \boldsymbol{A}^{\perp} \cap S^n(r). \tag{5}$$

Obviously, $\boldsymbol{A} \cap S^n(r) \subset \boldsymbol{A}$ itself is a k-dimensional sphere of radius r, and the great sphere polar to it is an $(n-k-1)$-sphere of radius r. One-dimensional great spheres are called *great circles*. In particular, the great 0-spheres are pairs of diametrically opposite or antipodal points; for purely formal reasons, because of $F(S^n(r)) = \boldsymbol{o} \cap S^n(r) = \emptyset$, we consider the empty set as a (-1)-dimensional great sphere. On S^2 with the usual geographic coordinates, the 0-sphere consisting of North and South Pole is the polar of the equator.

Exercise 2. Prove: a) The map

$$S : \boldsymbol{A} \in \mathfrak{P}^n \longmapsto S(\boldsymbol{A}) := \boldsymbol{A} \cap S^n(r)$$

is a G_n-isomorphy from the lattice $\mathfrak{P}^n$ of subspaces in elliptic space onto the lattice of great spheres contained in $S^n(r)$; moreover, the corresponding polarities satisfy

$$F(S(\boldsymbol{A})) = S(F(\boldsymbol{A})).$$

b) Extending the projection (4) by $p(M) := [M]$ to arbitrary subsets $M \subset S^n(r)$ we obtain $p(S(\boldsymbol{A})) = \boldsymbol{A}$, and $S(p(M))$ is the smallest great sphere containing M. – c) The points $\mathfrak{x}_i \in S^n(r)$, $i = 0, \ldots, k, k \leq n$, are said to be *in general position* if they are not contained in any great $(k-1)$-sphere. Given $k+1$ points in general position, there is a unique great k-sphere containing them.

The great k-spheres play the part of k-planes in spherical geometry; they are particular subspheres: A set $\Sigma^k \subset S^n(r)$ is called a ***k-subsphere*** if there is a Euclidean $(k+1)$-plane $\boldsymbol{H}^{k+1} \subset \boldsymbol{E}^{n+1}$ such that $\Sigma^k = S^n(r) \cap \boldsymbol{H}^{k+1}$, $0 \leq k \leq n$, and if, in addition, Σ^k contains at least two points. Thus, a subsphere is a great sphere if the Euclidean $(k+1)$-plane defining it contains the center of $S^k(r)$, i.e. the origin of the coordinate system. So, e.g., the circles of latitude for the geographic coordinates on a sphere S^2 are 1-subspheres, among which only the equator is a great circle. Some elementary properties of subspheres are stated in the following exercise:

Exercise 3. Prove that each k-subsphere $\Sigma^k \subset S^n(r)$ uniquely determines the $(k+1)$-plane $\boldsymbol{H}^{k+1}$ generating it. It is a hypersphere in $\boldsymbol{H}^{k+1}$, whose center $\mathfrak{m}$ is the foot of the perpendicular from the center $\boldsymbol{o}$ of $S^n(r)$ onto $\boldsymbol{H}^{k+1}$, and whose Euclidean radius is equal to $\sqrt{r^2 - c^2}$, where c denotes the distance from $\boldsymbol{o}$ to $\boldsymbol{H}^{k+1}$, cf. Exercise I.6.2.3.

Obviously, an arbitrary Euclidean $(k+1)$-plane cuts out a k-subsphere in $S^n(r)$ if and only if its distance $c \geq 0$ to the origin $\boldsymbol{o}$ is less than r; $c = 0$ just leads to the great k-spheres.

Example 1. The restriction of the canonical map p to an open hemisphere

$$S^n_+(r, \mathfrak{a}) := \{\mathfrak{x} \in S^n(r) | \langle \mathfrak{x}, \mathfrak{a} \rangle > 0\}, \qquad \mathfrak{a} \neq \mathfrak{o}, \tag{6}$$

is injective. Identifying antipodal points in the *great hypersphere* bounding the hemisphere,

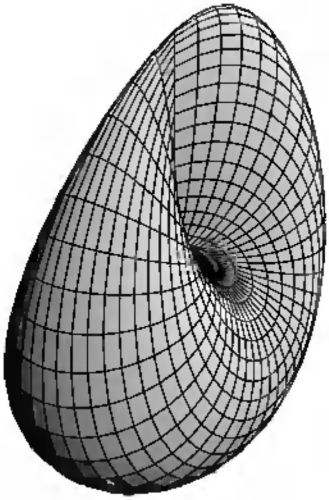


Fig. 2.2. A cross-cap.

$$S^{n-1}(r, \mathfrak{a}) := \{\mathfrak{x} \in S^n(r) | \langle \mathfrak{x}, \mathfrak{a} \rangle = 0\}, \qquad \mathfrak{a} \neq \mathfrak{o}, \tag{7}$$

defines a bijective correspondence with projective space; once a great hypersphere (7) is distinguished, it is frequently called the *equator*. By means of the bijection, or else considering p as a topological covering, topological properties of the projective space $\boldsymbol{P}^n$ can be derived from those of the n-sphere. Locally, more precisely, restricted to any hemisphere, p is a homeomorphism, and, transferring the differentiable structure, even a diffeomorphism. This way, the real projective space $\boldsymbol{P}^n$ becomes a compact n-dimensional differentiable manifold. For odd dimension n it is orientable, whereas it is non-orientable for even n (see, e.g., R. Sulanke. P. Wintgen [103], where an elementary criterion for orientability can also be found). The projective plane is thus a non-orientable, closed surface, that can only be embedded with self-intersections into the three-dimensional Euclidean space $\boldsymbol{E}^3$. Figure 2.2 shows a model of this surface that, apart from the self-intersections, also shows singularities; it is called a *cross-cap*. In Example 4 below we will come back to this representation. Another surface, also displaying singularities is the *Roman Surface* discovered by J. Steiner, Figure 2.3. A detailed discussion of various representations for the projective plane can be found in the book by F. Apery [3], which, in addition, contains numerous illuminating pictures, see also the anthology edited by G. Fischer [37]. Parameter representations and methods to generate these illustrations using the program *Mathematica* by S. Wolfram [114] are described in Alfred Gray's textbook [46]. *Boy's Surface* is a representation of the projective plane free of singularities with self-intersections in the Euclidean space $\boldsymbol{E}^3$, Figure 2.4. W. Boy defined and discussed it in his dissertation [20], which was stimulated by a question of

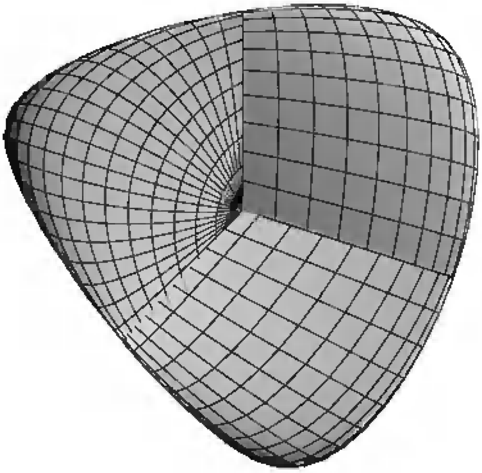


Fig. 2.3. J. Steiner's Roman Surface.

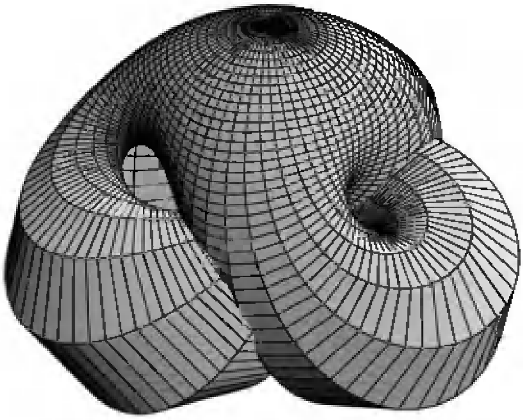


Fig. 2.4. Boy's Surface.

D. Hilbert, cf. also W. Boy [21]. The edges visible in this figure are caused by the numerical approximation; in reality, the surface is smooth. □

2.5.2 Distance and Angle

The most important invariant of the orthogonal group G_n for its linear action on the Euclidean vector space and of all transformation groups derived from it is the scalar product $\langle \mathfrak{x}, \mathfrak{y} \rangle$ determining the polarity F. Using this, the *spherical distance* of two points is defined as

$$\mathfrak{x}, \mathfrak{y} \in S^n(r) \longmapsto d(\mathfrak{x}, \mathfrak{y}) := r \arccos(\langle \mathfrak{x}, \mathfrak{y} \rangle / r^2), \tag{8}$$

i.e., the length of the shortest arc on a great circle through the points $\mathfrak{x}, \mathfrak{y}$. If these points are *antipodal*, $\mathfrak{y} = -\mathfrak{x}$, then the distance between them is independent of the chosen great circle and equal to $r\pi$. Obviously, the distance d has the following properties:

$$0 \leq d(\mathfrak{x}, \mathfrak{y}) \leq r\pi, \quad d(\mathfrak{x}, \mathfrak{y}) = 0 \Longleftrightarrow \mathfrak{x} = \mathfrak{y}, \quad d(\mathfrak{x}, \mathfrak{y}) = d(\mathfrak{y}, \mathfrak{x}) = d(-\mathfrak{x}, -\mathfrak{y}). \quad (9)$$

The triangle inequality, which any distance has to satisfy, will be derived below as a consequence of the law of cosines in spherical trigonometry. The *diameter of the sphere* $S^n(r)$, i.e. the supremum of all point distances $d(\mathfrak{x}, \mathfrak{y})$, is attained for any two antipodal points:

$$d(\mathfrak{x}, -\mathfrak{x}) = r\pi \qquad (\mathfrak{x} \in S^n(r)).$$

Exercise 4. Prove that two pairs $(\mathfrak{x}_1, \mathfrak{x}_2), (\mathfrak{y}_1, \mathfrak{y}_2) \in S^n(r) \times S^n(r)$ are SG_n-congruent (G_n-congruent) if and only if

$$d(\mathfrak{x}_1, \mathfrak{x}_2) = d(\mathfrak{y}_1, \mathfrak{y}_2).$$

In order to define the distance in elliptic geometry, consider again the covering (4). For each $g \in \boldsymbol{GL}(n+1, \mathbf{R})$ we have $g[\mathfrak{x}] = [g\mathfrak{x}]$ by the definition of projectivities, i.e., π is an equivariant map. Moreover, Relation (4) implies that p is an equivariant map for the action of G_n. Extending this action in the obvious way to sets of k-tuples and to power sets,

$$g(x_1, \ldots, x_k) := (gx_1, \ldots, gx_k), \tag{10}$$
$$g(A) := \{gx | x \in A\}, \tag{11}$$

we again obtain actions on these extensions, for which the natural maps are equivariant. A map $f : \boldsymbol{P}^n \times \ldots \times \boldsymbol{P}^n \to M$, M any non-empty set, is a G_n*-invariant* if it satisfies

$$f(g\boldsymbol{x}_1, \ldots, g\boldsymbol{x}_k) = f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) \tag{12}$$

for all $g \in G_n$, $\boldsymbol{x}_\kappa \in \boldsymbol{P}^n$, $\kappa = 1, \ldots, k$. For other transformation groups we will use a similar terminology. Immediately from the definitions we obtain

Lemma 1. *If $f : \boldsymbol{P}^n \times \ldots \times \boldsymbol{P}^n \to M$ is a G_n-invariant on the space of k-tuples of projective points, then the pull-back of the map,*

$$p^* f(\mathfrak{x}_1, \ldots, \mathfrak{x}_k) := f(p(\mathfrak{x}_1), \ldots, p(\mathfrak{x}_k)), \tag{13}$$

again is a G_n-invariant; moreover, the map $F := p^* f$ satisfies the symmetry conditions

$$F(\pm\mathfrak{x}_1, \ldots, \pm\mathfrak{x}_k) = F(\mathfrak{x}_1, \ldots, \mathfrak{x}_k). \tag{14}$$

If, conversely, F is a G_n-invariant on the space of k-tuples of points in the n-sphere $S^n(r)$ satisfying the symmetry condition (14), *then*

$$f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) := F(\mathfrak{x}_1, \ldots, \mathfrak{x}_k) \textit{ with } p(\mathfrak{x}_\kappa) = \boldsymbol{x}_\kappa,\ \kappa = 1, \ldots, k, \tag{15}$$

is a correctly defined G_n-invariant, i.e., it is independent of the chosen representatives $\mathfrak{x}_\kappa \in p^{-1}(\boldsymbol{x}_\kappa)$. □

Obviously, the spherical distance (8) does not satisfy symmetry condition (14). Replacing, however, in (8) the scalar product by its absolute value, we obtain a function defined on $S^n(r) \times S^n(r)$ satisfying (14). By Lemma 1 we can thus define the *distance* of two points in elliptic space as follows:

$$e(\boldsymbol{x}, \boldsymbol{y}) := r \arccos(|\langle \mathfrak{x}, \mathfrak{y} \rangle|) \text{ with } \boldsymbol{x} = [\mathfrak{x}], \boldsymbol{y} = [\mathfrak{y}], \mathfrak{x}, \mathfrak{y} \in S^n(1)). \tag{16}$$

From (9) we immediately obtain corresponding properties for the metric

$$0 \leq e(\boldsymbol{x}, \boldsymbol{y}) \leq r\pi/2, \quad e(\boldsymbol{x}, \boldsymbol{y}) = 0 \Longleftrightarrow \boldsymbol{x} = \boldsymbol{y}, \quad e(\boldsymbol{x}, \boldsymbol{y}) = e(\boldsymbol{y}, \boldsymbol{x}); \tag{17}$$

the triangle inequality will be proved below. We equip the *elliptic space* with the metric defined by (16) and denote it by $\boldsymbol{P}^n(r)$. Obviously, the map p is no *isometry*[1]: In fact, the diameter of the elliptic space is half the diameter of the sphere of radius r:

Exercise 5. a) Prove that in the elliptic space $\boldsymbol{P}^n(r)$ the *diameter*, i.e. the supremum of the distance function, satisfies

$$\sup\{e(\boldsymbol{x}, \boldsymbol{y}) |\ \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{P}^n(r)\} = r\pi/2.$$

b) The projection (4) $p : S^n(r) \to \boldsymbol{P}^n(r)$ satisfies:

$$e(p(\mathfrak{x}), p(\mathfrak{y})) = d(\mathfrak{x}, \mathfrak{y}) \Longleftrightarrow d(\mathfrak{x}, \mathfrak{y}) \leq r\pi/2 \qquad (\mathfrak{x}, \mathfrak{y}) \in S^n(r).$$

c) In any space E with a metric d, let $B(\boldsymbol{z}, \rho)$ denote the *open ball of radius ρ with center $\boldsymbol{z}$*:

$$B(\boldsymbol{z}, \rho) := \{\boldsymbol{x} \in E |\ d(\boldsymbol{x}, \boldsymbol{z}) < \rho\} \qquad (\rho > 0). \tag{18}$$

Prove that for each open ball $B(\boldsymbol{z}, \rho) \subset S^n(r)$ whose radius satisfies $\rho \leq r\pi/4$ the map $p : B(\boldsymbol{z}, \rho) \to p(B(\boldsymbol{z}, \rho))$ is an isometry, i.e.,

$$e(p(\mathfrak{x}), p(\mathfrak{y})) = d(\mathfrak{x}, \mathfrak{y}) \text{ for all } \mathfrak{x}, \mathfrak{y} \in B(\boldsymbol{z}, \rho),$$

and that this does not hold for $\rho > r\pi/4$.

In general, a surjective map $p : E \to P$ between two metric spaces is called a *local isometry* if for each point $\boldsymbol{z} \in E$ there is an open ball $B = B(\boldsymbol{z}, \epsilon)$ such that the restriction $q := p|B : B \to p(B)$ is an isometry. It is easy to prove that an isometry q is always bijective, and that the inverse map q^{-1} is also an isometry. Taking the open balls of a metric space as neighborhoods of their centers generates the *metric topology*, in which notions like convergence, open and closed sets etc. can be defined just as is familiar from real analysis, cf.

[1] As usual, a surjective map $q : E_1 \to E_2$ between two metric spaces is called an *isometry* if it preserves the distance, i.e. $d_2(q(x), q(y)) = d_1(x, y)$ for all $x, y \in E_1$.

e.g. W. Rinow [94], p. 28 ff. A very illuminating introduction into the real projective geometry of the plane and of three-dimensional space highlighting the metric aspects can be found in the book by H. Busemann and P. J. Kelly [25]

In a space E with a metric d let $S(\boldsymbol{z}, \rho)$ denote the *metric hypersphere of radius ρ with center $\boldsymbol{z}$*:

$$S(\boldsymbol{z}, \rho) := \{\boldsymbol{x} \in E |\ d(\boldsymbol{x}, \boldsymbol{z}) = \rho\} \qquad (\rho > 0).$$

Note that the Euclidean radius of a subsphere has to be distinguished from its spherical radius, which is determined by the metric d on the n-sphere:

Exercise 6. Prove: a) As in Exercise 3, let $\Sigma^k \subset S^n(r)$ be a k-subsphere with $c > 0$. Then it is contained in the metric sphere with center $\mathfrak{m}_o = \mathfrak{m}r/|\mathfrak{m}|$ and radius $\rho = r \arccos(c/r)$:

$$\Sigma^k = \{\mathfrak{x} \in S^n(r) \cap \boldsymbol{H}^{k+1} |\ d(\mathfrak{x}, \mathfrak{m}_o) = \rho\}. \tag{19}$$

b) If a $(k+1)$-plane $\boldsymbol{H}^{k+1}$ satisfies the equation $c = r$, then $\boldsymbol{H}^{k+1}$ is tangent to $S^n(r)$, and the intersection $S^n(r) \cap \boldsymbol{H}^{k+1}$ is the point of tangency. – c) If the subsphere is a great sphere ($c = 0$), then (19) holds for each point $\mathfrak{m}_o$ of the great sphere polar to it, and, moreover, $\rho = r\pi/2$. – d) Each k-subsphere Σ^k is the orbit of a subgroup conjugate to the subgroup $\boldsymbol{O}(k+1)$ in $\boldsymbol{O}(n+1)$. – e) Metric hyperspheres are $(n-1)$-subspheres, and vice versa.

Since a Euclidean vector space has no isotropic subspaces, we obtain from Proposition 4.2

Proposition 2. *Two finite point sequences $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)$, $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k)$ in the elliptic space $\boldsymbol{P}^n(r)$ are G_n-congruent if and only if the corresponding distances are equal:*

$$e(\boldsymbol{x}_i, \boldsymbol{x}_j) = e(\boldsymbol{y}_i, \boldsymbol{y}_j),\ \ i, j = 1, \ldots, k.$$

□

To prove this it suffices to note that the points can be represented by vectors in such a way that corresponding scalar products coincide. From the properties of Gram's determinant, cf. Corollary I.6.3.1 or (2.33), we conclude that the dimensions of corresponding subspaces spanned by the points satisfy Assumption a) from Proposition 4.1, so that we may apply it.

Exercise 7. Prove that a result similar to Proposition 2 holds for point sequences in the n-sphere. Hint. First prove that the linear spans $\mathfrak{L}(\mathfrak{x}_1, \ldots, \mathfrak{x}_h), \mathfrak{L}(\mathfrak{y}_1, \ldots, \mathfrak{y}_h)$, $h = 1, \ldots, k$, have equal dimension, then choose suitably adapted orthonormal frames to represent corresponding vectors by equal coordinates.

Using the polarity $F : \boldsymbol{A} \mapsto \boldsymbol{A}^\perp$ we define the notion dual to the distance of two points: The *angle* between hyperplanes $\boldsymbol{X}, \boldsymbol{Y}$ in the elliptic space $\boldsymbol{P}^n(r)$ is taken to be the normalized distance of their poles

$$\angle(\boldsymbol{X},\boldsymbol{Y}) := e(\boldsymbol{X}^{\perp},\boldsymbol{Y}^{\perp})/r. \tag{20}$$

The normalization spoils the duality somewhat, nevertheless, for historical and practical reasons, it is quite common. Since the polarity is a G_n-equivariant involution, the properties (17) of the metric immediately imply

$$0 \le \angle(\boldsymbol{X},\boldsymbol{Y}) \le \pi/2,\ \angle(\boldsymbol{X},\boldsymbol{Y}) = 0 \Longleftrightarrow \boldsymbol{X} = \boldsymbol{Y},\ \angle(\boldsymbol{X},\boldsymbol{Y}) = \angle(\boldsymbol{Y},\boldsymbol{X}). \tag{21}$$

Applying the polarity we obtain from Proposition 2

Corollary 3. *Two finite sequences of hyperplanes* $(\boldsymbol{X}_1,\dots,\boldsymbol{X}_k)$ *and* $(\boldsymbol{Y}_1,\dots,\boldsymbol{Y}_k)$ *in the elliptic space* $\boldsymbol{P}^n(r)$ *are* G_n*-congruent if and only if corresponding angles are equal:*

$$\angle(\boldsymbol{X}_i,\boldsymbol{X}_j) = \angle(\boldsymbol{Y}_i,\boldsymbol{Y}_j),\ i,j = 1,\dots,k.$$

□

Exercise 5, a) implies

$$\sup\{\angle(\boldsymbol{X},\boldsymbol{Y})|\ \boldsymbol{X},\boldsymbol{Y} \subset \boldsymbol{P}^n(r) \text{ hyperplanes}\} = \pi/2. \tag{22}$$

Exercise 8. Define the *angle between two great hyperspheres* to be the angle between the hyperplanes in elliptic space corresponding to them (cf. (20)). Prove the statement analogous to Corollary 3 for sequences of great hyperspheres of $S^n(r)$.

In contrast to elliptic geometry, in spherical geometry the notion of orientation can be defined and applied in an unrestricted way. We call a sphere $S^n(r)$ *oriented* if the Euclidean vector space $\boldsymbol{V}^{n+1}$ containing it is oriented; a *spherical* n*-simplex* $(\mathfrak{x}_0,\dots,\mathfrak{x}_n)$ is positively oriented if the volume of the cuboid formed by its vertices is positive:

$$[\mathfrak{x}_0,\dots,\mathfrak{x}_n] = \det(\mathfrak{x}_0,\dots,\mathfrak{x}_n) > 0;$$

here the determinant is to be taken for the coordinates with respect to the standard basis, which is assumed to be positive. A k-simplex $(\mathfrak{x}_0,\dots,\mathfrak{x}_k)$ with vertices in general position determines a great k-sphere and, at the same time, its orientation, if we consider it as positively oriented; if $(\mathfrak{y}_0,\dots,\mathfrak{y}_k)$ determines the same great k-sphere, then the simplices $(\mathfrak{x}_0,\dots,\mathfrak{x}_k)$, $(\mathfrak{y}_0,\dots,\mathfrak{y}_k)$ are *equally oriented* if the determinant of the basis transformation is positive:

$$\mathfrak{y}_j = \sum_0^k \mathfrak{x}_i\alpha_{ij}, \qquad \det(\alpha_{ij}) > 0. \tag{23}$$

Obviously, this is an equivalence relation with two equivalence classes, each representing one of the two possible orientations of the great k-sphere. If the n-sphere $S^n(r)$ is oriented, then the polarity F defines an involution between oriented great spheres: In the image $F(S^k(r)) = S_1^{n-k-1}(r)$, a generating simplex $(\mathfrak{y}_{k+1}, \ldots, \mathfrak{y}_n)$ is to be considered as positively oriented if it spans a positively oriented n-simplex $(\mathfrak{x}_0, \ldots, \mathfrak{x}_k, \mathfrak{y}_{k+1}, \ldots, \mathfrak{y}_n)$ in $S^n(r)$ together with a positively oriented simplex $(\mathfrak{x}_0, \ldots, \mathfrak{x}_k)$ in $S^k(r)$. It is easy to see that this definition does not depend on the choice of the simplices representing the orientations. The oriented great 0-spheres and, by means of the polarity F the oriented great hyperspheres as well, bijectively correspond to the points of $S^n(r)$. The pole of a great hypersphere spanned by $(\mathfrak{x}_0, \ldots, \mathfrak{x}_{n-1})$ with respect to the polarity F is easily computed using the vector product to be

$$\mathfrak{y} = \mathfrak{x}_0 \times \ldots \times \mathfrak{x}_{n-1} r / |\mathfrak{x}_0 \times \ldots \times \mathfrak{x}_{n-1}|. \tag{24}$$

Note that the vectors of a generating $(n-1)$-simplex are linearly independent, and compare Proposition I.6.3.2. The *angle between two oriented great hyperspheres* is defined as the normalized distance of their poles:

$$0 \leq \angle(S^{n-1}(r), S_1^{n-1}(r)) := d(F(S), F(S_1))/r \leq \pi. \tag{25}$$

Example 2. A great 0-sphere consists of two antipodal points; it is oriented by prescribing their order $S^0 = (-\mathfrak{x}, \mathfrak{x})$, i.e. by choosing $\mathfrak{x}$. An oriented great circle on the unit sphere S^2 in the Euclidean space $\boldsymbol{V}^3$ is determined by a pair $(\mathfrak{x}_0, \mathfrak{x}_1)$ of linearly independent unit vectors; it decomposes S^2 into two hemispheres, one of which contains the pole of the oriented great circle

$$\mathfrak{y} = \mathfrak{x}_0 \times \mathfrak{x}_1 / |\mathfrak{x}_0 \times \mathfrak{x}_1|;$$

we take this to be the *positive hemisphere*:

$$\Sigma_+(\mathfrak{y}) := \{\mathfrak{x} \in S^2 | \langle \mathfrak{x}, \mathfrak{y} \rangle > 0\},$$

and the opposite one, $\Sigma_- = -\Sigma_+$, as the *negative hemisphere.* The great circle as the polar of $\mathfrak{y}$ is determined by the condition

$$S^1(\mathfrak{y}) = \{\mathfrak{x} \in S^2 | \langle \mathfrak{x}, \mathfrak{y} \rangle = 0\};$$

it is oriented by a pair of its points $(\mathfrak{x}_0, \mathfrak{x}_1)$ satisfying

$$[\mathfrak{y}, \mathfrak{x}_0, \mathfrak{x}_1] > 0.$$

The definitions introduced here hold correspondingly for hyperspheres in an n-sphere. □

Example 3. We define a *biangle* as the region in the sphere S^2 bounded by the halves of two great circles intersecting in antipodal points. Two great circles decompose the sphere S^2 into four biangles, among which the two opposite ones are congruent. The four biangles can be determined by choosing one of the four possible combinations of orientations for the great circles bounding them, or as the intersection of the hemispheres defined by the great circles:

$$\Sigma_{1+} \cap \Sigma_{2+},\ \Sigma_{1+} \cap \Sigma_{2-},\ \Sigma_{1-} \cap \Sigma_{2-},\ \Sigma_{1-} \cap \Sigma_{2+}.$$

As the *angle β of the biangle* we take the angle $\beta = \alpha$ between the bounding great circles if the biangle is one of the four determined by the great circles, and $\beta = 2\pi - \alpha$ otherwise; in other words, β, $0 \leq \beta \leq 2\pi$, is the interior angle of the biangle. It is easy to show that the angle characterizes the biangle up to G_2-congruence. □

2.5.3 The Law of Cosines and the Triangle Inequality

In the notations common in elementary geometry let A, B, C be three points in general position on the sphere $S^n(r)$. To fix the concept we will always assume the three points to be arranged into a triple. *Spherical trigonometry* is concerned with the search for invariants of such point triples and related construction problems. Since each triple (A, B, C) uniquely determines a two-dimensional great sphere $S^2(r) \subset S^n(r)$, and since two such great spheres are always G_n-congruent, we may concentrate on the case $n = 2$.[2] The point pairs $(A, B), (B, C), (C, A)$ then define three great circles, the *side circles* of the triple subdividing the sphere S^2 into eight *triangular regions*. Each of these triangular regions is the intersection of three hemispheres determined by the great circles, cf. Figure 2.5. Out of these triangular regions we assign a unique particular one to the triple A, B, C in the following way: Let Σ_C denote the hemisphere bounded by the great circle $S^1(A, B)$ containing C; Σ_B and Σ_A are defined analogously. The intersection $\Sigma_A \cap \Sigma_B \cap \Sigma_C$ is called the *Euler triangle* with *vertices* A, B, C, cf. H.-G. Bigalke [13]. Note that each of the eight triangular regions occurring in the decomposition of the sphere is an Euler triangle for a corresponding triple formed by points from the triple A, B, C and their antipodes.

Exercise 9. Prove that a region bounded by three arcs of great circles in the sphere $S^2(r)$ is an Euler triangle if and only if these arcs are shorter than $r\pi$, and that this holds if and only if the interior angles of the region at all three vertices are less than π.

In contrast to the situation in the Euclidean plane, on the sphere there are different ways to associate with a point triple (A, B, C) a triangular region whose boundary consists of arcs of great circles, AB, BC, CA. One of these regions is the complement of the corresponding Euler triangle. Again we start from the side circles of the triple (A, B, C). As the *side* c we choose one of the *arcs* $AB \subset S^1(A, B)$ of the great circle through A and B into which it is

[2] In the sequel, the parameter r will only be included at important places; if it is omitted, then we refer to the considerations for arbitrarily fixed r.

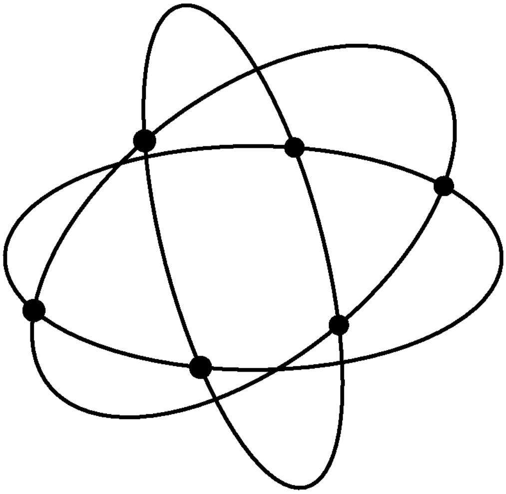


Fig. 2.5. Euler Triangles.

divided by these points. Let $b = CA$, $a = BC$ be the other, correspondingly chosen sides of the triangle. We suppose that the closed curve abc formed by juxtaposing these sides has no self-intersections. According to the Jordan Curve Theorem[3], its complement in S^2 has two components, whose common boundary it is; they are what we called *triangular regions* before. We consider one of these regions as distinguished and take a *spherical triangle* to be a triangular region with the vertices A, B, C and the sides a, b, c defined above. For formal reasons, it may sometimes be useful to consider a hemisphere as a triangle as well: Just choose any three different points on the bounding great circle as its vertices. To introduce orientations on the sides of the triangle and on the great circles they determine we simply take the order of the pairs $(A, B), (B, C), (C, A)$. Furthermore, among the triangles bounded by abc we distinguish the *right-hand* one from the *left-hand* one depending on whether, moving along abc on the sphere, the triangular region lies to the right or to the left of this boundary with respect to the orientation of the sphere. The *angle* α of a triangle at the vertex A is understood to be the angle of the biangle containing the triangular region formed by the adjacent side circles for the sides b and c; the angles β at the vertex B and γ at the vertex C are defined similarly.

As usual in elementary geometry, the side lengths of a triangle will be denoted by the same letter as the sides themselves:

$$c = d(A, B),\ a = d(B, C),\ b = d(C, A).$$

[3] Compare, e.g., W. Rinow [94], p. 400. The proposition for the plane stated there can be transferred to the sphere S^2 using stereographic projection or one-point compactification.

Recall Example 1; to compute the poles of the oriented sides on $S^2(r)$ we have to identify the points with their position vectors:

$$C' = \frac{A \times B}{|A \times B|} r, \; A' = \frac{B \times C}{|B \times C|} r, \; B' = \frac{C \times A}{|C \times A|} r. \tag{26}$$

According to Proposition I.6.3.2 and Formula I.(6.3.35) the norms of the vector products satisfy

$$|A \times B| = r^2 \sin(c/r), \; |B \times C| = r^2 \sin(a/r), \; |C \times A| = r^2 \sin(b/r). \tag{27}$$

By Definition (25), the angles at the vertices are equal to the distances of the normalized poles. In order to apply this definition we have to assume that the angles of the triangle are less than or equal to π; for Euler triangles this was shown in Example 2. From (17) and (14) we obtain

$$\cos\gamma = -\langle A', B' \rangle / r^2, \; \cos\alpha = -\langle B', C' \rangle / r^2, \; \cos\beta = -\langle C', A' \rangle / r^2. \tag{28}$$

The opposite sign in these equations refers to the fact that in order to define an angle the great circles have to be oriented from the vertex towards both the other points. Now insert equations (26), (27) into (28) and apply the following well-known formula (cf. I.(6.3.30)) from vector algebra:

$$\langle X_1 \times X_2, Y_1 \times Y_2 \rangle = \langle X_1, Y_1 \rangle \langle X_2, Y_2 \rangle - \langle X_1, Y_2 \rangle \langle X_2, Y_1 \rangle.$$

This leads to the *law of cosines for the sides* in spherical and elliptic geometry: *The side lengths a, b, c and the angle γ of an Euler triangle are related by*

$$\cos(c/r) = \cos(a/r)\cos(b/r) + \sin(a/r)\sin(b/r)\cos\gamma; \tag{29}$$

corresponding relations hold for the other angles and sides. From Formula (29) the metric properties of the distance functions d, e are easily derived:

Proposition 4. *The distance functions defined in* (8) *and* (16) *are G_n-invariant metrics on the n-dimensional sphere $S^n(r)$ and the elliptic space $\boldsymbol{P}^n(r)$, respectively.*

Proof. By (9) and (17), respectively, d and e have all the properties required of a distance function except the triangle inequality. Since the G_n-invariance is obvious from the very definitions, it suffices to prove the following lemma for the distance function d:

Lemma 5. *The distance function defined by* (8) *and* (16), *respectively, both satisfy the* triangle inequality

$$d(A, B) \le d(A, C) + d(B, C), \qquad A, B, C \in S^n(r), \tag{30}$$

$$e(A, B) \le e(A, C) + e(B, C), \qquad A, B, C \in \boldsymbol{P}^n(r). \tag{31}$$

In (30) *equality holds if and only if A, B, C lie on a common great circle, and C is a point between*[4] *A and B on an arc of length $d(A,B) \leq r\pi$; similarly, equality holds in* (31) *if and only if A, B, C lie on a common line, and C lies between A and B on a line segment of length $d(A,B) \leq r\pi/2$.*

Proof. Three points in general position on the sphere $S^n(r)$ determine a two-dimensional subsphere $S^2(r)$, and in this they define an Euler triangle whose side lengths satisfy

$$c = d(A,B),\ a = d(B,C),\ b = d(A,C).$$

Because of $a, b, c < r\pi$ and $0 < \gamma < \pi$, the law of cosines for the sides implies

$$\cos(c/r) = \cos(a/r)\cos(b/r) - \sin(a/r)\sin(b/r) + \sin(a/r)\sin(b/r)(1 + \cos\gamma),$$
$$\cos(c/r) > \cos((a+b)/r)).$$

Since $\cos x$ is monotonously decreasing for $0 \leq x \leq \pi$, we have $c < a + b$, and this is Equation (30) with $<$ instead of $\leq$. Because all distances are less than or equal to $r\pi/2$ in the elliptic plane, we find a point triple covering A, B, C in the covering sphere with the same distances. Thus, (31) immediately follows from (30). If equality holds in these inequalities, then the points are not in general position, i.e., they lie on a great circle or a line, respectively. Discussing the obvious cases separately then proves the lemma. $\square$

Corollary 6. *The side lengths a, b, c of an Euler triangle on the sphere $S^2(r)$ or in the elliptic plane $\boldsymbol{P}^2(r)$ always satisfy*

$$0 < a + b + c < 2\pi r.$$

Proof. Consider the complement of the Euler triangle (A, B, C) inside the biangle of the sphere determined by the sides CA, CB. It is itself an Euler triangle, namely $(A, B, -C)$. For the lengths of its sides we have

$$a' = \pi r - a,\ b' = \pi r - b,\ c' = c.$$

The inequality obtained from Lemma 5, $c < a' + b' = 2\pi r - (a + b)$, is nothing but the assertion. In the case of the elliptic plane, under the covering (4) the complementary Euler triangle in the biangle, $A, B, -C$, corresponds to the triangle A, B, C bounded by the complementary line segments of the sides a, b in the original triangle together with its side c. The side lengths of this triangle are again those in the last formula above. Hence Lemma 5 can be applied. $\square$

[4] We do not introduce the notion of betweenness formally. On curves equipped with with a real parameter t it is always used in the sense of the order transferred from $\mathbf{R}$ onto the curve via the parametrization.

Proposition 7. *The angles of an Euler triangle on the sphere or of an arbitrary triangle in the elliptic plane satisfy the inequalities*

$$\pi < \alpha + \beta + \gamma < 3\pi, \tag{32}$$

$$\alpha + \beta < \gamma + \pi, \quad \alpha + \gamma < \beta + \pi, \quad \beta + \gamma < \alpha + \pi. \tag{33}$$

Proof. Choose the orientations on the sides of the triangle so that, moving along the boundary in this direction, the triangular region lies to the left. The angles $\alpha' = \pi - \alpha, \beta' = \pi - \beta, \gamma' = \pi - \gamma$ are called the *exterior angles* of the triangle A, B, C. Since only the angles are relevant, we may as well suppose that the triangle lies on the unit sphere S^2. The poles $a^\perp, b^\perp, c^\perp$ of the oriented sides are the vertices of the *polar triangle*, whose oriented sides are the polars $A^\perp, B^\perp, C^\perp$ of the vertices in the original triangle. By Exercise 9, the polar triangle is also an Euler triangle. According to (20) the angles between the oriented sides, i.e. the exterior angles α', β', γ', coincide with the side lengths of the polar triangle. Hence Corollary 6 implies

$$0 < \alpha' + \beta' + \gamma' = 3\pi - (\alpha + \beta + \gamma) < 2\pi,$$

and this immediately leads to (33). Now the triangle inequality applied to the polar triangle yields $\gamma' < \alpha' + \beta'$, i.e. the first of the inequalities (32). The remaining ones are proved analogously. Since, by Exercise 8, the angles between great circles are equal to the angles between the lines in the elliptic plane corresponding to them under the covering (4), all the inequalities also hold in this case. □

Exercise 10. Prove that, for $r \to \infty$, the above law of cosines for the sides tends to the law of cosines from Euclidean geometry,

$$c^2 = a^2 + b^2 - 2ab\cos\gamma.$$

2.5.4 Excess, Curvature, and Surface Area

Equality (32) leads to an essential difference between elliptic and Euclidean geometry: The so-called *spherical excess* of the triangle Δ,

$$\epsilon(\Delta) := \alpha + \beta + \gamma - \pi, \tag{34}$$

is always positive, whereas, as is well-known, in Euclidean geometry the *angle sum is equal to* π for every triangle. The latter is an immediate consequence of the parallel postulate together with some simple propositions concerning the angles formed by a pair of parallel lines and a line intersecting them, cf. Figure 2.6.

Parallel Postulate *If $\boldsymbol{H} \subset \boldsymbol{E}^2$ is a line in the Euclidean plane, and $\boldsymbol{x} \in \boldsymbol{E}^2 \setminus \boldsymbol{H}$ is a point not lying on it, then there is a unique line $\boldsymbol{H}_1 \subset \boldsymbol{E}^2$ not intersecting $\boldsymbol{H}$.*

This line is known as the *parallel to $\boldsymbol{H}$ through $\boldsymbol{x}$* (cf. Corollary I.4.3.1). Obviously, in elliptic geometry as well as in spherical geometry, where great circles are taken to be lines, there can be no parallels: Two different lines in one plane always intersect. We will encounter a different negation of the parallel postulate in the next section dealing with hyperbolic geometry.

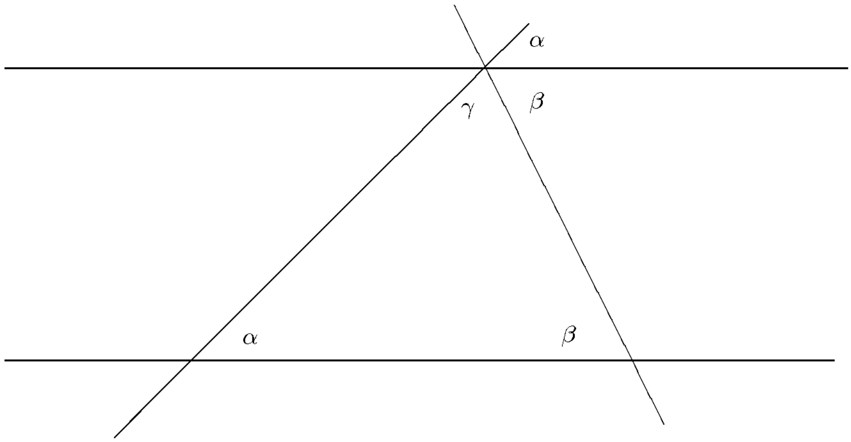


Fig. 2.6. Angles of a Euclidean triangle.

The excess can be used to define the elementary *surface area* in the two-dimensional spherical and elliptic geometries: The *surface area* of an Euler triangle Δ in the sphere $S^2(r)$ or in the elliptic plane $\boldsymbol{P}^2(r)$ is defined as the positive number

$$F(\Delta) := \epsilon(\Delta)r^2. \tag{35}$$

A subset M of a sphere or an elliptic plane will be called *elementary* if it has a *triangulation* $M = \bigcup_1^k \Delta_j$, i.e. a representation as the union of finitely many Euler triangles two of which have at most boundary elements (vertices or sides) in common. The area of an elementary set is then defined to be the sum of the areas of these triangles:

$$F(M) := \sum_1^k F(\Delta_j). \tag{36}$$

In order to justify this definition we have to prove that the sum in (36) does not depend on the chosen triangulation. First we will show that this is true for triangles.

Lemma 8. *For each triangulation $\Delta = \bigcup_{j=1}^f \Delta_j$ of a triangle Δ in the sphere S^2 or in the elliptic plane*

$$\epsilon(\Delta) = \sum_{j=1}^{f} \epsilon(\Delta_j). \tag{37}$$

Proof. Let e_{int} denote the number of vertices of the triangulation belonging to the interior of Δ, let e_{bd} be the number of vertices lying on the boundary of Δ (including the vertices of Δ itself), let k_{int} denote the number of edges of the triangulation contained in the interior of Δ, and, lastly, let k_{bd} be the number of edges on the boundary of Δ. Then the numbers e of all vertices and k of all edges of the triangulation satisfy

$$e = e_{int} + e_{rd}, \; k = k_{int} + k_{rd}.$$

According to *Euler's Polyhedral Formula* (cf. Exercise 13),

$$f - k + e = 1. \tag{38}$$

Since each interior edge belongs to precisely two triangles of the triangulation and each triangle Δ_j has three edges, we obtain

$$3f = 2k_{int} + k_{rd}. \tag{39}$$

Denoting the angles of Δ by α, β, γ yields

$$\sum_{j=1}^{f} \epsilon(\Delta_j) = \alpha + \beta + \gamma + e_{int}2\pi + (e_{rd} - 3)\pi - f\pi;$$

in fact, at the vertices of Δ the angles of the adjacent triangles Δ_j of the triangulation add up to α, β and γ, respectively. On the other hand, at the $e_{rd} - 3$ remaining vertices on the boundary of Δ their sum is π, and at each interior vertex it is 2π. Summarizing we arrive at

$$\sum_{j=1}^{f} \epsilon(\Delta_j) = \alpha + \beta + \gamma - \pi(f - 2e_{int} - e_{rd} + 3).$$

Now $e_{int} = e - e_{rd}$ implies

$$f - 2e_{int} - e_{rd} + 3 = f - 2e + e_{rd} + 3,$$

and because of $e_{rd} = k_{rd}$ a straightforward calculation using (38) leads to

$$f - 2e_{int} - e_{rd} + 3 = 3f - 2k + k_{rd} + 1.$$

Taking into account $k = k_{int} + k_{rd}$ equation (39) yields

$$f - 2e_{int} - e_{rd} + 3 = 3f - 2k_{int} - k_{rd} + 1 = 1;$$

this is nothing but the assertion. □

This lemma suffices as a justification for Definition (36):

Corollary 9. *The area of an elementary set M defined by (36) does not depend on the triangulation. The area is additive in the following sense: If $M = M_1 \cup M_2$ is the union of two elementary sets having at most boundary points in common, then $F(M) = F(M_1) + F(M_2)$.*

Proof. Let $M = \bigcup_1^f \Delta_j = \bigcup_1^{\hat{f}} \hat{\Delta}_k$ be two triangulations of an elementary set M. Then we find a common refinement of these triangulations, i.e. a triangulation $M = \bigcup_{j,k} \Delta_{j,k}$ of M such that the triangles Δ_j, $\hat{\Delta}_k$ are triangulated by the triangles of the refinement:

$$\Delta_j = \bigcup_k \Delta_{j,k}, \ \hat{\Delta}_k = \bigcup_j \Delta_{j,k}.$$

Summing over the triangles of the respective triangulations we obtain by Lemma 8:

$$\sum_{j=1}^{f} \epsilon(\Delta_j) = \sum_{j=1}^{f} \sum_k \epsilon(\Delta_{j,k}) = \sum_{k=1}^{\hat{f}} \sum_j \epsilon(\Delta_{j,k}) = \sum_{k=1}^{\hat{f}} \epsilon(\hat{\Delta}_k).$$

Multiplying by r^2 proves the first assertion. Additivity immediately follows from the definition. □

Remark. The excess $\alpha + \beta + \gamma - \pi$ quantitatively characterizes, how much the geometry of the triangle differs from that of a Euclidean triangle with the same angles; for this reason it is also called the *curvature* of the triangle. It is the homogeneity of the sphere that permits to make it the basis for the definition of the surface area: The curvature function or Gauß curvature of the sphere is constant, and the curvature of a region is proportional to its area. Frequently, even in elementary textbooks, cf. e.g. M. Berger [9] or H. G. Bigalke [13], the notion of surface area is based on its differential geometric definition. A synthetic definition of the surface area, based solely on the inner geometry of the surface determined by the metric can be found in Chapter X of the book [2] by A. D. Alexandrow. This book deals with convex surfaces which may even have singularities like vertices and edges; spheres are the simplest and, because of their perfect symmetry also most interesting examples for this type of surface. Discussing non-homogeneous surfaces the excess thus has to be interpreted as a measure for the curvature of a triangle instead of its area. For instance, in the case of a convex surface the excess of a triangle (bounded by shortest lines in the surface) is always non-negative. In greater detail the general relation to differential geometry is presented in M. Berger's book [9]. Using constructions developed in measure theory it is possible to start from elementary sets and introduce measures for area as well as curvature for general sets. If the corresponding surfaces are smooth, then the results coincide with those obtained by surface and curvature integrals in differential geometry. For the sphere $S^2(r)$ and the elliptic plane $\boldsymbol{P}^2(r)$ the Gauß curvature $K = 1/r^2$ is the density of the curvature measure. It is obvious that the curvature of a sphere is greater the smaller its radius. Assume that the surface area of a triangle Δ in $S^2(r)$ occurring in Formula (35) is constant and let its radius r tend to infinity, then its excess $\epsilon(\Delta)$ tends to zero: With increasing radius the local geometry of the sphere $S^2(r)$ approaches the Euclidean geometry. Hence it is possible to draw to scale plane maps of not too large regions of the earth's surface with acceptable precision.

We want to derive a formula for the area of a convex spherical polygon. A *spherical polygon* P is understood to be a simply connected domain of the sphere $S^2(r)$ bounded by finitely many arcs of great circles, its *sides* having at most one of their endpoints in common. These endpoints are called the *vertices* of the polygon. Let n denote the number of vertices of the polygon. Introducing formal vertices on one of its boundary arcs, any hemisphere or biangle can always be considered as a triangle. Hence we may suppose $n \geq 3$. To each vertex A_i of the polygon we assign the angle α_i to be that of the biangle determined by the sides meeting at A_i containing interior points of the polygon arbitrarily close to A_i. The angle at a formal vertex is thus always equal to π. The polygon P_n is called *convex* if for every two points $A, B \in P_n$ a shortest arc AB of a great circle belongs to P_n. The following lemma contains an expression for the surface area of a convex spherical polygon:

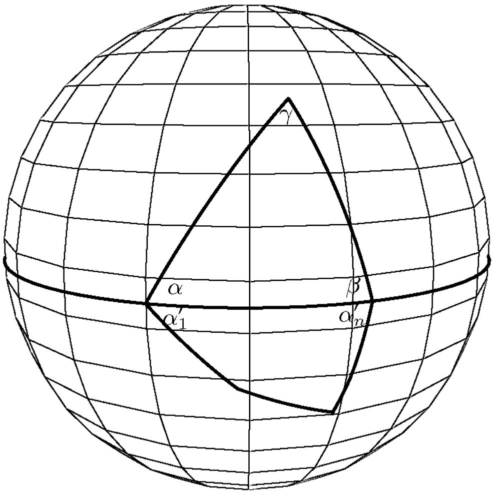


Fig. 2.7. Proof of Lemma 10.

Lemma 10. *Let P_n be a convex polygon in the sphere $S^2(r)$ with angles α_i at the vertices A_i, $i = 1, \ldots, n$. Then its area is*

$$F(P_n) = \left(\sum_{i=1}^{n} \alpha_i - (n-2)\pi \right) r^2. \tag{40}$$

Proof. Without loss of generality we may suppose $r = 1$. If the polygon is a triangle, then (40) coincides with (35). Assume that (40) holds for all integers j, $3 \leq j \leq n$. Consider a polygon P_{n+1} with $n+1$ vertices and draw the chord A_1A_n. Now we have to distinguish two cases: 1. The point A_{n+1} lies on the great circle Σ through A_1, A_n. In this case, A_{n+1} is just a formally added vertex with angle $\alpha_{n+1} = \pi$. As expected, we obtain for the area:

$$F(P_{n+1}) = \sum_{i=1}^{n+1} \alpha_i - (n-1)\pi = \sum_{i=1}^{n} \alpha_i - (n-2)\pi = F(P_n).$$

2. Assume that the point A_{n+1} does not lie on the great circle Σ. Let Σ_- denote the closed hemisphere bounded by Σ that does not contain A_{n+1}. The convexity of P_{n+1} implies that all the other vertices of P_{n+1} lie in Σ_-, and, moreover, that, as the intersection of P_{n+1} with the hemisphere Σ_-, the polygon P_n with vertices $(A_1, \ldots, A_n)$ is again convex. Let α_1', α_n' denote the angle of P_n at the vertex A_1, A_n, respectively. Assuming again $r = 1$, we conclude from (40) together with the Definitions (35) and (36) for the polygon P_n,

$$\begin{aligned} F(P_{n+1}) &= F(P_n) + F(\Delta(A_1, A_n, A_{n+1})) \\ &= \sum_{i=2}^{n-1} \alpha_i + \alpha_1' + \alpha_n' - (n-2)\pi + \alpha + \beta + \alpha_{n+1} - \pi \\ &= \sum_{i=1}^{n+1} \alpha_i - (n+1-2)\pi, \end{aligned}$$

since the angles of P_n and Δ at the common vertices A_1, A_n add up to the angle of P_{n+1}, cf. Figure 2.7, we obtain:

$$\alpha_1 = \alpha_1' + \alpha, \ \alpha_n = \alpha_n' + \beta.$$

On the other hand, the angle γ of Δ is nothing but the angle α_{n+1} of P_{n+1}. As the result we obtain Formula (40) for P_{n+1}. Since the proof works for every triangulation of P_n, and as the result does not depend on the triangulation but only on the invariants of the polygons in question, the proof of the lemma is complete. □

Remark. To compute the volume for polyhedra of higher dimension is much more involved. It seems that there does not even exist an elementary formula expressing the volume of a general n-simplex, $n \geq 3$, in a space of constant curvature. See J. Böhm, H. Hertel [16].

Exercise 11. Prove that the surface area of a biangle with angle α in $S^2(r)$ is equal to $2r^2\alpha$. The area of the whole sphere $S^2(r)$ is $4r^2\pi$.

Exercise 12. Define the notions of convex polygon, its angles, and its surface area for the elliptic plane $\boldsymbol{P}^2(r)$, and prove that Formula (40) continues to hold in this case. The area of the whole elliptic plane $\boldsymbol{P}^2(r)$ is equal to $2r^2\pi$.

Exercise 13. *Euler's Polyhedral Formula* is one of the roots of combinatorial topology. It does not only hold for triangulations in the strict sense above, but for arbitrary, even curvilinear decompositions of planar regions bounded by a simply closed curve into similar regions, where the edges are taken to be the arcs between two vertices. They are required to have at most vertices in common. Prove Euler's Polyhedral Formula (38). Hint. Taking away an edge separating two regions of the triangulation does not change the number (38). Removing a non-separating edge together with one vertex exclusively belonging to it, then the same holds. A proof can be found in H. G. Bigalke, [13], p. 310. (Note that in Proposition 10.1 of [13] the decomposition is one of the whole sphere. So the complement of Δ has to be taken into account to compute f, hence $f - k + e = 2$ for the decomposition of S^2. This formula holds for the numbers of faces, edges, and vertices of an arbitrary polyhedron homeomorphic to the sphere S^2.)

2.5.5 Spherical Trigonometry

In Section 3 we already proved the law of cosines for the sides and formulated the fundamental problem of spherical or elliptic trigonometry, to find invariants of triangles. Below we state several related results from spherical trigonometry as exercises. We confine the discussion to Euler triangles in the unit sphere S^2 and continue to use the notation familiar from elementary geometry.

Exercise 14. Prove the *law of cosines for the angles*:

$$\cos\gamma = -\cos\alpha\cos\beta + \sin\alpha\sin\beta\cos c.$$

Hint. Apply the law of cosines for the sides to the polar triangle.

Exercise 15. Prove the *law of sines* in spherical trigonometry:

$$\frac{\sin a}{\sin\alpha} = \frac{\sin b}{\sin\beta} = \frac{\sin c}{\sin\gamma}.$$

Hint. Prove first

$$\det(A, B, C) = \sin b \sin c \sin\alpha. \tag{41}$$

To this end, represent A, B, C in a suitably adapted orthonormal three-frame. Proving (41) also shows the first of the congruence theorems contained in the following exercise.

Exercise 16. Prove the following *congruence theorems* :
a) Two Euler triangles are congruent if and only if two corresponding sides and the angle between them coincide.
b) Two Euler triangles are congruent if and only if the corresponding sides are equal.
c) Two Euler triangles are congruent if and only if corresponding angles are equal.

While the first two congruence theorems hold literally also in Euclidean geometry, this is not true of assertion c): In spherical geometry similar triangles are always congruent. We recommend to look at further formulas and propositions from spherical trigonometry in a suitable collection, e.g. I. N. Bronstein, K. A. Semendjajew [23], Section 2.6.4; try and prove them using the tools at hand; see also M. Berger [9], Section 18.6.

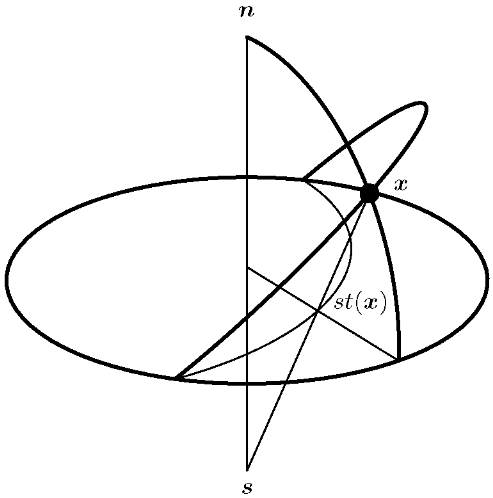


Fig. 2.8. Stereographic projection of $\boldsymbol{P}^2(\mathbf{R})$.

Example 4. The covering map (4) provided us with a picture of the real projective plane as the sphere with antipodal points identified, cf. Exercise 1 and Example 1.1.1. We obtain a representation for this plane that is well suited for detailed investigations by stereographically projecting the upper hemisphere $S_+ := S^2_+(\mathfrak{e}_2) \cup S^1(\mathfrak{e}_2)$ including the equator $S^1(\mathfrak{e}_2)$ *from the South Pole* $\boldsymbol{s} = -\mathfrak{e}_2$ *onto the equatorial plane* $\boldsymbol{E} : x^2 = 0$, cf. (6), (7). This projection is (cf. Figure 2.8)

$$st : \boldsymbol{x} \in S_+ \longmapsto \boldsymbol{y} = st(\boldsymbol{x}) := (\boldsymbol{s} \vee \boldsymbol{x}) \cap \boldsymbol{E}. \tag{42}$$

It has the following properties:

1. The image of the projective plane is the closed unit disk; antipodal points of the boundary, i.e. the equator, are to be identified.
2. The equator is the image of a projective line, in fact, of the polar for the North Pole $\boldsymbol{n} = [\mathfrak{e}_2]$. Since the projection p is injective on the open hemisphere $S^2_+(\mathfrak{e}_2)$ and all lines are projectively equivalent, the complement of a line in the real projective plane is homeomorphic to an open disk.
3. The lines through the North Pole $\boldsymbol{n}$ correspond to the great circles through $\mathfrak{e}_2$; they are mapped to diameters of the equator by st.

4. The lines not containing the North Pole $\boldsymbol{n}$ are mapped by st to arcs intersecting the equator in antipodal points.
5. If the projective plane is the elliptic plane, then the map st preserves angles: The angle between two lines in the elliptic plane is equal to the angles between their images in the Euclidean plane[5] $\boldsymbol{E}$.

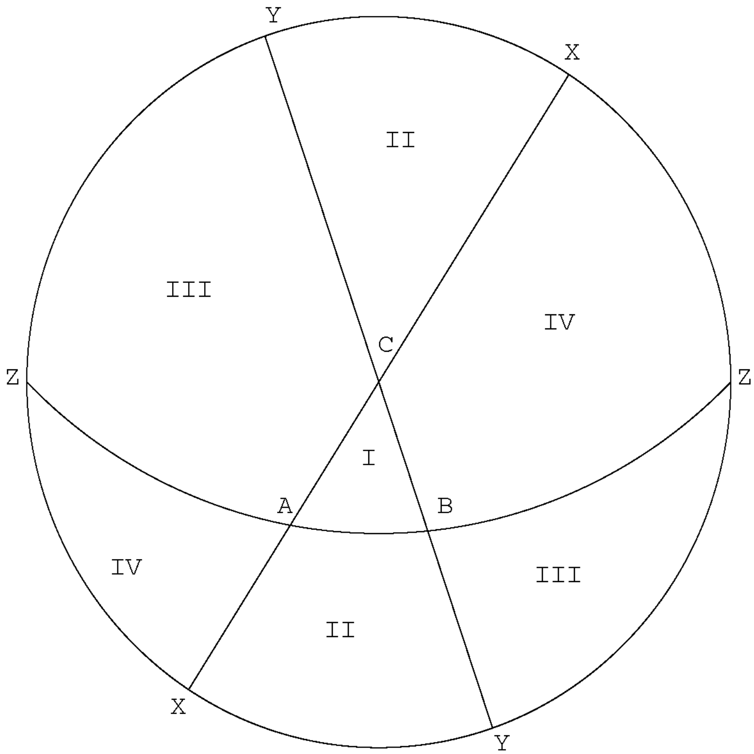


Fig. 2.9. Euler triangles in $\boldsymbol{P}^2(\mathbf{R})$.

To prove this statement it suffices to perform the calculations in Euclidean space $\boldsymbol{E}^3$, which we leave to the reader[6]. For two points in the elliptic plane (or in the elliptic space $\boldsymbol{P}^n$) there are two *shortest paths* connecting them if and only if their distance is equal to the diameter of the space, for the elliptic plane with $r = 1$, if and only if their distance is $\pi/2$, cf. Exercise 5. And these shortest paths again are the two line segments into which the points

[5] The angle between two curves at a point of intersection is defined as the angle between their tangents at this point.

[6] The n-dimensional stereographic projection will be discussed in greater detail in the next section, cf. Lemma 6.3, Exercise 5.4.

divide the line joining them. In general, a shortest path is a connecting curve of least length. The length of a curve is defined in differential geometry; the definition together with the triangle inequality immediately imply that, in elliptic geometry, shortest paths have to be line segments. The ***cut locus*** of a point $\boldsymbol{x}$ is the set of all points in a metric space for which there are at least two shortest paths joining them to $\boldsymbol{x}$. So, in the sphere S^n the cut locus of $\mathfrak{x}$ is the opposite point $-\mathfrak{x}$, whereas in elliptic space the cut locus of $\boldsymbol{x}$ is the polar $F(\boldsymbol{x})$. In the stereographic picture 2.8 the polar of the North Pole is the equator with antipodal points identified. The two line segments leading to a point $[\mathfrak{x}] = [-\mathfrak{x}] \in F(\boldsymbol{n})$ are represented by line segments from the center $\mathfrak{o} \in \boldsymbol{E}$ of the diameter determined by $\mathfrak{x}$ to $\mathfrak{x}$ and $-\mathfrak{x}$, respectively. Obviously, the stereographic projection is no isometry, the distances between pre-images and images are in general different.

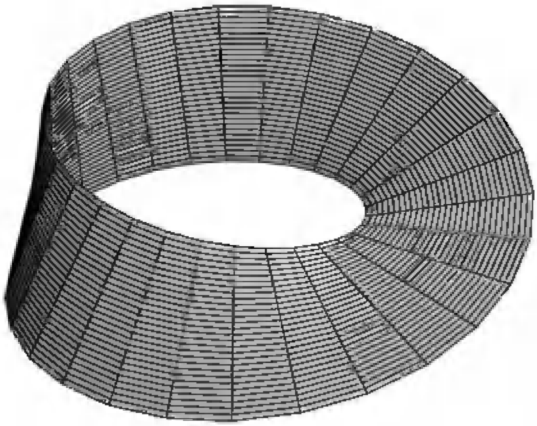


Fig. 2.10. A Möbius strip.

Now consider three points A, B, C of the elliptic plane in general position. Their connecting lines decompose the plane into four triangles I, II, III, IV, cf. Figure 2.9. Identifying, to begin with, the equator only in triangle II produces a *Möbius strip*, see Figure 2.10; this is a surface frequently called *one-sided,* i.e. non-orientable, whose boundary is a closed curve homeomorphic to a circle composed of the two arcs XY of the equator not yet identified. Performing, in addition, this identification inevitably results in self-intersections of the surface, and leads to a surface similar to the cross-cap, Figure 2.2. Applying a suitable motion we may always arrange the North Pole to represent the point C, $p(\mathfrak{e}_2) = C$; the side lines $A \vee C, B \vee C$ then turn out to be diameters in $\boldsymbol{E}$ intersecting at C under the angle γ. The points A, B divide the side line $A \vee B$ into two line segments with lengths $c = e(A, B) \leq \pi/2$ and $c' = \pi - c \geq \pi/2$, respectively; analogously for the other side lines. Since for line segments with lengths smaller than or equal to $\pi/2$ the arcs of great circles projected onto them from S_+ have the same length, — in fact, in this case the spherical distance coincides with the elliptic one — we always find an Euler triangle lying in S_+ whose vertices represent $A, B, C = \boldsymbol{n}$, and whose side lengths are just the distances of the vertices. If the side lengths a, b, c are all different, or if at most one of them is equal to $\pi/2$, then this triangle is even uniquely determined. In Figure 2.9 I is the image of such a triangle under the stereographic projection st. If, e.g., all the distances are equal to $\pi/2$, then the four triangles are all congruent, and their angles are $\pi/2$ as well; they correspond to the intersections of S_+ with the four upper octants of the chosen orthonormal coordinate systems for $\boldsymbol{E}^3$. Their stereographic projection leads to the simple Figure 2.11. This decomposition implies that the area of the elliptic plane is equal to 2π, hence for $\boldsymbol{P}^2(r)$ the area is $2r^2\pi$. □

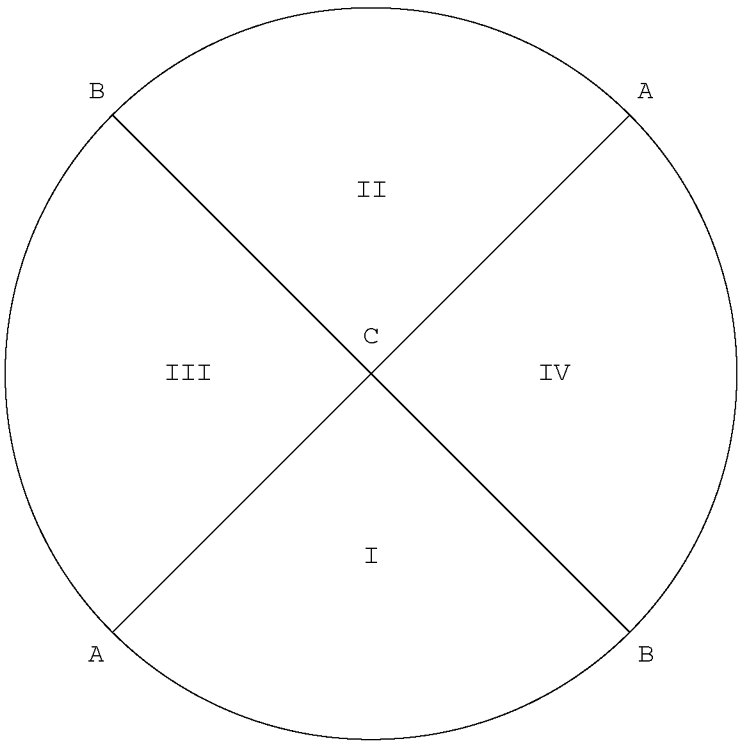


Fig. 2.11. Decomposition of $\boldsymbol{P}^2(\mathbf{R})$ into four congruent triangles.

2.5.6 The Metric Geometry of Elliptic Space

In this section we formulate some results belonging to the metric geometry of the n-sphere and the n-dimensional elliptic space in the form of exercises. Obviously, the set

$$Iso(M) := \{f : M \to M | f \text{ isometry}\}$$

of isometries of a metric space $[M, \rho]$ forms a group with the composition $\circ$ as operation, called the *isometry group* of the metric space.

Exercise 17. Prove that the isometry group of the n-sphere $S^n(r)$ (the elliptic space $\boldsymbol{P}^n(r)$) is isomorphic to the orthogonal group $\boldsymbol{O}(n+1)$ (the projective orthogonal group $\boldsymbol{PO}(n+1)$). (Hint. In order to show that each isometry f belongs to the corresponding group first prove that it preserves angles. Then consider a frame of $n+1$ mutually orthogonal vectors; the associated orthogonal n-simplex is transformed by f into another orthogonal n-simplex. On the other hand, there is an orthogonal transformation $g \in \boldsymbol{O}(n+1)$ coinciding with f on the vertices of the simplex. Since the coordinates of an arbitrary point are determined by the distances to the vertices of the simplex, which are preserved by f as well as g, $f = g$ follows.)

Exercise 18. A metric space $[M, \rho]$ is called *two-point homogeneous* if it has the following property: For two point pairs $(a, b), (a', b') \in M \times M$ with equal distances $\rho(a, b) = \rho(a', b')$ there is always an isometry $g \in Iso(M)$ such that $g(a) = a'$ and $g(b) = b'$. Obviously, every two-point homogeneous space is homogeneous. Prove that the Euclidean spaces $\boldsymbol{E}^n$, the elliptic spaces $\boldsymbol{P}^n(r)$, and the n-spheres $S^n(r)$ all are two-point homogeneous.

Let $[M, \rho]$ be a metric space, and let $B \subset M$ be a non-empty set. For each point $x \in M$ the *distance from x to B* is taken to be the infimum of all distances from x to the points of $b \in B$,

$$\rho(x, B) := \inf\{\rho(x, b) \,|\, b \in B\}.$$

Exercise 19. Let now $S_1^m \subset S^n$ be a great m-sphere in S^n, $r = 1$, $0 \leq m < n$. Prove: a) The spherical distance d from $\mathfrak{x} \in S^n$ to $S_1^m \subset S^n$ satisfies

$$d(\mathfrak{x}, S_1^m) = \arccos(|p_1(\mathfrak{x})|), \tag{43}$$

where p_1 is the orthogonal projection onto the $(m+1)$-plane $\boldsymbol{H}$ spanned by S_1^m, i.e., $0 \leq d(\mathfrak{x}, B) \leq \pi/2$. – b) For each point $\mathfrak{x} \in S^n \setminus (S_1 \cup S_1^{\perp})$ there is a unique point $\mathfrak{z} \in S_1$ with $d(\mathfrak{x}, \mathfrak{z}) = d(\mathfrak{x}, S_1)$. The great circle determined by $\mathfrak{x}, \mathfrak{z}$ is orthogonal to S_1 and intersects S_1 in the 0-sphere $\{\mathfrak{z}, -\mathfrak{z}\}$. – c) Formulate and prove statements analogous to a) and b) for the elliptic space.

The result of the last exercise, c), is

Corollary 11. *If $\boldsymbol{B}$ is an m-plane in the elliptic space, and $\boldsymbol{x} \in \boldsymbol{P}^n(r) \setminus \boldsymbol{B}^{\perp}$ is a point not belonging to the polar of $\boldsymbol{B}$, then there is a unique point $\boldsymbol{z} \in \boldsymbol{B}$ such that $e(\boldsymbol{x}, \boldsymbol{B}) = e(\boldsymbol{x}, \boldsymbol{z})$.* $\square$

Since in this case $e(\boldsymbol{x}, \boldsymbol{z}) < r\pi/2$, the line segment $\boldsymbol{xz} \subset \boldsymbol{x} \vee \boldsymbol{z}$ is also uniquely determined; it is called *the perpendicular from $\boldsymbol{x}$ onto $\boldsymbol{B}$*, and $\boldsymbol{z}$ is the *foot of the perpendicular*. Obviously, for $\boldsymbol{x} \in \boldsymbol{B}$ we have $\boldsymbol{x} = \boldsymbol{z}$.

Let again $[M, \rho]$ be a metric space, and let $B \subset M$ be a non-empty set. *The distance set $\delta(B, \epsilon)$ of radius* ϵ is defined as

$$\delta(B, \epsilon) := \{x \in M | \rho(x, B) = \epsilon\}. \tag{44}$$

Exercise 20. In the notations of Exercise 19, let the great m-sphere S_1^m lie in the $(m+1)$-plane $\boldsymbol{H}$ spanned by the vectors $\mathfrak{e}_0, \ldots, \mathfrak{e}_m$ from the orthonormal basis $(\mathfrak{e}_j), j = 0, \ldots, n$. Prove: a) For $\epsilon < \pi/2$ the distance set $\delta(S_1, \epsilon)$ is the intersection of S^n with the hypercone defined by

$$(\sum_{\mu=0}^{m} (x^\mu)^2) \cos^2 \epsilon - (\sum_{\nu=m+1}^{n} (x^\nu)^2) \sin^2 \epsilon = 0. \tag{45}$$

Deduce from this that the distance set is a generalized *torus*; in fact, as a hypersurface in S^n it can be represented by the map

$$h: \ (\mathfrak{x}, \mathfrak{y}) \in S_1^m(a) \times S_1^{\perp}(\sqrt{1-a^2}) \longmapsto \mathfrak{x} + \mathfrak{y} \in S^n \ \text{ with } a = \cos \epsilon.$$

This map is a homeomorphism. Here $S_1^m(a)$ denotes the hypersphere of radius a in $\boldsymbol{H}$, and $S_1^{\perp}(\sqrt{1-a^2})$ is the hypersphere of radius $\sqrt{1-a^2}$ in $\boldsymbol{H}^{\perp}$. For $n = 3$ and $m = 1$ we obtain the torus as the product manifold of two circles. – b) Together with $\mathfrak{x}$ also $-\mathfrak{x}$ satisfies equation (45); hence the distance set is invariant under this transformation. The covering $p: S^n \to \boldsymbol{P}^n$ maps it to the distance set of the m-plane $p(S_1^m)$ with respect to the elliptic metric, and the restriction of p to $\delta(S_1, \epsilon)$ again is a double covering.

Exercise 21. Prove that each of the distance sets $\delta(S_1^m, \epsilon)$ considered in Exercise 20 can also be interpreted as the distance set of an $(n - m - 1)$-dimensional great sphere. Which one is it, and what is the radius?

Exercise 22. Consider the distance set $D := \delta(S_1^m, \epsilon) \subset S^n$, and determine its isotropy group $G_D := \{g \in G_n | gD = D\}$. Prove that D can be generated in two different ways by a *generalized rotation* of certain subspheres Σ: $D = \bigcup_{h \in H} h\Sigma$, where H is the subgroup of all the elements $h \in SG_n$ leaving each vector in a certain subspace of the Euclidean vector space $\boldsymbol{W} \subset \boldsymbol{V}^{n+1}$ fixed: $h|\boldsymbol{W} = \mathrm{id}_{\boldsymbol{W}}$. The isotropy group G_D acts transitively on D; determine the isotropy group for this action.

Let $[M, \rho]$ be a metric space, let $B \subset M$ be a non-empty subset, and take $\epsilon > 0$. The set

$$U_\epsilon(B) := \{x \in M | \rho(x, B) < \epsilon\} \tag{46}$$

is called a *neighborhood of radius ϵ for B* .

Exercise 23. Prove that the neighborhood $U := U_\epsilon(\boldsymbol{A})$ of radius ϵ, $0 < \epsilon < \pi/2$, of a line $\boldsymbol{A}$ in the elliptic plane $\boldsymbol{P}^2, r = 1$, is a Möbius strip, cf. Figure 2.10. The complement $\boldsymbol{P}^2 \setminus U$ is homeomorphic to a closed disk. Hence, as a topological space the real projective plane is a closed surface obtained by gluing a disk to the boundary of a Möbius strip, which, in fact, is homeomorphic to a circle.

Exercise 24. Consider the distance set $D := \delta(S_1^m, \epsilon)$, $\epsilon < \pi/2$, of a great m-sphere $S_1 \subset S^n$. Take $k := \min\{m, n-m-1\}$. Prove: a) Through each point $\mathfrak{x} \in D$ there is a great k-sphere completely lying in D. – b) For all $l > k$ there is no great l-sphere lying in D. – c) For $k = 0$ the 0-sphere $\{\mathfrak{x}, -\mathfrak{x}\}$ is the only great sphere through $\mathfrak{x}$ lying in D. – d) If $n = 3$ and $m = 1$, then through each point $\mathfrak{x} \in D$ there are precisely two great circles in D. – e) In the cases different from c) and d) there are always infinitely many great k-spheres through $\mathfrak{x} \in D$ in D. (Hint. Recall Exercise 2.4 and Equation (2.10).)

In case d) of the preceding exercise let $S_2 \subset D$ be one of the great circles through $\mathfrak{x} \in D$. Since S_2 lies in the distance set D of radius ϵ for S_1, the distance $d(\mathfrak{y}, S_1)$ is the same for all $\mathfrak{y} \in S_2$, a property characterizing the lines parallel to a given one in Euclidean geometry. Hence S_2 is called a *Clifford parallel to* S_1. Note that this does not define an equivalence relation, because transitivity is violated: If S_2, S_2' are the two Clifford parallels through $\mathfrak{x}$ in D, then because of $d(\mathfrak{x}, S_2') = 0$ and $S_2 \neq S_2'$, the great circle S_2 is no Clifford parallel for S_2'. A detailed discussion of Clifford parallels and references to the literature, in particular also to generalizations, can be found in M. Berger [9], Section 18.8.

Exercise 25. Let S_2 be a Clifford parallel for the great circle S_1. Prove that S_1 is a Clifford parallel for S_2. (Note that the definition is not symmetric!)

Remark. The notions from metric geometry discussed in this section have far-reaching generalizations. A particularly successful development of metric geometry originated in the work on surface theory by A. D. Alexandrow [2] and his numerous collaborators. A systematic treatment is contained in the monography by W. Rinow [93]. Quite general results concerning the cut locus on convex surfaces can be found in the thesis by J. Kunze [71]. It is known that the cut loci of points in convex surfaces do not contain closed curves. For the elliptic plane the contrary holds: Although the local geometry is spherical, hence certainly convex, the cut locus of each point is a closed curve, in fact, as shown above, its polar.

2.5.7 Angle between Subspaces and Distance of Great Spheres

In this section we will deal with the task to find complete systems of invariants for pairs (S_1^k, S_2^m) of great spheres in the n-sphere, where great spheres of different dimension are permitted. For all k, m with $0 \leq k \leq m < n$ we consider the set of all such pairs with the action of G_n induced by its action on the n-sphere, as described in (10), (11). Since the bijective relations between great spheres and vector subspaces as well as between these and projective subspaces are G_n-isomorphisms of transformation groups, it suffices to find complete systems of invariants for the action of the orthogonal group $\boldsymbol{O}(n)$ on the corresponding products $G_{n,k} \times G_{n,m}$ of Graßmann manifolds. To solve this problem we consider the angle between subspaces, cf. H. Reichardt, [91], § V.3. The general result for unitary spaces was already formulated in Exercise I.6.4.5. First we discuss this case purely algebraically and will then derive the corresponding geometric consequences. For the sake of simplicity and in accordance with the notations used in I.6.4 we now suppose:

Assumption A. *Let $\boldsymbol{V}^n$ be an n-dimensional real or complex unitary vector space with a linear action of the orthogonal group for $K = \mathbf{R}$, and of the unitary group $G := \boldsymbol{U}(n)$ for $K = \mathbf{C}$, preserving the Euclidean or the positive definite Hermitean scalar product, respectively* (cf. Section I.6.1). *This action induces, as described more generally in* (10), (11), *an action of G over the products $G_{n,k} \times G_{n,m}$ of Graßmann manifolds:*

$$g(\boldsymbol{U}^k, \boldsymbol{W}^m) = (g\boldsymbol{U}^k, g\boldsymbol{W}^m),\ g \in G,\ 1 \le k \le m < n. \tag{47}$$

Let now $\boldsymbol{V}^n = \boldsymbol{W} \oplus \boldsymbol{W}^\perp$ be the direct orthogonal decomposition of the vector space $\boldsymbol{V}^n$ defined by the subspace $\boldsymbol{W}$, let $pr : \boldsymbol{V} \to \boldsymbol{W}$ be the orthogonal projection onto the first component, and let $q := pr|\boldsymbol{U} : \boldsymbol{U} \to \boldsymbol{W}$ be its restriction to the subspace $\boldsymbol{U}$; moreover, let $q^* : \boldsymbol{W} \to \boldsymbol{U}$ denote the adjoint operator (Definition I.6.4.1) corresponding to the dual map.

Proposition 12. *Suppose that Assumption* A *holds. For each pair of subspaces $(\boldsymbol{U}^k, \boldsymbol{W}^m)$ there is a unique self-adjoint operator $a := q^* \circ q \in \mathrm{End}(\boldsymbol{U}^k)$. Its k eigenvalues λ_α are real and, suitably labelled, they satisfy the inequalities*

$$1 \ge \lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_k \ge 0. \tag{48}$$

The sequence of these eigenvalues is a complete system of invariants for the action of G on $G_{n,k} \times G_{n,m}$.

Proof. Corollary I.6.4.1 implies $a^* = (q^* \circ q)^* = q^* \circ (q^*)^* = q^* \circ q = a$; hence a is self-adjoint. By Proposition I.6.4.2 a is diagonalizable, and all eigenvalues are real; moreover, there exists an orthonormal basis $(\mathfrak{b}_\kappa), \kappa = 1, \ldots, k$, for $\boldsymbol{U}^k$ consisting of eigenvectors of a. If $\mathfrak{b}$, $|\mathfrak{b}| = 1$, is an eigenvector, then

$$1 \ge \lambda = \langle a\mathfrak{b}, \mathfrak{b} \rangle = \langle q\mathfrak{b}, q\mathfrak{b} \rangle \ge 0.$$

The first inequality follows from the fact that, as the orthogonal projection of $\mathfrak{b}$, the vector $q\mathfrak{b}$ cannot have a larger norm than $\mathfrak{b}$. Hence, suitably labelling the eigenvalues we obtain the inequalities (48). A simultaneous transformation (47) of the subspaces by a unitary or orthogonal transformation leads to the following relations; pr_1, q_1, a_1 are the maps associated with the pair $(g\boldsymbol{U}^k, g\boldsymbol{W}^m)$:

$$pr_1 \circ g = g \circ pr,\ q_1 = g \circ q \circ g^{-1}|\boldsymbol{U}_1,\ a_1 = g|\boldsymbol{U} \circ a \circ g^{-1}|\boldsymbol{U}_1.$$

Since $g|\boldsymbol{U} : \boldsymbol{U} \to \boldsymbol{U}_1$ is unitary, the eigenvectors of a are transformed into the eigenvectors of a_1 with the same eigenvalues. Hence the sequence (48) of eigenvalues consists of G-invariants. We now show that this system of invariants is complete, i.e., pairs of subspaces with coinciding sequences (48) of eigenvalues can be transformed into one another by a map $g \in G$. To this end we adapt an orthonormal basis $(\mathfrak{a}_j)$ for $\boldsymbol{V}^n$ to the pair $(\boldsymbol{U}^k, \boldsymbol{W}^m)$ so that, with respect to this basis, the position of the subspaces is expressed by the eigenvalues. Then, for two pairs of subspaces with coinciding eigenvalues the adapted bases determine a unique unitary transformation mapping them, and hence the pairs of subspaces as well, into one another. Since the basis was chosen so that $\boldsymbol{W} = \mathfrak{L}(\mathfrak{a}_1, \ldots, \mathfrak{a}_m)$ holds, $\boldsymbol{W}^\perp$ is the linear span of the remaining basis vectors. Consider, as above, a fixed orthonormal basis for $\boldsymbol{U}^k$ consisting of eigenvectors of a, $(\mathfrak{b}_\kappa), \kappa = 1, \ldots, k,$. Decompose the vectors of this basis into their components with respect to $\boldsymbol{W}$ and $\boldsymbol{W}^\perp$:

$$\mathfrak{b}_\kappa = q\mathfrak{b}_\kappa + q^\perp \mathfrak{b}_\kappa. \tag{49}$$

The orthonormality of this basis $(\mathfrak{b}_\kappa)$ implies

$$\langle q\mathfrak{b}_\mu, q\mathfrak{b}_\nu\rangle = \langle a\mathfrak{b}_\mu, \mathfrak{b}_\nu\rangle = \lambda_\mu \delta_{\mu\nu},\ \mu, \nu = 1, \ldots, k, \tag{50}$$

$$\langle q^\perp\mathfrak{b}_\mu, q^\perp\mathfrak{b}_\nu\rangle = \langle \mathfrak{b}_\mu, \mathfrak{b}_\nu\rangle - \langle q\mathfrak{b}_\mu, q\mathfrak{b}_\nu\rangle = (1 - \lambda_\mu)\delta_{\mu\nu}. \tag{51}$$

If $\lambda_\mu = 1$, then by (50) $\langle q\mathfrak{b}_\mu, q\mathfrak{b}_\mu\rangle = \langle \mathfrak{b}_\mu, \mathfrak{b}_\mu\rangle = 1$, and hence $q\mathfrak{b}_\mu = \mathfrak{b}_\mu \in \boldsymbol{U} \cap \boldsymbol{W}$. Obviously, $\boldsymbol{U} \cap \boldsymbol{W}$ is the eigensubspace for the eigenvalue 1. We set

$$\mathfrak{a}_\alpha := \mathfrak{b}_\alpha \text{ for } \alpha = 1, \ldots, d := \dim \boldsymbol{U} \cap \boldsymbol{W}. \tag{52}$$

Let now $\mathfrak{b}_\beta, \beta = d+1, \ldots, s$, be the vectors of the eigenbasis whose components in (49) both are non-zero; hence $k - s$ is the multiplicity of the eigensubspace for the eigenvalue $\lambda = 0$, which is the kernel of a as well as that of q. Therefore, according to (50) and (51),

$$\lambda_\gamma = 0, \mathfrak{b}_\gamma = q^\perp \mathfrak{b}_\gamma \in \boldsymbol{U} \cap \boldsymbol{W}^\perp \iff \gamma = s + 1, \ldots, k. \tag{53}$$

Note that the cases $d = 0$ or $s = k$ both are possible. For $\rho = d + 1, \ldots, s$ both components in (49) are different from zero; we normalize them and set

$$\mathfrak{a}_\rho := q\mathfrak{b}_\rho/|q\mathfrak{b}_\rho| \in \boldsymbol{W},\ \mathfrak{a}_{m+\rho-d} := q^\perp\mathfrak{b}_\rho/|q^\perp\mathfrak{b}_\rho| \in \boldsymbol{W}^\perp, \rho = d + 1, \ldots, s. \tag{54}$$

By (53) we find additional basis vectors of $\boldsymbol{W}^\perp$ by

$$\mathfrak{a}_{m-d+\sigma} := \mathfrak{b}_\sigma \in \boldsymbol{W}^\perp \text{ for } \sigma = s + 1, \ldots, k. \tag{55}$$

The vectors defined in (52), (54), and (55) are orthonormal by (50) and (51). For $s < m$ or $k - d < n - m$ we complete $\mathfrak{a}_1, \ldots, \mathfrak{a}_s$ to an orthonormal basis for $\boldsymbol{W}^m$ and $\mathfrak{a}_{m+1}, \ldots, \mathfrak{a}_{m+k-d}$ to an orthonormal basis for $\boldsymbol{W}^\perp$. This concludes the adaptation process for the basis for $\boldsymbol{V}^n$. Summarizing we have:

1. The vectors $\mathfrak{a}_1, \ldots, \mathfrak{a}_m$ span the subspace $\boldsymbol{W}$.
2. The vectors $\mathfrak{b}_1, \ldots, \mathfrak{b}_k$ with the basis representations

$$\begin{aligned}
\mathfrak{b}_\alpha &= \mathfrak{a}_\alpha \text{ for } \alpha = 1, \ldots, d, \\
\mathfrak{b}_\rho &= \mathfrak{a}_\rho \cos\varphi_\rho + \mathfrak{a}_{m+\rho-d} \sin\varphi_\rho \text{ for } \rho = d+1, \ldots, s, \\
\mathfrak{b}_\sigma &= \mathfrak{a}_{m-d+\sigma} \text{ for } \sigma = s+1, \ldots, k,
\end{aligned}$$

span the subspace $\boldsymbol{U}$. Moreover, taking into account (51), the *angles* φ_κ, $0 \le \varphi_\kappa \le \pi/2$, *between* $\boldsymbol{U}$ *and* $\boldsymbol{W}$ are uniquely determined by the equations

$$\langle q\mathfrak{b}_\kappa, q\mathfrak{b}_\kappa \rangle = \lambda_\kappa = \cos^2\varphi_\kappa, \ \kappa = 1, \ldots, k. \tag{56}$$

Obviously, d, s and the angles φ_κ are uniquely determined by the sequence (48) of eigenvalues, and they again determine the form of the adapted basis. Because of the usually necessary basis completion as well as the multiplicities of the eigenvalues, this itself is not uniquely determined. For another pair of subspaces, $(\hat{\boldsymbol{U}}^k, \hat{\boldsymbol{W}}^m)$, with the same eigenvalues (48) we find an orthonormal basis $(\hat{\mathfrak{a}}_i)$ for $\boldsymbol{V}^n$ adapted to it, satisfying the properties 1., 2. for $(\hat{\boldsymbol{U}}^k, \hat{\boldsymbol{W}}^m)$, correspondingly. It is straightforward that the unitary transformation uniquely determined by $g\mathfrak{a}_i = \hat{\mathfrak{a}}_i$, $i = 1, \ldots, n$, maps the pair $(\boldsymbol{U}, \boldsymbol{W})$ into $(\hat{\boldsymbol{U}}^k, \hat{\boldsymbol{W}}^m)$. □

Example 5. Now let again (S_1^k, S_2^m), $0 \le k \le m < n$, be a pair of great subspheres of the n-sphere $S^n, r = 1$. Since the dimensions of the corresponding vector spaces exceed those of the spheres by one, we again start the labelling of indices by zero. According to (19), (43), the quadratic form

$$Q_2(\mathfrak{x}) := \cos^2(d(\mathfrak{x}, S_2^m)) = \langle q\mathfrak{x}, q\mathfrak{x} \rangle$$

is the cosine square of the distance from the point $\mathfrak{x}$ to the great m-sphere S_2. The restriction of Q_2 to the great k-sphere S_1 is just the quadratic form whose eigenvalues $\lambda_0 \ge \lambda_1 \ge \ldots \ge \lambda_k$ form the system of invariants determined in Proposition 12. From the extremal properties of the eigenvalues, see e.g. Exercise I.6.5.4, we conclude that φ_0 is the minimum of the distances from the points $\mathfrak{x} \in S_1$ to the great m-sphere S_2. It is equal to zero if and only if $\lambda_0 = 1$, i.e. if $S_1 \cap S_2 \ne \emptyset$. The angles φ_κ are the stationary values of the distance function $d(\mathfrak{x}, S_2^m)$ under the condition $\mathfrak{x} \in S_1$, and φ_k is the maximal distance of points $\mathfrak{x} \in S_1$ to S_2. Definition (16) for the distance in elliptic space implies that the considerations in this example can immediately be transferred to distances of points in a k-plane to an m-plane in elliptic space. □

Exercise 26. Consider two hyperspheres $S_1, S_2 \subset S^n$. Prove that there is at most one eigenvalue different from one, actually λ_{n-1}, and the corresponding angle φ_{n-1} coincides with that between the hyperspheres as defined in (20).

Exercise 27. Consider all position possibilities for great subspheres in the dimensions $n = 2$ and $n = 3$. Discuss the occurring angles as well as distances. Moreover, investigate how the complete systems of invariants for the pairs $(S_1, S_2), (S_1^\perp, S_2)$, $(S_1, S_2^\perp), (S_1^\perp, S_2^\perp)$ are mutually related.

Exercise 28. Consider the action of the Euclidean group $\boldsymbol{E}(n)$ on the set of pairs $(\boldsymbol{H}^k, \boldsymbol{M}^m)$ consisting of a k- and an m-plane in the Euclidean point space $\boldsymbol{E}^n$. Find a complete system invariants for this action.

2.5.8 Quadrics

In Section 1.9.6 the classification of quadrics in the real projective space $\boldsymbol{P}^n$ was reduced to the classification of real polar maps and hence of symmetric bilinear forms, cf. Example 1.9.2. In elliptic geometry two such forms are equivalent if and only if they can be transformed into one another by an orthogonal transformation of the associated Euclidean vector space; as is well-known this is decided by the Eigendecomposition Theorem, cf. e.g. Section I.6.4. The classification of quadrics in elliptic and spherical geometry essentially proceeds along the same line as in the Euclidean case. Now, however, the facts that the coordinates are homogeneous and the symmetric bilinear form is determined only up to a non-zero factor have to be taken into account. Since the covering map (4) is equivariant, we can deduce the elliptic classification from the spherical: For each elliptic quadric the spherical one that is determined by the same bilinear form is a double cover. Hence elliptic quadrics are equivalent with respect to the action of the orthogonal group $\boldsymbol{O}(n+1)$ if and only if their coverings are, and this again is the case if and only if the defining bilinear forms are transformed into one another by an orthogonal transformation and a multiplication by $\kappa \neq 0$. In this section we consider quadrics in S^n. To do so we will work in model (2) for the spherical geometry with radius $r = 1$; as described before, this way the classification of quadrics in elliptic geometry is accomplished as well. Let the quadric Q be given by the symmetric bilinear form

$$b(\mathfrak{x}, \mathfrak{y}) = \Sigma_{i,j=0}^{n} \beta_{ij} x^i x^j \text{ mit } \beta_{ij} = \beta_{ji};$$

here we suppose $b \neq 0$. The x^i, $i = 0, \ldots, n$, are always the coordinates of the vectors representing the points in an orthonormal coordinate system for the Euclidean vector space $\boldsymbol{V}^{n+1}$, thus homogeneous coordinates for the points as well. The *spherical quadric* $Q = Q_b$ corresponding to the bilinear form b is the intersection of the cone determined by $b(\mathfrak{x}, \mathfrak{x}) = 0$ with the unit hypersphere, cf. Figures 2.12, 2.13:

$$Q = \{\mathfrak{x} \in \boldsymbol{V}^{n+1} | \ b(\mathfrak{x}, \mathfrak{x}) = 0, |\mathfrak{x}| = 1\} \tag{57}$$

In elliptic geometry we may interpret this equation as the definition for the set of representatives, uniquely determined up to multiplication by ± 1,

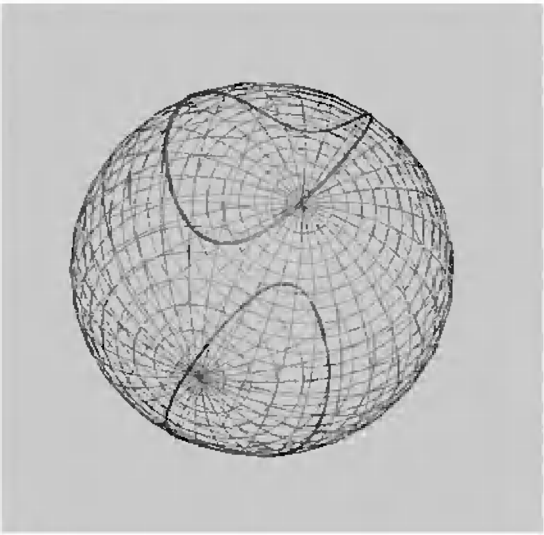


Fig. 2.12. A spherical ellipse.

for the points of the quadric; hence together with the spherical the elliptic classification of quadrics is accomplished as well. As always in this transition, we just have to identify the antipodal points of the spherical quadric to pass to the corresponding elliptic quadric. Definition (57) together with the Eigendecomposition Theorem (cf. Proposition I.6.5.1) immediately lead to

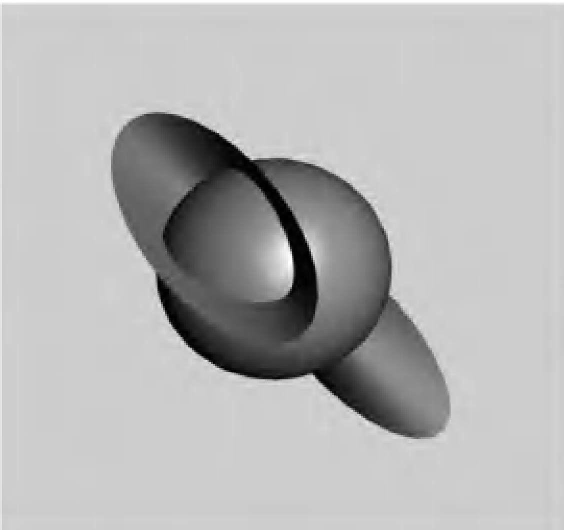


Fig. 2.13. Intersection of the sphere with an elliptic cone.

Proposition 13. *The spherical quadrics are invariant under reflections in the center of the hypersphere* S^n *containing them; hence* $Q = -Q$. *For each symmetric bilinear form* b *there is an orthonormal basis* $(\mathfrak{a}_i)$ *of the Euclidean vector space* $\boldsymbol{V}^{n+1}$ *such that its matrix satisfies*

$$b(\mathfrak{a}_i, \mathfrak{a}_j) = \beta_{ij} = \lambda_i \delta_{ij},\ i, j = 0 \ldots, n. \tag{58}$$

The numbers λ_i *are the eigenvalues of the self-adjoint operator corresponding to* b. *By* r *we denote the rank and by* l *the index of the bilinear form* b. *Multiplying, if necessary, the form by* -1, *we may suppose* $0 \leq l \leq r/2$. *Labelling the non-zero eigenvalues* λ_i *in a monotonously increasing way the bilinear form defining the quadric* Q *satisfies* (58) *as well as*

$$\begin{gathered} \lambda_0 \leq \lambda_1 \leq \ldots \leq \lambda_{l-1} < 0 < \lambda_l \leq \ldots \leq \lambda_{r-1} \neq 0, \\ \lambda_r = \ldots = \lambda_n = 0, \\ 0 < r \leq n+1,\ 0 \leq l \leq r/2. \end{gathered} \tag{59}$$

Dividing by λ_{r-1} *formulas* (58), (59) *together with* $\lambda_{r-1} = 1$ *yield a uniquely determined normal form for* b; *under this condition the numbers* $\lambda_0, \ldots, \lambda_{r-2}$ *form a complete system of invariants for the polar maps corresponding to* b, *and hence for the quadrics in spherical and elliptic geometry as well.* □

In the exercises and examples to follow we want to apply Proposition 13 to the geometry of quadrics. For this we may always suppose that the quadric is given by a bilinear form in the normal form described in Proposition 13.

Exercise 29. Prove: a) For $l = 0$ the quadric is an $(n-r)$-dimensional great sphere; for $r = n+1$ it is empty. Different positive values of $\lambda_\rho, \rho = 0, \ldots, r-1$, always lead to the same quadric; the example shows that non-equivalent symmetric bilinear forms may determine the same quadric. – b) If $r < n+1$, i.e. if the quadric Q is degenerate, then it contains the great $(n-r)$-sphere Σ determined by the kernel of the corresponding polar map. – c) The intersection $Q \cap \Sigma^\perp$ of the quadric Q with the great sphere polar to Σ is a non-degenerate quadric in $\Sigma^\perp$.

Example 6. In this example we consider a degenerate quadric $Q \subset \boldsymbol{P}^n$ of rank $r < n+1$ in elliptic space. Denote by $\boldsymbol{B}$ the vertex space of Q, i.e. the kernel of the polar map F corresponding to b:

$$\boldsymbol{B} = \ker F = [\mathfrak{a}_r, \ldots, \mathfrak{a}_n] = \boldsymbol{a}_r \vee \ldots \vee \boldsymbol{a}_n.$$

In other words, $\boldsymbol{B}$ is an $(n-r)$-dimensional projective subspace spanned by the last $n-r+1$ points in a polar simplex corresponding to a normal form for Q. Denote by $\boldsymbol{A} = \boldsymbol{B}^\perp$ its polar. By Exercise 29 the intersection $\hat{Q} := Q \cap \boldsymbol{A}$ is a non-degenerate quadric. If it is not empty, then by Exercise 1.9.7

$$Q = \bigcup_{\boldsymbol{x} \in \hat{Q}} \boldsymbol{x} \vee B;$$

in fact, $\boldsymbol{B}, \boldsymbol{A}$ are complementary subspaces in $\boldsymbol{P}^n$. This describes the structure of degenerate quadrics: *Degenerate quadrics are cones whose generators are the joins of the vertex space with the points of a non-degenerate quadric in the polar of the vertex space, or, in the case $l = 0$, the vertex space itself.* □

Exercise 30. Prove the result corresponding to the preceding example in spherical geometry and discuss, in particular, all degenerate quadrics in dimensions $n = 2, 3$.

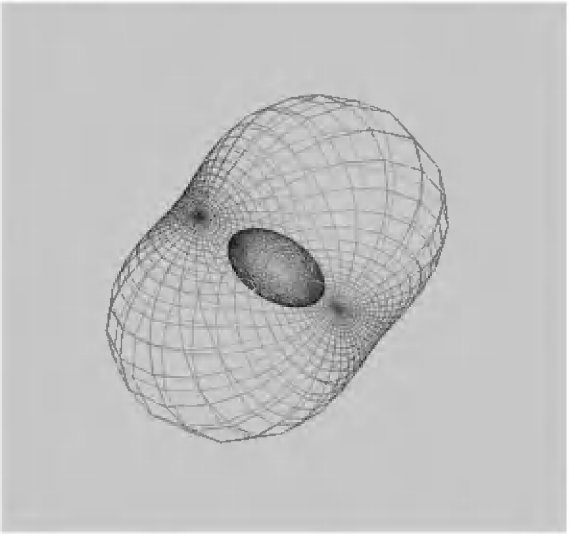


Fig. 2.14. Stereographic projection of a spherical ellipsoid.

Example 7. Now we consider spherical non-empty, non-degenerate quadrics; so suppose $l \geq 1$ and $r = n + 1$. The points different from the vertex of the cone in the Euclidean vector space $\boldsymbol{V}^{n+1}$ determined by the bilinear form b are the solutions of the equation between two positive quantities

$$-\lambda_0(x^0)^2 - \ldots - \lambda_{l-1}(x^{l-1})^2 = \lambda_l(x^l)^2 + \ldots + \lambda_{n-1}(x^{l-1})^2 + (x^n)^2. \quad (60)$$

Since the eigensubspaces are orthogonal, the vector space $\boldsymbol{V}^{n+1}$ splits into the sum of a pair of in orthogonal subspaces:

$$\boldsymbol{V}^{n+1} = \boldsymbol{W}_-^l \oplus \boldsymbol{W}_+^{n+1-l}, \ \boldsymbol{W}_- = [\mathfrak{a}_0, \ldots, \mathfrak{a}_{l-1}], \ \boldsymbol{W}_+ = [\mathfrak{a}_l, \ldots, \mathfrak{a}_n].$$

This implies: *Each non-degenerate, non-empty quadric in the sphere S^n or in the elliptic space $\boldsymbol{P}^n$ defines a pair of pole and polar assigned to the quadric Q in an invariant way, which is determined by the eigensubspaces of negative and positive eigenvalues, respectively.* Since the equation (60) is homogeneous, it suffices to solve the equation for a common value, i.e., the equation decouples into the two following equations for separate variables:

$$-\lambda_0(x^0)^2 - \ldots - \lambda_{l-1}(x^{l-1})^2 = 1,\ \lambda_l(x^l)^2 + \ldots + \lambda_{n-1}(x^{l-1})^2 + (x^n)^2 = 1.$$

Each of these equations determines a hyperellipsoid $Q_- \subset \boldsymbol{W}_-, Q_+ \subset \boldsymbol{W}_+$ in the associated subspace, and *the map*

$$\varphi : (\mathfrak{x}_0, \mathfrak{x}_1) \in Q_- \times Q_+ \longmapsto (\mathfrak{x}_0, \mathfrak{x}_1)/\sqrt{|\mathfrak{x}_0|^2 + |\mathfrak{x}_1|^2} \in Q \tag{61}$$

***is a bijective representation of the quadric** $Q \subset S^n$ **as the direct product of two hyperellipsoids with dimensions** $l-1, n-l$.* Moreover, it is a double cover of the corresponding quadric in elliptic space. The classification we just obtained refines the projective classification of quadrics described in Example 1.9.2; the new invariants occurring now are the eigenvalues, which, just as in Euclidean geometry, determine the metric properties of the quadric. Figure 2.12 shows the example of a spherical ellipse; this is the only type of a non-empty, non-degenerate quadric in S^2. Already S^3 eludes any direct Euclidean presentation, hence we use stereographic projection to depict examples for both the types of non-empty, non-degenerate quadrics in the three-sphere in Figures 2.14, 2.15. □

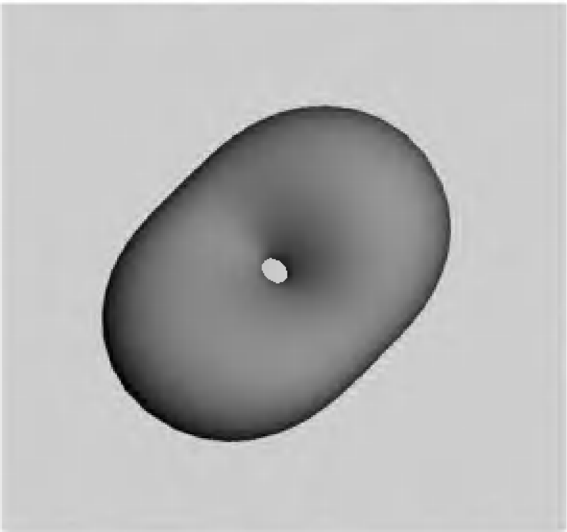


Fig. 2.15. Stereographic projection of a spherical hyperboloid.

Exercise 31. Prove the bijectivity of the map φ defined in (61). Complete the classification of non-degenerate quadrics for the dimensions $n = 2, 3$. Find geometric interpretations for the eigenvalues in all the resulting cases.

2.6 Hyperbolic Geometry

In this section we denote by $\boldsymbol{P}^n$ the real n-dimensional projective space and assume that the associated real vector space $\boldsymbol{V}^{n+1}$ is equipped with a non-degenerate, symmetric bilinear form $\langle , \rangle$ of index one. In general, a real vector space of dimension N with such a bilinear form as scalar product is called an *N-dimensional Minkowski space*; as the space-time world, four-dimensional Minkowski space forms the basis for the special theory of relativity[1] describing electrodynamics, cf. H. Minkowski [77][2]. According to 1.9.4, see also Example 2.1.3 and Table 2.1, the scalar product has the normal form

$$\langle \mathfrak{x}, \mathfrak{y} \rangle = -x^0 y^0 + \sum_{j=1}^{n} x^j y^j. \tag{1}$$

A basis $(\mathfrak{e}_i), i = 0, \ldots, n$, for which the scalar product has the normal form (1), is *pseudo-orthonormal*; i.e., with respect to it we have

$$\langle \mathfrak{e}_i, \mathfrak{e}_k \rangle = \epsilon_i \delta_{ik} \text{ with } \epsilon_0 = -1 \text{ and } \epsilon_j = 1 \text{ for } j = 1, \ldots, n. \tag{2}$$

The corresponding homogeneous or inhomogeneous point or vector coordinates are called *pseudo-orthogonal*. The isotropy group of the scalar product is the pseudo-orthogonal group $\boldsymbol{O}(1, n)$; it transforms the pseudo-orthonormal bases simply transitively among themselves. Orthogonality with respect to $\langle , \rangle$ defines the polarity $F := F_{n+1,1}$ of the projective space; its isotropy group is the projective pseudo-orthogonal group $G_n := \boldsymbol{PO}(1, n)$ generated by the pseudo-orthogonal group, cf. (1.24). The set of isotropic vectors in $\boldsymbol{V}^{n+1}$ forms the *isotropic cone;* the generators of the cone represent the quadric Q determined by F; it is a hyperellipsoid:

$$\boldsymbol{x} = [\mathfrak{x}] \in Q \Longleftrightarrow \langle \mathfrak{x}, \mathfrak{x} \rangle = 0, \qquad \mathfrak{x} \in \boldsymbol{V}, \mathfrak{x} \neq \mathfrak{o}. \tag{3}$$

Since all hyperellipsoids are projectively equivalent, we may as well use the term hypersphere $Q = S^{n-1} \subset \boldsymbol{P}^n$; in inhomogeneous coordinates the representation of Q is

$$\sum_{j=1}^{n} (x^j)^2 = 1;$$

in fact, for each solution of (3) the inequality $x^0 \neq 0$ has to hold, so that we may set $x^0 = 1$. The isotropic cone (3) splits the vector space $\boldsymbol{V}$ into two regions: one containing the *space-like vectors*,

$$\boldsymbol{V}_+ := \{\mathfrak{x} \in \boldsymbol{V} | \langle \mathfrak{x}, \mathfrak{x} \rangle > 0\}, \tag{4}$$

[1] A very clear presentation of the mathematical foundations for the special theory of relativity is contained in P. K. Raschewski [90], Chapter IV.

[2] A reprint of this paper with a commenting supplement can be found in J. Böhm, H. Reichardt [17].

and another one, consisting of two connected components, with the *time-like vectors*, cf. Figure 2.16:

$$\boldsymbol{V}_- := \{\mathfrak{x} \in \boldsymbol{V} \,|\, \langle \mathfrak{x}, \mathfrak{x} \rangle < 0\}, \tag{5}$$

Obviously, these regions are invariant under the action of $\boldsymbol{O}(1, n)$. Projectively, this division of $\boldsymbol{V} \setminus \{\mathfrak{o}\}$ corresponds to the decomposition of projective space into

$$\boldsymbol{P}^n = A(Q) \cup Q \cup I(Q). \tag{6}$$

Here $A(Q) := \pi(\boldsymbol{V}_+)$ is the outer and $I(Q) := \pi(\boldsymbol{V}_-)$ the inner region of the quadric Q, cf. Example 3.1. According to Corollary 3.1 the group G_n acts transitively on each of the sets $Q, A(Q), I(Q)$.

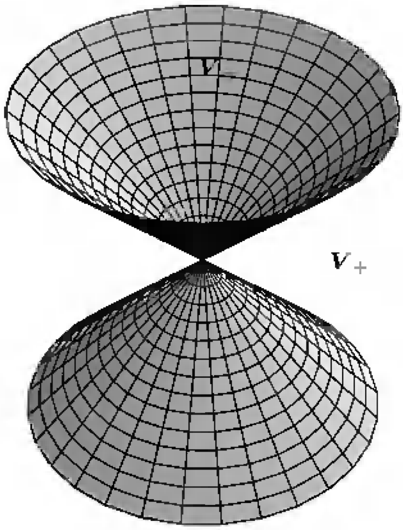


Fig. 2.16. The isotropic cone.

Definition 1. The n-dimensional *hyperbolic space* $\boldsymbol{H}^n := I(Q) = \pi(\boldsymbol{V}_-)$ is the inner region of Q; *hyperbolic geometry* studies the geometric properties of objects in $\boldsymbol{H}^n$ invariant under the action of the projective pseudo-orthogonal group $G_n = \boldsymbol{PO}(1, n)$. □

The action of G_n on the hypersphere Q defines the *Möbius geometry* on the hypersphere to be discussed below, whereas, as we will see in Example 1, the geometry determined by the action of G_n on the outer region $A(Q)$ may be interpreted as the geometry of hyperplanes in hyperbolic space.

This section deals with hyperbolic geometry, which is frequently called *non-Euclidean geometry*. Historically, hyperbolic geometry originated in the century-long unavailing effort to derive the *parallel postulate* discussed in the previous section from the other axioms of Euclidean geometry. As we will explain below, through each point in the hyperbolic plane not on a given line there are infinitely many parallels to this line. Hyperbolic geometry was invented at about the same time by C. F. Gauß, J. Bolyai, and N. I. Lobatschewski. The history of this discovery including some of the original papers is described in H. Reichardt [92]. The term "non-Euclidean" is to attribute to the strong contrast between this geometry and the Euclidean, and perhaps also to the fact that no other not Euclidean geometry had been discovered earlier. The geometry on the sphere, discussed in the previous section and known long before, is also not Euclidean; in contrast to the hyperbolic, however, it is realized as the inner geometry of a surface in Euclidean space, and can thus in a way be studied within the Euclidean framework. It is a deep result from differential geometry, proved by D. Hilbert [52] in 1901, that there can be no global, singularity-free realization of the hyperbolic plane as a surface in three-dimensional Euclidean space. The invention of hyperbolic geometry contradicted philosophical prejudices raising Euclidean geometry and all of mathematics based upon it to an a priori given form of our visual perception (I. Kant). The "foundations of geometry", which can only be indicated here, were for the first time systematically, and in a way finally, presented in the book by D. Hilbert [53]. An interesting description of these connections can also be found in the contribution by W. Klingenberg [66]. As there are a lot of non-Euclidean geometries by now, it is perhaps better to follow the suggestion by F. Klein [60] (see also [92]) and use the term hyperbolic geometry.

2.6.1 Models of Hyperbolic Space

Now we consider the hyperbolic spaces $\boldsymbol{H}^n$, $n > 0$; in case $n = 2$ it is also called the *hyperbolic plane* $\boldsymbol{H}^2$, and for $n = 1$ the *hyperbolic line*. For each point $\boldsymbol{x} \in \boldsymbol{H}^n$ there are precisely two normalized representatives

$$\pm\mathfrak{x} \in \boldsymbol{V} \text{ with } \langle \mathfrak{x}, \mathfrak{x} \rangle = -1, \tag{7}$$

belonging to the upper and lower sheet $H \subset \boldsymbol{V}^{n+1}$ of the n-hyperboloid defined by $\langle \mathfrak{x}, \mathfrak{x} \rangle = -1$, respectively; they are determined by (7) and $x^0 = -\langle \mathfrak{e}_0, \mathfrak{x} \rangle > 0$ or $x^0 = -\langle \mathfrak{e}_0, \mathfrak{x} \rangle < 0$.

Let H_+ denote the upper sheet of the hyperboloid H, determined by $x^0 > 0$, cf. Figure 2.17, and denote by $\boldsymbol{E}^n := \mathfrak{L}(\mathfrak{e}_1, \ldots, \mathfrak{e}_n)$ the linear span of the basis vectors orthogonal to $\mathfrak{e}_0$. Since for each vector $\mathfrak{x}_o \in \boldsymbol{E}^n$ there exists a unique vector $q(\mathfrak{x}_o) \in H_+$ lying above it,

$$q : \mathfrak{x}_o \in \boldsymbol{E}^n \longmapsto q(\mathfrak{x}_o) := \mathfrak{e}_0 \sqrt{1 + \langle \mathfrak{x}_o, \mathfrak{x}_o \rangle} + \mathfrak{x}_o \in H_+, \tag{8}$$

the map q is bijective; this is a parameter representation of H_+ relating it homeomorphically to $\boldsymbol{E}^n$. Since, on the other hand, each point $\boldsymbol{x}$ in the hyperbolic space $\boldsymbol{H}^n$ has precisely one normalized representative in H_+, the map

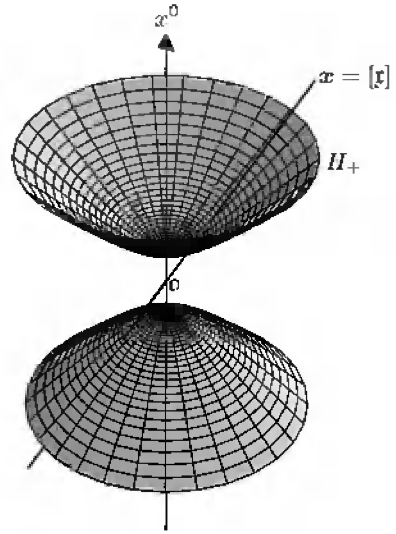


Fig. 2.17. The pseudo-Euclidean model.

$$p : \mathfrak{x} \in H_+ \longmapsto p(\mathfrak{x}) := [\mathfrak{x}] \in \boldsymbol{H}^n \tag{9}$$

is also bijective. Using it to transfer the topology defined by the parameter representation (8) onto $\boldsymbol{H}^n$ implies:

Lemma 1. *The hyperbolic space* $\boldsymbol{H}^n$ *is homeomorphic to the Euclidean space* $\boldsymbol{E}^n$. *The map* $h := p \circ q : \boldsymbol{E}^n \to \boldsymbol{H}^n$ *defined by* (8), (9) *is a bijective parametrization of* $\boldsymbol{H}^n$ *describing this homeomorphism.* □

The map p defined by (9) corresponds to the equally denoted map (5.4), which is only a local homeomorphism. Like there, one can introduce a parameter analogous to the radius r by defining the n-hyperboloid $H(r)$ and its upper sheet:

$$H(r) := \{\mathfrak{x} \in \boldsymbol{V} \,|\, \langle \mathfrak{x}, \mathfrak{x} \rangle = -r^2\},\ H_+(r) := \{\mathfrak{x} \in H(r) | - \langle \mathfrak{e}_0, \mathfrak{x} \rangle > 0\}. \tag{10}$$

Nothing essential will change this way. As the spheres $S^n(r)$ are orbits of the group $\boldsymbol{O}(n+1)$, the n-hyperboloids $H(r) \subset \boldsymbol{V}^{n+1}$ turn out to be orbits for the action of the pseudo-orthogonal group $\boldsymbol{O}(1, n)$. It is easy to see that a transformation $g \in \boldsymbol{O}(1, n)$ maps $H_+(r)$ to itself if and only if in the matrix of g with respect to an orthonormal basis

$$\gamma_{00} := -\langle \mathfrak{e}_0, g\mathfrak{e}_0 \rangle > 0. \tag{11}$$

Since one of the transformations $g, -g \in \boldsymbol{O}(1,n)$ always has this property, and as, according to (1.24), the projective pseudo-orthogonal group is the quotient group

$$G_n := \boldsymbol{O}(1,n)/\{\pm I_{n+1}\},$$

G_n is isomorphic to the isotropy group $\boldsymbol{O}(1,n)^+$ of H_+ in $\boldsymbol{O}(1,n)$; we will henceforth identify this isotropy group with G_n. In the theory of relativity the invariance of H_+ means that the time orientation is preserved, i.e. past and future cannot be interchanged; what appears to correspond to reality. Obviously, any transformation satisfying $g\mathfrak{e}_0 = -\mathfrak{e}_0$ does not leave H_+ invariant. The sign of the determinant $\det(g) \in G_n$ decides, whether g preserves the orientation of the hypersurface H_+, and hence also that of the hyperbolic space $\boldsymbol{H}^n$, or not[3].

Exercise 1. a) Prove by computation: If $a, b \in \boldsymbol{O}(1,n)$, $c = a \circ b$, and

$$-\langle \mathfrak{e}_0, a\mathfrak{e}_0\rangle > 0, -\langle \mathfrak{e}_0, b\mathfrak{e}_0\rangle > 0,$$

then $-\langle \mathfrak{e}_0, c\mathfrak{e}_0\rangle > 0$. – b) The relation $\langle \mathfrak{x}, \mathfrak{y}\rangle < 0$ is an equivalence relation on the n-hyperboloid H defined by (7); the equivalence classes are H_+ and H_-.

Exercise 2. Prove that the isotropy group of a point $x \in \boldsymbol{H}^n$ in hyperbolic space is isomorphic to the orthogonal group $\boldsymbol{O}(n)$. (Hint. Representing the transformation $g \in G_n$ by a pseudo-orthogonal matrix with respect to the standard basis $(\mathfrak{e}_i)$, $i = 0, \ldots, n$, chosen above, G_n appears as the subgroup $\boldsymbol{O}(1,n)^+$ of matrices $(a_{ij}) \in \boldsymbol{O}(1,n)$ satisfying $a_{00} > 0$ (Exercise 1); then the isotropy group of $\mathfrak{e}_0$ under the action of G_n is the subgroup of $\boldsymbol{O}(1,n)^+$ for which $a_{00} = 1$.) Because of the transitivity of G_n on $\boldsymbol{H}^n$ (Corollary 3.1) the hyperbolic space is the quotient space

$$\boldsymbol{H}^n \cong \boldsymbol{O}(1,n)^+/\boldsymbol{O}(n).$$

Using the bijectivity of the map (9) we identify H_+ with $\boldsymbol{H}^n$ and choose the vectorial or the projective representation, as appears convenient. The structures defined on H_+ by the scalar product are invariant, since the scalar product has this property. In analogy to elliptic geometry, these structures may be used to describe a hyperbolic metric and trigonometry, so that H_+ proves to be a very suitable model for hyperbolic space; we will call it the *pseudo-Euclidean model of hyperbolic space* [4].

The term "model" used here has its origin in the historic development of non-Euclidean geometry. The starting point was a purely axiomatic theory

---

[3] One can prove that the group $\boldsymbol{O}(1,n)$ has four connected components, which are characterized by the property that the transformation g preserves the time orientation and/or the space orientation or reverses them, see e.g. P. K. Raschewski [90], or (4.33) for $n = 1$.

[4] Using differential-geometric methods one can prove that the pseudo-Euclidean scalar product on $H_+(r)$ induces a Riemannian metric of constant negative curvature $-1/r^2$, cf. P. K. Raschewski [90], § 118.

in which all axioms of Euclidean geometry apart from the parallel postulate were supposed to hold; the parallel postulate was replaced by the following

Negation of the parallel postulate. *Through each point $\boldsymbol{x}$ of the hyperbolic plane $\boldsymbol{H}^2$ and every line $\boldsymbol{h}$ not containing $\boldsymbol{x}$ there are at least two lines $\boldsymbol{g}_1, \boldsymbol{g}_2 \subset \boldsymbol{H}^2$ in this plane not intersecting $\boldsymbol{h}$, $\boldsymbol{x} = \boldsymbol{g}_1 \wedge \boldsymbol{g}_2$.*

As with all axiomatic theories, the question arises, whether the theory is free of contradictions. This question can be answered by finding a structure within a contradiction-free theory satisfying the required axioms. Such a structure is called a *model* for the axiomatic theory. In this sense, Euclidean geometry as constructed in real coordinates is a model proving that the system of axioms formulated in "Euclid's Elements" is free of contradictions supposing that computing with real numbers is contradiction-free. This assumption, however, is not at all proved within mathematics, more precisely: within the framework of mathematical logic and the foundations of mathematics. Because of the fundamental role the real numbers are playing in mathematics, no mathematician, and all the more no physicist, will have the slightest doubt that the real numbers are free of contradictions. The possibility to test their properties by measurements based on physical theories, which are formulated with their help, provides empirical evidence that they indeed are contradiction-free. In this sense, we will develop the hyperbolic space $\boldsymbol{H}^n$ defined as the inner region of a hypersphere or the corresponding vector set H_+ as models for hyperbolic geometry, without discussing its axioms in any detail. This model goes back to F. Klein [60]. A detailed description of the axiomatic references can be found in the paper by W. Klingenberg [66].

Exercise 3. Prove that the pseudo-Euclidean model H_+ for the hyperbolic line is the upper branch of a hyperbola in $\boldsymbol{V}^2$ with the asymptotes defined by $\langle \mathfrak{x}, \mathfrak{x} \rangle = 0$. Defining

$$t \in \mathbf{R} \longmapsto \mathfrak{x}(t) := \mathfrak{e}_0 \cosh(t) + \mathfrak{e}_1 \sinh(t) \in H_+ \tag{12}$$

yields a bijective parameter representation of the hyperbolic line.

Definition 2. A *hyperbolic k-plane $\boldsymbol{B}^k$*, $0 \leq k \leq n$, is understood to be a projective k-plane whose intersection with the inner region $\boldsymbol{H}^n = I(Q)$ of the quadric Q is not empty. Formally, the empty set will also be considered as a hyperbolic k-plane of dimension $k = -1$. As usual we speak of points ($k = 0$), lines ($k = 1$), and hyperplanes ($k = n - 1$). If not stated otherwise, the points or subplanes of a hyperbolic k-plane will always meant to be the elements of the intersection $\boldsymbol{B}^k \cap \boldsymbol{H}^n$; the points from $\boldsymbol{B}^k \cap Q$ are called *points at infinity* of $\boldsymbol{B}^k$, and those from $\boldsymbol{B}^k \cap A(Q)$ are called their *outer points*. The set of hyperbolic k-planes will be denoted by $H_{n,k}$. □

Orthogonality arguments, the considerations in Section 1.9.8, and the transitivity statements in Example 2.1 immediately lead to:

Lemma 2. *A projective k-plane M^k is a hyperbolic k-plane if and only if the associated vector space $\boldsymbol{W}^{k+1}$ is a $(k+1)$-dimensional pseudo-Euclidean subspace of the pseudo-Euclidean vector space $\boldsymbol{V}^{n+1}$ defining the geometry, and this again is the case if and only if the polar $F(M^k)$ lies completely in the outer region $A(Q)$ of the quadric. The hyperbolic k-planes are hyperbolic spaces of dimension k. The group G_n acts transitively on the set $H_{n,k}$ of hyperbolic k-planes.* □

Example 1. The hyperbolic points $\boldsymbol{x} \in \boldsymbol{H}^n$ correspond to the one-dimensional time-like subspaces; hence their polars $\boldsymbol{x}^\perp$ are projective hyperplanes lying completely in $A(Q)$, since they correspond to Euclidean subspaces $\boldsymbol{W}^n \subset \boldsymbol{V}^{n+1}$. On the other hand, each point in the outer region $\boldsymbol{x} \in A(Q)$ corresponds to a one-dimensional space-like subspace whose orthogonal complement is pseudo-Euclidean and hence determines a hyperbolic hyperplane, and vice versa. If $\boldsymbol{x} = F(\boldsymbol{X}) \in A(Q)$ is the pole of a hyperbolic hyperplane, then for each of its representatives $\mathfrak{x}$, $\boldsymbol{x} = [\mathfrak{x}]$, we have the inequality $\langle \mathfrak{x}, \mathfrak{x} \rangle > 0$; there are always two normalized representatives $\boldsymbol{x} = [\mathfrak{x}] = [-\mathfrak{x}]$ satisfying the equation

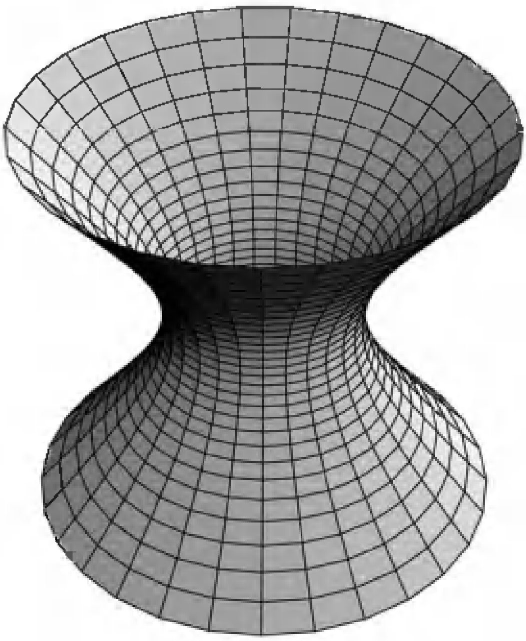


Fig. 2.18. The hyperboloid $\langle \mathfrak{x}, \mathfrak{x} \rangle = 1$.

$$\langle \mathfrak{x}, \mathfrak{x} \rangle = -(x^0)^2 + \sum_{i=1}^{n} (x^i)^2 = 1 \tag{13}$$

describing a one-sheeted n-hyperboloid Q_1 in $\boldsymbol{V}^{n+1}$, cf. Figure 2.18. The space $A(Q)$ of all hyperbolic hyperplanes can be obtained from the hypersurface (13) by identifying antipodal vectors $\mathfrak{x}, -\mathfrak{x}$. This identification yields a correspondence between a vector $\mathfrak{x} \in Q_1$ and an *oriented hyperbolic hyperplane* as follows: A point $\boldsymbol{y} \in \boldsymbol{H}^n$ lies in the *upper half-space of the hyperplane* $\boldsymbol{X} = F(\boldsymbol{x})$ if its uniquely determined representative $\mathfrak{y} \in H_+$ satisfies the inequality $\langle \mathfrak{y}, \mathfrak{x} \rangle > 0$; the *lower half-space* determined by the oriented hyperplane is analogously defined by $\langle \mathfrak{y}, \mathfrak{x} \rangle < 0$. Reversing the orientation, i.e. replacing $\mathfrak{x}$ by $-\mathfrak{x}$, upper and lower half-space are interchanged. Algebraically, these orientations can also be expressed by the orientations of the bases in the associated vector spaces. Since the hyperplane $\boldsymbol{X} = F(\boldsymbol{x})$ is the polar of $\boldsymbol{x}$, its points satisfy

$$\boldsymbol{y} = [\mathfrak{y}] \in \boldsymbol{X} \iff \langle \mathfrak{y}, \mathfrak{x} \rangle = 0, \quad (\boldsymbol{X} = F(\boldsymbol{x}), \; \boldsymbol{x} = [\mathfrak{x}]).$$

As in the Euclidean and the spherical geometries each hyperbolic hyperplane decomposes hyperbolic space into two disjoint, open regions whose common boundary it is. The space of oriented, hyperbolic hyperplanes bijectively corresponds to the n-hyperboloid (13). □

Exercise 4. a) Prove that under the action of G_n the space of oriented hyperbolic hyperplanes is isomorphic to the quotient space

$$Q_1 \cong \boldsymbol{O}(1,n)^+/\mathbf{0}(1,n-1)^+.$$

(Proceed in a similar way as in Exercise 2). – b) Represent the sets of hyperbolic k-planes $H_{n,k}$ as quotient spaces of the group G_n.

A second model for the hyperbolic space $\boldsymbol{H}^n$, likewise going back to F.Klein, is obtained by normalizing the representatives $[\mathfrak{x}]$ of its points $\boldsymbol{x}$ differently: Decompose $\mathfrak{x}$ into its time- and space-like components with respect to the standard basis chosen at the beginning,

$$\mathfrak{x} = \mathfrak{e}_0 x^0 + \mathfrak{x}_1 \text{ with } \langle \mathfrak{x}_1, \mathfrak{e}_0 \rangle = 0.$$

Then by Definition 1 and (5), because of

$$\langle \mathfrak{x}, \mathfrak{x} \rangle = -(x^0)^2 + \langle \mathfrak{x}_1, \mathfrak{x}_1 \rangle < 0$$

and $\langle \mathfrak{x}_1, \mathfrak{x}_1 \rangle \geq 0$ we always have $x^0 \neq 0$, and hence

$$h : \boldsymbol{x} = [\mathfrak{e}_0 x^0 + \mathfrak{x}_1] \in \boldsymbol{H}^n \longmapsto h(\boldsymbol{x}) := \mathfrak{x}_1 / x^0 \in D^n \tag{14}$$

is a correct definition, i.e. independent of the chosen representative; here D^n denotes the n-dimensional open ball of radius 1 in the Euclidean vector space $\boldsymbol{E}^n = \mathfrak{L}(\mathfrak{e}_1, \ldots, \mathfrak{e}_n)$,

$$D^n := \{\mathfrak{y} \in E^n | \langle \mathfrak{y}, \mathfrak{y} \rangle < 1\}. \tag{15}$$

Obviously, the map h is bijective (even a homeomorphism with respect to the usual topologies); its inverse is

$$h^{-1} : \mathfrak{y} \in D^n \longmapsto \boldsymbol{x} = [\mathfrak{e}_0 + \mathfrak{y}] \in \boldsymbol{H}^n. \tag{16}$$

This completes the construction of the second model for hyperbolic space, which we want to call *Klein's model*, sometimes the name *projective model* is also used. It has the advantage to display the subspace structure of hyperbolic geometry particularly clearly: Think of D^n as embedded into the hyperplane $x^0 = 1$, as is suggested by the representation in (16). The hyperbolic k-planes are thus simply the intersections of D^n with the $(k+1)$-dimensional vector subspaces defining them. Looked at from within the Euclidean space $\boldsymbol{E}^n$ the hyperbolic k-planes appear as the intersections of Euclidean k-planes with the open ball D^n. Figure 2.19 shows the relation between both models for the hyperbolic plane.

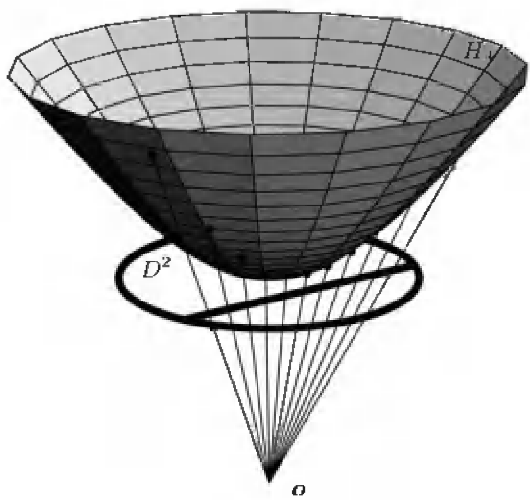


Fig. 2.19. The pseudo-Euclidean and Klein's model for $\boldsymbol{H}^2$.

Example 2. In Klein's model for the hyperbolic plane $\boldsymbol{H}^2$ the lines appear as chords of the open unit disk. The line $\boldsymbol{B}$ in $\boldsymbol{H}^2$ with the parameter representation (12) is mapped by h to the diameter

$$h([\mathfrak{e}_0 \cosh(t) + \mathfrak{e}_1 \sinh(t)]) = \mathfrak{e}_1 \tanh(t) \in \boldsymbol{E}^2, \ -\infty < t < \infty. \tag{17}$$

For t tending to $\pm\infty$ the points move towards the unit circle $\pm\mathfrak{e}_1$. If $\boldsymbol{a} = [\mathfrak{a}], \mathfrak{a} \in D^2$, is a point not lying on $\boldsymbol{B}$, then the two chords through

$\mathfrak{a}$ and the two points $\pm\mathfrak{e}_1$ of the boundary circle do not meet the line $\boldsymbol{B}$; they are called the *boundary parallels to* $\boldsymbol{B}$ *through* $\boldsymbol{a}$. They decompose the pencil of lines into two angle regions, one of them consisting of lines intersecting $\boldsymbol{B}$, the other formed by the *lines parallel* to $\boldsymbol{B}$, cf. Figure 2.20. In the pseudo-Euclidean model for $\boldsymbol{H}^2$ the line $\boldsymbol{B}$ corresponds to the upper branch of the hyperbola H_+, obtained by central projection of $\mathfrak{e}_0 + h(\boldsymbol{B})$ from the origin $\mathfrak{o} \in \boldsymbol{V}^3$ onto H_+. Obviously, in a hyperbolic space $\boldsymbol{H}^n$ with $n > 2$ there are, moreover, *skew lines*, i.e. two lines not lying in a plane. In dimensions $n > 2$, the notions parallel and boundary parallel only apply to lines that are not skew. Hence, like before through each point $\boldsymbol{a} \in \boldsymbol{H}^n \setminus \boldsymbol{B}$ there are precisely two boundary parallels to the line $\boldsymbol{B}$ and infinitely many parallels; in addition, however, there are infinitely many lines skew to $\boldsymbol{B}$. □

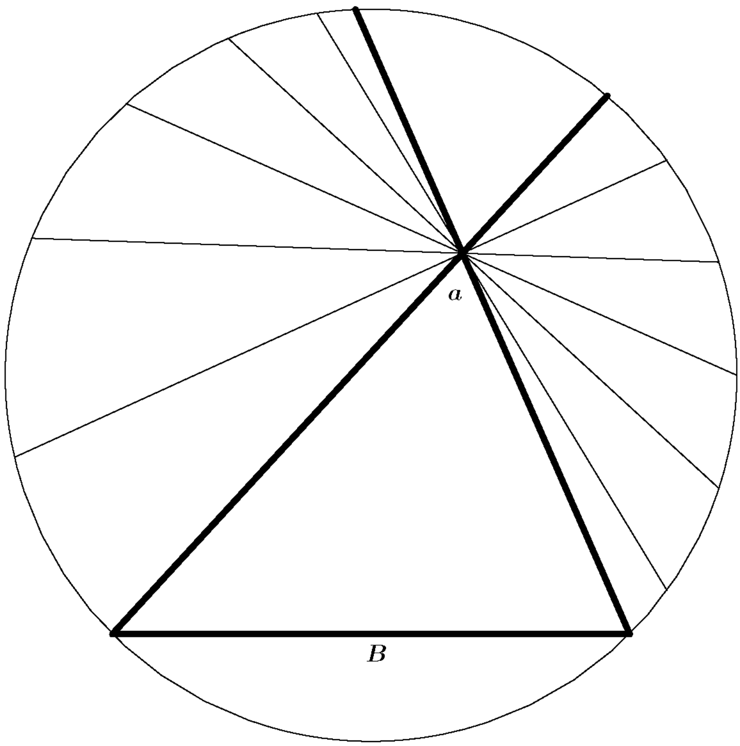


Fig. 2.20. Parallels in Klein's model for $\boldsymbol{H}^2$.

In the sequel we will repeatedly need the general *stereographic projection*, which, from the systematic point of view, belongs to Euclidean geometry; for $n = 2$ we already used it in Example 5.4. To define it consider the unit hypersphere $S^n \subset \boldsymbol{E}^{n+1}$ in the Euclidean $(n + 1)$-dimensional vector space with the origin $\mathfrak{o}$ as its center; the Euclidean scalar product will be denoted by $(\mathfrak{x}, \mathfrak{y})$. Choose an orthonormal standard basis $(\mathfrak{e}_i)$, i.e. one for which

$$(\mathfrak{e}_i, \mathfrak{e}_j) = \delta_{ij}, \; i, j = 0, \ldots, n.$$

The point $\boldsymbol{s} := -\mathfrak{e}_0$ will be called the *South Pole*, and the hyperplane $\boldsymbol{E}^n := \mathfrak{L}(\mathfrak{e}_1, \ldots, \mathfrak{e}_n)$ defined by $x^0 = 0$ the *equatorial hyperplane.* The *stereographic projection* $st : S^n \setminus \{\boldsymbol{s}\} \to \boldsymbol{E}^n$ is defined as the map (cf. (5.42))

$$st : \boldsymbol{x} \in S^n \setminus \{\boldsymbol{s}\} \longmapsto \boldsymbol{y} = st(\boldsymbol{x}) := (\boldsymbol{s} \vee \boldsymbol{x}) \cap \boldsymbol{E}^n \in \boldsymbol{E}^n, \; (n \geq 1). \tag{18}$$

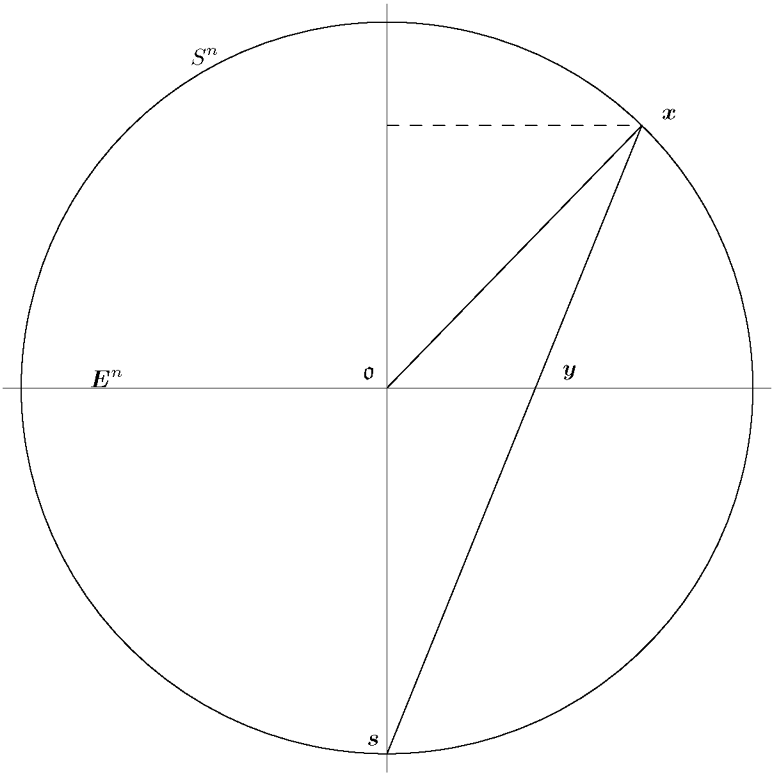


Fig. 2.21. Stereographic projection.

Hence the image $\boldsymbol{y}$ of $\boldsymbol{x}$ is the intersection of the line connecting $\boldsymbol{x}$ to the south pole with the equatorial hyperplane, schematically represented in Figure 2.21. It is obvious that the image point $\boldsymbol{y}$ moves towards infinity if $\boldsymbol{x}$ tends to the south pole $\boldsymbol{s}$. Viewing the Euclidean space as compactified by a point denoted ∞, whose neighborhoods are defined as the complements of closed sets in $\boldsymbol{E}^n$, and setting $st(\boldsymbol{s}) = \infty$ the stereographic projection is extended to a homeomorphism from S^n onto this *one-point compactification* of $\boldsymbol{E}^n$. We have

Lemma 3. *The stereographic projection* (18) *is a homeomorphism. It maps k-subspheres $A^k \subset S^n \setminus \{\boldsymbol{s}\}$ to k-spheres of $\boldsymbol{E}^n$; k-subspheres containing the south pole $\boldsymbol{s}$ are mapped to k-planes in $\boldsymbol{E}^n$. Moreover, the map is conformal.*

Recall that a map is called *conformal* if it preserves angles. A classical result from differential geometry proved by C. F. Gauß, the "Theorema Egregium", implies that there is no isometric map, i.e. one preserving distances, between spheres and planes; hence conformity is the best one can have. Therefore, for $n = 2$ stereographic projection is applied in cartography and complex function theory; in the latter it establishes the relation between the Riemann sphere and the extended Gauß number plane. To prove Lemma 3 we recommend that the reader proceeds as explained in the following exercise.

Exercise 5. Let $x_i, i = 0, \dots, n$, be the coordinates of $\boldsymbol{x}$, and let $y_j, j = 1, \dots, n$, be the coordinates of $\boldsymbol{y} = st(\boldsymbol{x})$ with respect to the standard basis and the origin $\mathfrak{o}$. a) Expressing (18) in coordinates yields

$$st: \quad y_j = \frac{x_j}{1 + x_0}, \; j = 1, \dots, n. \tag{19}$$

b) To obtain formulas for the inverse $\boldsymbol{x} = st^{-1}(\boldsymbol{y})$ compute

$$st^{-1}: \quad x_0 = \frac{1 - (\mathfrak{y}, \mathfrak{y})}{1 + (\mathfrak{y}, \mathfrak{y})}, \; x_i = \frac{2y_i}{1 + (\mathfrak{y}, \mathfrak{y})}, \; i = 1, \dots, n. \tag{20}$$

Here $\mathfrak{y}$ denotes the position vector of $\boldsymbol{y}$. – c) Prove the assertion in Lemma 3 concerning the images of k-subspheres. (Hint. First consider the case $k = n - 1$ and use the fact that each hypersphere can be represented as the intersection of S^n with a hyperplane. Let the hyperplane be described by its Hessian normal form $(\mathfrak{n}, \mathfrak{x}) = p$, cf. I.(6.3.37), where $\mathfrak{n} \in S^n$ is a unit normal vector and $p, 0 \le p < 1$, its distance from the origin. Note, moreover, that each k-subsphere can be represented as the intersection of finitely many hyperspheres.) – d) Prove that st is a conformal map. (Hint. First consider the case n=2. At the arbitrary point $\boldsymbol{x} \in S^2$ let $\mathfrak{s}_1, \mathfrak{s}_2$ be the unit tangent vectors to the meridian and the circle of latitude through this point, respectively. Compute the images of these vectors using the parametrizations of these curves by arc length s and differentiation with respect to s. Show that they are orthogonal and both are multiplied by the same factor, which only depends on the point $\boldsymbol{x}$. For $n = 2$ this is nothing but conformity. For $n > 2$ take an arbitrary orthonormal basis $\mathfrak{s}_j, j = 1, \dots, n$, of the tangent space $T_{\boldsymbol{x}} S^n$. Since $\boldsymbol{x}$ and two of the basis vectors $\mathfrak{s}_i, \mathfrak{s}_j, i \ne j$, determine a great 2-sphere $S^2 \subset S^n$, the conformity in this general case is a direct consequence of that for $n = 2$.)

Example 3. We want to describe three more models for hyperbolic space. To do so we start from Klein's model. Consider the open unit ball D^n as embedded into the hyperplane $\boldsymbol{E}^n = \mathfrak{L}(\mathfrak{e}_1, \dots, \mathfrak{e}_n) \subset \boldsymbol{E}^{n+1}$ with normal vector $\mathfrak{e}_0$ and take only the upper part S^n_+, $x_0 > 0$, of the unit hypersphere, in whose equatorial hyperplane D^n lies. Then for each point $\boldsymbol{x} \in D^n$ there is a unique point $\boldsymbol{x}_+ \in S^n_+$ lying above it, which is orthogonally projected to $\boldsymbol{x}$. The hyperbolic lines correspond to semicircles orthogonally intersecting the boundary S^{n-1} of D^n. Each hyperbolic k-plane is generated by a Euclidean k-plane $\boldsymbol{M}^k \subset \boldsymbol{E}^n$ intersecting the boundary S^{n-1} in a $k-1$-sphere; then the only non-zero angle formed by its lift to S^n_+ with $\boldsymbol{E}^n$ is again $\pi/2$. The upper half S^n_+ of the hypersphere with the structure in which k-dimensional hemispheres are the hyperbolic k-planes is again a model for the hyperbolic geometry; we will call it the *spherical model.* Applying stereographic projection to the upper half of the hypersphere S^n_+ we obtain a further model for hyperbolic space within the open ball D^n, in which the hyperbolic lines are represented by semicircles or diameters orthogonally intersecting the boundary S^{n-1}. Analogously for the hyperbolic k-planes: they are mapped to k-dimensional hemispheres orthogonal to S^{n-1}. We will call this model, also first described by F. Klein, the *conformal disk model for hyperbolic space.* A fifth model goes back to H. Poincaré: First we think of D^n as lying in the hyperplane of the meridian determined by $x_1 = 0$, then we consider the hemisphere defined by $x_1 > 0$ with the hyperbolic structure just described, and, finally, we apply the stereographic projection *st*. Altogether we obtain the *Poincaré model*: the open half-space of $\boldsymbol{E}^n$ determined by $x_1 > 0$. Because of conformity and circle invariance, the hyperbolic lines appear as semicircles or half-lines orthogonal to the boundary, i.e. the hyperplane $x_1 = 0$ in $\boldsymbol{E}^n$, and the hyperbolic k-planes become the k-dimensional hemispheres or k-dimensional half-spaces orthogonal to it in the described sense. Figure 2.22 is a transformation of Figure 2.20 to the Poincaré model for the hyperbolic plane.

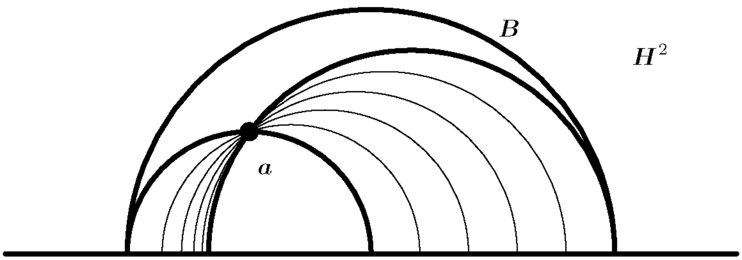


Fig. 2.22. Parallels in the Poincaré model for $\boldsymbol{H}^2$.

Figure 2.23 shows a family of circle tangents in the conformal model for the hyperbolic plane. It is remarkable that each of the tangents belongs to a quadrangle of tangents in the family whose vertices belong to Q, i.e. they are points at infinity. In each of these quadrangles any two adjacent sides are boundary parallel, whereas opposite sides are parallel. To every line the family contains several parallels. This picture clearly shows the considerable differences to the Euclidean as well as the elliptic geometry particularly clearly: In Euclidean geometry, for each tangent to a circle there is a unique second parallel tangent, whereas in elliptic geometry there are no parallels at all. □

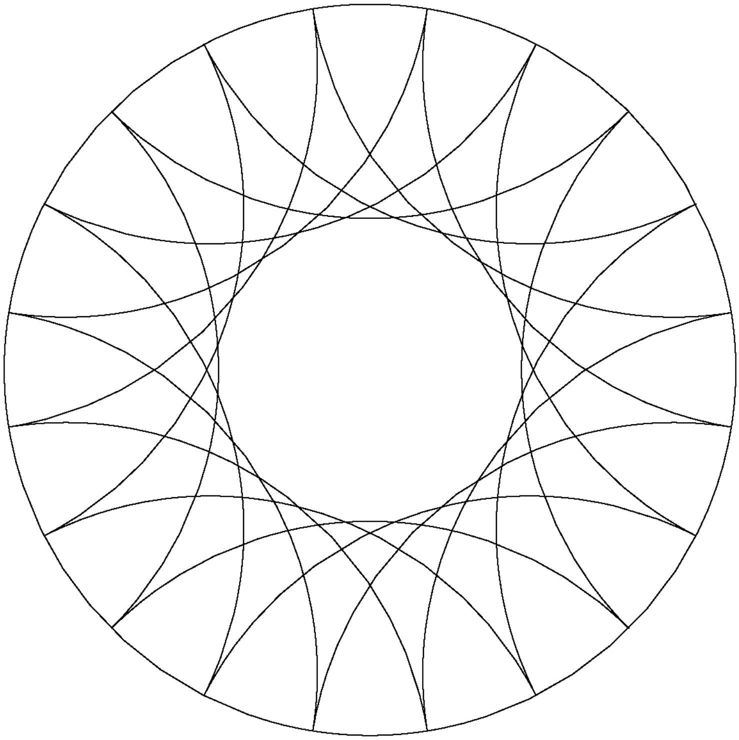


Fig. 2.23. Circle tangents in the conformal disk model for $\boldsymbol{H}^2$.

2.6.2 Distance and Angle

Now we return to the original definition of hyperbolic space as the inner region of the quadric Q; we describe the point $\boldsymbol{x} \in \boldsymbol{H}^n$ by its unique normalized representative $\boldsymbol{x} = [\mathfrak{x}], \mathfrak{x} \in H^n_+, \langle \mathfrak{x}, \mathfrak{x} \rangle = -1$. The fundamental invariant for the hyperbolic geometry of radius r is the *distance of two points* which, in analogy to the elliptic distance (5.16), is defined by

$$h(\boldsymbol{x}, \boldsymbol{y}) := r \operatorname{arcosh}(|\langle \mathfrak{x}, \mathfrak{y} \rangle|), \ \boldsymbol{x} = [\mathfrak{x}], \boldsymbol{y} = [\mathfrak{y}], \ \mathfrak{x}, \mathfrak{y} \in H^n_+, \tag{21}$$

see also Exercise 6. It is easy to prove that h has the following properties of a metric:

$$h(\boldsymbol{x},\boldsymbol{y}) = h(\boldsymbol{y},\boldsymbol{x}), h(\boldsymbol{x},\boldsymbol{y}) \geq 0, h(\boldsymbol{x},\boldsymbol{y}) = 0 \Longleftrightarrow \boldsymbol{x} = \boldsymbol{y}. \tag{22}$$

Below the triangle inequality for h will be derived as a consequence of the hyperbolic law of cosines. For simplicity, we will from now on suppose that the radius is equal to one, $r = 1$.

Exercise 6. a) Prove $1 < |\langle \mathfrak{x}, \mathfrak{y} \rangle|$ for $\mathfrak{x}, \mathfrak{y} \in H_+^n$, $\mathfrak{x} \neq \mathfrak{y}$. Hint. Apply (2.33), (2.34) to the pseudo-Euclidean subspace $[\mathfrak{x}, \mathfrak{y}] \subset \boldsymbol{V}^{n+1}$. – b) Prove (22).

Just like Proposition 5.2 in elliptic geometry this leads to

Proposition 4. *Two finite point sequences $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)$, $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k)$ in the n-dimensional hyperbolic space $\boldsymbol{H}^n$ are G_n-congruent if and only if the distances of corresponding point pairs coincide:*

$$h(\boldsymbol{x}_i, \boldsymbol{x}_j) = h(\boldsymbol{y}_i, \boldsymbol{y}_j) \text{ for } i, j = 1, \ldots, k.$$

Proof. The equality of distances implies the equality for the scalar products of the representatives. Next we have to verify the assumptions of Proposition 4.2. Since in a pseudo-Euclidean vector space of index 1 any subspace containing a time-like vector is not isotropic, Assumption c) of Proposition 4.2 holds. For the same reason we may apply Formula (2.33) for Gram's determinant, implying Assumption a) (of Proposition 4.1). Finally, because of $\kappa = 1$, Assumption d) is trivial, and this proves the assertion. □

Example 4. Let us now consider two hyperbolic hyperplanes $\boldsymbol{X}, \boldsymbol{Y}$, $\boldsymbol{X} \cap I(Q) \neq \emptyset$, $\boldsymbol{Y} \cap I(Q) \neq \emptyset$; if there is no danger of confusion, the last two conditions will not be mentioned explicitly, i.e., we will just speak of hyperbolic hyperplanes $\boldsymbol{X}, \ldots$, or, more generally, hyperbolic k-planes, whenever these conditions are satisfied. The polars of these hyperplanes $\boldsymbol{x} = \boldsymbol{X}^\perp, \boldsymbol{y} = \boldsymbol{Y}^\perp$ are always points in the outer region, whose normalized representatives $\pm\mathfrak{x}$, $\pm\mathfrak{y}$ are determined up to sign and lie on the one-sheeted hyperboloid Q_1, cf. (13), Example 1. If the hyperbolic hyperplanes $\boldsymbol{X}, \boldsymbol{Y}$ are oriented, then their normalized representatives are uniquely determined, and so we have a uniquely determined invariant:

$$I(\boldsymbol{X}, \boldsymbol{Y}) := \langle \mathfrak{x}, \mathfrak{y} \rangle, \ [\mathfrak{x}] = \boldsymbol{X}^\perp, \ [\mathfrak{y}] = \boldsymbol{Y}^\perp, \mathfrak{x}, \mathfrak{y} \in Q_1. \tag{23}$$

For non-oriented hyperplanes $\boldsymbol{X}, \boldsymbol{Y}$ we have to do without the sign and obtain the invariant

$$|I|(\boldsymbol{X}, \boldsymbol{Y}) := |\langle \mathfrak{x}, \mathfrak{y} \rangle|, \ [\mathfrak{x}] = \boldsymbol{X}^\perp, \ [\mathfrak{y}] = \boldsymbol{Y}^\perp, \mathfrak{x}, \mathfrak{y} \in Q_1. \tag{24}$$

Using representing vectors which are not normalized we have the following general formulas

$$I(\boldsymbol{X}, \boldsymbol{Y}) = \frac{\langle \mathfrak{x}, \mathfrak{y} \rangle}{\sqrt{\langle \mathfrak{x}, \mathfrak{x} \rangle \langle \mathfrak{y}, \mathfrak{y} \rangle}}, \ |I|(\boldsymbol{X}, \boldsymbol{Y}) = \frac{|\langle \mathfrak{x}, \mathfrak{y} \rangle|}{\sqrt{\langle \mathfrak{x}, \mathfrak{x} \rangle \langle \mathfrak{y}, \mathfrak{y} \rangle}}. \tag{25}$$

Now suppose $\boldsymbol{X} \neq \boldsymbol{Y}$ and distinguish three cases:

1. $|I|(\boldsymbol{X},\boldsymbol{Y}) < 1$. This holds if and only if $\mathfrak{x},\mathfrak{y}$ span a Euclidean subspace, and this again is satisfied if and only if $\boldsymbol{X}\wedge\boldsymbol{Y} = (\boldsymbol{x}\vee\boldsymbol{y})^{\perp}$ is a pseudo-Euclidean subspace, i.e., the hyperplanes intersect in a hyperbolic $(n-2)$-plane $\boldsymbol{X}\cap\boldsymbol{Y}$. The *angle α between the intersecting hyperplanes* is then uniquely determined by

$$\cos\alpha = I(\boldsymbol{X},\boldsymbol{Y}),\ 0<\alpha\le\pi, (\boldsymbol{X},\boldsymbol{Y} \text{ oriented}),$$
$$\cos\alpha = |I|(\boldsymbol{X},\boldsymbol{Y}),\ 0<\alpha\le\pi/2, (\boldsymbol{X},\boldsymbol{Y} \text{ not oriented}).$$

2. $|I|(\boldsymbol{X},\boldsymbol{Y}) = 1$. This holds if and only if $\mathfrak{x},\mathfrak{y}$ span an isotropic subspace, and this again holds if and only if $\boldsymbol{X}\wedge\boldsymbol{Y} = (\boldsymbol{x}\vee\boldsymbol{y})^{\perp}$ is an isotropic subspace, i.e., the hyperplanes intersect Q in a tangent $(n-2)$-plane $\boldsymbol{X}\cap\boldsymbol{Y}$; $\boldsymbol{X},\boldsymbol{Y}$ are said to be *boundary parallel.*
3. $|I|(\boldsymbol{X},\boldsymbol{Y}) > 1$. This holds if and only if $\mathfrak{x},\mathfrak{y}$ span a pseudo-Euclidean subspace, and this again holds if and only if $\boldsymbol{X}\wedge\boldsymbol{Y} = (\boldsymbol{x}\vee\boldsymbol{y})^{\perp}$ is a Euclidean subspace, i.e. $\boldsymbol{X}\wedge\boldsymbol{Y}$ is an $(n-2)$-plane lying completely in the outer region; $\boldsymbol{X},\boldsymbol{Y}$ are said to be *parallel hyperplanes.*

Obviously, for two pairs $(\boldsymbol{X}_\alpha,\boldsymbol{Y}_\alpha), \boldsymbol{X}_\alpha \neq \boldsymbol{Y}_\alpha, \alpha = 1,2$, of different hyperplanes to be G_n-congruent equality of the invariants $|I|$ is necessary. To prove its sufficiency one can proceed as follows. Starting from the equality of the scalar products for the normalized representatives of the poles one finds an isomorphism between the two-dimensional subspaces spanned by them. Then one applies E. Witt's Theorem or an elementary construction for adapted pseudo-orthonormal bases to show that equality of the invariants $|I|(\boldsymbol{X}_1,\boldsymbol{Y}_1) = |I|(\boldsymbol{X}_2,\boldsymbol{Y}_2)$ implies G_n-congruence of the pairs. In particular, two boundary parallel pairs are always G_n-congruent, whereas, just like in Euclidean geometry, two parallel hyperplanes have an invariant. □

Exercise 7. a) Find for each of the cases in Example 4 and every $n > 0$ pairs of hyperbolic hyperplanes satisfying the corresponding conditions. – b) Explain the elementary construction mentioned in the final paragraph of Example 4 in detail. – c) Show by adapting the frames that two pairs of boundary parallel hyperplanes (case 2 of Example 4) are always G_n-congruent.

Example 5. Let $\boldsymbol{X},\boldsymbol{Y}$ be two hyperbolic half-lines starting from their point of intersection $\boldsymbol{a}\in\boldsymbol{H}^n$. In accordance with Equation (12) let these half-lines in the model H^n_+ be described by

$$\mathfrak{b}(t) = \mathfrak{a}\cosh(t) + \mathfrak{b}_1\sinh(t),\ t>0,$$
$$\mathfrak{c}(s) = \mathfrak{a}\cosh(s) + \mathfrak{c}_1\sinh(s),\ s>0,$$

where

$$\mathfrak{a}, \mathfrak{b}(t), \mathfrak{c}(s) \in H^n_+, -\langle\mathfrak{a},\mathfrak{a}\rangle = \langle\mathfrak{b}_1,\mathfrak{b}_1\rangle = \langle\mathfrak{c}_1,\mathfrak{c}_1\rangle = 1, \langle\mathfrak{a},\mathfrak{b}_1\rangle = \langle\mathfrak{a},\mathfrak{c}_1\rangle = 0.$$

Since $\mathfrak{b}_1,\mathfrak{c}_1$ are orthogonal to the time-like unit vector $\mathfrak{a}$, they span a Euclidean subspace; moreover, because of the bijectivity of Equation (8) they are uniquely determined space-like unit vectors. As the *angle between the half-lines* with intersection point $\boldsymbol{a}$ we take the number uniquely defined by

$$\cos\alpha = \langle \mathfrak{b}_1, \mathfrak{c}_1 \rangle, 0 \leq \alpha \leq \pi.$$

We will show that this definition is compatible with that of Example 4, Case 1. To see this, consider the hyperbolic plane spanned by $\boldsymbol{a}$ and the two half-lines. The corresponding pseudo-Euclidean subspace $\boldsymbol{W}^3$ is that spanned by $\mathfrak{a}, \mathfrak{b}_1, \mathfrak{c}_1$. We take it to be oriented, so that a vector product is defined. The orthogonality relations imply that $\mathfrak{a} \times \mathfrak{b}_1, \mathfrak{a} \times \mathfrak{c}_1$ are normalized space-like unit vectors representing the poles of the lines $\boldsymbol{X}, \boldsymbol{Y}$, which are oriented by the parametrizations above. From Formula (2.37), in which we have to set

$$a = \det(\langle \mathfrak{e}_i, \mathfrak{e}_j \rangle) = -1,$$

we obtain

$$I(\boldsymbol{X}, \boldsymbol{Y}) = \langle \mathfrak{a} \times \mathfrak{b}_1, \mathfrak{a} \times \mathfrak{c}_1 \rangle = -\langle \mathfrak{a}, \mathfrak{a} \rangle \langle \mathfrak{b}_1, \mathfrak{c}_1 \rangle = \langle \mathfrak{b}_1, \mathfrak{c}_1 \rangle.$$

In fact, by construction we have $\langle \mathfrak{a}, \mathfrak{b}_1 \rangle = \langle \mathfrak{a}, \mathfrak{c}_1 \rangle = 0$. □

Example 6. Next we want to prove that the spherical, and hence Klein's and Poincaré's models for hyperbolic space as well, are conformal. Since a proof relying solely on linear algebra would, on the one hand, be too complicated, and, on the other hand, just produce particular results, this time we will apply differential-geometric methods. Consider the upper sheet H_+ of the hyperboloid representing $\boldsymbol{H}^n$ defined by (7) and $x^0 > 0$, and take the coordinates x^i, $i = 1, \ldots, n$, as parameters. Then the parameter representation of this hypersurface is

$$\mathfrak{x} : (x^i) \in D^n \longmapsto \mathfrak{x}(x^i) := (\sqrt{1 + \sum_1^n (x^i)^2}, x^1, \ldots, x^n) \in H_+,$$

where we inserted the following consequence of (7)

$$x^0 = \sqrt{1 + \sum_1^n (x^i)^2}.$$

The metric induced by means of the differential $d\mathfrak{x}$ from the pseudo-Euclidean scalar product in the tangent spaces of H_+ is positive definite because of

$$\langle \mathfrak{x}, \mathfrak{x} \rangle = -1, \ \langle \mathfrak{x}, d\mathfrak{x} \rangle = 0.$$

The Riemannian metric it defines is described by the G_n-invariant arc element

$$\langle d\mathfrak{x}, d\mathfrak{x} \rangle = \frac{1}{(x^0)^2}((x^0)^2 \sum_1^n (dx^j)^2 - (\sum_1^n x^j dx^j)^2).$$

Using

$$\frac{d\mathfrak{b}}{dt} = \mathfrak{a} \sinh t + \mathfrak{b}_1 \cosh t$$

and the corresponding equation for the derivative of $\mathfrak{c}(s)$ we obtain the following formula for the angle between the lines in Example 5 at their point of intersection, $\mathfrak{a}$, i.e. for $s = t = 0$, expressed in this arc element

$$\langle \frac{d\mathfrak{b}}{dt}(0), \frac{d\mathfrak{c}}{ds}(0) \rangle = \langle \mathfrak{b}_1, \mathfrak{c}_1 \rangle = \cos\alpha.$$

This is what we already computed before. Now consider the spherical model S^n_+ for H^n_+ described in Example 3; performing the computations according to the construction, the image of $\mathfrak{x}(x^i)$ in this parameter representation is described by

$$\mathfrak{x}_S(x^i) = \frac{1}{x^0}(1, x^1, \ldots, x^n), \ ((x^i) \in D^n).$$

The arc element of the hypersphere $S^n \subset \boldsymbol{E}^{n+1}$ is to be computed starting from the Euclidean scalar product $(,)$. Differentiating the last equation and taking the scalar product yields

$$(d\mathfrak{x}_S, d\mathfrak{x}_S) = \frac{1}{(x^0)^2}\langle d\mathfrak{x}, d\mathfrak{x}\rangle.$$

This shows the conformity of the map $\mathfrak{x} \in H^n_+ \longmapsto \mathfrak{x}_S \in S^n_+$. According to Lemma 3 the stereographic projection is also conformal, hence the composition of these conformal maps immediately leads to: *The conformal disk model and Poincaré's model for the hyperbolic spaces are angle-preserving images of hyperbolic geometry.* □

Example 7. Consider the *polar coordinates* u, v on the hyperbolic plane; they are the parameters in the representation

$$\mathfrak{x}(u, v) := \mathfrak{e}_0 \cosh u + (\mathfrak{e}_1 \sin v + \mathfrak{e}_2 \cos v) \sinh u$$

of the pseudo-Euclidean model H^2_+. Figure 2.24 shows the images of the orthogonal coordinate lines in Poincaré's model. In the Klein's model the cartesian coordinates x, y determine a net consisting of hyperbolic lines, which are not orthogonal in the hyperbolic sense. Figures 2.25 and 2.26 show this net in the conformal disk and Poincaré's model, respectively. □

2.6.3 Distance and Angles as Cross Ratios

In the previous section we proved that the hyperbolic space is *two-point homogeneous*, i.e. for any two point pairs in $\boldsymbol{H}^n$ there is a transformation $g \in G_n$ mapping one into the other if and only if they have the same distance; it is the distance alone that determines them up to congruence. On the other hand,

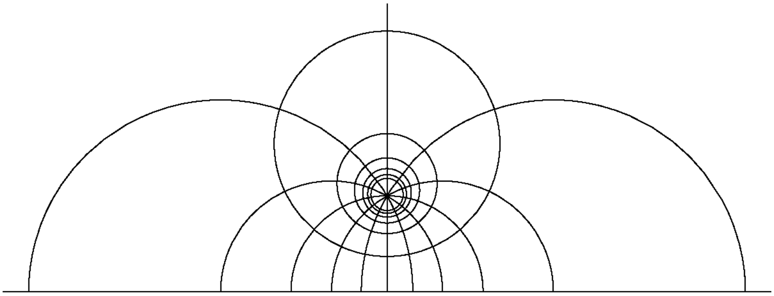


Fig. 2.24. Polar coordinates in Poincaré's model for $\boldsymbol{H}^2$.

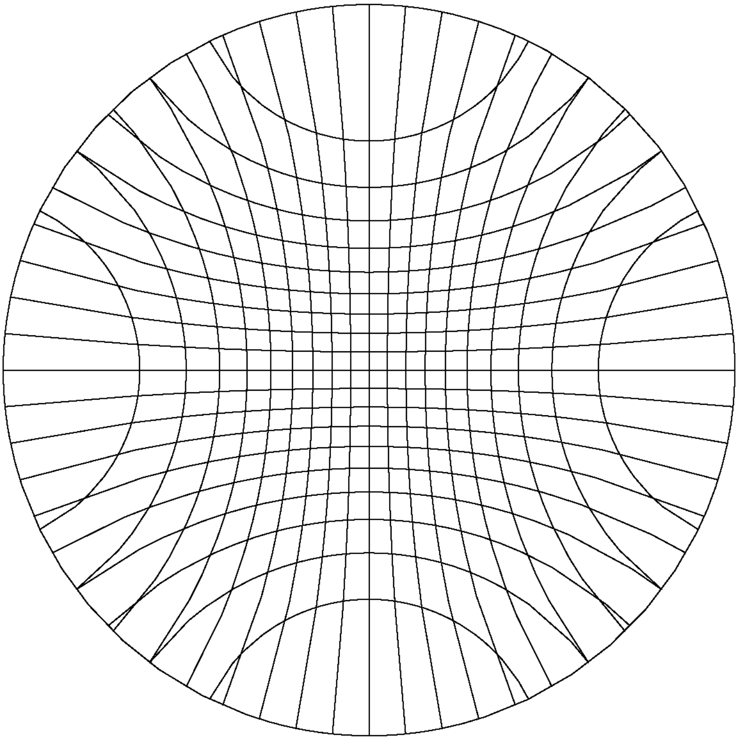


Fig. 2.25. Cartesian coordinate lines x, y in the conformal disk model for $\boldsymbol{H}^2$.

we can assign a projective invariant to any pair of different points in hyperbolic space in a natural way: In fact, if $\boldsymbol{x}_Q, \boldsymbol{y}_Q$ are the points of intersection of the connecting line $\boldsymbol{x} \vee \boldsymbol{y}$ with the quadric Q, we can form the cross ratio (CR) $(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{x}_Q, \boldsymbol{y}_Q)$, which still depends on the order of the points. Since it is a projective invariant, it appears natural to ask, whether there is a relation between this CR and the distance of the points. Already in 1859 A. Cayley

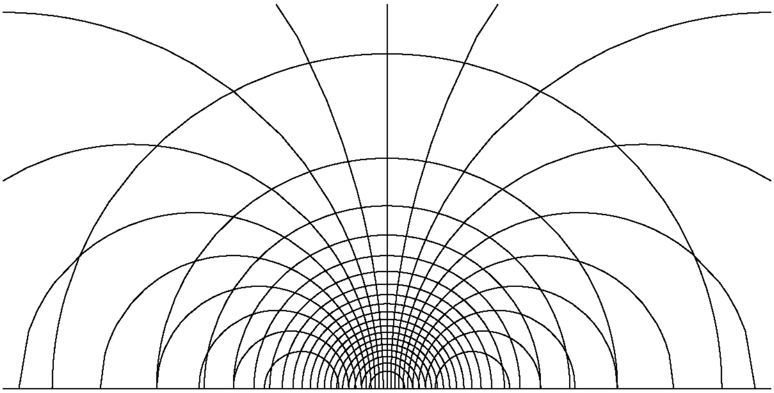


Fig. 2.26. Cartesian coordinate lines x, y in Poincaré's model for $\boldsymbol{H}^2$.

[29] studied this problem quite generally, and later F. Klein [60] resumed the discussion.

Proposition 5. *Let $\boldsymbol{x} \neq \boldsymbol{y}$ be two points in the hyperbolic space $\boldsymbol{H}^n(r)$, and let $\boldsymbol{x}_Q, \boldsymbol{y}_Q$ be the points of intersection of the connecting line $\boldsymbol{x} \vee \boldsymbol{y}$ with the absolute Q. Then the distance of the points is related to the cross ration by*

$$h(\boldsymbol{x}, \boldsymbol{y}) = \frac{r}{2} |\ln(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{x}_Q, \boldsymbol{y}_Q)|. \tag{26}$$

Proof. First we note that the right-hand side of (26) does not change, if $\boldsymbol{x}$ and $\boldsymbol{y}$ or $\boldsymbol{x}_Q$ and $\boldsymbol{y}_Q$ are interchanged, since by (1.4.7) the CR is replaced by its inverse. For the proof of the proposition we pass to the complex extensions of both the vector space $\boldsymbol{V}^{n+1}$ and the scalar product $\langle , \rangle$. Now we prove a more general lemma allowing for more applications.

Lemma 6. *Let $\boldsymbol{P}^n_{\mathbf{C}}$ be the complex orthogonal space with the non-degenerate quadric $Q_{\mathbf{C}}$ (Example 3.2). Let the connecting line $\boldsymbol{x} \vee \boldsymbol{y}$ of the two points $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{P}^n_{\mathbf{C}} \setminus Q_{\mathbf{C}}$ be a secant with the points of intersection $\{\boldsymbol{x}_Q, \boldsymbol{y}_Q\} = (\boldsymbol{x} \vee \boldsymbol{y}) \cap Q_{\mathbf{C}}$ (Example 1.9.8). Then by (4.17) the invariant $Sq(\boldsymbol{x}, \boldsymbol{y}) \in \mathbf{C}$ is defined; the solution set of the equation*

$$Sq(\boldsymbol{x}, \boldsymbol{y}) = \cos^2 \varphi, \ \varphi \in \mathbf{C}, \tag{27}$$

is $\Omega = \{\pm\varphi_0 + k\pi |\ k \in \mathbf{Z}\}$, where $\varphi_0 \in \mathbf{C}$ is the solution that is, up to sign, uniquely determined by (27) and the condition $0 \leq \mathrm{Re}(\varphi_0) < \pi$. The number

$$k(\boldsymbol{x}, \boldsymbol{y}) := |\,\mathrm{Im}(\varphi)| \tag{28}$$

does not depend on the choice of $\varphi \in \Omega$, and, moreover,

$$k(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2} |\ln |(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{x}_Q, \boldsymbol{y}_Q)||. \tag{29}$$

The real part of φ and the argument α of the CR *satisfy*

$$\alpha(\boldsymbol{x},\boldsymbol{y}) := \arg(\boldsymbol{x},\boldsymbol{y};\boldsymbol{x}_Q,\boldsymbol{y}_Q) = 2\,\mathrm{Re}(\varphi) \bmod 2\pi. \tag{30}$$

Proof. That Ω is the solution set and the statement concerning (28) follow from the properties of the complex function $\cos z$. To prove (29) we first compute the parameter on the line $\boldsymbol{x} \vee \boldsymbol{y}$ for the points of intersection $\boldsymbol{x}_Q, \boldsymbol{y}_Q$. Since $\boldsymbol{x}, \boldsymbol{y}$ do not belong to $Q_{\mathbb{C}}$, we find normalized representatives $\mathfrak{x}, \mathfrak{y}$, $\boldsymbol{x} = [\mathfrak{x}]$, $\boldsymbol{y} = [\mathfrak{y}]$, $\langle \mathfrak{x}, \mathfrak{x} \rangle = \langle \mathfrak{y}, \mathfrak{y} \rangle = 1$. The ansatz $\mathfrak{z}(t) := \mathfrak{x}t + \mathfrak{y} \in Q_{\mathbb{C}}$ leads to the quadratic equation $\langle \mathfrak{z}(t), \mathfrak{z}(t) \rangle = 0$ with the solutions

$$t_{1,2} = -\langle \mathfrak{x}, \mathfrak{y} \rangle \pm \sqrt{\langle \mathfrak{x}, \mathfrak{y} \rangle^2 - 1}. \tag{31}$$

Note that $t_1 \neq t_2$. In fact, if this were not the case we had $\langle \mathfrak{x}, \mathfrak{y} \rangle = \pm 1$; the line $\boldsymbol{x} \vee \boldsymbol{y}$ would then correspond to an isotropic vector space and hence be a tangent to $Q_{\mathbb{C}}$ with the point of contact determined by $t_1 = t_2$.

From Formula (1.4.16) we obtain for the CR

$$(\boldsymbol{x},\boldsymbol{y};\boldsymbol{z}(t_1),\boldsymbol{z}(t_2)) = t_2/t_1. \tag{32}$$

Choosing an arbitrary complex solution φ for $\langle \mathfrak{x}, \mathfrak{y} \rangle = \cos\varphi$, inserting it into (31) and the result into (32) a simple computation yields

$$(\boldsymbol{x},\boldsymbol{y};\boldsymbol{x}_Q,\boldsymbol{y}_Q) = \mathrm{e}^{2\,\mathrm{i}\,\varphi}\,. \tag{33}$$

When $\boldsymbol{x}_Q$ and $\boldsymbol{y}_Q$ are interchanged, then according to (1.4.7) the CR is replaced by its inverse, i.e. φ by $-\varphi$. Hence φ is only determined up to sign by $\boldsymbol{x}, \boldsymbol{y}$. We now write the CR in trigonometric form as $\rho\,\mathrm{e}^{\mathrm{i}\,\alpha}$ and use the decomposition

$$\varphi = \xi(\boldsymbol{x},\boldsymbol{y}) + \mathrm{i}\,\eta(\boldsymbol{x},\boldsymbol{y})$$

of φ into real and imaginary part. Inserting this into (33) leads to the formulas (29) and (30) of the lemma. □

We conclude the proof of Proposition 5. Since the connecting line $\boldsymbol{x} \vee \boldsymbol{y} \subset \boldsymbol{H}^n$ is a secant of the real quadric Q, the solutions t_1, t_2 from (31) are real and different; hence we have

$$\cos^2(\varphi) = Sq(\boldsymbol{x},\boldsymbol{y}) = \langle \mathfrak{x}, \mathfrak{y} \rangle^2 > 1.$$

Thus φ_0 is purely imaginary, $\varphi_0 = \mathrm{i}\,\psi$, and by (33) the CR is positive. Taking into account (21), (28), (29) with $h(\boldsymbol{x},\boldsymbol{y}) = r\,k(\boldsymbol{x},\boldsymbol{y})$ we conclude from (33) the assertion of Proposition 5. □

We return now to elliptic geometry and again consider the extension of the elliptic space to the complex orthogonal projective space $\boldsymbol{P}^n_{\mathbb{C}}$ with the bilinear extension of the scalar product. Because of $\boldsymbol{x} \neq \boldsymbol{y}$ the normalized representatives satisfy $\langle \mathfrak{x}, \mathfrak{y} \rangle^2 < 1$, and hence the solutions (31) are complex conjugate. Thus in equation (33) for the CR φ is real, and Definition (5.16) (with r=1) for the distance in elliptic geometry implies

Corollary 7. *For the distance of two points $\boldsymbol{x}, \boldsymbol{y}$ in the elliptic space with parameter r the following equation holds*

$$e(\boldsymbol{x}, \boldsymbol{y}) = \frac{r}{2} |\ln(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{x}_Q, \boldsymbol{y}_Q)| = r|\varphi|. \tag{34}$$

Here $\boldsymbol{x}_Q, \boldsymbol{y}_Q$ denote the points of intersection of the complex extension $(\boldsymbol{x} \vee \boldsymbol{y})_{\mathbf{C}}$ for the connecting line with the absolute $Q_{\mathbf{C}}$ of the complex extension for the elliptic space. The logarithm is meant to be the principal value of the complex logarithm. □

By Definition (5.20) the angle between two elliptic hyperplanes is reduced to the distance of its poles; hence there is a corresponding formula for this as well. The duality of points and hyperplanes, lines and hyperplane pencils with an $(n-2)$-plane as support, auto-polar points $\boldsymbol{x}_Q \in Q_{\mathbf{C}}$ and tangent hyperplanes leads to

Corollary 8. *The angle of two hyperplanes $\boldsymbol{X}, \boldsymbol{Y}$ in the elliptic space satisfies*

$$\angle(\boldsymbol{X}, \boldsymbol{Y}) = \frac{1}{2} |\ln(\boldsymbol{X}, \boldsymbol{Y}; \boldsymbol{X}_Q, \boldsymbol{Y}_Q)| = |\varphi|. \tag{35}$$

Here $\boldsymbol{X}_Q, \boldsymbol{Y}_Q$ denote the tangent hyperplanes for the complex extension of the hyperplane pencil with support $(\boldsymbol{X} \wedge \boldsymbol{Y})_{\mathbf{C}}$ to the absolute $Q_{\mathbf{C}}$ for the complex extension of elliptic space. The logarithm is meant to be the principal value of the complex logarithm. □

Example 8. Corollary 8 correspondingly also holds for the angle of two intersecting hyperplanes $\boldsymbol{X}, \boldsymbol{Y}$ in hyperbolic space; in fact, according to Example 4, Case 1, the poles of these planes span a line in the outer region of the quadric, whose complex extension again intersects the complex quadric $Q_{\mathbf{C}}$ in two complex conjugate points, the polars of which are the (complex) tangent hyperplanes to $Q_{\mathbf{C}}$. If the hyperplanes are parallel but not boundary parallel (Case 3 of Example 4), then their intersection $\boldsymbol{X} \wedge \boldsymbol{Y}$ lies in the outer region $A(Q)$ of the quadric; hence its polar $(\boldsymbol{X} \wedge \boldsymbol{Y})^{\perp} = \boldsymbol{x} \vee \boldsymbol{y}$, $\boldsymbol{x} = \boldsymbol{X}^{\perp}, \boldsymbol{y} = \boldsymbol{Y}^{\perp}$, is a secant of the quadric, where now $\boldsymbol{x}, \boldsymbol{y}$ are outer points. Since the CR is real again, the angle $\varphi = \mathrm{i}\,\psi$, $\psi \in \mathbf{R}$, in (33) has to be purely imaginary. Thus the CR is positive and determines via

$$|\psi| = \frac{1}{2} |\ln(\boldsymbol{X}, \boldsymbol{Y}; \boldsymbol{X}_Q, \boldsymbol{Y}_Q)| = \operatorname{arcosh} |I|(\boldsymbol{X}, \boldsymbol{Y}) \tag{36}$$

a kind of *distance* for the parallel hyperbolic hyperplanes. This notation is generalized and justified in Proposition 25. □

Example 9. In the pseudo-orthonormal standard basis we consider the vectors $\mathfrak{n}(t) := \mathfrak{e}_0 \cos(t) + \mathfrak{e}_1 \sin(t)$. The corresponding points $\boldsymbol{n}(t) = [\mathfrak{n}(t)]$ oscillate back and forth on the line segment between the inner point $[\mathfrak{e}_0]$ and the outer point $[\mathfrak{e}_1]$ of the quadric Q. The polar hyperplanes $\boldsymbol{N}(t) = \boldsymbol{n}(t)^{\perp}$ satisfy:

1. For $0 \leq t < \pi/4$ the vector $\mathfrak{n}(t)$ is time-like, hence the polar is an outer hyperplane; the scalar product restricted to its vector space is positive definite and determines an elliptic geometry in the corresponding projective hyperplane.
2. For $\pi/4 < t < \pi/2$ the vector $\mathfrak{n}(t)$ is space-like, hence the polar is a hyperbolic hyperplane; the scalar product restricted to its vector space is pseudo-orthogonal of index 1 and determines a hyperbolic geometry in the corresponding projective hyperplane.
3. The point $\boldsymbol{p} := \boldsymbol{n}(\pi/4)$ lies on the quadric Q, and its polar $\boldsymbol{T} := \boldsymbol{N}(\pi/4)$ is the tangent hyperplane to Q at this point. The polar is spanned by

$$\boldsymbol{T} = [\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_n] \text{ with } \mathfrak{a}_1 := \mathfrak{e}_0 + \mathfrak{e}_1, \mathfrak{a}_i = \mathfrak{e}_i \text{ for } i = 2, \ldots, n.$$

The bilinear form induced on the associated n-dimensional vector space $\boldsymbol{W}^n$ is *positive semidefinite* of rank $n-1$ and obviously symmetric. With respect to the basis $(\mathfrak{a}_j), j = 1 \ldots, n$, it has the coordinate representation

$$\langle \mathfrak{x}, \mathfrak{y} \rangle = x^2 y^2 + \ldots + x^n y^n .$$

The next exercise contains some properties of the geometry induced on a tangent hyperplane. □

Exercise 8. Prove under the assumptions and using the notation of Example 9, 3: a) The isotropy group $G_{\boldsymbol{T}}$ of $\boldsymbol{T}$ under the action of G_n coincides with the subgroup of the projective group preserving the polar map determined by the scalar product $\langle , \rangle|_{\boldsymbol{T}\times\boldsymbol{T}}$. – b) The point of contact $\{\boldsymbol{p}\} = Q \cap \boldsymbol{T}$ is fixed for the action of $G_{\boldsymbol{T}}$, and $G_{\boldsymbol{T}}$ acts transitively on the complement $\boldsymbol{T}\backslash\{\boldsymbol{p}\}$. – c) The group $G_{\boldsymbol{T}}$ acts transitively on the bundle

$$\tau_k(\boldsymbol{p}) := \{\boldsymbol{B}^k \subset \boldsymbol{T} |\ \boldsymbol{p} \in \boldsymbol{B}^k\}$$

as well as on the set

$$M_k(\boldsymbol{p}) := \{\boldsymbol{B}^k \subset \boldsymbol{T} |\ \boldsymbol{p} \notin \boldsymbol{B}^k\}$$

of k-planes not containing $\boldsymbol{p}$. – d) Let $\boldsymbol{x} \neq \boldsymbol{y}$ be two points of $\boldsymbol{T}$ different from $\boldsymbol{p}$. Then by (4.17) their invariant $\mathrm{Sq}(\boldsymbol{x}, \boldsymbol{y})$ is well-defined; two such points are $G_{\boldsymbol{T}}$-congruent if and only if these invariants coincide. The points $\boldsymbol{x}, \boldsymbol{y}$ lie on a line through $\boldsymbol{p}$ if and only if $\mathrm{Sq}(\boldsymbol{x}, \boldsymbol{y}) = 1$. — The geometry discussed in this exercise is the elementary example of an *isotropic geometry*. Isotropic geometry was studied and applied by Karl Strubecker in various ways. A complete list of references together with a recognition of his merits can be found in the obituary by K. Leichtweiß [74].

Example 10. We still want to show how the *Euclidean geometry* fits into the scheme described here. To this end, we start from the elliptic geometry, i.e., we consider a real vector space $\boldsymbol{V}^{n+1}$ with the positive definite scalar product $\langle , \rangle$, the orthonormal standard basis $(\mathfrak{e}_j), j = 0, \ldots, n$, and the associated elliptic space $\boldsymbol{P}^n$. Let $\boldsymbol{W}^n := [\mathfrak{e}_1, \ldots, \mathfrak{e}_n]$ denote the subspace spanned by the last n basis vectors, and let $\mathfrak{x}_o \in \boldsymbol{W}^n$ denote the orthogonal projection of

$\mathfrak{x} = \mathfrak{e}_0 x^0 + \mathfrak{x}_o \in \boldsymbol{V}^{n+1}$ onto this subspace. We now define the scalar product over $\boldsymbol{V}$ depending on the parameter $t \in \mathbf{R}$:

$$\langle \mathfrak{x}, \mathfrak{y} \rangle_t := -(x^0 y^0 \cos t + \langle \mathfrak{x}_o, \mathfrak{y}_o \rangle \sin t). \tag{37}$$

For $0 < t < \pi/2$ the corresponding quadric $Q_t : \langle \mathfrak{x}, \mathfrak{x} \rangle_t = 0$ is imaginary; hence the geometry is elliptic, and for $-\pi/2 < t < 0$ it is a real ellipsoid, whose inner region is a model for hyperbolic geometry. For $t = 0$ the quadric degenerates into the "double counted hyperplane" $(x^0)^2 = 0$ with the vector space $\boldsymbol{W}^n$. We choose this as the (improper) hyperplane at infinity of an affine space $\boldsymbol{A}^n \subset \boldsymbol{P}^n$. Since for $t \to 0$, $t > 0$, the factor $\sin t$ cancels for the points in this hyperplane the invariant

$$\cos \varphi = \frac{\langle \mathfrak{x}_o, \mathfrak{y}_o \rangle_t}{\sqrt{\langle \mathfrak{x}_o, \mathfrak{x}_o \rangle_t \langle \mathfrak{y}_o, \mathfrak{y}_o \rangle_t}}$$

actually does not depend on t, and defines the angle φ between the vectors in the usual way. The improper hyperplane is thus equipped with an elliptic geometry in a natural way. Its group is the orthogonal group $\boldsymbol{O}(n)$ acting on $\boldsymbol{W}^n$, and this preserves the distance defined for the proper points as usual: Consider their representatives $\mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}^{n+1}$ normalized by $x^0 = y^0 = 1$, then $\mathfrak{x} - \mathfrak{y} \in \boldsymbol{W}^n$, and $\boldsymbol{O}(n)$ as well as the group of translations indeed preserve

$$\rho(\boldsymbol{x}, \boldsymbol{y}) := \sqrt{\langle \mathfrak{x} - \mathfrak{y}, \mathfrak{x} - \mathfrak{y} \rangle}.$$

We thus have the analytic model for Euclidean geometry, which in this context, because of its bordering position between hyperbolic and elliptic geometry, is also called *parabolic geometry.* According to O. Giering [44] the absolute of Euclidean space consists of the degenerate quadric $(x^0)^2 = 0$ together with a given empty quadric $Q_{\mathbf{C}}$ in its vertex space (its "cusp"), which thus lies in the complex projective space belonging to the complex extension $\boldsymbol{W}^n_{\mathbf{C}}$. Considering, at the same time, real projective spaces and their complex extensions he widely generalizes this scheme in the notion *Cayley-Klein geometries.* The classification he accomplishes in low dimensions is based on the classification of real and complex polarities. □

2.6.4 The Hyperbolic Law of Cosines and Hyperbolic Metric

Let $A, B, C \in \boldsymbol{H}^n$ be three hyperbolic points in general position. For the purposes of trigonometry we may again suppose $n = 2$. Represent the sides of the triangle starting at A as in Example 5, where the pseudo-orthonormal three-frame is suitably adapted: Set

$$A = [\mathfrak{a}] = [\mathfrak{e}_0], \tag{38}$$

$$B = [\mathfrak{b}] = [\mathfrak{e}_0 \cosh c + \mathfrak{e}_1 \sinh c], \tag{39}$$

$$C = [\mathfrak{c}] = [\mathfrak{e}_0 \cosh b + \mathfrak{d}_1 \sinh b], \tag{40}$$

$$\mathfrak{d}_1 = \mathfrak{e}_1 \cos \alpha + \mathfrak{e}_2 \sin \alpha. \tag{41}$$

Equation (41) is a consequence of the fact that $\mathfrak{d}_1$ is orthogonal to $\mathfrak{e}_0$; according to Example 5 the angle α of the triangle at the vertex A, $0 < \alpha < \pi$, is uniquely determined by $\cos\alpha = \langle \mathfrak{e}_1, \mathfrak{d}_1 \rangle$, and $\mathfrak{e}_2$ is uniquely determined by (41); more generally, the three-frame is uniquely determined by the triangle via (38)–(41). Here, by (21) the numbers b, c are the lengths of the sides in the triangle, which, as usual, are denoted correspondingly:

$$h(A, B) = c,\ h(A, C) = b,\ h(B, C) = a.$$

Note that the scalar product of two vectors from H^n_+ is always negative, more precisely, we have

$$\langle \mathfrak{a}, \mathfrak{b} \rangle < -1 \text{ for all } \mathfrak{a}, \mathfrak{b} \in H^n_+,\ \mathfrak{a} \neq \mathfrak{b}. \tag{42}$$

In fact, we may always choose the orthonormal vectors $\mathfrak{e}_0, \mathfrak{e}_1$ so that (38) and (39) hold for $A = [\mathfrak{a}], B = [\mathfrak{b}]$, which implies $\langle \mathfrak{a}, \mathfrak{b} \rangle = -\cosh c < -1$. Now the side length a is easily computed from (21), (40), and (41):

$$\langle \mathfrak{b}, \mathfrak{c} \rangle = -\cosh a = -\cosh c \cosh b + \langle \mathfrak{e}_1, \mathfrak{d}_1 \rangle \sinh c \sinh b.$$

Moreover, (41) implies the *hyperbolic law of cosines*

Proposition 9. *Consider the notations for a triangle common in elementary geometry. Then, in the hyperbolic geometry of radius $r = 1$ the length of the side a opposite to the vertex A satisfies*

$$\cosh a = \cosh b \cosh c - \sinh b \sinh c \cos\alpha. \tag{43}$$

□

Exercise 9. Prove that formula (43) also holds for degenerate triangles ($\alpha = 0, \pi$).

Corollary 10. *The hyperbolic distance satisfies the triangle inequality for any three points $A, B, C \in \boldsymbol{H}^n$:*

$$h(B, C) \leq h(B, A) + h(A, C); \tag{44}$$

equality holds in (44) *if and only if A lies on the hyperbolic line $B \vee C$ between*[5] *B and C. Hence by* (22) *h is a G_n-invariant metric on $\boldsymbol{H}^n$. The isometry group of this metric is the group G_n.*

Proof. Because of $-\cos\alpha \leq 1$, Equation (43) is an immediate consequence of the addition formulas for cosh

$$\cosh a \leq \cosh(b + c);\ \cosh a = \cosh(b + c) \Longleftrightarrow \alpha = \pi.$$

[5] We do not introduce the notion of betweenness formally. On curves equipped with with a real parameter t it is always used in the sense of the order transferred from $\mathbf{R}$ onto the curve via the parametrization.

The last condition, however, just expresses the position of the points asserted in the corollary. We now prove the last assertion. If g is a projective transformation preserving the metric, then any linear map generating it has to be conformally pseudo-orthogonal; by the definition of G_n we may suppose that this linear transformation belongs to G_n. To prove that every isometry of $\boldsymbol{H}^n$ is generated by a map $g \in G_n$ we introduce an orthogonal coordinate system as follows: Set

$$\mathfrak{a}_0 := \mathfrak{e}_0,\, \mathfrak{a}_i := \mathfrak{e}_0\sqrt{2} + \mathfrak{e}_i \text{ for } i = 1, \ldots, n.$$

Obviously, the edges of the simplex starting at $\boldsymbol{a}_0 = [\mathfrak{a}_0]$ are orthogonal, the $\mathfrak{a}_i$ are normalized representatives of its vertices $\boldsymbol{a}_i = [\mathfrak{a}_i] \in \boldsymbol{H}^n$, and by (21) the distances of the vertices can be computed from

$$\langle \mathfrak{a}_0, \mathfrak{a}_i \rangle = -\sqrt{2} \text{ for } i > 0,\ \langle \mathfrak{a}_i, \mathfrak{a}_j \rangle = -2 \text{ for } i, j > 0, i \neq j.$$

If, conversely, $\boldsymbol{a}_0, \ldots, \boldsymbol{a}_n$ are $n+1$ points whose normalized representatives $\mathfrak{a}_i$ satisfy these equations, then

$$\mathfrak{e}_0 := \mathfrak{a}_0;\ \mathfrak{e}_i := \mathfrak{a}_i - \mathfrak{a}_0\sqrt{2} \text{ for } i = 1, \ldots, n,$$

uniquely determines a pseudo-orthonormal frame $(\mathfrak{e}_i)$, $i = 0, \ldots, n$. We call such point sequences *orthogonal coordinate simplices* and conclude: *For any two orthogonal coordinate simplices* $(\boldsymbol{a}_i)$, $(\boldsymbol{b}_i)$ *there is a unique transformation* $g \in G_n$ *mapping them into one another, i.e.* $g(\boldsymbol{a}_i) = \boldsymbol{b}_i$. g *is determined by an appropriate assignment between the pseudo-orthonormal bases corresponding to the point sequences.* Now let $\boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{H}^n$ be an arbitrary point; by ξ^i we denote the coordinates of its normalized representatives with respect to the basis $(\mathfrak{a}_i)$. Then the distances $h_i := h(\boldsymbol{x}, \boldsymbol{a}_i)$ satisfy:

$$\langle \mathfrak{x}, \mathfrak{a}_0 \rangle = -\cosh h_0 = -\xi^0 - \sqrt{2}\sum_{i=1}^{n} \xi^i,$$

$$\langle \mathfrak{x}, \mathfrak{a}_j \rangle = -\cosh h_j = -\sqrt{2}\xi^0 - \xi^j - 2 \sum_{i=1, i\neq j}^{n} \xi^i.$$

The determinant of the matrix for this system of equations is equal to -1. Hence the coordinates (ξ^i) can be uniquely expressed in terms of the distances (h_i), and vice versa. Let now $f : \boldsymbol{H}^n \to \boldsymbol{H}^n$ be an arbitrary isometry. Then $(\boldsymbol{b}_i) := (f(\boldsymbol{a}_i))$ is an orthogonal coordinate simplex as well. The corresponding map $g \in G_n$ is obviously an isometry. Hence the distances satisfy

$$h(g(\boldsymbol{x}), \boldsymbol{b}_i) = h(\boldsymbol{x}, \boldsymbol{a}_i) = h(f(\boldsymbol{x}), \boldsymbol{b}_i).$$

So all the distances from $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$, respectively, to the vertices $\boldsymbol{b}_i$ of the orthogonal coordinate simplex coincide; hence this has to hold for the coordinates of these points as well, implying $f = g \in G_n$. $\square$

For every two hyperbolic points $A, B \in \boldsymbol{H}^n$, $A \neq B$, there is a unique hyperbolic connecting line $A \vee B \cap I(Q)$ and on it (with the notations (38), (39)) the uniquely determined connecting line segment with the parameter representation

$$\mathfrak{x}(t) = \mathfrak{e}_0 \cosh t + \mathfrak{e}_1 \sinh t,\ 0 \leq t \leq c = h(A, B). \tag{45}$$

For $t \to \pm\infty$ the corresponding point on the line tends to its point of intersection with the quadric Q, which thus in hyperbolic geometry has to be considered as the "infinite" of the hyperbolic space. The triangle inequality again implies that the connecting line segment is the shortest curve between two points, which, moreover, is uniquely determined here. Hence in hyperbolic geometry the cut locus is empty for each point. In the next proposition we will determine the distance from a point to a k-plane, and this will lead to the notion of the *perpendicular* familiar from Euclidean geometry.

Proposition 11. *Consider a k-plane $\boldsymbol{A}^k \subset \boldsymbol{H}^n$ and a point $\boldsymbol{a} \in \boldsymbol{H}^n$ in hyperbolic space. Then there is a unique point $\boldsymbol{b} \in \boldsymbol{A}^k$ such that*

$$h(\boldsymbol{a}, \boldsymbol{b}) = \min\{h(\boldsymbol{a}, \boldsymbol{x}) |\ \boldsymbol{x} \in \boldsymbol{A}^k\}.$$

For $\boldsymbol{a} \notin \boldsymbol{A}^k$ each line $\boldsymbol{c} \subset \boldsymbol{A}^k$ through the point $\boldsymbol{b}$ is orthogonal to $\boldsymbol{a} \vee \boldsymbol{b}$.

Proof. Let the pseudo-orthonormal basis $(\mathfrak{e}_i)$ of the vector space $\boldsymbol{V}^{n+1}$ be chosen so that the subspace $\boldsymbol{W}^{k+1}$ associated with the k-plane $\boldsymbol{A}^k$ is spanned by the vectors $\mathfrak{e}_0, \ldots, \mathfrak{e}_k$ (transitivity of G_n on the set of hyperbolic k-planes). We decompose the representative $\mathfrak{a} \in H^n_+$ for the point $\boldsymbol{a}$ into its orthogonal components

$$\mathfrak{a} = \mathfrak{a}_0 + \mathfrak{a}_1,\ \mathfrak{a}_0 \in \boldsymbol{W}, \mathfrak{a}_1 \in \boldsymbol{W}^\perp.$$

Because of $\langle \mathfrak{a}, \mathfrak{a} \rangle = -1$, there is a unique number $c \geq 0$ such that $\langle \mathfrak{a}_0, \mathfrak{a}_0 \rangle = -\cosh^2 c$; the vector $\mathfrak{b} := \mathfrak{a}_0 / \cosh c \in H^n_+$ is the normalized representative for the point $\boldsymbol{b} = [\mathfrak{a}_0]$ in $\boldsymbol{A}^k$. We show that the point $\boldsymbol{b}$ satisfies the assertion of the proposition. In fact, for $\boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{A}^k, \mathfrak{x} \in H^n_+$, we have by (21) and (42)

$$\begin{aligned} \cosh h(\boldsymbol{a}, \boldsymbol{x}) = -\langle \mathfrak{a}, \mathfrak{x} \rangle = -\langle \mathfrak{a}_0, \mathfrak{x} \rangle = -\langle \mathfrak{b}, \mathfrak{x} \rangle \cosh c = \cosh h(\boldsymbol{b}, \boldsymbol{x}) \cosh c \\ \geq \cosh c. \end{aligned}$$

Equality holds if and only if $\cosh h(\boldsymbol{b}, \boldsymbol{x}) = 1$, i.e. $h(\boldsymbol{b}, \boldsymbol{x}) = 0$, and hence $\boldsymbol{x} = \boldsymbol{b}$ (cf. (22)). On the other hand,

$$\cosh h(\boldsymbol{a}, \boldsymbol{b}) = -\langle \mathfrak{a}, \mathfrak{b} \rangle = -\langle \mathfrak{a}_0, \mathfrak{b} \rangle = \cosh c,$$

implying the first assertion. To prove the second we may suppose $\boldsymbol{b} = [\mathfrak{e}_0]$ without loss of generality. To compute the angle we determine the vector $\mathfrak{d}_1$ uniquely defined by

$$\mathfrak{a} = \mathfrak{e}_0 \cosh c + \mathfrak{d}_1 \sinh c, \boldsymbol{a} = [\mathfrak{a}], \ \langle \mathfrak{e}_0, \mathfrak{d}_1 \rangle = 0, \ \langle \mathfrak{d}_1, \mathfrak{d}_1 \rangle = 1, \ c > 0.$$

The case $c = h(\boldsymbol{a}, \boldsymbol{b}) = 0$ is only possible for $\boldsymbol{a} \in \boldsymbol{A}^k$, and then nothing remains to be proved. An arbitrary line through $\boldsymbol{b}$ and $\boldsymbol{x} \in \boldsymbol{A}^k$ has the parameter representation

$$\mathfrak{x}(t) = \mathfrak{e}_0 \cosh(t) + \mathfrak{x}_1 \sinh(t) \text{ mit } \mathfrak{x} = \mathfrak{x}(h(\boldsymbol{b}, \boldsymbol{x})), \langle \mathfrak{e}_0, \mathfrak{x}_1 \rangle = 0, \langle \mathfrak{x}_1, \mathfrak{x}_1 \rangle = 1.$$

According to Example 5 the angle then satisfies $\cos \alpha = \langle \mathfrak{d}_1, \mathfrak{x}_1 \rangle$. The distance $h(\boldsymbol{a}, \boldsymbol{x}(t))$ is computed from (21) by

$$\begin{aligned} f(t) &:= \langle \mathfrak{a}, \mathfrak{x}(t) \rangle \\ &= \langle \mathfrak{e}_0 \cosh c + \mathfrak{d}_1 \sinh c, \mathfrak{e}_0 \cosh(t) + \mathfrak{x}_1 \sinh t \rangle \\ &= -\cosh c \cosh t + \langle \mathfrak{d}_1, \mathfrak{x}_1 \rangle \sinh c \sinh t \\ &= -\cosh c \cosh t + \cos \alpha \sinh c \sinh t. \end{aligned}$$

This function has to be stationary at $t = 0$; because of

$$f'|_{t=0} = \cos \alpha \sinh c = 0$$

and $c > 0$ we immediately conclude $\alpha = \pi/2$. □

As usual in metric spaces we call the number $h(\boldsymbol{A}^k, \boldsymbol{a}) := h(\boldsymbol{b}, \boldsymbol{a})$ the *distance from the point $\boldsymbol{a}$ to the k-plane $\boldsymbol{A}^k$*, the segment between $\boldsymbol{a}$ and $\boldsymbol{b}$ is called the *perpendicular* from $\boldsymbol{a} \notin \boldsymbol{A}^k$ onto $\boldsymbol{A}^k$, and $\boldsymbol{b}$ is called the *foot* of the perpendicular. By means of a suitably adapted basis it is easy to show

Corollary 12. *Two pairs $(\boldsymbol{A}^k, \boldsymbol{a})$, $(\boldsymbol{B}^k, \boldsymbol{b})$ of hyperbolic k-planes $\boldsymbol{A}^k, \boldsymbol{B}^k$ and points $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{H}^n$ are G_n-congruent if and only if the distances satisfy $h(\boldsymbol{A}^k, \boldsymbol{a}) = h(\boldsymbol{B}^k, \boldsymbol{b})$.* □

Exercise 10. Let $\boldsymbol{A}^1$ be a line in the hyperbolic plane. Prove that the set

$$M(\boldsymbol{A}^1, c) := \{\boldsymbol{x} \in \boldsymbol{H}^2 |\, h(\boldsymbol{A}^1, \boldsymbol{x}) = c\}, \quad (c > 0),$$

is decomposed into two G_n-congruent arcs meeting the line $\boldsymbol{A}^1$ in its points at infinity $\boldsymbol{A}^1 \cap Q$. In contrast to Euclidean geometry, these arcs are no lines (cf. Figure 2.27). They are called the *equidistants of the line $\boldsymbol{A}^1$*. Find a parameter representation for these arcs. What do the similarly defined equidistants in spherical and in elliptic geometry look like? By Exercise 5.23 the ϵ-neighborhood of the line $\boldsymbol{A}^1$ in the elliptic plane $\boldsymbol{P}^2(r)$:

$$U(\boldsymbol{A}^1, \epsilon) := \{\boldsymbol{x} \in \boldsymbol{P}^2(r) |\, e(\boldsymbol{A}^1, x) < \epsilon\}, \quad (0 < \epsilon < r\pi/2),$$

is a Möbius strip.

Exercise 11. Let $\boldsymbol{A}^k$ be a hyperbolic k-plane, and let $\boldsymbol{a} \notin \boldsymbol{A}^k, \boldsymbol{b} \in \boldsymbol{A}^k$ be two points such that the line $\boldsymbol{a} \vee \boldsymbol{b}$ is orthogonal to every line $\boldsymbol{B}^1 \subset \boldsymbol{A}^k$ with $\boldsymbol{b} \in \boldsymbol{B}^1$. Prove that the line segment between $\boldsymbol{a}$ and $\boldsymbol{b}$ is the perpendicular from $\boldsymbol{a}$ onto $\boldsymbol{A}^k$.

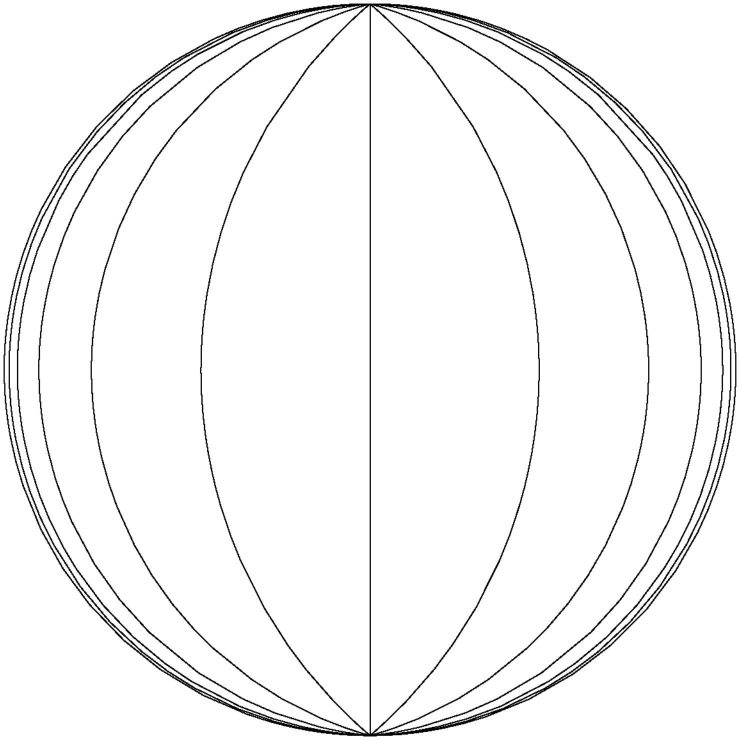


Fig. 2.27. Equidistants of a line in the conformal disk model.

Exercise 12. Prove: a) For every two lines $\boldsymbol{A} \neq \boldsymbol{B}$ in the hyperbolic plane $\boldsymbol{H}^2$ there exists a line $\boldsymbol{C}$ orthogonal to both if and only if they are parallel (and not boundary parallel). (Hint. Apply the angle definition from Example 4, Case 1, and consider the polar of the point of intersection $\boldsymbol{A} \wedge \boldsymbol{B}$ lying in the outer region $A(Q)$.) – b) The *common perpendicular* $\boldsymbol{C}$ of two parallel lines is uniquely determined. – c) If $\boldsymbol{a} = \boldsymbol{A} \wedge \boldsymbol{C}, \boldsymbol{b} = \boldsymbol{B} \wedge \boldsymbol{C}$ are the points of intersection of the lines with the common perpendicular, then $h(\boldsymbol{A}, \boldsymbol{B}) := h(\boldsymbol{a}, \boldsymbol{b})$ is called the *distance of the parallel lines* $\boldsymbol{A}, \boldsymbol{B}$. Prove that this coincides with their metric distance:

$$h(\boldsymbol{A}, \boldsymbol{B}) = \min\{h(\boldsymbol{x}, \boldsymbol{y}) | \, \boldsymbol{x} \in \boldsymbol{A}, \boldsymbol{y} \in \boldsymbol{B}\}.$$

d) Two pairs of parallel lines $(\boldsymbol{A}, \boldsymbol{B}), (\boldsymbol{A}_1, \boldsymbol{B}_1)$ are G_n-congruent if and only if they have equal distances. – e) $h(\boldsymbol{A}, \boldsymbol{B}) = r \operatorname{arcosh}(|I|(\boldsymbol{A}, \boldsymbol{B}))$ (see Example 4).

Exercise 13. Prove the following statement generalizing Proposition 11: Let $\boldsymbol{V}^{n+1}$ be an arbitrary vector space with scalar product, let $\boldsymbol{P}^n$ be the associated projective space, and let F be the auto-correlation determined by the scalar product. If $\boldsymbol{A}^k$ is a non-isotropic subspace, then for each point $\boldsymbol{a} \notin \boldsymbol{A}^k \cup F(\boldsymbol{A}^k)$ there are uniquely determined points $\boldsymbol{b}_0 \in \boldsymbol{A}^k, \boldsymbol{b}_1 \in F(\boldsymbol{A}^k)$ such that the line $\boldsymbol{a} \vee \boldsymbol{b}_0$ ($\boldsymbol{a} \vee \boldsymbol{b}_1$) is orthogonal to $\boldsymbol{A}^k$ (to $F(\boldsymbol{A}^k)$) in the sense described in Proposition 11. (Hint. If $\boldsymbol{W} \subset \boldsymbol{V}$ is the subspaces associated with $\boldsymbol{A}^k$, then consider the decomposition of a representative $\mathfrak{a}$ for $\boldsymbol{a}$ into its orthogonal components, $\mathfrak{a} = \mathfrak{b}_0 + \mathfrak{b}_1, \mathfrak{b}_0 \in \boldsymbol{W}, \mathfrak{b}_1 \in \boldsymbol{W}^\perp$.)

Exercise 14. Conclude from Exercise 13: For $\boldsymbol{a} \in Q$ and a hyperbolic k-plane $\boldsymbol{A}^k \subset \boldsymbol{H}^n$ in the hyperbolic geometry there is a unique point $\boldsymbol{b} \in \boldsymbol{A}^k$ such that the line $\boldsymbol{a} \vee \boldsymbol{b}$ is orthogonal to $\boldsymbol{A}^k$ in the sense described in Proposition 11.

Exercise 15. Consider two lines $\boldsymbol{A} \neq \boldsymbol{B}$ in the hyperbolic plane $\boldsymbol{H}^2$. Prove that there is a unique *symmetry axis* $\boldsymbol{S}$, i.e. a hyperbolic line for which there is an isometry $g \in G_2$ satisfying the following conditions:

$$g(\boldsymbol{A}) = \boldsymbol{B}, g(\boldsymbol{s}) = \boldsymbol{s} \text{ for all } \boldsymbol{s} \in \boldsymbol{S}, g^2 = \mathrm{id}_{\boldsymbol{H}^2}\,.$$

2.6.5 Hyperbolic Trigonometry

In this section we again consider triangles in the hyperbolic plane and use the traditional notation. Recall the proof of the law of cosines (Proposition 9). There we adapted a basis $(\mathfrak{e}_i), i = 0, 1, 2$, to a given triangle, cf. (38)–(41). Using this it is easy to prove the *hyperbolic law of sines*:

Proposition 13. *The side lengths a, b, c and the angles α, β, γ of a non-degenerate triangle A, B, C in the hyperbolic plane $\boldsymbol{H}^2$ satisfy*

$$\frac{\sin\alpha}{\sinh a} = \frac{\sin\beta}{\sinh b} = \frac{\sin\gamma}{\sinh c}. \tag{46}$$

Proof. From equations (38)–(41) one computes the volume of the cuboid spanned by the representatives $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}$

$$v(\mathfrak{a}, \mathfrak{b}, \mathfrak{c}) = \sin\alpha \sinh b \sinh c = \sin\beta \sinh a \sinh c = \sin\gamma \sinh a \sinh b,$$

where the last two equations are obtained by cyclic permutations of the vertices A, B, C. Division by $\sinh a \sinh b \sinh c$ yields the assertion. □

Exercise 16. Let (A, B, C) be a right triangle in the hyperbolic plane with the angle $\gamma = \pi/2$. – a) Prove:

$$\sin\alpha = \frac{\sinh a}{\sinh c}, \; \cosh a = \frac{\cosh c}{\cosh b}, \; \cos\alpha = \frac{\tanh b}{\tanh c}. \tag{47}$$

– b) If, for fixed points A and C, the point B moves towards infinity on the line $C \vee B$, then the side lengths tend to infinity, and the angle β tends to zero. – c) In the situation of b) the angle α tends to a limit ω satisfying the equation

$$\cos\omega = \tanh b.$$

(Hint. In addition to (47) apply the hyperbolic law of cosines (43).)

The last equation implies that the angle ω only depends on the length b of the perpendicular from A onto the line $B \vee C$; it is called the *parallel angle*, since it is nothing but the angle of the boundary parallel through A for the line $C \vee B$. The function

$$\omega := \Pi(b) = \arccos(\tanh(b))$$

is often called the *Lobatschewski function* of hyperbolic geometry. The definition above implies that the parallel angle is a monotonously decreasing function of the distance b, and for $b > 0$ it is always less than $\pi/2$. Moreover, for b tending to infinity it approaches zero. Note that the same definition for the parallel angle in Euclidean geometry would yield the constant value $\pi/2$. The parallel angle was introduced by N. I. Lobatschewski in his invention of hyperbolic geometry, cf. A. P. Norden[79] and H. Reichardt [92].

The *hyperbolic law of cosines for angles* dual to the law of cosines cannot be proved as easily as in elliptic geometry by just relying on the duality. In fact, in hyperbolic geometry the projective duality is violated, since the polar of a hyperbolic point is an outer line, i.e. something not hyperbolic; nevertheless, the hyperbolic object corresponding to an outer line is the pencil of lines through its pole ($n = 2$). Here we quote a proof taken from the interesting and clearly written book [79] by A. P. Norden[6].

Proposition 14. *In the traditional notations, a non-degenerate triangle in the hyperbolic plane (parameter $r = 1$) satisfies the* law of cosines for angles*:*

$$\cos\alpha = -\cos\beta\cos\gamma + \sin\beta\sin\gamma\cosh a. \tag{48}$$

Proof. By the hyperbolic law of cosines (43) we have

$$\cosh c = \cosh a \cosh b - \sinh a \sinh b \cos\gamma.$$

Replacing in (43) $\cosh c$ by this expression yields

$$\cosh a = \cosh^2 b \cosh a - \cosh b \sinh b \sinh a \cos\gamma - \sinh b \sinh c \cos\alpha.$$

We apply the hyperbolic law of sines to the last term in this equation,

$$\sinh b \sinh c \cos\alpha = \sinh b \sinh a \sin\gamma \cot\alpha,$$

then insert the result, and, after a straightforward computation, obtain

[6] For many years A. P. Norden held the geometry chair at the University of Kasan, the very same University where N. I. Lobatschewski had worked, who was the first (in the Kasan Messenger, 1829) to publish his discovery of non-Euclidean geometry. The paper of the Hungarian mathematician J. Bolyai only appeared in the year 1832. C. F. Gauß mentioned his considerations concerning non-Euclidean geometry exclusively in letters. It is generally thought that the foundations of non-Euclidean geometry were laid by these three geometers independently.

$$\cosh a \sinh^2 b = \sinh a \sinh b(\cosh b \cos\gamma + \cot\alpha \sin\gamma).$$

Dividing by $\sinh a \sinh b$ leads to

$$\coth a \sinh b = \cosh b \cos\gamma + \cot\alpha \sin\gamma.$$

To the left-hand side of this equation we again apply the hyperbolic law of sines:

$$\coth a \sinh b = \cosh a \sin\beta / \sin\alpha.$$

Inserting and multiplying by $\sin\alpha$ yields

$$\cosh a \sin\beta = \cos\alpha \sin\gamma + \sin\alpha \cos\gamma \cosh b.$$

Similarly,

$$\cosh b \sin\alpha = \cos\beta \sin\gamma + \sin\beta \cos\gamma \cosh a.$$

Inserting this equation into the last but one, we obtain

$$\cosh a \sin\beta = \cos\alpha \sin\gamma + \cos\gamma \cos\beta \sin\gamma + \sin\beta(1 - \sin^2\gamma) \cosh a.$$

A simple calculation then results in (48). □

Exercise 17. Prove that two triangles in the hyperbolic space $\boldsymbol{H}^n$, for which corresponding sides have equal lengths or corresponding angles are equal, are G_n-congruent. Formulate and prove further congruence theorems.

Now we turn to the essential characteristics of hyperbolic geometry distinguishing it from the elliptic as well as the Euclidean one. In differential geometry the following property is shown to be characteristic of Riemannian manifolds with negative curvature:

Proposition 15. *Let A, B, C be three non-collinear points in the hyperbolic space $\boldsymbol{H}^n$. Then the sum of the angles in the triangle they span satisfies*

$$\alpha + \beta + \gamma < \pi.$$

P r o o f. Obviously, we may suppose $n = 2$. First consider a right triangle with the angle $\gamma = \pi/2$. By the law of cosines for angles, (48), we have

$$\cos\beta = \sin\alpha \cosh b > 0, \text{ hence } \beta < \pi/2.$$

This implies

$$\begin{aligned}\cos(\alpha + \beta) &= \cos\alpha \cos\beta - \sin\alpha \sin\beta, \\ &= \sin\alpha(\cos\alpha \cosh b - \sin\beta), \\ &= \sin\alpha(\cos\alpha \cosh b - \sqrt{1 - \sin^2\alpha \cosh^2 b}).\end{aligned}$$

Now

$$\sqrt{1-\sin^2\alpha\cosh^2 b} < \sqrt{1-\sin^2\alpha} = \cos\alpha,$$

and hence

$$\cos(\alpha+\beta) > \sin\alpha\cos\alpha(\cosh b - 1) > 0,$$

i.e. $\alpha+\beta < \pi/2$. Thus the assertion holds for right triangles. Next consider an arbitrary triangle A, B, C. Then the base point D of the perpendicular from C onto the line $A \vee B$ either lies between A and B, or to the left of A, or to the right of B. In the first case we decompose the triangle A, B, C into two right ones, A, D, C and D, B, C. Then its angle γ is the sum of the angles γ_1, γ_2 of both right triangles meeting at the vertex C. Hence

$$\alpha+\beta+\gamma = \alpha+\gamma_1+\beta+\gamma_2 < \pi.$$

Suppose now, e.g., that D lies to the left of A. The right triangle D, A, C has the angles $\pi/2, \pi-\alpha, \gamma_1$, and $\pi-\alpha < \pi/2$ implies $\alpha > \pi/2$. Then it is easy to see — e.g. assigning to each line in the pencil through A a parameter on this line via the points of intersection with the side $B \vee C$ — that the base point of the perpendicular from A onto the side $B \vee C$ has to lie between B and C. Again, this triangle is decomposed into two right ones, and then the assertion is proved as above. The case that D lies to the right of B is treated analogously. □

Exercise 18. Already in Example 3 we showed that in the conformal disk model hyperbolic lines intersect the quadric Q orthogonally. From this one may conclude that the angle between two boundary parallel lines is always zero. Let $\boldsymbol{x}(t) = [\mathfrak{e}_0\cosh t + \mathfrak{e}_1\sinh t] \in \boldsymbol{A}^1$ be a parameter representation of the line $\boldsymbol{A}^1$, and let $\boldsymbol{b} \notin \boldsymbol{A}^1$ be a point not lying on it. Compute the angle $\alpha(t)$ between the lines $\boldsymbol{A}^1$ and $\boldsymbol{b} \vee \boldsymbol{x}(t)$ and show that $\alpha(t)$ tends to zero for $t \to \infty$.

In the hyperbolic plane *polygons* and *convex polygons* are defined just as in spherical or elliptic geometry, cf. Exercise 5.12. This time, however, we allow the vertices to lie lie at infinity, i.e. on the bounding quadric Q. A polygon is called *simply, doubly, etc.* or, in general, *k-fold asymptotic* if one, two, etc., or k of its vertices belong to Q. Since all the sides adjacent to an *asymptotic vertex* $\boldsymbol{a} \in Q$ of the polygon are boundary parallel, the angle at this vertex is zero (Exercise 18).

Exercise 19. Consider non-degenerate triangles in the hyperbolic plane. Prove: a) Two threefold asymptotic triangles are always G_2-congruent. – b) Two doubly asymptotic triangles are G_2-congruent if they form the same angle at their non-asymptotic vertex. – c) For simply asymptotic triangles let, e.g., the vertex $C \in Q$ be asymptotic. Since the sides a, b are half-lines, i.e. have infinite length, a simply asymptotic triangle of this kind has as its essential invariants only the angles α, β, and the side length c. Prove that two such simply asymptotic triangles are G_2-congruent if and only if they coincide in two mutually corresponding of these invariants. (Hint for c). Prove first: If $C = [\mathfrak{e}_0 + \mathfrak{e}_2] \in Q$, and

$$\boldsymbol{A}^1 : \boldsymbol{x}(t) = [\mathfrak{e}_0\cosh t + \mathfrak{e}_1\sinh t]$$

is a hyperbolic line, then for the oriented angle $\alpha(t)$ from $\boldsymbol{A}^1$ to $\boldsymbol{x}(t)\vee C$ the following equation holds

$$\cos\alpha(t) = -\tanh(t),\ -\infty < t < \infty,\ 0 < \alpha < \pi. \tag{49}$$

This statement holds for an arbitrary hyperbolic line $\boldsymbol{A}^1$ and any point $C \in Q$ at infinity, $C \notin \boldsymbol{A}^1$.)

The definition of the area for elementary sets in Section 5 can almost literally be transferred from spherical or elliptic to plane hyperbolic geometry. The only difference is that now the excess $\alpha + \beta + \gamma - \pi$ is negative for every hyperbolic triangle Δ; therefore, the area of a triangle in the hyperbolic plane with parameter r is defined to be the quantity

$$F(\Delta) := (\pi - (\alpha + \beta + \gamma))r^2. \tag{50}$$

Similar to Lemma 5.8 and Corollary 5.9 one proves

Proposition 16. *If M is an elementary set, and $(\Delta_i), i = 1, \ldots, k$, is any triangulation of M, then the definition of the* area *for M*

$$F(M) := \sum_{i=1}^{k} F(\Delta_i) \tag{51}$$

is independent of the chosen triangulation. □

Obviously, the area is an additive, monotonous, and G_2-invariant function on the system of elementary sets. Similar to what is done in the theory of the Lebesgue measure, this area can be extended to a large class of measurable sets in a G_2-invariant way. It is not difficult to see that the hyperbolic plane $\boldsymbol{H}^2$ itself is not an elementary set. The next exercise implies that its measure is infinite:

Exercise 20. Prove: a) Formula (50) also makes sense for asymptotic triangles. Each trebly asymptotic triangle has area πr^2. – b) In the hyperbolic plane there are elementary sets of arbitrarily large area.

Exercise 21. Prove: a) In the hyperbolic plane there are no rectangles. – b) There exist G_2-congruent, fourfold asymptotic quadrangles.

Exercise 22. In analogy to (5.40) prove that the area of a convex hyperbolic polygon $P_n, n \geq 3$, with the angles α_i at the vertices A_i, $i = 1, \ldots, n$, satisfies

$$F(P_n) = ((n-2)\pi - \sum_{i=1}^{n} \alpha_i)r^2. \tag{52}$$

In particular, each convex n-fold asymptotic n-angle has area $(n-2)\pi r^2$.

Exercise 23. In the complement $\boldsymbol{H}^2 \setminus M$ of each elementary set M there are elementary sets of arbitrarily large area. Find examples for this.

2.6.6 Hyperspheres, Equidistants, and Horospheres

As a metric space the hyperbolic space $\boldsymbol{H}^n$ contains the *metric hyperspheres* $S^{n-1}(\boldsymbol{z}, r)$ of *radius* r with *center* $\boldsymbol{z} = [\mathfrak{z}]$; its points $\boldsymbol{x} = [\mathfrak{x}]$ satisfy the equation

$$\langle \mathfrak{x}, \mathfrak{z} \rangle = -\cosh r, \ (\langle \mathfrak{x}, \mathfrak{x} \rangle = -1, \langle \mathfrak{z}, \mathfrak{z} \rangle = -1). \tag{53}$$

Hence the notions circle ($n = 2$), sphere ($n = 3$), and hypersphere are defined just as in elementary Euclidean geometry; and like there we have

Proposition 17. *The isotropy group $G_{\boldsymbol{z}} \subset G_n$ of the center $\boldsymbol{z}$ for the hypersphere $S^{n-1} = S(\boldsymbol{z}, r)$ acts transitively on S^{n-1}; it is isomorphic to the orthogonal group $\boldsymbol{O}(n)$. Thus the isotropy group of a point $\boldsymbol{x}_0 \in S^{n-1}$ under the action of $G_{\boldsymbol{z}}$ is again isomorph to $\boldsymbol{O}(n-1)$, and, as homogeneous spaces, the hyperbolic hyperspheres are equivariantly isomorphic to the Euclidean hyperspheres: $S^{n-1} \cong \boldsymbol{O}(n)/\boldsymbol{O}(n-1)$. The hyperspheres are projective hyperellipsoids, each of whose tangent hyperplanes orthogonally intersects the line in the pencil through the center $\boldsymbol{z}$ to its point of contact.*

Proof. As G_n acts transitively on the hyperbolic points, we may suppose $\mathfrak{z} = \mathfrak{e}_0$. The isotropy group of $[\mathfrak{e}_0]$ in G_n is the orthogonal group of the Euclidean vector space $[\mathfrak{e}_1, \ldots, \mathfrak{e}_n]$, which acts transitively on the unit sphere S_1^{n-1} of this vector space in the usual linear way. Since it preserves the distance, and as the points $\boldsymbol{x} \in S^{n-1}$ can be represented bijectively and equivariantly in the form

$$\mathfrak{e}_1 \in S_1^{n-1} \longmapsto \boldsymbol{x} = [\mathfrak{e}_0 \cosh r + \mathfrak{e}_1 \sinh r] \in S^{n-1}, \tag{54}$$

the first four statements of the proposition are immediate. Passing to the homogeneous coordinates of $\boldsymbol{x}$ in equation (53) and using $\mathfrak{z} = \mathfrak{e}_0$ we obtain

$$-(x^0)^2(\sinh r)^2 + (\sum_{i=1}^{n}(x^i)^2)(\cosh r)^2 = 0.$$

This is the equation of a hyperquadric of index 1. The reflection in a radial line $\boldsymbol{z} \vee \boldsymbol{x}$ leaves each point of this line fixed, hence $\boldsymbol{z}$ as well. Since it is an isometry, the hypersphere S^{n-1} is mapped into itself. As $\boldsymbol{x}$ also remains fixed, the tangent hyperplane at $\boldsymbol{x}$ has to be mapped into itself, too. Thus it is orthogonal to the radial line $\boldsymbol{z} \vee \boldsymbol{x}$. □

Thus, just as in Euclidean geometry, the hyperspheres are orbits of the isotropy group of their centers. It is certainly a promising principle to consider the elementary geometry of a given transformation group $[G, M]$ as the study of G-invariant properties of the orbits of suitable subgroups $H \subset G$ in M. As will soon become clear, in this respect hyperbolic geometry offers more opportunities than the Euclidean or elliptic ones, since the pseudo-orthogonal groups have a richer subgroup structure as compared to the orthogonal one of the same dimension. Consider now a hyperbolic hyperplane $\boldsymbol{B}^{n-1} \subset \boldsymbol{H}^n$. Its pole $\boldsymbol{b} = \boldsymbol{B}^\perp$ is an outer point. The group $G_{n-1} \subset G_n$ is the corresponding isotropy group in G_n, and at the same time it is the isotropy group of $\boldsymbol{B}$; again it consists of two components of $\boldsymbol{O}(1, n-1)$. Choosing $\boldsymbol{b} = [\mathfrak{e}_n]$, the vector space associated with $\boldsymbol{B}$ is $\boldsymbol{W}^n = [\mathfrak{e}_0, \dots, \mathfrak{e}_{n-1}]$. The line $\boldsymbol{b} \vee \boldsymbol{x} = [\mathfrak{x}, \mathfrak{e}_n]$, $\boldsymbol{x} \in \boldsymbol{B}$, is orthogonal to each of the lines $\boldsymbol{x} \vee \boldsymbol{y} = [\mathfrak{x}, \mathfrak{v}]$ through $\boldsymbol{x}$ lying in $\boldsymbol{B}$, which we may always consider as spanned by a unit vector $\mathfrak{v} \in \boldsymbol{W} = [\mathfrak{e}_n]^\perp$ orthogonal to $\mathfrak{x}$. Hence the family of hyperbolic lines $\{\boldsymbol{b} \vee \boldsymbol{x}\}_{\boldsymbol{x} \in \boldsymbol{B}}$ consists of the *normals* to the hyperplane $\boldsymbol{B}$. They are parallel, and all go through the same outer point $\boldsymbol{b}$. Any two of these normals have a common perpendicular, the connecting line from their point of intersection to $\boldsymbol{B}$. Obviously, $\boldsymbol{B}$ is the orbit of $\boldsymbol{x}_0 = [\mathfrak{e}_0]$ under the action of G_{n-1}. We may consider the orbit of the point

$$\boldsymbol{x}_r := [\mathfrak{e}_0 \cosh r + \mathfrak{e}_n \sinh r] \in \boldsymbol{b} \vee \boldsymbol{x}_0$$

as described by the parameter representation

$$\boldsymbol{y}_r(\boldsymbol{x}) := g[\mathfrak{e}_0 \cosh r + \mathfrak{e}_n \sinh r] = [\mathfrak{x} \cosh r + \mathfrak{e}_n \sinh r] \tag{55}$$

with $\mathfrak{x} \in \boldsymbol{W}$, $\langle \mathfrak{x}, \mathfrak{x} \rangle = -1$. Since G_{n-1} consists of isometries, the hyperbolic distance

$$h(\boldsymbol{x}, \boldsymbol{y}_r(\boldsymbol{x})) = h(g(\boldsymbol{x}_0), g(\boldsymbol{x}_r)) = r$$

is constant; hence the hypersurface described by (55) is called the *equidistant* $A(r, \boldsymbol{B})$ of the hyperplane $\boldsymbol{B}$ with distance r. Note that negative values are also allowed for r; they define the *opposite equidistants*.In Exercise 10 we already introduced the equidistants of a line in the hyperbolic plane and learnt already that they are no lines. Obviously, the equidistants of a hyperplane in hyperbolic geometry are no hyperplanes. We now want to show that the tangent hyperplane $T_{\boldsymbol{y}} A(r, \boldsymbol{B})$ of the equidistant at the point $\boldsymbol{y}$ is orthogonal to the line $\boldsymbol{b} \vee \boldsymbol{y}$. To see this we insert the parameter representation for the normalized representative of a point in the hyperplane $\boldsymbol{B}$

$$\mathfrak{x}(t, \mathfrak{x}_S) = \mathfrak{e}_0 \cosh t + \mathfrak{x}_S \sinh t \text{ with } \mathfrak{x}_S \in [\mathfrak{e}_1, \dots, \mathfrak{e}_{n-1}],\ \langle \mathfrak{x}_S, \mathfrak{x}_S \rangle = 1 \tag{56}$$

into representation (55): $\mathfrak{y}_r(t, \mathfrak{x}_S) = \mathfrak{x}(t, \mathfrak{x}_S) \cosh r + \mathfrak{e}_n \sinh r$. Then we compute the differential:

$$d\mathfrak{y}_r(t, \mathfrak{x}_S) = d\mathfrak{x}(t, \mathfrak{x}_S) \cosh r = ((\mathfrak{e}_0 \sinh t + \mathfrak{x}_S \cosh t) dt + \sinh t\, d\mathfrak{x}_S) \cosh r.$$

Since the vectors $\mathfrak{x}_S$ lie in the unit hypersphere of a Euclidean vector space, we have $\langle \mathfrak{x}_S, d\mathfrak{x}_S \rangle = 0$, and a straightforward computation leads to $\langle \mathfrak{y}_r(t, \mathfrak{x}_S), d\mathfrak{y}_r(t, \mathfrak{x}_S) \rangle = 0$, the asserted orthogonality. The arc element for the equidistant

$$\langle d\mathfrak{y}_r, d\mathfrak{y}_r \rangle = \langle d\mathfrak{x}, d\mathfrak{x} \rangle \cosh^2 r \tag{57}$$

differs from the arc element of the hyperbolic hyperplane only by the constant factor $\cosh^2 r$, so that its inner (metric) geometry itself is hyperbolic. As is easy to verify, the isotropy group of the point $\boldsymbol{x}_r \in A(r, \boldsymbol{B})$ for the action of G_{n-1} is conjugate to the orthogonal group $\boldsymbol{O}(n-1)$. We summarize what we proved concerning the equidistants in the following proposition:

Proposition 18. *The normals of a hyperbolic hyperplane $\boldsymbol{B}$ are parallel; they all go through the same outer point, the pole $\boldsymbol{b} = \boldsymbol{B}^{\perp}$ of the hyperplane. The orthogonal trajectories of the family of normals are the hyperplane itself and its equidistants. Each equidistant is an orbit for the isotropy group of the hyperplane; as a homogeneous space it is equivariant to the hyperbolic hyperplane:*

$$A(r, \boldsymbol{B}^{n-1}) \cong G_{n-1}/\boldsymbol{O}(n-1).$$

Moreover, its inner geometry is hyperbolic.

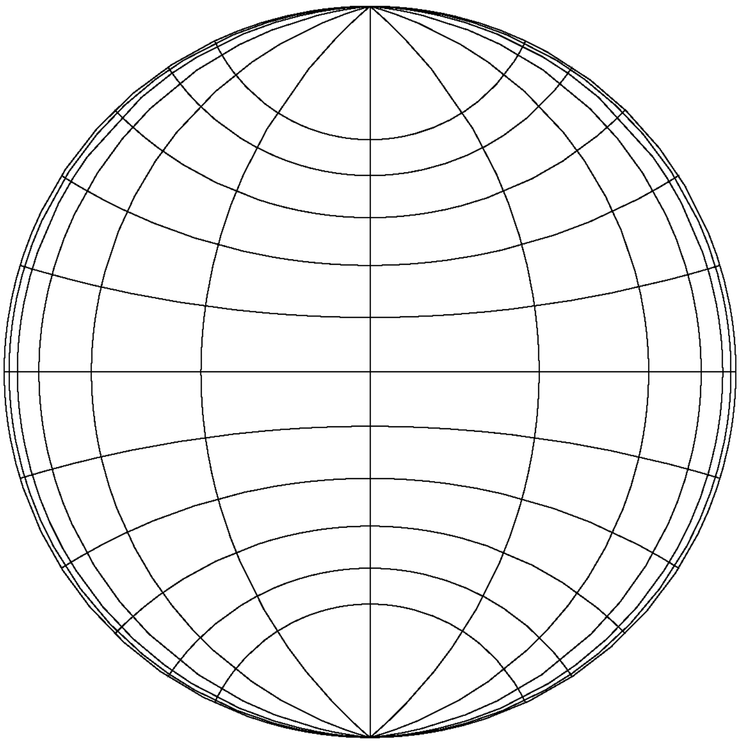


Fig. 2.28. Equidistants as orthogonal trajectories for the normals of a line in the conformal disk model (cf. Figure 2.27).

Exercise 24. Prove that the equidistants of arbitrary hyperbolic hyperplanes with the same distance parameter r are G_n-congruent.

Let now $\boldsymbol{x} \in Q$ be an arbitrary point of the hyperquadric Q; for transitivity reasons we may suppose $\boldsymbol{x} = [\mathfrak{e}_0 - \mathfrak{e}_n]$. In order to determine the isotropy group of this point, it is useful to introduce a different type of frame, which we will also need in later applications. For each pseudo-orthonormal basis $(\mathfrak{e}_i)$, $i = 0, \ldots, n$, we introduce a new basis by

$$\mathfrak{a}_0 := (\mathfrak{e}_0 - \mathfrak{e}_n)/\sqrt{2},\ \mathfrak{a}_i := \mathfrak{e}_i, i = 1, \ldots, n-1,\ \mathfrak{a}_n := (\mathfrak{e}_n + \mathfrak{e}_0)/\sqrt{2}; \tag{58}$$

it will be called *isotropic orthogonal.* A basis $(\mathfrak{a}_j)$ is isotropic orthogonal if and only if the matrix containing its scalar products has the form

$$(\langle \mathfrak{a}_j, \mathfrak{a}_k \rangle) = B := \begin{pmatrix} 0 & \mathfrak{o}' & -1 \\ \mathfrak{o} & I_{n-1} & \mathfrak{o} \\ -1 & \mathfrak{o}' & 0 \end{pmatrix}; \tag{59}$$

here I_{n-1} denotes the unit matrix and $\mathfrak{o}$ the zero column vector of order $n-1$, so $\mathfrak{o}'$ is the corresponding zero row vector. Conversely, setting

$$\mathfrak{e}_0 := (\mathfrak{a}_0 + \mathfrak{a}_n)/\sqrt{2},\ \mathfrak{e}_i := \mathfrak{a}_i, i = 1, \ldots, n-1,\ \mathfrak{e}_n := (\mathfrak{a}_n - \mathfrak{a}_0)/\sqrt{2}, \tag{60}$$

we assign a pseudo-orthonormal basis to every isotropic orthogonal one. The elements a in the isotropy group $H_{\boldsymbol{x}}$ of a point $\boldsymbol{x} = [\mathfrak{a}_0]$ have to satisfy the condition $a(\mathfrak{a}_0) = \mathfrak{a}_0 \lambda^{-1}$, $\lambda \in \mathbf{R}^*$. Starting with the following ansatz for the form of its block matrix

$$a = \begin{pmatrix} \lambda^{-1} & \mathfrak{a}' & c \\ \mathfrak{o} & A & \mathfrak{d} \\ 0 & \mathfrak{b}' & \mu, \end{pmatrix}$$

and taking into account that, because of $a \in G_n$, it preserves the scalar product, i.e.

$$a' B a = B,$$

a straightforward computation involving block matrices shows

Lemma 19. *With respect to an isotropic orthogonal basis* $(\mathfrak{a}_i)$ *the isotropy group* $H_{\boldsymbol{x}}$ *of the point* $\boldsymbol{x} = [\mathfrak{a}_0]$ *consists of all transformations in the group* G_n *which have a matrix of the following form in this basis:*

$$a(A, \mathfrak{a}, \lambda) := \begin{pmatrix} \lambda^{-1} & \lambda^{-1}\mathfrak{a}'A & \langle \mathfrak{a}, \mathfrak{a} \rangle / 2\lambda \\ \mathfrak{o} & A & \mathfrak{a} \\ 0 & \mathfrak{o}' & \lambda \end{pmatrix} \quad \textit{with } A \in \boldsymbol{O}(n-1), \mathfrak{a} \in \mathbf{R}^{n-1}, \lambda > 0. \tag{61}$$

It is easy to see that, in contrast to the isotropy group of the hyperbolic as well as the outer points, the isotropy group $H_{\boldsymbol{x}}$ acts transitively on hyperbolic space,

$$H_{\boldsymbol{x}}[\mathfrak{e}_0] = \boldsymbol{H}^n. \tag{62}$$

This can be verified by a computation, or else one could argue as follows: The one-parameter subgroup $\boldsymbol{O}(1,1) = \{a(E, \mathfrak{o}, \mathrm{e}^t)\}$ acts transitively on the hyperbolic line $[\mathfrak{e}_0, \mathfrak{e}_n]$:

$$a(E, \mathfrak{o}, \mathrm{e}^t)\mathfrak{e}_0 = a(E, \mathfrak{o}, \mathrm{e}^t)(\mathfrak{a}_0 + \mathfrak{a}_n)/\sqrt{2} = (\mathfrak{a}_0 \,\mathrm{e}^{-t} + \mathfrak{a}_n \,\mathrm{e}^t)/\sqrt{2}$$

$$= \mathfrak{e}_0 \cosh t + \mathfrak{e}_n \sinh t,$$

and the isotropy group $H_{\boldsymbol{x}}$ acts transitively on the subpencil of all hyperbolic lines through $\boldsymbol{x}$, so that $[\mathfrak{e}_0]$ can be transformed into any other point $\boldsymbol{y} \in \boldsymbol{H}^n$. In order to obtain interesting orbits, we consider the subgroup

$$E(n-1) := \{a(A, \mathfrak{a}, 1) |\, A \in \boldsymbol{O}(n-1), \mathfrak{a} \in \mathbf{R}^{n-1}\}.$$

A calculation using block matrices of the form (61) shows that

$$a(A, \mathfrak{a}, \lambda)a(B, \mathfrak{b}, \mu) = a(AB, A\mathfrak{b} + \mathfrak{a}\mu, \lambda\mu). \tag{63}$$

This implies that $E(n-1)$ is a normal subgroup in $H_{\boldsymbol{x}}$, since it is the kernel of the homomorphism $a(A, \mathfrak{a}, \lambda) \in H_{\boldsymbol{x}} \mapsto \lambda \in \mathbf{R}^*$,. Comparing this to the representation (I.6.2.42) of the Euclidean groups immediately shows that $E(n-1)$ is isomorphic to the Euclidean group of the $(n-1)$-dimensional Euclidean space. Formulas (58), (60), (61) applied to elements $\boldsymbol{z} = E(n-1)\boldsymbol{y}, \boldsymbol{y} := [\mathfrak{e}_0]$ of the orbit yield

$$\boldsymbol{z} = a(A, \mathfrak{a}, 1)\boldsymbol{y} = [\mathfrak{e}_0(1 + \langle \mathfrak{a}, \mathfrak{a}\rangle/4) + \mathfrak{a}/\sqrt{2} - \mathfrak{e}_n\langle \mathfrak{a}, \mathfrak{a}\rangle/4],\ \mathfrak{a} \in \mathbf{R}^{n-1}. \tag{64}$$

Hence $\boldsymbol{z} = \boldsymbol{y}$ holds if and only if $\mathfrak{a} = \mathfrak{o}$; thus the isotropy group of $\boldsymbol{y}$ under the action of $E(n-1)$ is the orthogonal group $\boldsymbol{O}(n-1)$. This orbit $E(n-1)\boldsymbol{y}$ is called a *horosphere* or *limit sphere*; *horocycle* or *limit circle* in the case $n = 2$. Since G_n is transitive on Q, we may choose an arbitrary point from Q as $\boldsymbol{x}$ and define the restricted isotropy group $H_{1,\boldsymbol{x}} \subset H_{\boldsymbol{x}}$ isomorphic to $E(n-1)$ as above by $\lambda = 1$. Since the subpencil of hyperbolic lines through the arbitrary point $\boldsymbol{x} \in Q$ simply covers the hyperbolic space $\boldsymbol{H}^n$, for every point pair $\boldsymbol{x} \in Q, \boldsymbol{y} \in \boldsymbol{H}^n$ there is a horosphere $H_{1,\boldsymbol{x}}\boldsymbol{y}$. From the considerations so far we conclude

Proposition 20. *All horospheres in hyperbolic space are G_n-congruent. If $H_{1,\boldsymbol{x}}\boldsymbol{y}$ is a horosphere, then choosing $\mathfrak{e}_0, \mathfrak{e}_n$ with $\boldsymbol{y} = [\mathfrak{e}_0]$, $\boldsymbol{x} = [\mathfrak{e}_0 - \mathfrak{e}_n]$ Equation* (64) *becomes a parameter representation for this hypersurface. The horospheres are equivariant to the $(n-1)$-dimensional Euclidean space $\boldsymbol{E}^{n-1} \cong E(n-1)/\boldsymbol{O}(n-1)$.* □

It is remarkable that, as a subgroup in the isometry group of hyperbolic space, the Euclidean group and its orbits appear as Euclidean spaces of codimension 1: This way, Euclidean geometry is embedded into non-Euclidean geometry. This also holds for the (inner) Riemannian metric generated by the

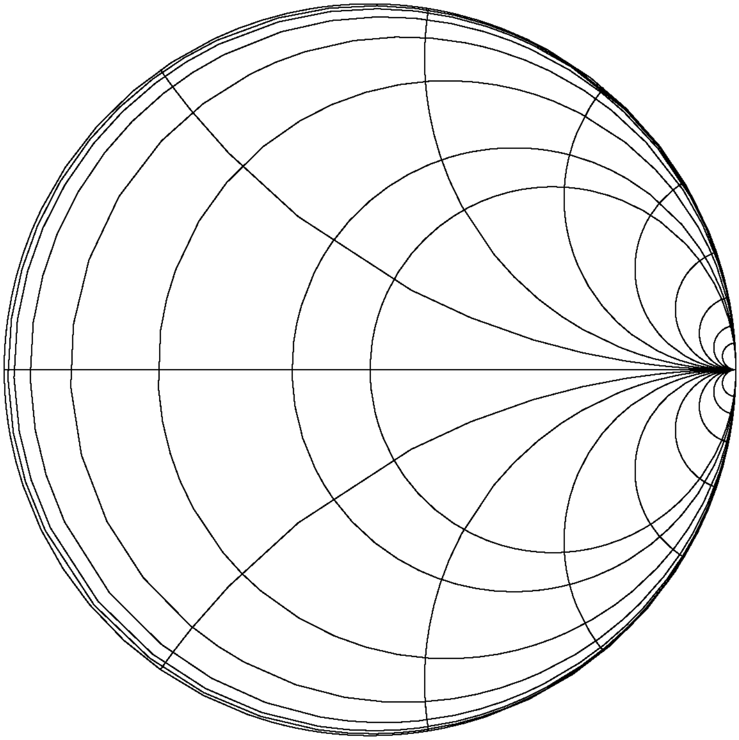


Fig. 2.29. Horocycles in the conformal disk model.

hyperbolic metric on the orbits. Taking derivatives of the normalized representative $\mathfrak{z}$ for $\boldsymbol{z}$ with respect to the components a^i of

$$\mathfrak{a} = \sum_{i=1}^{n-1} \mathfrak{e}_i a^i$$

(64) implies the equations

$$\partial\mathfrak{z}/\partial a^i = (\mathfrak{e}_0 - \mathfrak{e}_n)a^i/2 + \mathfrak{e}_i/\sqrt{2}, \tag{65}$$

$$\langle \partial\mathfrak{z}/\partial a^i, \partial\mathfrak{z}/\partial a^j \rangle = \delta_{ij}/2.$$

This proves the assertion. Moreover, from (65) we see

$$\langle \mathfrak{z}, \partial\mathfrak{z}/\partial a^i \rangle = 0 \text{ and } \langle \mathfrak{e}_0 - \mathfrak{e}_n, \partial\mathfrak{z}/\partial a^i \rangle = 0.$$

This implies (cf. Figure 2.29)

Corollary 21. *The horospheres $H_{1,\boldsymbol{xy}}$ are the orthogonal trajectories for the pencil of boundary parallel hyperbolic lines through the point at infinity $\boldsymbol{x} \in Q$.* □

2.6.7 Stationary Angles

Let $\boldsymbol{A}^l, \boldsymbol{B}^m \subset \boldsymbol{H}^n$ be an l-plane and an m-plane in the n-dimensional hyperbolic space, respectively. We want to determine a complete system of invariants for pairs like $(\boldsymbol{A}^l, \boldsymbol{B}^m)$ with $0 \leq l \leq m < n$. The case of points, $l = m = 0$, is the topic of Proposition 4, and hyperplanes, $l = m = n-1$, were discussed in Example 4. In order not to have to distinguish too many cases we introduce:

Condition A. *There is no hyperplane* $\boldsymbol{L}^{n-1} \subset \boldsymbol{H}^n$ *containing* $\boldsymbol{A}^l \cup \boldsymbol{B}^m$.

In other words, we suppose that the set $\boldsymbol{A}^l \cup \boldsymbol{B}^m$ generates the whole space. Now we consider all pairs of hyperplanes $\boldsymbol{L}^{n-1}, \boldsymbol{M}^{n-1} \subset \boldsymbol{H}^n$ with $\boldsymbol{A}^l \subset \boldsymbol{L}^{n-1}$, $\boldsymbol{B}^m \subset \boldsymbol{M}^{n-1}$ and determine the relative extrema of the invariant $I(\boldsymbol{L}, \boldsymbol{M})$ leading us to the notion of *stationary angles*. Similar to the Euclidean or elliptic geometries they form the complete system of invariants we are looking for.

Let $\boldsymbol{U}^{l+1}, \boldsymbol{W}^{m+1}$ be the pseudo-Euclidean subspaces associated with $\boldsymbol{A}^l$ and $\boldsymbol{B}^m$, respectively. The oriented hyperplanes containing $\boldsymbol{A}^l$ are determined by the unit vectors in the orthogonal complement $\boldsymbol{U}^\perp$, and analogously for the hyperplanes containing $\boldsymbol{B}^m$:

$$\mathfrak{u} \in \boldsymbol{U}^\perp, \langle \mathfrak{u}, \mathfrak{u} \rangle = 1, \ \mathfrak{w} \in \boldsymbol{W}^\perp, \langle \mathfrak{w}, \mathfrak{w} \rangle = 1. \tag{66}$$

Hence, according to (23) we have to find the relative extrema of the scalar product $\langle \mathfrak{u}, \mathfrak{w} \rangle$ under the constraint (66). Since the arguments $\mathfrak{u}, \mathfrak{w}$ vary on unit spheres in Euclidean vector spaces, the existence of minima and maxima is guaranteed. Consider the function

$$f(\mathfrak{u}, \mathfrak{w}, \lambda, \mu) = \langle \mathfrak{u}, \mathfrak{w} \rangle - \lambda(\langle \mathfrak{u}, \mathfrak{u} \rangle - 1) - \mu(\langle \mathfrak{w}, \mathfrak{w} \rangle - 1), \ \mathfrak{u} \in \boldsymbol{U}^\perp, \ \mathfrak{w} \in \boldsymbol{W}^\perp \tag{67}$$

with the Lagrange multipliers λ, μ. Differentiating with respect to $\mathfrak{u}, \mathfrak{w}$ we obtain the necessary conditions for the relative extrema:

$$\begin{aligned} \langle d\mathfrak{u}, \mathfrak{w} \rangle - 2\lambda \langle d\mathfrak{u}, \mathfrak{u} \rangle &= \langle d\mathfrak{u}, \mathfrak{w} - 2\mathfrak{u}\lambda \rangle = 0, \\ \langle d\mathfrak{w}, \mathfrak{u} \rangle - 2\mu \langle d\mathfrak{w}, \mathfrak{w} \rangle &= \langle d\mathfrak{w}, \mathfrak{u} - 2\mathfrak{w}\mu \rangle = 0. \end{aligned}$$

Since $d\mathfrak{u} \in \boldsymbol{U}^\perp, d\mathfrak{w} \in \boldsymbol{W}^\perp$ can vary freely, these conditions are satisfied if and only if

$$\mathfrak{w} - 2\mathfrak{u}\lambda \in \boldsymbol{U} \text{ and } \mathfrak{u} - 2\mathfrak{w}\mu \in \boldsymbol{W}. \tag{68}$$

Let $pr_{\boldsymbol{U}}, pr_{\boldsymbol{U}^\perp}$ denote the projections resulting from the direct decomposition

$$\boldsymbol{V} = \boldsymbol{U} \oplus \boldsymbol{U}^\perp,$$

then (68) is equivalent to

$$pr_{\boldsymbol{U}^\perp} \mathfrak{w} = 2\mathfrak{u}\lambda \text{ and } pr_{\boldsymbol{W}^\perp} \mathfrak{u} = 2\mathfrak{w}\mu. \tag{69}$$

Define the operator A as follows:

$$p := pr_{\boldsymbol{U}^\perp}|\boldsymbol{W}^\perp,\ q := pr_{\boldsymbol{W}^\perp}|\boldsymbol{U}^\perp,\ A := q \circ p \in \mathrm{End}(\boldsymbol{W}^\perp). \tag{70}$$

Then (69) implies: *If* $\mathfrak{u}, \mathfrak{w}$ *is a solution of the extremal problem for the function* (67)*, then* $\mathfrak{w} \in \boldsymbol{W}^\perp$ *is an eigenvector for the operator* A*:*

$$A\mathfrak{w} = \mathfrak{w}4\lambda\mu. \tag{71}$$

This completes the analytic part of the considerations, which essentially serves to motivate the definition of the operator A; obviously, this operator is assigned to the pair $\boldsymbol{A}^l, \boldsymbol{B}^m$ in an invariant, more precisely, equivariant way. Since A is an endomorphism of the Euclidean vector space $\boldsymbol{W}^\perp$, its diagonalizability immediately follows from

Lemma 22. *The operator* A *defined by* (70) *is self-adjoint, more precisely,*

$$p^* = q,\ q^* = p,\ A^* = A. \tag{72}$$

The eigenvalues of A *are non-negative. The eigensubspace of* A *for the eigenvalue* 0 *is the kernel of* p*.*

Proof. The formulas (72) immediately follow from the definition of the adjoint map together with the relation

$$\langle \mathfrak{u}, \mathfrak{w}\rangle = \langle q\mathfrak{u}, \mathfrak{w}\rangle = \langle \mathfrak{u}, p\mathfrak{w}\rangle$$

holding for all $\mathfrak{u} \in \boldsymbol{U}^\perp, \mathfrak{w} \in \boldsymbol{W}^\perp$. So A is a self-adjoint operator in a Euclidean vector space, hence diagonalizable with real eigenvalues. If $\mathfrak{w}$ is an eigenvector for the eigenvalue α, then by (72)

$$\langle \mathfrak{w}, A\mathfrak{w}\rangle = \langle \mathfrak{w}, \mathfrak{w}\rangle\alpha = \langle p\mathfrak{w}, p\mathfrak{w}\rangle,$$

immediately implying the last two assertions. □

Hence by $\dim \boldsymbol{W}^\perp = n - m$ there exist $n - m$ not necessarily different eigenvalues α_i for A, which we label according to their magnitude:

$$\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_{n-m} \geq 0. \tag{73}$$

Let r_A be the rank of the operator A, i.e. (if $r_a < n - m$)

$$\alpha_{r_A+1} = \ldots = \alpha_{n-m} = 0.$$

We fix an orthonormal eigenbasis $(\mathfrak{a}_i)$ for A with $A\mathfrak{a}_i = \mathfrak{a}_i\alpha_i$. Then (72) implies

$$\langle \mathfrak{a}_i, A\mathfrak{a}_j\rangle = \langle p\,\mathfrak{a}_i, p\,\mathfrak{a}_j\rangle\alpha_j = \delta_{ij}\alpha_j.$$

Hence

$$\mathfrak{b}_i := p\,\mathfrak{a}_i/\sqrt{\alpha_i} \text{ for } i = 1, \ldots, r_A \tag{74}$$

are r_A orthonormal vectors from $\boldsymbol{U}^\perp$, which we complete to an orthonormal basis $(\mathfrak{b}_j), j = 1, \ldots, n-l$, for $\boldsymbol{U}^\perp$. Then

$$\langle \mathfrak{b}_i, \mathfrak{a}_j \rangle = \delta_{ij}\sqrt{\alpha_j} \text{ holds for } i = 1, \ldots, n-l,\ j = 1, \ldots, n-m. \tag{75}$$

In fact, because of $\mathfrak{b}_i \in \boldsymbol{U}^\perp$ we have

$$\langle \mathfrak{b}_i, \mathfrak{a}_j \rangle = \langle \mathfrak{b}_i, p\,\mathfrak{a}_j \rangle = \delta_{ij}\sqrt{\alpha_i}$$

for $i = 1, \ldots, r_A$ by the definition of the $\mathfrak{b}_i$. Since the remaining vectors $\mathfrak{b}_\rho$ belong to the image of p and are obviously also orthogonal to $\boldsymbol{U}$, in addition,

$$\langle \mathfrak{b}_\rho, \mathfrak{a}_j \rangle = \langle \mathfrak{b}_\rho, p\,\mathfrak{a}_j \rangle = 0 = \delta_{\rho,j}\sqrt{\alpha_j}$$

for $\rho = r_A + 1, \ldots, n-m$. From (75) and the following rule generally holding for adjoint maps,

$$(\operatorname{Im} p)^\perp = \operatorname{Ker} p^*,$$

it is easy to deduce

$$\operatorname{Im} p = [\mathfrak{b}_1, \ldots, \mathfrak{b}_{r_A}],\ \operatorname{Ker} p = [\mathfrak{a}_{r_A+1}, \ldots, \mathfrak{a}_{n-m}] = \boldsymbol{U} \cap \boldsymbol{W}^\perp \tag{76}$$

$$\operatorname{Im} q = [\mathfrak{a}_1, \ldots, \mathfrak{a}_{r_A}],\ \operatorname{Ker} q = [\mathfrak{b}_{r_A+1}, \ldots, \mathfrak{b}_{n-l}] = \boldsymbol{W} \cap \boldsymbol{U}^\perp, \tag{77}$$

and

$$q(\mathfrak{b}_i) = \mathfrak{a}_i\sqrt{\alpha_i} \text{ for } i = 1, \ldots, n-m. \tag{78}$$

The qualitative properties of the pair $\boldsymbol{A}^l, \boldsymbol{B}^m$ are reflected in the magnitude of α_1. The extremal properties of the eigenvalues, cf. (73) and (75), imply

$$\sqrt{\alpha_1} = \max\{\langle \mathfrak{u}, \mathfrak{w} \rangle | \mathfrak{u} \in \boldsymbol{U}^\perp, \langle \mathfrak{u}, \mathfrak{u} \rangle = 1, \mathfrak{w} \in \boldsymbol{W}^\perp, \langle \mathfrak{w}, \mathfrak{w} \rangle = 1\}. \tag{79}$$

Exercise 25. Prove (81) by a direct estimate.

Lemma 23. *If Condition* A *is satisfied, then* α_1 *distinguishes the following cases:*

1. $\alpha_1 > 1 \Leftrightarrow \boldsymbol{A}^l \cap \boldsymbol{B}^m$ *is a projective subspace lying in the outer region* $A(Q)$ *or the nopoint;*
2. $\alpha_1 = 1 \Leftrightarrow \boldsymbol{A}^l \cap \boldsymbol{B}^m$ *is a projective subspace tangent to* Q*;*
3. $\alpha_1 < 1 \Leftrightarrow \boldsymbol{A}^l \cap \boldsymbol{B}^m$ *is a hyperbolic projective subspace and different from the nopoint.*

Proof. Let $\alpha_1 > 1$. Then by (75) the linear span $[\mathfrak{a}_1, \mathfrak{b}_1]$ is a pseudo-Euclidean subspace of $\boldsymbol{U}^\perp + \boldsymbol{W}^\perp$, and hence $\boldsymbol{U}^\perp + \boldsymbol{W}^\perp$ is also pseudo-Euclidean. Consequently, the intersection $\boldsymbol{U} \cap \boldsymbol{W} = (\boldsymbol{U}^\perp + \boldsymbol{W}^\perp)^\perp$ is Euclidean or equal to $\{\mathfrak{o}\}$, hence the associated projective subspace $\boldsymbol{A}^l \cap \boldsymbol{B}^m$ lies in $A(Q)$, or it only consists of the nopoint. Conversely, if this is the case, then there is a time-like vector $\mathfrak{x} = \mathfrak{u} + \mathfrak{w}, \mathfrak{u} \in \boldsymbol{U}^\perp, \mathfrak{w} \in \boldsymbol{W}^\perp$. Let $\mathfrak{u}_0, \mathfrak{w}_0$ be the corresponding normalized vectors. Since the linear span $[\mathfrak{u}, \mathfrak{w}] = [\mathfrak{u}_0, \mathfrak{w}_0]$ is pseudo-Euclidean, the determinant of the matrix of scalar products $1 - \langle \mathfrak{u}_0, \mathfrak{w}_0 \rangle^2$ has to be negative. The fact that α_1 is maximal implies

$$1 < \langle \mathfrak{u}_0, \mathfrak{w}_0 \rangle^2 \leq \alpha_1$$

proving Case 1. Let $\alpha_1 = 1$. Then by (75) we have $\langle \mathfrak{a}_1, \mathfrak{b}_1 \rangle = 1$. Thus the vector $\mathfrak{z} := \mathfrak{a}_1 - \mathfrak{b}_1 \in \boldsymbol{U}^\perp + \boldsymbol{W}^\perp$ is isotropic. Actually, if $\mathfrak{z} = \mathfrak{o}$, then $\mathfrak{a}_1 = \mathfrak{b}_1 \in \boldsymbol{U}^\perp \cap \boldsymbol{W}^\perp$, in contradiction to Condition A being equivalent to either of the statements

$$\boldsymbol{U} + \boldsymbol{W} = \boldsymbol{V},\ \boldsymbol{U}^\perp \cap \boldsymbol{W}^\perp = \{\mathfrak{o}\}. \tag{80}$$

Moreover, (75) implies that $\mathfrak{z}$ is orthogonal to each of the generating vectors $\mathfrak{a}_i$, $\mathfrak{b}_j$ of $\boldsymbol{U}^\perp + \boldsymbol{W}^\perp$; hence this subspace is isotropic. The same then also holds for its orthogonal complement $\boldsymbol{U} \cap \boldsymbol{W}$, and thus $\boldsymbol{A}^l \cap \boldsymbol{B}^m$ is tangent to Q, cf. Corollary 1.9.19. Conversely, if this is the case, then $\boldsymbol{U}^\perp + \boldsymbol{W}^\perp$ has to be isotropic. Thus we find an isotropic vector $\mathfrak{y} = \mathfrak{u} + \mathfrak{w} \in \boldsymbol{U}^\perp + \boldsymbol{W}^\perp$, $\mathfrak{u} \in \boldsymbol{U}^\perp$, $\mathfrak{w} \in \boldsymbol{W}^\perp$, orthogonal to $\boldsymbol{U}^\perp + \boldsymbol{W}^\perp$. In particular,

$$\langle \mathfrak{y}, \mathfrak{u} \rangle = \langle \mathfrak{u}, \mathfrak{u} \rangle + \langle \mathfrak{w}, \mathfrak{u} \rangle = 0,\ \langle \mathfrak{y}, \mathfrak{w} \rangle = \langle \mathfrak{u}, \mathfrak{w} \rangle + \langle \mathfrak{w}, \mathfrak{w} \rangle = 0,$$

implying

$$\frac{\langle \mathfrak{u}, \mathfrak{w} \rangle^2}{\langle \mathfrak{u}, \mathfrak{u} \rangle \langle \mathfrak{w}, \mathfrak{w} \rangle} = 1.$$

Since α_1 is maximal, the relation $\alpha_1 \geq 1$ has to hold. By what we proved in Case 1, $\alpha_1 > 1$ implies that the space $\boldsymbol{U}^\perp + \boldsymbol{W}^\perp$ is pseudo-Euclidean. Hence $\alpha_1 = 1$. This proves Assertion 2. Since both sides of the equivalences stated in the lemma are complete disjunctions, the third assertion follow from the first two; thus the lemma is proved. □

The result of this lemma gives rise to the following definition:

Definition 3. Two hyperbolic subspaces $\boldsymbol{A}^l, \boldsymbol{B}^m$ are called *parallel* if the section $\boldsymbol{A}^l \wedge \boldsymbol{B}^m$ of the associated projective subspaces lies in the outer region $A(Q)$ and, moreover, does not only consist of the nopoint; they are called *boundary parallel* if this section is tangent, and *intersecting* if it is a hyperbolic subspace. Finally, recall that two projective subspaces $\boldsymbol{A}^l, \boldsymbol{B}^m \subset \boldsymbol{P}^n$ are *skew* if their intersection is the nopoint. □

Obviously, the eigenvalues $\alpha_i, i = 1, \ldots, n - m$, are invariants of the pair $\boldsymbol{A}^l, \boldsymbol{B}^m$, since the operator A is equivariantly assigned to this pair. In order to prove that this system of invariants is complete, we construct a basis for $\boldsymbol{V}$ such that the following holds:

1. The basis $(\mathfrak{b}_j), j = 0, \ldots, n$, is pseudo-orthonormal.
2. $(\mathfrak{b}_j), j = 1, \ldots, n - l$, is a basis for $\boldsymbol{U}^\perp$.
3. For the eigenvalues α_i and eigenvectors $\mathfrak{a}_i$ of A the relations (73), (74), and (75) hold.
4. The coordinate representation of the eigenvectors in the basis $(\mathfrak{b}_j)$ is canonical in the sense that the coordinates of $\mathfrak{a}_i$ are uniquely determined by the eigenvalue α_i.

Such a basis will be called *adapted* to the pair $\boldsymbol{A}^l, \boldsymbol{B}^m$. Proving the existence of adapted bases suffices to prove the completeness for the system of invariants: If $\boldsymbol{A}'^l, \boldsymbol{B}'^m$ is a pair of hyperbolic subspaces with coinciding invariants (73), and $(\mathfrak{b}'_j)$ denotes a basis adapted to this pair, then the pseudo-orthogonal transformation of $\boldsymbol{V}$ determined by $g(\mathfrak{b}_i) = \mathfrak{b}'_i, i = 1, \ldots, n+1$, generates an equally denoted hyperbolic isometry such that $g\boldsymbol{A}^l = \boldsymbol{A}'^l$, $g\boldsymbol{B}^m = \boldsymbol{B}'^m$.

Before we start the construction of an adapted basis, we introduce the following notation: Let $\boldsymbol{M}(\mathfrak{m}), \langle \mathfrak{m}, \mathfrak{m} \rangle > 0$, be the hyperbolic hyperplane corresponding to the pole $[\mathfrak{m}] \in A(Q)$. The quantities (cf. (75))

$$\sqrt{\alpha_i} = \langle \mathfrak{a}_i, \mathfrak{b}_i \rangle = |I|(\boldsymbol{M}(\mathfrak{a}_i), \boldsymbol{M}(\mathfrak{b}_i))$$

are called the *stationary invariants* of the pair $\boldsymbol{A}^l, \boldsymbol{B}^m$. If $\alpha_1 > 1$, then we set

$$\sqrt{\alpha_1} = \cosh \beta_1 \tag{81}$$

and call β_1 the *maximal distance* of hyperplanes containing $\boldsymbol{A}^l$ and $\boldsymbol{B}^l$ (cf. (36)). Eventually, if $\alpha_i \leq 1$, then we set

$$\sqrt{\alpha_i} = \cos \beta_i, \ (0 \leq \beta_i \leq \pi/2) \tag{82}$$

and call β_i the *stationary angles* of the pair $\boldsymbol{A}^l, \boldsymbol{B}^m$ (cf. Example 4).

We already constructed the vectors $\mathfrak{b}_1, \ldots, \mathfrak{b}_{n-l} \in \boldsymbol{U}^\perp$ of the adapted basis; they satisfy the Conditions 2. and 3. In order to adapt the remaining vectors, i.e. a pseudo-orthonormal basis for $\boldsymbol{U}^{l+1}$, $(\mathfrak{b}_0, \mathfrak{b}_j), j = n-l+1, \ldots, n$, to the situation, consider the projections of the eigenvectors (recall (74))

$$pr_{\boldsymbol{U}} \mathfrak{a}_i = \mathfrak{a}_i - \mathfrak{b}_i \sqrt{\alpha_i}, \ i = 1, \ldots, n-m,$$

onto the space $\boldsymbol{U}$. Relations (74) and (75) imply

$$\langle pr_{\boldsymbol{U}} \mathfrak{a}_i, pr_{\boldsymbol{U}} \mathfrak{a}_j \rangle = \delta_{ij}(1 - \alpha_j) \text{ for } i, j = 1, \ldots, n-m. \tag{83}$$

Hence these projections are orthogonal; we normalize those different from $\mathfrak{o}$ in accordance with the cases considered in Lemma 23 and complete them orthogonally to form the adapted basis we were looking for.

1. Case. *If* $\alpha_1 > 1$, *then this eigenvalue has multiplicity 1, and all remaining eigenvalues are less than one,* $\alpha_i < 1, i = 2, n-m$. *There is an adapted basis* $\mathfrak{b}_i$ *for* $\boldsymbol{V}$ *with*

$$\mathfrak{a}_1 = \mathfrak{b}_1 \cosh \beta_1 + \mathfrak{b}_0 \sinh \beta_1, \tag{84}$$

$$\mathfrak{a}_i = \mathfrak{b}_i \cos \beta_i + \mathfrak{b}_{n-l+i-1} \sin \beta_i, \ i = 2, \ldots, n-m. \tag{85}$$

Proof. According to (83) the projection $pr_{\boldsymbol{U}} \mathfrak{a}_1$ is time-like; normalization gives rise to the basis vector $\mathfrak{b}_0$ and shows the validity of (84). Since $\boldsymbol{V}$ does not contain two orthogonal time-like vectors, $\alpha_2 \leq 1$. If now $\alpha_2 = 1$, then (83) implies that the projection $pr_{\boldsymbol{U}} \mathfrak{a}_2$ had to be isotropic and orthogonal to $\mathfrak{b}_0$, or else the null vector. Since the subspace orthogonal to $\mathfrak{b}_0$ is Euclidean, this projection then has to be the null vector. But this again means $\mathfrak{a}_2 \in \boldsymbol{U}^\perp \cap \boldsymbol{W}^\perp$ contradicting Condition A, cf. (80). Normalizing the projections $pr_{\boldsymbol{U}}\mathfrak{a}_i, i \geq 2$, leads to orthonormal vectors $\mathfrak{b}_{n-l+i-1},\ i \geq 2$, where the largest index is

$$2n - (m + l + 1) = n - d,\ d := \dim \boldsymbol{U} \cap \boldsymbol{W}.$$

For $d > 0$ we complete the set of vectors $\mathfrak{b}_i$ defined up to now by choosing d more vectors, so that all together form a pseudo-orthonormal basis for $\boldsymbol{V}$ satisfying conditions 1–4.

2. Case. *If $\alpha_1 = 1$, then this eigenvalue has multiplicity 1. There is an adapted basis $\mathfrak{b}_i$ for $\boldsymbol{V}$ such that*

$$\mathfrak{a}_1 = \mathfrak{b}_1 + \mathfrak{b}_0 + \mathfrak{b}_{n-l+1}, \tag{86}$$

$$\mathfrak{a}_i = \mathfrak{b}_i \cos\beta_i + \mathfrak{b}_{n-l+i} \sin\beta_i,\ i = 2, \ldots, n - m. \tag{87}$$

Proof. By (83), in this case $pr_{\boldsymbol{U}}(\boldsymbol{W}^\perp)$ is isotropic. If, in addition, $\alpha_2 = 1$, then we have two isotropic vectors $\mathfrak{a}_1 - \mathfrak{b}_1$, $\mathfrak{a}_2 - \mathfrak{b}_2$, in this projection. They are linearly dependent, since $\boldsymbol{V}$ is pseudo-Euclidean of index 1, cf. Definition 2.2.3 and the remark following it. This implies

$$\mathfrak{c} = \mathfrak{a}_1\lambda + \mathfrak{a}_2\mu = \mathfrak{b}_1\lambda + \mathfrak{b}_2\mu$$

with $\lambda\mu \neq 0$, i.e. $\mathfrak{o} \neq \mathfrak{c} \in \boldsymbol{U}^\perp \cap \boldsymbol{W}^\perp$. But this contradicts (80). According to (83) we determine the vectors $\mathfrak{b}_i$ by normalizing the projections of $\mathfrak{a}_i, i \geq 2$, so that (87) holds. The orthogonal complement of the linear span $[\mathfrak{b}_{n-l+2}, \ldots, \mathfrak{b}_{2n-l-m}]$ in $\boldsymbol{U}$ is pseudo-Euclidean. Hence we may represent the isotropic vector $\mathfrak{a}_1 - \mathfrak{b}_1 \in \boldsymbol{U}$ by choosing $\mathfrak{b}_0, \mathfrak{b}_{n-l+1} \in \boldsymbol{U}$ so that (86) holds.

3. Case. *If $\alpha_1 < 1$, then there is an adapted basis $\mathfrak{b}_i$ for $\boldsymbol{V}$ with*

$$\mathfrak{a}_i = \mathfrak{b}_i \cos\beta_i + \mathfrak{b}_{n-l+i} \sin\beta_i,\ i = 1, \ldots, n - m. \tag{88}$$

Proof. The proof follows immediately from (83) by normalizing the projections of the vectors $\mathfrak{a}_i$ onto $\boldsymbol{U}$.

Finally, consider the case that Condition A is not satisfied. Let $k + 1 = \dim \boldsymbol{U} + \boldsymbol{W}$, and let $k < n$ be the dimension of the hyperbolic subspace $\boldsymbol{A}^l \vee \boldsymbol{B}^m$. Then we have

$$df(\boldsymbol{A}^l, \boldsymbol{B}^m) := \dim \boldsymbol{U}^\perp \cap \boldsymbol{W}^\perp = n - k > 0.$$

We call $d = df(\boldsymbol{A}^l, \boldsymbol{B}^m)$ the *deficit* of the subspaces $\boldsymbol{A}^l, \boldsymbol{B}^m$. Obviously, the deficit is an invariant of these subspaces; for the pair $\boldsymbol{A}^l, \boldsymbol{B}^m$, considered as subspaces of $\boldsymbol{A}^l \vee \boldsymbol{B}^m$, Condition A is satisfied, and we obtain $k - m = n - m - df(\boldsymbol{A}^l, \boldsymbol{B}^m)$ eigenvalues as essential invariants of this pair. We remark that all vectors from $\boldsymbol{U}^\perp \cap \boldsymbol{W}^\perp$ are eigenvectors of A for the eigenvalue 1, which always occurs for $d > 0$. Summarizing we formulate:

Proposition 24. *If Condition A is satisfied, then for each pair of hyperbolic subspaces $\boldsymbol{A}^l, \boldsymbol{B}^m \subset \boldsymbol{H}^n$, $0 \leq l \leq m < n$, there is an adapted basis for $\boldsymbol{V}$ with the properties stated in the Cases 1–3. In general, two pairs with the same eigenvalues α_i and equal deficit are congruent; these quantities form a complete system of invariants for the action of the isometry group G_n on the product set $H_{n,l} \times H_{n,m}$ of these pairs. If $\mathfrak{a}_i$ is an eigenvector for an eigenvalue $\alpha_i > 0$, then the hyperplanes $\boldsymbol{M}(\mathfrak{a}_i) \supset \boldsymbol{B}^m$, $\boldsymbol{M}(p(\mathfrak{a}_i)) \supset \boldsymbol{A}^l$ realize the stationary invariant*

$$|I|(\boldsymbol{M}(\mathfrak{a}_i), \boldsymbol{M}(p(\mathfrak{a}_i))) = \sqrt{\alpha_i}, \quad (\alpha_i > 0).$$

Each eigenvector $\mathfrak{w}$ for the eigenvalue $\alpha_\kappa = 0$ belongs to $\boldsymbol{U}$; the hyperplane $\boldsymbol{M}(\mathfrak{w})$ is orthogonal to all hyperplanes $\hat{\boldsymbol{M}}$ containing $\boldsymbol{A}^l$. [7] □

As an application of this classification we now want to prove a generalization of Proposition 11:

Proposition 25. *Let $\boldsymbol{A}^l, \boldsymbol{B}^m \subset \boldsymbol{H}^n$, $0 \leq l \leq m < n$, be two subspaces of hyperbolic space. There exists a common perpendicular $\boldsymbol{x}_0 \vee \boldsymbol{y}_0$, $\boldsymbol{x}_0 \in \boldsymbol{A}^l \setminus \boldsymbol{B}^m, \boldsymbol{y}_0 \in \boldsymbol{B}^m \setminus \boldsymbol{A}^l$, for $\boldsymbol{A}^l, \boldsymbol{B}^m$ if and only if the largest invariant satisfies $\alpha_1 > 1$, i.e. if the subspaces are skew or parallel. In this case, the hyperbolic distance satisfies*

$$h(\boldsymbol{x}_0, \boldsymbol{y}_0) = \min\{h(\boldsymbol{a}, \boldsymbol{b}) |\ \boldsymbol{a} \in \boldsymbol{A}^l, \boldsymbol{b} \in \boldsymbol{B}^m\} = \beta_1 \textit{ with } \alpha_1 = \cosh \beta_1.$$

Here orthogonality is meant as formulated in Proposition 11 for the perpendicular. The points $\boldsymbol{x}_0, \boldsymbol{y}_0$ for which the minimal distance is attained are unique.

P r o o f. Considering only the subspace $\boldsymbol{A}^l \vee \boldsymbol{B}^m$ Condition A is satisfied. If the subspaces intersect, and if $\boldsymbol{z} \in \boldsymbol{A}^l \cap \boldsymbol{B}^m$ is a hyperbolic point, then, supposing the existence of a common perpendicular, the triple $(\boldsymbol{x}_0, \boldsymbol{y}_0, \boldsymbol{z})$ forms a hyperbolic triangle whose angle sum is larger than π contradicting Proposition 15. Exercise 12 implies that boundary parallel subspaces ($\alpha_1 = 1$) also have no common perpendicular. Now consider the case $\alpha_1 > 1$ of parallel or skew subspaces; for $l = 0, \boldsymbol{A}^0 = \boldsymbol{a}$, skew just means that $\boldsymbol{a} \notin \boldsymbol{B}^m$, and we are again in the situation of Proposition 11. In order to complete the general case now, we also need a suitable basis for $\boldsymbol{W}$ that is adapted to the situation. This again will be constructed by considering projections of the vectors $\mathfrak{b}_i \in \boldsymbol{U}^\perp, i = 1, \ldots, n-l$. We have

$$pr_{\boldsymbol{W}} \mathfrak{b}_i = \mathfrak{b}_i - pr_{\boldsymbol{W}^\perp} \mathfrak{b}_i,\ i = 1, \ldots, n-l.$$

Relation (75) implies

$$pr_{\boldsymbol{W}^\perp} \mathfrak{b}_i = \mathfrak{a}_i \sqrt{\alpha_i}, \alpha_i = 0 \text{ for } i = r_A + 1, \ldots, n-l.$$

[7] Essentially the same classification was originally proved within the framework of Möbius geometry, cf. R. Sulanke [102].

This last arrangement takes into account that $\mathfrak{b}_i \in \boldsymbol{W} \cap \boldsymbol{U}^\perp$, where $i = r_A + 1, \ldots, n-l$, cf. (77). Hence

$$pr_{\boldsymbol{W}} \mathfrak{b}_i = \mathfrak{b}_i - \mathfrak{a}_i \sqrt{\alpha_i} \text{ for } i = 1, \ldots, n-l. \tag{89}$$

A straightforward calculation involving the scalar products of these projections leads to

$$\langle \mathfrak{b}_i - \mathfrak{a}_i \sqrt{\alpha_i}, \mathfrak{b}_j - \mathfrak{a}_j \sqrt{\alpha_j} \rangle = \delta_{ij}(1 - \alpha_i), \ i, j = 1, \ldots, n-l. \tag{90}$$

In particular, the projection of $\mathfrak{b}_1$ is time-like; so we define

$$\mathfrak{a}_0 := -(\mathfrak{b}_1 - \mathfrak{a}_1 \cosh \beta_1)/\sinh \beta_1 = \mathfrak{b}_0 \cosh \beta_1 - \mathfrak{b}_1 \sinh \beta_1, \tag{91}$$

where the last equation follows from (81) and (84). Taking into account (82) and (85) for $i = 1, \ldots, n-l-1$, the normalization of the remaining projections leads to

$$\mathfrak{a}_{n-m+i} := \frac{\mathfrak{b}_{i+1} - \mathfrak{a}_{i+1}\sqrt{\alpha_{i+1}}}{\sqrt{1 - \alpha_{i+1}}} = \mathfrak{b}_{i+1} \sin \beta_{i+1} - \mathfrak{b}_{n-l+i} \cos \beta_{i+1}. \tag{92}$$

If necessary, we again complete this basis to a pseudo-orthonormal basis for $\boldsymbol{W}$. Now define $\boldsymbol{x}_0 := [\mathfrak{b}_0] \in \boldsymbol{A}^l$, $\boldsymbol{y}_0 := [\mathfrak{a}_0] \in \boldsymbol{B}^m$. Since $\langle \mathfrak{a}_0, \mathfrak{b}_0 \rangle = -\cosh \beta_1$, the distance of these points is $h(\boldsymbol{x}_0, \boldsymbol{y}_0) = \beta_1$. We show that the line $\boldsymbol{x}_0 \vee \boldsymbol{y}_0$ is orthogonal to each line through $\boldsymbol{y}_0$ in $\boldsymbol{B}^m$. With the initial point $\boldsymbol{y}_0$ the line $\boldsymbol{x}_0 \vee \boldsymbol{y}_0$ has the parameter representation

$$\mathfrak{x}(t) = \mathfrak{a}_0 \cosh t + \frac{\mathfrak{b}_0 - \mathfrak{a}_0 \cosh \beta_1}{\sinh \beta_1} \sinh t.$$

An arbitrary line through $\boldsymbol{y}_0$ in $\boldsymbol{B}^m$ has the representation

$$\mathfrak{z}(s) = \mathfrak{a}_0 \cosh s + \mathfrak{a}^\perp \sinh s, \ \mathfrak{a}^\perp \in \boldsymbol{W}, \langle \mathfrak{a}^\perp, \mathfrak{a}^\perp \rangle = 1, \langle \mathfrak{a}_0, \mathfrak{a}^\perp \rangle = 0. \tag{93}$$

It is easy to conclude from (91) and (92) the following

$$\langle \mathfrak{b}_0 - \mathfrak{a}_0 \cosh \beta_1, \mathfrak{a}^\perp \rangle = 0.$$

The angle definition for intersecting lines (Example 5) implies orthogonality. Analogously one proves the orthogonality of $\boldsymbol{x}_0 \vee \boldsymbol{y}_0$ and each line through $\boldsymbol{x}_0$ in $\boldsymbol{A}^l$. To show the asserted distance property consider, in analogy to (93), the parameter representation

$$\mathfrak{y}(t) = \mathfrak{b}_0 \cosh t + \mathfrak{b}^\perp \sinh t, \ \mathfrak{b}^\perp \in \boldsymbol{U}, \langle \mathfrak{b}^\perp, \mathfrak{b}^\perp \rangle = 1, \langle \mathfrak{b}_0, \mathfrak{b}^\perp \rangle = 0.$$

Now we investigate the function describing the distance of the (essentially arbitrary) points $[\mathfrak{y}(t)] \in \boldsymbol{A}^l$, $[\mathfrak{z}(s)] \in \boldsymbol{B}^m$:

$$f(s,t) := -\langle \mathfrak{y}(t), \mathfrak{z}(s)\rangle = \cosh\beta_1 \cosh s \cosh t - \cos\gamma \sinh s \sinh t$$

$$= \cosh h([\mathfrak{y}(t)], [\mathfrak{z}(s)]).$$

There we could set

$$\cos\gamma = \langle \mathfrak{a}^{\perp}, \mathfrak{b}^{\perp}\rangle,$$

since both these unit vectors are orthogonal to $\mathfrak{a}_0$ and $\mathfrak{b}_0$, and consequently span a Euclidean subspace. An elementary argument shows that this function has an extremum only for $s = t = 0$, which, moreover, is the asserted minimum. Furthermore, the *foot points* $\boldsymbol{x}_0, \boldsymbol{y}_0$ are uniquely determined. □

On the basis of Proposition 25 we define the *distance of parallel* or *skew hyperbolic subspaces* as $h(\boldsymbol{A}^l, \boldsymbol{B}^m) := h(\boldsymbol{x}_0, \boldsymbol{y}_0) = \beta_1$. For intersecting subspaces the distance is obviously zero; as usual, the distance of two sets A, B in a space with metric ρ is defined as follows:

$$\rho(A,B) := \inf\{\rho(a,b) |\; a \in A, b \in B\}. \tag{94}$$

In this sense, for *boundary parallel subspaces* $\boldsymbol{A}^l, \boldsymbol{B}^m$ *the distance is zero,* $h(\boldsymbol{A}^l, \boldsymbol{B}^m) = 0$. (The proof is left to the reader as an exercise; it suffices to consider boundary parallel lines in the plane.)

2.6.8 Quadrics

A *quadric in hyperbolic space* $\boldsymbol{H}^n$ is defined geometrically as the intersection of a projective quadric B with the hyperbolic space: $B_H := B \cap \boldsymbol{H}^n$. Obviously, such a quadric may be empty; already the intersection Q_H of the hypersphere Q defining the hyperbolic space with $\boldsymbol{H}^n$ is, by the definition of hyperbolic space as its inner region, empty; the points of B belonging to Q are called the *limit points* or the points at infinity of B. To determine the points of B_H amounts to finding the hyperbolic points of the projective quadric B, i.e. the time-like solutions of the quadratic equation

$$b(\mathfrak{x}) := b(\mathfrak{x},\mathfrak{x}) = \Sigma_{i,j=0}^{n} b_{ij} x^i x^j = 0,\ \langle \mathfrak{x},\mathfrak{x}\rangle = -1, (b_{ij} = b_{ji}); \tag{95}$$

b is a symmetric, real bilinear form with the matrix (b_{ij}). Since the equation is homogeneous, it suffices to consider the equation only for the normalized vectors $\mathfrak{x}$, i.e. time-like unit vectors. Equation (95) suggest to consider a quadric as the intersection of the cone in the vector space $\boldsymbol{V}^{n+1}$ defined by $b(\mathfrak{x},\mathfrak{x}) = 0$ with the pseudo-Euclidean model H_+ for the hyperbolic space, cf. (10) with $r = 1$. To classify these quadrics with respect to congruence under the action of the group G_n it is not enough to use the Eigendecomposition Theorem as in the Euclidean, the spherical, or the elliptic geometries. This time it is necessary to apply the results concerning the normal forms of symmetric endomorphisms to the endomorphisms associated with a symmetric bilinear form, cf. Exercise 2.17. The variety of algebraic possibilities resulting from this as well as from the pseudo-Euclidean structure on the vector space correspond to the intricate geometric situation. In fact, now we have to investigate the position of an arbitrary quadric with respect to the absolute, which itself is a quadric, namely a hypersphere. It is an interesting, but in higher dimensions not quite simple task to establish the relations between the geometric and the algebraic properties.

Example 11. Central quadrics. A non-empty quadric $B \neq \emptyset$ is called *central* if there is a point $\boldsymbol{m} \in \boldsymbol{H}^n$ such that the reflection $s_{\boldsymbol{m}}$ in this point maps the quadric into itself: $s_{\boldsymbol{m}}(B) \subset B$; $\boldsymbol{m}$ is called a *center* of the quadric B. Note that in this definition all points of B are taken into account, i.e. those possibly lying in the outer region of Q as well. To describe $s_{\boldsymbol{m}}$ algebraically we choose the pseudo-orthonormal basis $(\mathfrak{e}_i)$ so that $\boldsymbol{m} = [\mathfrak{e}_0]$ is represented by the time-like unit vector $\mathfrak{e}_0$; then

$$s_{\boldsymbol{m}} : \boldsymbol{x} = [\mathfrak{x}] = [\mathfrak{e}_0 x^0 + \mathfrak{x}_1] \longmapsto s_{\boldsymbol{m}}(\boldsymbol{x}) = [\mathfrak{e}_0 x^0 - \mathfrak{x}_1] \text{ with } \mathfrak{x}_1 \in [\mathfrak{e}_0]^\perp. \tag{96}$$

Now we may identify $[\mathfrak{e}_0]^\perp = \boldsymbol{E}^n$ with the n-dimensional Euclidean vector space; then the isotropy group of $\mathfrak{e}_0$ under the action of G_n is the orthogonal group $\boldsymbol{O}(n)$, which acts simply transitively on the orthonormal frames of $\boldsymbol{E}^n$. Let

$$b_1 : \mathfrak{x}_1 \in \boldsymbol{E}^n \longmapsto b_1(\mathfrak{x}_1) = \Sigma_{i,j=1}^n b_{ij} x^i x^j \in \mathbf{R}$$

be the restriction of b on the Euclidean space $\boldsymbol{E}^n$. It is obvious that the equation

$$b(\mathfrak{x}) := b_{00}(x^0)^2 + b_1(\mathfrak{x}_1) = 0 \text{ for } \mathfrak{x} = \mathfrak{e}_0 x^0 + \mathfrak{x}_1, \langle \mathfrak{e}_0, \mathfrak{x}_1 \rangle = 0 \tag{97}$$

determines a central quadric in $\boldsymbol{H}^n$, if there are solutions for (97) at all. In a moment we will prove that every central quadric can be represented in the form (97), which will easily lead to a classification of central quadrics. □

Lemma 26. *A point $\boldsymbol{z} \in \boldsymbol{H}^n$ is a center of the non-empty quadric B if and only if B can be represented by an equation of the form (97) with $\boldsymbol{z} = [\mathfrak{e}_0]$.*

Proof. Example 11 shows that the condition is sufficient. Now we consider a point $\boldsymbol{z} \in \boldsymbol{H}^n$. Assuming there is no representation for B in the form (97) with the given z we will see that the point cannot be the center of B. To this end we take a pseudo-orthonormal coordinate system with $\boldsymbol{z} = [\mathfrak{e}_0]$. In the notations of Example 11 we then have for the equation of the arbitrary quadric B

$$b(\mathfrak{x}) = b_{00}(x^0)^2 + 2x^0\omega(\mathfrak{x}_1) + b_1(\mathfrak{x}_1),$$

where the linear form ω is defined by

$$\omega(\mathfrak{x}_1) := \Sigma_{j=1}^n b_{0j} x^j =: \langle \mathfrak{v}, \mathfrak{x}_1 \rangle.$$

The second equation determines the vector $\mathfrak{w}$ in the Euclidean vector space $\boldsymbol{E}^n = [\mathfrak{e}_0]^\perp$ corresponding to the form ω. Choosing a suitable orthonormal basis for the Euclidean vector space $\boldsymbol{E}^n$ we may suppose that the symmetric bilinear form b_1 is diagonal (Eigendecomposition Theorem, cf. Proposition I.6.4.2). Next we pass to Klein's model for hyperbolic space; to do so we normalize the representatives $\mathfrak{x}$ of the points by requiring $x^0 = 1$ (cf. (16)). The quadric B then consists of all points $\boldsymbol{x} = [\mathfrak{e}_0 + \mathfrak{x}_1]$ satisfying

$$b_{00} + 2\langle \mathfrak{v}, \mathfrak{x}_1 \rangle + \Sigma_{j=1}^n \lambda_j (x^j)^2 = 0,\ \mathfrak{x}_1 \in [\mathfrak{e}_0]^\perp. \tag{98}$$

Such a representation exists for each quadric; we call it *a representation of the quadric B related to $\boldsymbol{z} \in \boldsymbol{H}^n$*. Since we assume that there is no representation of the form (97), we conclude $\mathfrak{v} \neq \mathfrak{o}$. Now we decompose the space $\boldsymbol{E}^n$ into three mutually orthogonal subspaces, where we do not exclude trivial subspaces:

$$\begin{aligned} \boldsymbol{U}_+ &:= [(\mathfrak{e}_i)_{\lambda_i > 0}], \\ \boldsymbol{U}_- &:= [(\mathfrak{e}_i)_{\lambda_i < 0}], \\ \boldsymbol{U}_0 &:= [(\mathfrak{e}_i)_{\lambda_i = 0}]. \end{aligned}$$

We label the vectors in the basis so that the first k belong to positive, the following $m-k$ to negative eigenvalues, and the remaining span the defect subspace of b_1; i.e., they correspond to the eigenvalue zero:

$$\boldsymbol{U}_+ = [\mathfrak{e}_1, \ldots, \mathfrak{e}_k],\ \boldsymbol{U}_- = [\mathfrak{e}_{k+1}, \ldots, \mathfrak{e}_m],\ \boldsymbol{U}_0 = [\mathfrak{e}_{m+1}, \ldots, \mathfrak{e}_n],$$

here $m = \mathrm{rk}(b_1)$ and $m - k = \mathrm{index}(b_1)$. For $m = 0$ the assertion is trivial. Now let $m > 0$. We decompose the vector $\mathfrak{v} = \mathfrak{v}_1 + \mathfrak{v}_0$ into its orthogonal components $\mathfrak{v}_1 \in \boldsymbol{U}_+ \oplus \boldsymbol{U}_-$, $\mathfrak{v}_0 \in \boldsymbol{U}_0$ and show: *If $\mathfrak{v}_0 \neq \mathfrak{o}$, then $\boldsymbol{z}$ is no center for B*. To see this note that there is $\mathfrak{a} \in \boldsymbol{U}_+ \oplus \boldsymbol{U}_-$ satisfying $a := b_{00} + b_1(\mathfrak{a}) \neq 0$ Since $\mathfrak{v}_0 \neq \mathfrak{o}$, we find a vector $\mathfrak{b} \in \boldsymbol{U}_0$ such that $\mathfrak{x}_1 = \mathfrak{a} + \mathfrak{b}$ corresponds to a point of B, hence

$$a + 2\langle \mathfrak{v}, \mathfrak{x}_1 \rangle = b_{00} + b_1(\mathfrak{a}) + 2\langle \mathfrak{v}_1, \mathfrak{a} \rangle + 2\langle \mathfrak{v}_0, \mathfrak{b} \rangle = 0.$$

Because of $a \neq 0$ we also have $\langle \mathfrak{v}, \mathfrak{x}_1 \rangle \neq 0$, and since replacing $\mathfrak{x}_1$ by $-\mathfrak{x}_1$ the value of a does not change, the corresponding point does not belong to B; $\boldsymbol{z}$ is no center for B.

Now we consider the case that $\mathfrak{v}_0 = \mathfrak{o}$, equation (98) for $\mathfrak{x}_1$ only depends on the first m variables $x^\alpha, \alpha = 1, \ldots, m$. Rewriting it slightly it takes the form

$$A + \Sigma_{\alpha=1}^m \lambda_\alpha (x^\alpha + a^\alpha / \lambda_\alpha)^2 = 0,$$

where (a^α) are the coordinates of the vector $\mathfrak{v}$, and we have set

$$A := b_{00} - \Sigma_{\alpha=1}^m (a^\alpha)^2 / \lambda_\alpha.$$

We suppose $A \geq 0$ and order the equation according to the sign of its terms:

$$A + \Sigma_{\alpha=1}^{k} \lambda_\alpha (x^\alpha + a^\alpha/\lambda_\alpha)^2 = \Sigma_{\alpha=k+1}^{m} - \lambda_\alpha (x^\alpha + a^\alpha/\lambda_\alpha)^2. \qquad (99)$$

In case $A < 0$, we subtract this number and obtain in each case an equation containing only non-negative summands on both sides. For $k = m$ the right-hand side is equal to zero. Since B is not empty, in this case A has to be zero, and the only solution is the vector $\mathfrak{x}_1 \in \boldsymbol{U}_+$ with coordinates

$$x^\alpha = -a^\alpha/\lambda_\alpha \text{ für } \alpha = 1, \ldots, m.$$

Since, by assumption, $\mathfrak{v} \neq \mathfrak{o}$, the vector $-\mathfrak{x}_1$ is no solution, and hence $\boldsymbol{z}$ is no center for B. In the case $k < m$ we find a solution for any value of the left-hand side; in particular, this holds for the value

$$A_1 := A + \Sigma_{\alpha=1}^{k} (a^\alpha)^2/\lambda_\alpha,$$

which occurs for $x^\alpha = 0, \alpha = 1, \ldots, k$. If $A_1 = 0$, then $a^\alpha = 0$, $\alpha = 1, \ldots, k$. The only solution is then

$$\mathfrak{x}_1 = -\Sigma_{\alpha=k+1}^{m} \mathfrak{e}_\alpha a^\alpha/\lambda_\alpha \neq \mathfrak{o},$$

since $\mathfrak{v} \neq \mathfrak{o}$. In this case, however, $-\mathfrak{x}_1$ is no solution, hence $\boldsymbol{z}$ is no center. Let now $A_1 > 0$. Then the equation to determine the remaining components of $\mathfrak{x}_1$,

$$A_1 = -\Sigma_{\alpha=k+1}^{m} \lambda_\alpha (x^\alpha + a^\alpha/\lambda_\alpha)^2,$$

is the equation for a hyperellipsoid in the Euclidean vector space $\boldsymbol{U}_-$. If not all a^α are zero, $\alpha = k+1, \ldots, m$, then we certainly find a solution

$$\mathfrak{x}_1 = \Sigma_{\alpha=k+1}^{m} \mathfrak{e}_\alpha x^\alpha \text{ mit } \langle \mathfrak{v}, \mathfrak{x}_1 \rangle = \Sigma_{k+1}^{m} a^\alpha x^\alpha \neq 0.$$

But then again, $-\mathfrak{x}_1$ cannot be a solution and $\boldsymbol{z}$ no center. All that remains is to consider the case $a^\alpha = 0, \alpha = k+1, \ldots, m$. Then we write (99) in the form

$$\Sigma_{\alpha=1}^{k} \lambda_\alpha (x^\alpha + a^\alpha/\lambda_\alpha)^2 = -\Sigma_{\alpha=k+1}^{m} \lambda_\alpha (x^\alpha)^2 - A =: A_2$$

and fix the $x^\alpha, \alpha = k+1, \ldots, m$, so that $A_2 > 0$. This time the solution is a hyperellipsoid in the Euclidean vector space $\boldsymbol{U}_+$. But since $\mathfrak{v} \neq \mathfrak{o}$, not all $a^\alpha, \alpha = 1, \ldots, k$, can be equal to zero as well, we again find a solution $\mathfrak{x}_1 \in \boldsymbol{U}_+$ for which $-\mathfrak{x}_1$ does not solve the equation. Rewriting the equation as described the case $A < 0$ is treated along the same lines. We leave this to the reader. □

As already carried out in Euclidean geometry we want to describe the geometric properties of a quadric, e.g. to be central, in terms of its invariants. As we know, rank and index of a symmetric real bilinear form are invariant for all linear transformations, hence for the pseudo-orthogonal $g \in G_n$ as well. A symmetric bilinear form and the corresponding quadric are called *degenerate* if the rank is less than $n+1$. The quadric does not change, when its equation is multiplied by a number $\kappa \in \mathbf{R}^*$; so we may always suppose

$$\operatorname{index}(b) \le \operatorname{rk}(b)/2. \tag{100}$$

Furthermore, there still is the possibility to suitably normalize b by multiplying it with a positive number. However, we do not want to generally require this normalization; instead we will vary it according to the specific situation. In Lemma 2.20 we associated a linear endomorphism with each biform in a vector space with scalar product; moreover, this map is a G_n-isomorphism, and in our case linear as well (cf. Exercise 2.17). Hence the G_n-invariants of the associated endomorphism are, at the same time, invariants of the symmetric bilinear form, and, having agreed upon its normalization, also invariants of the corresponding quadric. Equation (2.49) for the matrix (a^i_k) of the associated linear map and the form (2) of Gram's matrix immediately imply that, for a pseudo-orthonormal basis, it is simply obtained by multiplying the first row of the matrix (b_{ij}) for the symmetric bilinear form by -1:

$$\begin{pmatrix} a^0_0 & \mathfrak{a}^{0\prime} \\ \mathfrak{a}_0 & A_{nn} \end{pmatrix} = \begin{pmatrix} -1 & \mathfrak{o}' \\ \mathfrak{o} & I_n \end{pmatrix} \begin{pmatrix} b^0_0 & \mathfrak{v}' \\ \mathfrak{v} & B_{nn} \end{pmatrix} = \begin{pmatrix} -b_{00} & -\mathfrak{v}' \\ \mathfrak{v} & B_{nn} \end{pmatrix}. \tag{101}$$

Here $A_{nn} = B_{nn}$ is a symmetric square matrix of order n, I_n is the unit matrix, and the vectors denoted by $'$ are the corresponding row vectors. From this formula it is obvious that (b_{ij}) is diagonal with respect to a pseudo-orthonormal basis if and only if this basis consists of eigenvectors for the associated operator; in fact, both matrices are simultaneously diagonal or not, respectively. Note that, in general, the matrix of the symmetric operator a is not symmetric; this only holds for $\mathfrak{v} = \mathfrak{o}$. Lemma 26 and the form of the associated operator imply

Corollary 27. *A non-empty quadric $B \subset \boldsymbol{P}^n(\mathbf{R})$ determined by a symmetric bilinear form b is central if and only if the corresponding associated endomorphism a has a time-like eigenvector, and this again holds if and only if a has a pseudo-orthonormal basis of eigenvectors.*

Proof. Lemma 26 and Formula (101) immediately imply that $\mathfrak{e}_0$ is an eigenvector for a, and vice versa. If $\mathfrak{e}_0$ is a time-like eigenvector, then the subspace $[e_0]$ is invariant for a. By Exercise 2.17, a is symmetric, and so according to Lemma 2.21 the Euclidean subspace $\boldsymbol{E} := [\mathfrak{e}_0]^\perp$ is invariant for a. By (101) $a|\boldsymbol{E}$ is a self-adjoint endomorphism in a Euclidean vector space. According to the Eigendecomposition Theorem, Proposition I.6.4.2, a has only real eigenvalues, and, moreover, there is an orthonormal basis of eigenvectors for $a|\boldsymbol{E}$; hence, completed by $\mathfrak{e}_0$ it is the pseudo-orthonormal basis of eigenvectors for a we were looking for. The converse is trivial. □

Proposition 28. *Let $B \subset \boldsymbol{P}^n(\mathbf{R})$ be a non-empty, non-degenerate central quadric in hyperbolic projective space. Then there is a unique symmetric bilinear form b of rank $n+1$ such that B is the solution set of the corresponding quadratic equation*

$$b(\mathfrak{x}) = -(x^0)^2 + \lambda_1(x^1)^2 + \ldots + \lambda_n(x^n)^2 = 0. \tag{102}$$

The index $k+1$ of b satisfies the inequality $k \leq (n-1)/2$, (x^i) are the coordinates of $\mathfrak{x}$ in a pseudo-orthonormal basis of eigenvectors for the endomorphism a associated with b, and

$$-1 \leq \lambda_1 \leq \ldots \leq \lambda_n. \tag{103}$$

Two of these quadrics are hyperbolically congruent if and only if the sequences of eigenvalues for the associated endomorphisms coincide, provided that the fixed normalizations are taken into account. Such a quadric contains a hyperbolic point $\boldsymbol{x} \in B_H = B \cap \boldsymbol{H}^n$ if and only if $\lambda_n > 1$. It only consists of hyperbolic points if and only if $\lambda_1 > 1$.

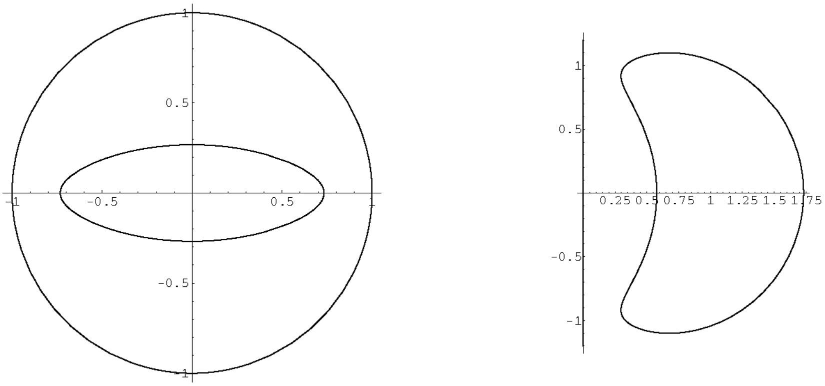


Fig. 2.30. Ellipse in the conformal disk and in Poincaré's model, $\lambda_1 = 1.1, \lambda_2 = 4$.

Proof. By (100) we may suppose that the index $k+1$ of b satisfies the inequality $k \leq (n-1)/2$, since b is non-degenerate. The form b cannot be definite, since then the quadric B would be empty. According to Corollary 27 the form b has a pseudo-orthonormal eigenbasis. We order the eigenvalues by their magnitude and divide the form by $|\lambda_0|$; then relations (102) and (103) hold. Note that all eigenvalues are different from zero, since b is non-degenerate. By Lemma 2.20 the map $b \mapsto a$, assigning the associated linear endomorphism to each symmetric bilinear form, is a G_n-map, and by Exercise 2.17, d) a linear isomorphism as well. Hence the eigenvalues are G_n-invariants of the bilinear form b, and, because of the uniqueness we accomplished by the normalization, of the quadric B as well. If two quadrics have the same system of eigenvalues satisfying (105), then they are G_n-congruent; the transformation $g \in G_n$ transforming the corresponding pseudo-orthonormal bases into one another realizes the congruence (assignment by equal coordinates). The quadric B contains hyperbolic points if and only if $B \cap H_+ \neq \emptyset$; H_+ again denotes the pseudo-Euclidean model of hyperbolic geometry (cf. Figure 2.17), described by the equation of time-like unit vectors:

$$-(x^0)^2 + (x^1)^2 + \ldots + (x^n)^2 = -1. \tag{104}$$

Subtracting (102) from (104) yields the equation

$$(1 - \lambda_1)(x^1)^2 + \ldots + (1 - \lambda_n)(x^n)^2 = -1,$$

which all points of the intersection $B \cap H_+$ satisfy. If now $\lambda_n \leq 1$, then, by (103), the left-hand side of the last equation contains only non-negative terms, and hence there can be no real solutions. If, on the other hand, $\lambda_n > 1$, then using

$$(x^0)^2 = \lambda_n/(\lambda_n - 1),\ x^1 = \ldots = x^{n-1} = 0,\ (x^n)^2 = 1/(\lambda_n - 1)$$

we obtain two solutions of both equations determined by $x^0 > 0$ and the roots $x^n = \pm\sqrt{1/(\lambda_n - 1)}$. To show the last claim we look at the quadric B in Klein's model, i.e., in $\boldsymbol{V}^{n+1}$ we consider the intersection of the cone (102) with the hyperplane $x^0 = 1$. It is described by the equation

$$\lambda_1(x^1)^2 + \ldots + \lambda_n(x^n)^2 = 1. \tag{105}$$

If $\lambda_1 > 1$, then, because of (103), for $i = 1, \ldots, n$ we have $\lambda_i > 1$, the solution of the equation is a hyperellipsoid which according to

$$\Sigma_{i=1}^n (x^i)^2 < \Sigma_{i=1}^n \lambda_i (x^i)^2 = 1$$

completely lies in the interior of the n-dimensional ball D^n of radius 1. If $\lambda_1 \leq 1$, then for $\lambda_1 < 0$ it is a hyperboloid, which at any rate contains points outside of D^n. For $0 < \lambda_1 \leq 1$ it is a hyperellipsoid with largest semi-axis $\sqrt{(1/\lambda_1)} \geq 1$, which thus contains at least two limit points (for $\lambda_1 = 1$) or points lying in the outer region of D^n. □

This easily leads to the classification of non-degenerate central quadrics in hyperbolic geometry: Proposition 28 immediately implies

Corollary 29. *Under the assumptions of Proposition 28 the sequence* $(\lambda_1, \ldots, \lambda_n)$ *of coefficients occurring in Equation (102) forms a complete system of invariants in hyperbolic geometry for the class of central quadrics considered there. If* $\lambda_i \neq 1$ *for* $i = 1, \ldots, n$*, then there is just the single center* $\boldsymbol{z} = [\mathfrak{e}_0]$ *for the quadric. In the general case, the centers of such a central quadric form a hyperbolic* $(d-1)$*-plane, where* d *denotes the dimension of the eigensubspace for* a *corresponding to the eigenvalue 1.*

Proof. To prove the second assertion note that by (101) and (102) the zeroth eigenvalue of the endomorphism a associated with b is always $a_{00} = 1$ under the assumption of Proposition 28. Thus, because of $\lambda_i \neq 1$, the corresponding eigensubspace is always orthogonal to $\mathfrak{e}_0$ and hence Euclidean. If the eigensubspace for the eigenvalue 1 of a, which is always pseudo-Euclidean, has dimension $d > 1$, then each time-like unit vector in this space determines a center, and all these centers together are a hyperbolic $(d-1)$-plane. □

Example 12. Already in Proposition 17 we saw that the metric *hyperspheres* are quadrics. Their defining equation derived there (formula line following (54)) implies that their invariants all coincide; more precisely,

$$\lambda_i = (\coth(r))^2, \; i = 1, \ldots, n,$$

where r denotes the hyperbolic radius of the hypersphere. Obviously, the center is uniquely determined. Figure 2.30 shows hyperbolic ellipses having a unique center, just like the circles. The *equidistants* are central quadrics as well. Equations (55) describe the time-like unit vectors representing the points of an equidistant. Hence the corresponding cone in the pseudo-Euclidean vector space has the parameter representation $\mathfrak{y}(\boldsymbol{x}, s) = \mathfrak{y}_r(\boldsymbol{x})s$, where by (55) we have to set $\boldsymbol{x} = [\mathfrak{x}]$ with $\mathfrak{x} \in \boldsymbol{W} = [\mathfrak{e}_n]^\perp$ and $\langle \mathfrak{x}, \mathfrak{x} \rangle = -1$. Eliminating the coordinates of $\boldsymbol{x}$ and s in this parameter representation implies that the equidistant with distance r to the hyperplane $x^n = 0$ is a central quadric, whose normalized equation has the following form:

$$-(y^0)^2 + (y^1)^2 + \ldots + (y^{n-1})^2 + (y^n)^2(\coth r)^2 = 0.$$

The invariants for these equidistants are thus $\lambda_1 = \ldots = \lambda_{n-1} = 1$, all points in the hyperplane determining them are centers, cf. Figure 2.27, and the last invariant is a function of the distance parameter. All equidistants with the same distance parameter are G_n-congruent, which was clear anyway (cf. Exercise 24), but is again evident from this representation. Finally, Figure 2.31 shows a central quadric not completely belonging to the hyperbolic plane. In the projective plane these are ellipses whose hyperbolic points have two branches orthogonally intersecting infinity, i.e. the boundary. As in the Euclidean situation their tangents become asymptotes for the curve point moving towards infinity. Hence it is certainly justified to call them hyperbolas in non-Euclidean geometry. We recommend that the reader proves the orthogonality; a proof as well as a collection of formulas to work with in the hyperbolic plane and to produce the pictures included here can be found in the Mathematica notebook noeuklid.nb. □

Example 13. The *horospheres* are also quadrics; they are a first example for quadrics whose associated endomorphisms do not have a basis of eigenvectors. From the definition of a horosphere it is obvious that they cannot be central quadrics: In fact, they have a unique limit point, which, for symmetry reasons, is impossible in the case of central quadrics (cf. Figure 2.29). In order to derive a quadratic equation for the horosphere (64) we multiply this equation by the parameter s and thus pass to the equation for the corresponding hypercone.

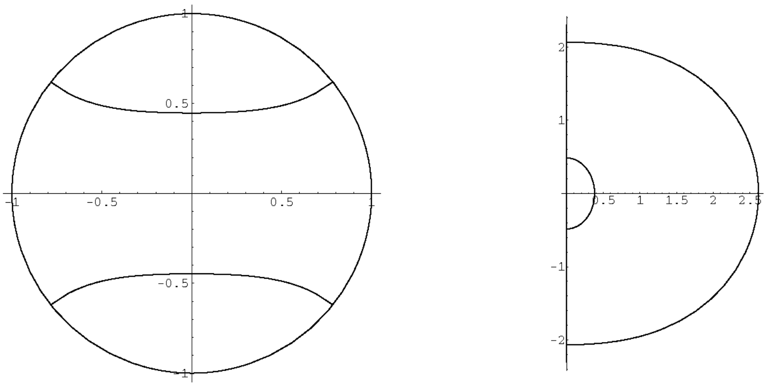


Fig. 2.31. Non-Euclidean hyperbolas, $\lambda_1 = 0.5, \lambda_2 = 1.8$.

Then we eliminate the parameters $\mathfrak{a}, s$. The coordinates (z^i) of the points in the horosphere satisfy the equation

$$\Sigma_{i=1}^{n-1}(z^i)^2 + 2(z^n)^2 + 2z^0 z^n = 0, \tag{106}$$

which one can also easily verify directly. Hence the matrix for the associated endomorphism in the pseudo-orthonormal standard basis is

$$a = \begin{pmatrix} 0 & \mathfrak{o}' & -1 \\ \mathfrak{o} & I_{n-1} & \mathfrak{o} \\ 1 & \mathfrak{o}' & 2 \end{pmatrix}.$$

The characteristic equation of this matrix is

$$\chi_a(t) = (1-t)^{n+1}.$$

Thus the only eigenvalue is 1, it has multiplicity $n+1$. There are, however, only n linearly independent eigenvectors, which are even mutually orthogonal:

$$a(\mathfrak{e}_i) = \mathfrak{e}_i \text{ for } i = 1, \ldots, n-1, \ a(\mathfrak{e}_n - \mathfrak{e}_0) = \mathfrak{e}_n - \mathfrak{e}_0.$$

But the last eigenvector is isotropic, it corresponds to the limit point of the horosphere. This may be seen by dividing the representing vector on the right-hand side of (64) by $\langle \mathfrak{a}, \mathfrak{a} \rangle$ and then letting $\mathfrak{a}$ tend to infinity. In the basis

$$\mathfrak{a}_i := \mathfrak{e}_i \text{ for } i = 1, \ldots, n-1, \ \mathfrak{a}_n := \mathfrak{e}_0, \ \mathfrak{a}_{n+1} := \mathfrak{e}_n - \mathfrak{e}_0,$$

which is no longer pseudo-orthonormal, the matrix for a has the Jordan normal form

$$\begin{pmatrix} I_{n-1} & 0 \\ 0 & \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \end{pmatrix}.$$

This describes the invariants for the horospheres, which are all G_n-congruent, as already mentioned. At the same time this provides an example for Exercise 2.19, which was to show that the eigenvector for a Jordan cell of order k>1 always has to be isotropic. If, conversely, the operator associated with any quadric has the same Jordan normal form as that of a horosphere, then this quadric itself is a horosphere; as explained before, it suffices to adapt a pseudo-orthonormal basis $(\mathfrak{e}_i)$ to the Jordan basis $(\mathfrak{a}_i)$; then the equation of the quadric in the orthonormal basis has the form (106), which is the equation of a horosphere. □

We called the points in the intersection $B \cap Q$ of the quadric B with the absolute Q the *limit points of the quadric* B. The formula

$$b(\mathfrak{a}, \mathfrak{x}) = \langle a(\mathfrak{a}), \mathfrak{x} \rangle = \langle \mathfrak{a}, \mathfrak{x} \rangle \lambda, \text{ if } a(\mathfrak{a}) = \mathfrak{a}\lambda, \tag{107}$$

immediately leads to

Corollary 30. *The isotropic eigenvectors for the endomorphism a associated with a quadric B determined by the symmetric bilinear form b are limit points for the quadric.* □

The example of non-Euclidean hyperbolas, cf. Figure 2.31, shows that the converse of this corollary does not hold. Misusing the terminology we will frequently speak of the Jordan normal form, the Jordan basis, the eigenvectors etc. of a quadric to mean the corresponding quantities of its associated endomorphism. To classify the quadrics it is crucial to decide, which Jordan normal forms may at all occur as normal forms for quadrics. For each of these types of normal forms the eigenvalues appear as parameters, which, taking into account a normalization convention, are invariants of the quadric. Now we will proceed with the classification of non-degenerate, plane quadrics. The variety of possibilities already occurring in this simple case gives an impression of the difficulty to be expected of the classification in arbitrary dimensions. A report on the results of the classification in the n-dimensional hyperbolic space can be found in Section 3.7.3 of the book [95] by B. A. Rosenfeld. There J. L. Coolidge [33] is also quoted, a first systematic presentation of hyperbolic geometry treating the classification of quadrics in the hyperbolic space $\boldsymbol{H}^3$ as well.

Now let a be the endomorphism associated with a plane quadric B. Its characteristic polynomial has degree 3 and real coefficients; hence a has at least one real eigenvalue λ. For non-degenerate quadrics all eigenvalues are different from zero, and there are only the following possibilities:

1. There is a time-like eigenvector. Then the quadric is central.
2. All eigenvalues are real, but there is no time-like eigenvector.
3. There are a real eigenvalue λ as well as two non-real, conjugate complex eigenvalues $\mu, \bar{\mu}$

Obviously, these cases exclude each other; in fact, in Case 3 λ can have no time-like eigenvector, since then, according to Corollary 27, all eigenvalues were real. In Case 1, according to the same corollary, the quadrics are classified by Proposition 28. Considering Case 2 we distinguish the following excluding subcases:

2.1. There is a unique Jordan cell of dimension 3.
2.2. There is a Jordan cell of dimension 2.

Three one-dimensional Jordan cells cannot occur in case 2, cf. the next Exercise 26. We will show that quadrics of the types 2.1, 2.2, and 3 exist and are classified by the corresponding Jordan normal forms.

Exercise 26. Let a be a symmetric endomorphism of a pseudo-Euclidean vector space $\boldsymbol{V}^n$ of index 1. Suppose that a is diagonalizable, i.e. a has n linearly independent eigenvectors. Then a has a time-like eigenvector as well. Hint. First prove that two linearly independent isotropic eigenvectors for a have to belong to the same eigenvalue.

Example 14. We consider the family of symmetric bilinear forms $bj3(\lambda)$ depending on a real parameter $\lambda \in \mathbf{R}$, which, in pseudo-orthonormal coordinates, are determined by a matrix of the same form denoted by the same symbol

$$bj3(\lambda) := \begin{pmatrix} -\lambda & 0 & 1 \\ 0 & \lambda & 1 \\ 1 & 1 & \lambda \end{pmatrix}. \tag{108}$$

Thus with the notations $z = x^0, x = x^1, y = x^2$ the equation of this quadric is

$$\lambda(x^2 + y^2 - z^2) + 2y(x + z) = 0, \;\; \lambda \neq 0. \tag{109}$$

For $\lambda = 0$ this is just a degenerate quadric. In the same coordinates the associated operator has the matrix

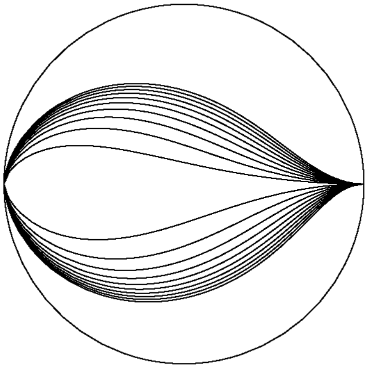
$$\begin{pmatrix} \lambda & 0 & -1 \\ 0 & \lambda & 1 \\ 1 & 1 & \lambda \end{pmatrix}.$$

So the Jordan normal form of this matrix is of the form we wanted

$$\begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}. \tag{110}$$

We call these quadrics *J3-curves*; Figure 2.32 shows some of these curves in the conformal and in Poincaré's model. We may suppose $\lambda > 0$ to hold in (109); multiplying by -1 and reflecting in the x-axis $(x, y, z) \mapsto (x, -y, z)$ we again obtain Equation (109) with positive λ and a G_2-congruent curve. Now we only consider orientation preserving transformations. Then each sign of λ

in (108) corresponds to one class of $J3$-curves, and the reflection transforms one into the other. Figure 2.32 displays together with each curve its mirror image. Since the eigenvalues of the associated endomorphisms are G-invariants for the bilinear forms, no two of the quadrics determined by (109) (in pseudo-orthonormal coordinates) with different values of $|\lambda|$ can be G_2-congruent. To classify the quadrics of Type 2.1 it remains to prove that each quadric whose associated endomorphism has the Jordan normal form (110) is G_2-congruent to a $J3$-curve (109). □



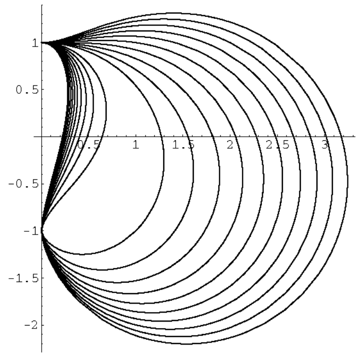


Fig. 2.32. $J3$-curves, $\lambda = \pm j/2, j = 1, 2, \ldots, 10$.

So let B be a non-degenerate quadric of type 2.1. We want to find a pseudo-orthonormal basis for the pseudo-Euclidean vector space $\boldsymbol{V}^3$ in which the quadric is described by an equation of the form (109). By assumption, there is a *Jordan basis* $(\mathfrak{a}_i)$ for $\boldsymbol{V}^3$ in which the matrix of the operator a associated with B has the Jordan normal form (110); obviously, this basis is not pseudo-orthonormal, since the eigenvector $\mathfrak{a}_3$ belonging to this basis is isotropic (cf. Exercise 2.19). We will see this again in a moment. The symmetry property of a,

$$\langle a(\mathfrak{x}), \mathfrak{y} \rangle = \langle \mathfrak{x}, a(\mathfrak{y}) \rangle \text{ for all } \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}, \tag{111}$$

implies for each Jordan basis of a cell with dimension $k > 1$

$$\langle \mathfrak{a}_{i+1}, \mathfrak{a}_j \rangle = \langle \mathfrak{a}_i, \mathfrak{a}_{j+1} \rangle, \quad i, j = 1, \ldots, k, \tag{112}$$

where we have to set $\mathfrak{a}_{k+1} = \mathfrak{o}$. This immediately follows from the form of the matrix (110) together with (111):

$$\langle a(\mathfrak{a}_i), \mathfrak{a}_j \rangle = \langle \mathfrak{a}_i, \mathfrak{a}_j \rangle \lambda + \langle \mathfrak{a}_{i+1}, \mathfrak{a}_j \rangle = \langle \mathfrak{a}_i, a(\mathfrak{a}_j) \rangle = \langle \mathfrak{a}_i, \mathfrak{a}_j \rangle \lambda + \langle \mathfrak{a}_i, \mathfrak{a}_{j+1} \rangle .$$

Setting $j = k$ in (112) and using $\mathfrak{a}_{k+1} = \mathfrak{o}$ yields

$$\langle \mathfrak{a}_i, \mathfrak{a}_k \rangle = 0 \text{ for } i = 2, \ldots, k; \tag{113}$$

in particular, we arrive at the asserted isotropy $\langle \mathfrak{a}_k, \mathfrak{a}_k \rangle = 0$. Moreover, for $i = k-2, j = k-1$ we have:

$$\langle \mathfrak{a}_{k-1}, \mathfrak{a}_{k-1} \rangle = \langle \mathfrak{a}_{k-2}, \mathfrak{a}_k \rangle, \quad k \geq 3. \tag{114}$$

This has a consequence, which very much restricts the possibilities occurring in the classification of symmetric endomorphisms in Minkowski spaces of arbitrary dimension:

Corollary 31. *If λ is a real eigenvalue of a symmetric endomorphism a in an n-dimensional pseudo-Euclidean vector space of index 1, then the cells $\Delta^k(\lambda)$ in its Jordan normal form have at most dimension $k = 3$.*

Proof. If $k > 3$, then (113) and (114) implied $\langle \mathfrak{a}_{k-1}, \mathfrak{a}_{k-1} \rangle = 0$. But then the subspace $[\mathfrak{a}_{k-1}, \mathfrak{a}_k]$ were totally isotropic, and this cannot occur in Minkowski space, cf. Exercise 2.4. □

In our case, $k = 3$, (114) only implies $\langle \mathfrak{a}_2, \mathfrak{a}_2 \rangle = \langle \mathfrak{a}_1, \mathfrak{a}_3 \rangle \neq 0$; in fact, for the same reason as in the proof of Corollary 31 $\mathfrak{a}_2$ cannot be isotropic. More precisely, since the index of $\boldsymbol{V}$ is equal to 1, the vector $\mathfrak{a}_2$ orthogonal to $\mathfrak{a}_3$ by (113) is space-like, hence $\mu := \langle \mathfrak{a}_2, \mathfrak{a}_2 \rangle > 0$. As $(\mathfrak{a}_i/\sqrt{\mu})$ is also a Jordan basis, after this normalization we may suppose

$$\langle \mathfrak{a}_2, \mathfrak{a}_2 \rangle = \langle \mathfrak{a}_1, \mathfrak{a}_3 \rangle = 1. \tag{115}$$

Moreover, for each Jordan basis we have

$$\langle \mathfrak{a}_2, \mathfrak{a}_3 \rangle = \langle \mathfrak{a}_3, \mathfrak{a}_3 \rangle = 0, \quad (k = 3). \tag{116}$$

It is easy to verify that for each $\nu \in \mathbf{R}$ the transformation

$$\hat{\mathfrak{a}}_1 := \mathfrak{a}_1 + \mathfrak{a}_2\nu, \ \hat{\mathfrak{a}}_2 := \mathfrak{a}_2 + \mathfrak{a}_3\nu, \ \hat{\mathfrak{a}}_3 := \mathfrak{a}_3$$

again defines a Jordan basis for a, which also satisfies (115). Together with (116) this implies

$$\langle \hat{\mathfrak{a}}_1, \hat{\mathfrak{a}}_2 \rangle = \langle \mathfrak{a}_1, \mathfrak{a}_2 \rangle + 2\nu,$$

and choosing $\nu = -\langle \mathfrak{a}_1, \mathfrak{a}_2 \rangle / 2$ we can achieve that the scalar product is zero, $\langle \hat{\mathfrak{a}}_1, \hat{\mathfrak{a}}_2 \rangle = 0$. Returning to the former notations there now is a Jordan basis for which, apart from (115), (116), we also have

$$\langle \mathfrak{a}_1, \mathfrak{a}_2 \rangle = 0. \tag{117}$$

Lastly, for each $\mu \in \mathbf{R}$ the transformation

$$\hat{\mathfrak{a}}_1 := \mathfrak{a}_1 + \mathfrak{a}_3\mu, \ \hat{\mathfrak{a}}_2 := \mathfrak{a}_2, \ \hat{\mathfrak{a}}_3 := \mathfrak{a}_3$$

also transforms a Jordan basis again into a Jordan basis, and Equations (115)–(117) are preserved. The equality

$$\langle \hat{\mathfrak{a}}_1, \hat{\mathfrak{a}}_1 \rangle = \langle \mathfrak{a}_1, \mathfrak{a}_1 \rangle + 2\mu$$

implies that, suitably choosing μ, passing to the old notations we may arrange that

$$\langle \mathfrak{a}_1, \mathfrak{a}_1 \rangle = -1. \tag{118}$$

This proves the first part of the following lemma.

Lemma 32. *For each symmetric endomorphism a in the Minkowski space $\boldsymbol{V}^3$ whose Jordan normal form consists of a single Jordan cell $\Delta^3(\lambda)$ there is a Jordan basis satisfying* (115)–(118). *Starting from such a basis the relations*

$$\mathfrak{e}_0 := \mathfrak{a}_1,\ \mathfrak{e}_1 := \mathfrak{a}_1 + \mathfrak{a}_3,\ \mathfrak{e}_2 := \mathfrak{a}_2 \tag{119}$$

define a pseudo-orthonormal basis for $\boldsymbol{V}^3$. □

The orthogonality relations for the basis $(\mathfrak{e}_i)$ are easily verified using (115)–(118). obviously, the assumption implies $\lambda \in \mathbf{R}$. The matrix for the basis transformation defined by (119) is

$$(\mathfrak{e}_i) = (\mathfrak{a}_i)T \text{ mit } T = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Since the operator a has the matrix (110) in the Jordan basis $(\mathfrak{a}_i)$, its matrix in the basis $(\mathfrak{e}_i)$ is

$$\begin{pmatrix} \lambda & 0 & -1 \\ 0 & \lambda & 1 \\ 1 & 1 & \lambda \end{pmatrix} = T^{-1} \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix} T.$$

Passing to the associated bilinear form we obtain the matrix $bj3(\lambda)$ defined by (108) for the $J3$-curve (109). Hence we obtain the result we aimed at:

Proposition 33. *Each non-degenerate quadric of type* 2.1 *in the hyperbolic plane is G_2-congruent to a unique $J3$-curve* (109) *with $\lambda > 0$. The $J3$-curves with parameters $\lambda, -\lambda$ are transformed into one another by the reflection in the line connecting their limit points.* □

Now let B be a non-degenerate quadric of Type 2.2. We want to adapt a pseudo-orthonormal basis to B in such a way that the suitably normalized equation determining B has a normal form still to be found depending only on the eigenvalues λ, μ of the associated operator a. This allows to conclude that two quadrics of this type with equal eigenvalues are congruent. For each Jordan basis $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}$ of the operator a we have

$$a(\mathfrak{a}) = \mathfrak{a}\lambda + \mathfrak{b}, a(\mathfrak{b}) = \mathfrak{b}\lambda, a(\mathfrak{c}) = \mathfrak{c}\mu. \tag{120}$$

A Jordan basis cannot be pseudo-orthonormal, since by Exercise 2.19 $\mathfrak{b}$ is isotropic. If $\mu \neq \lambda$, then by Proposition 2.24 $\mathfrak{c}$ is orthogonal to $\mathfrak{b}$ and hence space-like, since a pseudo-Euclidean vector space of index 1 cannot contain an isotropic vector orthogonal to a time-like or isotropic vector and linearly independent of it. But even in the case $\lambda = \mu$ the vector $\mathfrak{c}$ is space-like, since the symmetry property of a implies

$$\langle a(\mathfrak{a}), \mathfrak{c} \rangle = \langle \mathfrak{a}\mu + \mathfrak{b}, \mathfrak{c} \rangle = \langle \mathfrak{a}, a(\mathfrak{c}) \rangle = \langle \mathfrak{a}, \mathfrak{c} \rangle \mu,$$

and hence $\langle \mathfrak{b}, \mathfrak{c} \rangle = 0$. Then one argues as before. We normalize $\mathfrak{c}$ and set $\mathfrak{c} = \mathfrak{e}_2$. Since a is symmetric, according to Proposition 2.24 the pseudo-Euclidean subspace $[\mathfrak{e}_2]^\perp$ is also invariant under a, and $\langle \mathfrak{a}, \mathfrak{e}_2 \rangle = 0$ implies

$$\langle \mathfrak{b}, \mathfrak{e}_2 \rangle = \langle a(\mathfrak{a}) - \mathfrak{a}\lambda, \mathfrak{e}_2 \rangle = 0.$$

Hence $[\{\mathfrak{a}, \mathfrak{b}\}]$ is a pseudo-Euclidean vector space, and because of the isotropy of $\mathfrak{b}$ we have $\langle \mathfrak{a}, \mathfrak{b} \rangle \neq 0$. The following is easy to verify: If $\{\hat{\mathfrak{a}}, \hat{\mathfrak{b}}\}$ is another Jordan basis for $a|[\mathfrak{a}, \mathfrak{b}]$, then

$$\hat{\mathfrak{a}} = \mathfrak{a}s + \mathfrak{b}t, \ \hat{\mathfrak{b}} = \mathfrak{b}s, \ (s \neq 0).$$

Since $\langle \hat{\mathfrak{a}}, \hat{\mathfrak{b}} \rangle = \langle \mathfrak{a}, \mathfrak{b} \rangle s^2$, the sign of $\langle \mathfrak{a}, \mathfrak{b} \rangle$ is invariant under the transformation. We choose s so that $\langle \hat{\mathfrak{a}}, \hat{\mathfrak{b}} \rangle = \pm 1$. Then we compute the unique t solving the equation

$$\langle \hat{\mathfrak{a}}, \hat{\mathfrak{a}} \rangle = \langle \mathfrak{a}, \mathfrak{a} \rangle s^2 + 2\langle \mathfrak{a}, \mathfrak{b} \rangle st = -1.$$

Thus we adapted a pseudo-orthonormal basis to the operator a in such a way that

$$\mathfrak{a} = \mathfrak{e}_0, \mathfrak{b} = \mathfrak{e}_0 + \mathfrak{e}_1, \mathfrak{c} = \mathfrak{e}_2 \text{ for } \langle \mathfrak{a}, \mathfrak{b} \rangle < 0, \tag{121}$$

and

$$\mathfrak{a} = \mathfrak{e}_0, \mathfrak{b} = -\mathfrak{e}_0 + \mathfrak{e}_1, \mathfrak{c} = \mathfrak{e}_2 \text{ for } \langle \mathfrak{a}, \mathfrak{b} \rangle > 0, \tag{122}$$

is a Jordan basis for the operator a in each case. For the matrix of the operator a in the pseudo-orthonormal basis $(\mathfrak{e}_j), j = 0, 1, 2$ this implies

$$(\alpha_i^j) = \begin{pmatrix} \lambda + 1 & -1 & 0 \\ 1 & \lambda - 1 & 0 \\ 0 & 0 & \mu \end{pmatrix} \text{ if } \langle \mathfrak{a}, \mathfrak{b} \rangle < 0, \tag{123}$$

and

$$(\alpha_i^j) = \begin{pmatrix} \lambda - 1 & -1 & 0 \\ 1 & \lambda + 1 & 0 \\ 0 & 0 & \mu \end{pmatrix} \text{ if } \langle \mathfrak{a}, \mathfrak{b} \rangle > 0. \tag{124}$$

The matrices of the associated symmetric bilinear forms are

$$(\beta_i^j) = \begin{pmatrix} -\lambda - 1 & 1 & 0 \\ 1 & \lambda - 1 & 0 \\ 0 & 0 & \mu \end{pmatrix} \text{ or } (\beta_i^j) = \begin{pmatrix} -\lambda + 1 & 1 & 0 \\ 1 & \lambda + 1 & 0 \\ 0 & 0 & \mu \end{pmatrix}. \tag{125}$$

We now intersect the cones in the vector space $\boldsymbol{V}^3$ corresponding to these bilinear forms with the coordinate plane $x^0 = 1$ and rewrite the variables as $x^1 = x, x^2 = y$. This way we obtain the normal forms for the equation of the quadric in Klein's model we were looking for:

$$-\lambda - 1 + 2x + (\lambda - 1)x^2 + \mu y^2 = 0 \text{ if } \langle \mathfrak{a}, \mathfrak{b} \rangle < 0, \tag{126}$$

and

$$1 - \lambda + 2x + (\lambda + 1)x^2 + \mu y^2 = 0 \text{ if } \langle \mathfrak{a}, \mathfrak{b} \rangle > 0. \tag{127}$$

We call the solutions to these equations *J2-curves*; more precisely, these are the solutions of (126) with $J2^-(\lambda, \mu)$ and those of (127) with $J2^+(\lambda, \mu)$. Since $\mu \neq 0$, they are very simply expressed algebraically in terms of roots, e.g. in the case of (126) as

$$y = \pm\sqrt{(\lambda + 1 - 2x + (1 - \lambda)x^2)/\mu}. \tag{128}$$

The normalization for the equations is the requirement $\beta_{01} = 1$ determining the only mixed coefficient. We ask under which condition two $J2$-curves are congruent. This holds if and only if their normalized equations can be transformed into one another by a pseudo-orthogonal transformation of the coordinates. Hence no curves in any of the families $J2^+$, $J2^-$ can be congruent, since they differ in the eigenvalues of the associated operators. Substituting in (126) $x \to -x$ and multiplying the resulting equation by -1 yields Equation (127) for the parameter value $-\lambda, -\mu$, which leads to the congruence

$$J2^+(-\lambda, -\mu) \simeq J2^-(\lambda, \mu). \tag{129}$$

What we discuss here is the *algebraic congruence*, since the equivalence of equations in question is that under pseudo-orthogonal coordinate transformations. To the existence of real or hyperbolic points, respectively, whose coordinates satisfy these equations, we will return later. Obviously, algebraic equivalence of the equations implies the geometric congruence of their solutions. As already mentioned in Euclidean geometry the converse does not necessarily hold. Summarizing we formulate

Proposition 34. *For all numbers $\lambda, \mu \in \mathbf{R}$ satisfying $\lambda\mu \neq 0$ there is a unique equation for a quadric in the normal form* (126) *or* (127), *respectively, whose associated operators have these numbers as their eigenvalues. $J2$-quadrics are congruent if and only if corresponding eigenvalues of the operators associated with the normal forms* (126) *or* (127) *are equal. Quadrics of the families* (126), (127) *only differ by a reflection in the y-axis, where* (129) *holds.* □

The algebraically simple form of the solutions allows to decide, whether the quadric contains hyperbolic points. To this end we need a series of elementary observations: We need to find the positive domains for the radicands of (128), have to determine the intersection with the interval $-1 < x < 1$, and must then investigate, if the image points with coordinates (x, y) belong to the unit circle for the resulting sets. These considerations can be found in the Mathematica notebook noeuklid2D.nb on the homepage of R. Sulanke. This notebook contains many tools for the investigation of the hyperbolic plane; the modules to produce the figures reproduced in this section can be obtained there as well. Here we only present the results of these considerations and, at the end of this section, the corresponding pictures computed in the conformal model.

Now we still have to find the quadrics of Type 3 whose associated operators have two conjugate complex eigenvalues $\lambda \neq \bar{\lambda}$ and a real eigenvalue $\mu \neq 0$. Consider the complex extension $\boldsymbol{V}_c$ of the pseudo-Euclidean vector space $\boldsymbol{V}^3$. Bilinearly extending the pseudo-Euclidean scalar product $\boldsymbol{V}_c^3$ becomes a complex Euclidean vector space, cf. Lemma 1.10.15, Example 1.10.9, and Example 2.2.5. The linear extension of the symmetric linear endomorphism a, uniquely determined by Lemma 1.10.5 (and denoted the same way), is a symmetric linear endomorphism of the complex Euclidean vector space $\boldsymbol{V}_c$; this is immediate from the linearity of the extensions considered, if the symmetry property is expressed in a basis for the real form $\boldsymbol{V}$, which, at the same time, is a basis for $\boldsymbol{V}_c$. Since the eigenvalues of the endomorphism are different, according to Proposition 2.24 the corresponding eigensubspaces are orthogonal, one-dimensional complex subspaces of $\boldsymbol{V}_c^3$, and, since the scalar product is non-degenerate, they are not isotropic. Dividing by the real eigenvalue μ normalizes the equation of the quadric and the associated operator so that we may suppose $\mu = 1$. We then find a real eigenvector $\mathfrak{c} \in \boldsymbol{V}$, which has to be space-like: If it were time-like, then we had Case 1 of a central quadric; moreover, it cannot be isotropic according to what we just mentioned. Next we normalize it and take it to be the last vector $\mathfrak{e}_2$ in a pseudo-orthonormal basis for $\boldsymbol{V}^3$. The complex extension of $\boldsymbol{W}^2 := [\mathfrak{e}_2]^\perp \subset \boldsymbol{V}$ is the complex orthogonal complement $W_c^2 = [\mathfrak{e}_2]_c^\perp \subset \boldsymbol{V}_c$, which is spanned by the eigenvectors $\mathfrak{a}, \mathfrak{b}$ for the eigenvalues $\lambda, \bar{\lambda}$; moreover, these are not isotropic. We normalize $\mathfrak{a}$ so that $\langle \mathfrak{a}, \mathfrak{a} \rangle = -2$. Let

$$\lambda = \alpha + \beta\,\mathrm{i}, \mathfrak{a} = \mathfrak{x} + \mathfrak{y}\,\mathrm{i},\ \alpha, \beta \in \mathbf{R},\ \mathfrak{x}, \mathfrak{y} \in \boldsymbol{W},$$

be its decompositions into the real components. Passing to the complex conjugate quantities, $a(\mathfrak{a}) = \mathfrak{a}\lambda$ implies that $\mathfrak{b} := \mathfrak{x} - \mathfrak{y}\,\mathrm{i}$ is an eigenvector for the eigenvalue $\bar{\lambda}$. From

$$\langle \mathfrak{a}, \mathfrak{a} \rangle = \langle \mathfrak{x}, \mathfrak{x} \rangle - \langle \mathfrak{y}, \mathfrak{y} \rangle + 2\,\mathrm{i}\langle \mathfrak{x}, \mathfrak{y} \rangle = -2,$$
$$\langle \mathfrak{a}, \mathfrak{b} \rangle = \langle \mathfrak{x}, \mathfrak{x} \rangle + \langle \mathfrak{y}, \mathfrak{y} \rangle = 0$$

we immediately conclude the orthogonality relations

$$-\langle \mathfrak{x}, \mathfrak{x} \rangle = \langle \mathfrak{y}, \mathfrak{y} \rangle = 1,\ \langle \mathfrak{x}, \mathfrak{y} \rangle = 0.$$

We set $\mathfrak{e}_0 := \mathfrak{x}, \mathfrak{e}_1 := \mathfrak{y}$ and represent a in this basis taking into account the relations holding for an eigensystem; then

Proposition 35. *For each non-degenerate quadric B of Type 3 in the hyperbolic plane there is a pseudo-orthonormal basis (e_j) for the associated vector space such that, having normalized the equation to have the real eigenvalue 1, the matrix of the associated operator a has the form*

$$\begin{pmatrix} \alpha & \beta & 0 \\ -\beta & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{130}$$

where $\lambda = \alpha + \mathrm{i}\,\beta, \beta \neq 0$, denotes a complex eigenvalue of a. Thus the normal form of the equation for such a quadric B in Klein's model is

$$\alpha(x^2 - 1) - 2\beta x + y^2 = 0. \tag{131}$$

Two such quadrics are congruent if and only if their normal forms coincide up to the sign of β. All these quadrics contain hyperbolic points. □

It is straightforward to see the solutions from the simple form of Equation (131). Moreover, one notes that replacing β by $-\beta$ the corresponding curve is obtained from the first by a reflection in the y-axis; it is also equivalent to exchanging λ and $\bar{\lambda}$. The last assertion in the proposition is obtained by discussing the domains where the solutions are real as well as their magnitude; this is explained in the Mathematica notebook noeuklid2D.nb already cited. It also contains a module $qc[\alpha, \beta]$, which computes the quadric corresponding to the parameters α and β, more precisely, its hyperbolic points, in the conformal model; for brevity we call them *qc-curves*. The Figures 2.44, 2.45 were plotted using this module.

2.6.9 Pictures of Hyperbolic Quadrics

In this subsection we collect pictures of some quadrics discussed in the last subsection.

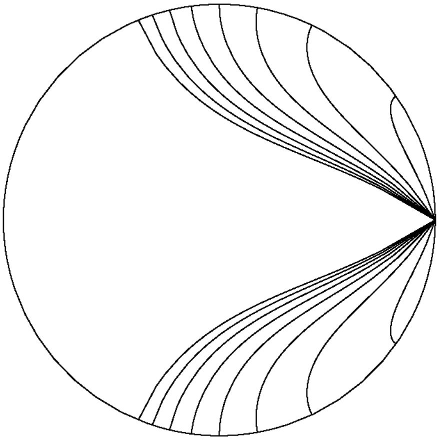


Fig. 2.33. $J2^-$-curves, $\lambda = 1$, $\mu = 1.1, 1.4, 1.7, \ldots, 3.2$.

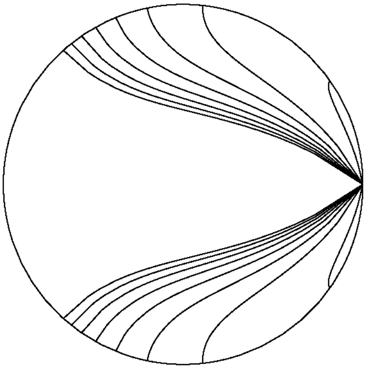


Fig. 2.34. $J2^-$-curves, $\lambda = 3$, $\mu = 3.1, 3.8, 4.5, \ldots, 8$.

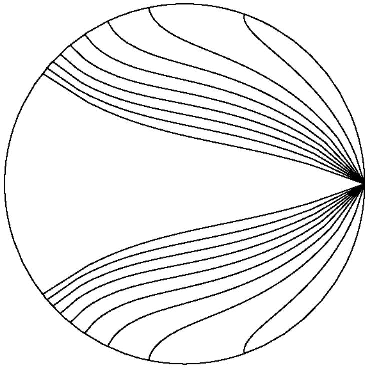


Fig. 2.35. $J2^{-}$-curves, $\lambda = 1.5, 2.5, \ldots, 9.5$, $\mu = 10$.

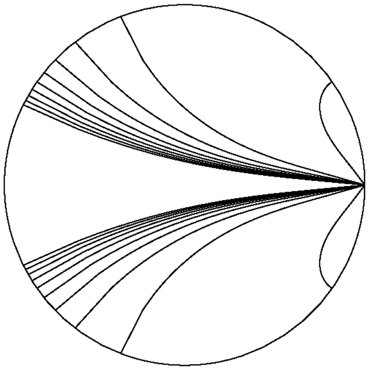


Fig. 2.36. $J2^{-}$-curves, $\lambda = 0.2, \mu = .3, 2.3, 4.3, \ldots, 18.3$.

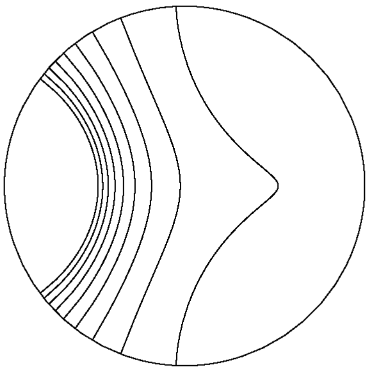


Fig. 2.37. $J2^{-}$-curves, $\lambda = -.1, -1.1, -2.1, \ldots, -8.1, \mu = 1$.

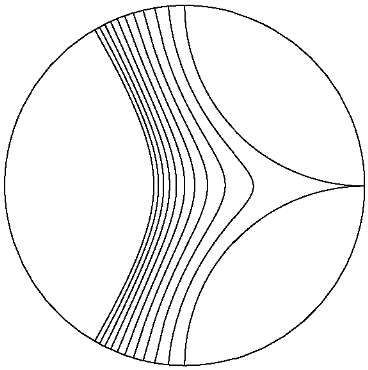


Fig. 2.38. $J2^-$-curves, $\lambda = 0, -0.2, -0.4, -0.6, \ldots, -2, \mu = 1$.

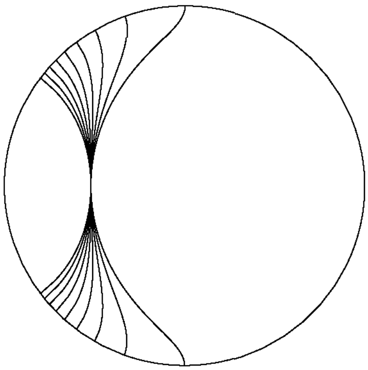


Fig. 2.39. $J2^-$-curves, $\lambda = -10, \mu = -1, -2, \ldots, -9$.

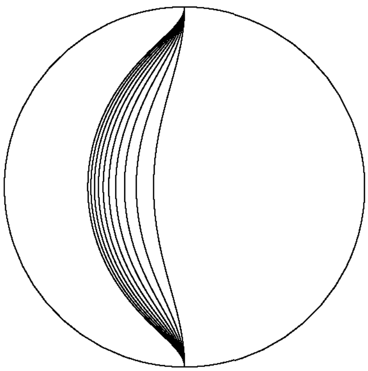


Fig. 2.40. $J2^-$-curves, $\lambda = -2, \ldots, -11, \mu = \lambda + 1$.

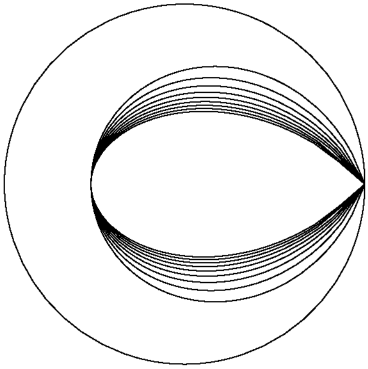


Fig. 2.41. $J2^{-}$-curves, $\lambda = -10, \mu = -11, -2, \ldots, -19$.

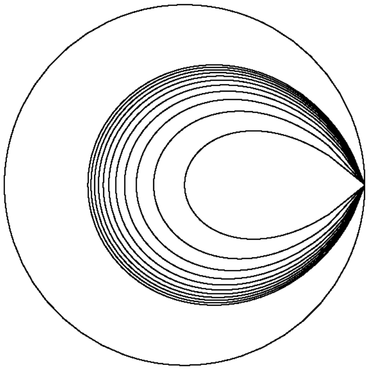


Fig. 2.42. $J2^{-}$-curves, $\lambda = -1, \ldots, -11, \mu = \lambda - 1$.

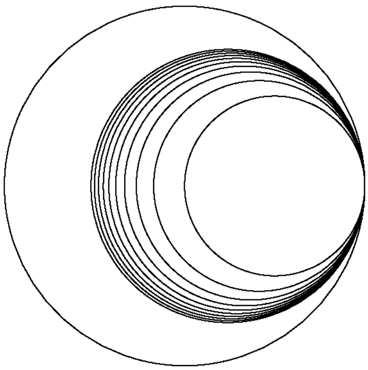


Fig. 2.43. $J2^{-}$-curves, $\lambda = \mu = -1, -2, \ldots, -10$, for $\lambda = \mu = -1$ the limit circle.

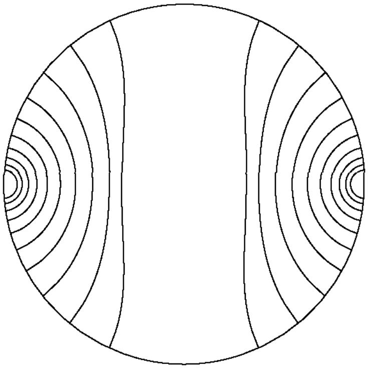


Fig. 2.44. qc-curves, $\lambda = 2^j \pm \mathrm{i}\,\sqrt{j}, j = 1, \ldots, 10$.

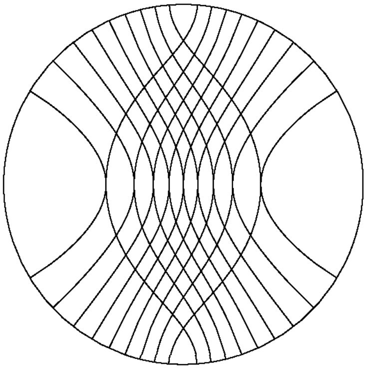


Fig. 2.45. qc-curves, $\lambda = \mathrm{e}^{2\pi\,\mathrm{i}\,j/10}, j = \pm 1, \ldots, \pm 9$.

2.7 Möbius Geometry

As already mentioned in the introduction to Section 2.6, from the projective perspective, *Möbius geometry* is nothing but the geometry on the hypersphere in projective space determined by the action of the isotropy group for this hypersphere. By Corollary 3.1, this group, in the context of Möbius geometry called the *Möbius group,* acts transitively on the hypersphere and leaves, as will be shown in a moment, the angle between intersecting circles invariant; using methods from differential geometry it can be proved that this group coincides with the conformal group of the hypersphere for the Riemannian metric induced by the standard embedding of the spheres into Euclidean space. Historically, the Möbius group was first introduced as the group of globally defined conformal maps on the Riemann sphere S^2 within the framework of complex function theory. On the other hand, the Möbius group is the isometry group for the hyperbolic metric defined on the inner region of the hypersphere, cf. Corollary 6.10. In the context of projective geometry this leads to a close and natural relation with hyperbolic geometry that will be applied in Section 7.1 to transfer the angle invariants from hyperbolic to Möbius geometry.

2.7.1 Spheres in Möbius Space

In accordance with Section 2.6 we agree to use the following notations: Let $\boldsymbol{V} = \boldsymbol{V}^{n+2}$ be the $(n+2)$-dimensional pseudo-Euclidean vector space equipped with the scalar product $\langle , \rangle$ of index 1, and let $\boldsymbol{P}^{n+1}$ be the associated projective space. The *n-dimensional Möbius space* S^n is understood to be the hypersphere Q for the scalar product $\langle , \rangle$ (cf. (6.3)) in the $(n+1)$-dimensional real projective space $\boldsymbol{P}^{n+1}$. Thus, we essentially retain the notations of Section 2.6, the dimension, however, is increased by 1 in order to keep the letter n for the dimension of the geometric space to be studied. In Lemma 6.19 we already described the isotropy group of a point $\boldsymbol{x} \in S^n$ with respect to an isotropic orthogonal basis in terms of block matrices; hence

Proposition 1. *The Möbius space S^n is a homogeneous space for the action induced from the projective action of $G_{n+1} = \boldsymbol{O}(1, n+1)^+ \subset \boldsymbol{O}(1, n+1)$. As a quotient space it is isomorphic to*

$$S^n \cong G_{n+1}/CE(n), \tag{1}$$

where $CE(n)$ denotes the conformal Euclidean group, *which is embedded into G_{n+1} via the matrix representation* (6.61) *(with $n-1$ replaced by n).* □

To justify this notation we first define the *group of similarities* for the Euclidean space $\boldsymbol{E}^n$: A *similarity* is a product of a Euclidean transformation with a dilation. These transformations form a subgroup of the affine group for the Euclidean space $\boldsymbol{E}^n$; using methods from differential geometry and, in case $n = 2$, function theory one can prove that it coincides with the group of all conformal (i.e. angle-preserving) transformations of $\boldsymbol{E}^n$, $n \geq 2$. The group $CE(n)$ defined in Proposition 1 is isomorphic to the group of similarities:

Exercise 1. Prove that the group described by the matrices (6.61) is isomorphic to the group of similarities for the Euclidean space $\boldsymbol{E}^n$. (Hint. Use the following model for the Möbius space S^n: Consider the unit hypersphere cut out of the isotropic cone $\langle \mathfrak{x}, \mathfrak{x} \rangle = 0$ by the Euclidean hyperplane

$$\boldsymbol{E}^{n+1} : -\langle \mathfrak{e}_0, \mathfrak{x} \rangle = -\langle \mathfrak{a}_0 + \mathfrak{a}_{n+1}, \mathfrak{x} \rangle / \sqrt{2} = 1$$

in the pseudo-orthogonal vector space $\boldsymbol{V}$. Here and in the sequel let $(\mathfrak{e}_i)$ denote a pseudo-orthonormal basis, and let $(\mathfrak{a}_i)$, $i = 0, \ldots, n+1$, be the isotropic orthogonal basis associated to it via (6.58). Take the fixed point $\boldsymbol{s} := [\mathfrak{a}_0] = [\mathfrak{e}_0 - \mathfrak{e}_{n+1}] \in S^n$ as the South Pole for the stereographic projection $st : S^n \setminus \{\boldsymbol{s}\} \to \boldsymbol{E}^n$ from S^n onto the equatorial plane $\boldsymbol{E}^n \subset \boldsymbol{E}^{n+1}$ of S^n determined by $\langle \mathfrak{e}_{n+1}, \mathfrak{x} \rangle = 0$ (cf. (6.18) and Exercise 6.5). Assigning $g \in CE(n) \longmapsto st \circ g \circ st^{-1}$ describes the isomorphy we were looking for. Here, for any element $g = a(A, \mathfrak{a}, \lambda)$ the entries $A, \mathfrak{a}, \lambda$ correspond to the rotations or rotoinversions, translations, or dilations in the group of similarities of $\boldsymbol{E}^n$, respectively.)

We may also identify the group of similarities with the conformal Euclidean group. Obviously, each similarity preserves the angles between intersecting lines, and hence it is conformal. The next two exercises contain geometric characterizations for the conformal Euclidean group.

Exercise 2. Prove that for $n \geq 2$ the conformal Euclidean group $CE(n)$ coincides with the group of bijective maps of $\boldsymbol{E}^n$ transforming pairs of orthogonally intersecting lines into the same kind of pairs. (Hint. Apply Proposition 1.5.6 and consider the images of lines intersecting orthogonally at the origin with direction vectors $\mathfrak{e}_i + \mathfrak{e}_j, \mathfrak{e}_i - \mathfrak{e}_j$ for $i \neq j$.) Obviously, this immediately implies that for $n \geq 2$ the conformal Euclidean group $CE(n)$ coincides with the group of bijective maps of $\boldsymbol{E}^n$ transforming lines into lines and preserving angles between intersecting lines.

Exercise 3. Prove that for $n \geq 2$ the conformal Euclidean group $CE(n)$ is the group of bijective maps of $\boldsymbol{E}^n$ transforming lines into lines and circles into circles. (Hint. Recall Thales' Theorem and apply the preceding exercise.)

Using isotropic orthogonal coordinates is often advisable to concentrate on the points of S^n as the simplest basic objects. Apart from the points the most important elementary objects of Möbius geometry are the m-spheres, for the investigation of which we will again use pseudo-orthonormal coordinates because of their close relation to the hyperbolic $(m+1)$-planes. As an *m-dimensional subsphere* $S^m \subset S^n$, for short *m-sphere*, we consider the intersection of a hyperbolic $(m+1)$-plane $\boldsymbol{B}^{m+1}$ with S^n. More precisely: If $\boldsymbol{W} \subset \boldsymbol{V}$ is an $(m+2)$-dimensional pseudo-Euclidean subspace, and $\boldsymbol{B}^{m+1}(\boldsymbol{W})$ is the corresponding projective subspace, then the associated m-sphere be defined as

$$S^m = S^m(\boldsymbol{W}) := \boldsymbol{B}^{m+1}(\boldsymbol{W}) \cap S^n, \; m = 0, 1, \ldots, n. \tag{2}$$

So 0-spheres are just arbitrary two-element sets $\{\boldsymbol{x}, \boldsymbol{y}\}, \boldsymbol{x}, \boldsymbol{y} \in S^n, \boldsymbol{x} \neq \boldsymbol{y}$; 1-spheres are *circles,* and $(n-1)$-spheres are *hyperspheres* of S^n. As non-degenerate quadrics of index 1 in $(m+1)$-dimensional projective spaces the m-spheres themselves are m-dimensional Möbius spaces as well. Denote by $S_{n,m}, m = 0, \ldots, n$, the set of m-spheres in the Möbius space S^n. Obviously, the map

$$F: \boldsymbol{B}^{m+1} \in H_{n+1,m+1} \longmapsto S^m = \boldsymbol{B}^{m+1} \cap S^n \in S_{n,m} \tag{3}$$

is bijective and an equivariant isomorphism from the transformation group $H_{n+1,m+1}$ of hyperbolic $(m+1)$-planes onto $S_{n,m}$, i.e.

$$F(g\boldsymbol{B}^{m+1}) = gF(\boldsymbol{B}^{m+1}) \text{ for all } g \in G_{n+1}, \boldsymbol{B}^{m+1} \in H_{n+1,m+1}.$$

This correspondingly also holds for the transformation groups of G_{n+1} defined by extending the action to power sets, product sets, etc. The transitivity properties for the actions of G_{n+1} in hyperbolic geometry immediately imply:

Corollary 2. *The Möbius group G_{n+1} acts transitively on the sets $S_{n,m}$ of m-spheres in the Möbius space S^n, $m = 0, 1, \ldots, n$.* □

The Möbius group was defined by restricting the isometry group G_{n+1} of hyperbolic space to the bounding quadric $Q = S^n$. The following proposition states geometric characterizations of this group.

Proposition 3. *The Möbius group G_{n+1}, $n \geq 2$, is the group of all bijective maps from S^n onto itself transforming circles into circles. Moreover, it is the group of bijective maps from S^n onto itself transforming hyperspheres into hyperspheres.*

Proof. According to Corollary 2 any Möbius transformation $g \in G_{n+1}$ transforms circles among themselves. Let now, conversely, $F : S^n \to S^n$ be a bijective map transforming circles into circles. Use the model for S^n described in Exercise 1 and the stereographic projection st. Since, by Proposition 1, the group G_{n+1} acts transitively on S^n, we find a Möbius transformation $g \in G_{n+1}$ such that $g(F(\boldsymbol{s})) = \boldsymbol{s}$. If we are able to show that the map $F_o = g \circ F$ is a Möbius transformation, then $F = g^{-1} \circ F_o$ is a Möbius transformation as well; without loss of generality we may thus suppose that F leaves the South Pole fixed: $F(\boldsymbol{s}) = \boldsymbol{s}$. Now consider the map $f := st \circ F \circ st^{-1}$, which is a bijection of the Euclidean space $\boldsymbol{E}^n$ onto itself. Since F transforms any circle through $\boldsymbol{s}$ into a circle through $\boldsymbol{s}$, which again are bijectively related to the lines in $\boldsymbol{E}^n$ via st, the map f maps the set of lines in $\boldsymbol{E}^n$ bijectively onto itself. Hence, by Proposition 1.5.6 it has to be an affine map. According to Lemma 6.3 the circles in Euclidean space are mapped by st^{-1} to the circles in S^n not passing through $\boldsymbol{s}$. Since this set is also mapped bijectively onto itself by F, the map f transforms the set of circles in $\boldsymbol{E}^n$ bijectively onto itself. Hence by Exercise 3 the map f is a similarity. So according to Exercise 1 F belongs to the conformal Euclidean subgroup $CE(n) \subset G_{n+1}$. The proof of the second assertion proceeds along the same lines; we leave it to the reader. □

Example 1. Examples of Möbius transformations that do not belong to the conformal Euclidean group are the *reflections in hyperspheres* also called *inversions*. Frequently, the term *inversive geometry* is used instead of Möbius geometry, compare e.g. the very instructive introductory text with this title by J. B. Wilker [111]. In the Euclidean space compactified by the point ∞ to the n-sphere S^n the reflection s in the hypersphere Σ^{n-1} with center $\boldsymbol{m} = \boldsymbol{o} + \mathfrak{m}$ and radius r is defined as follows:

$$s(\boldsymbol{x}) = s(\boldsymbol{m} + \mathfrak{x}) := \boldsymbol{m} + \frac{\mathfrak{x} r^2}{\langle \mathfrak{x}, \mathfrak{x} \rangle} \text{ for } \boldsymbol{x} \neq \boldsymbol{m}, \infty;\ s(\boldsymbol{m}) := \infty,\ s(\infty) := \boldsymbol{m}. \quad (4)$$

Hence the distance from the image $\boldsymbol{y} = s(\boldsymbol{x})$ to the center is inversely proportional to the distance from the pre-image to the center:

$$|\boldsymbol{y} - \boldsymbol{m}||\boldsymbol{x} - \boldsymbol{m}| = r^2,$$

and the image point $\boldsymbol{y}$ lies on the radial line through $\boldsymbol{x} \neq \boldsymbol{m}, \infty$; the last two points are interchanged. It is easy to verify that s is involutive, and that the set of fixed points of s coincides with the hypersphere Σ. Let us show that s corresponds to a Möbius transformation of S^n. Applying a similarity one may always arrange that the hypersphere defining s is the unit hypersphere of the Euclidean space $\boldsymbol{E}^n$ with the origin $\boldsymbol{o}$ as its center. Then $\mathfrak{x}$ is the position vector of the point $\boldsymbol{x}$, and the position vector $\mathfrak{y}$ of the image point $\boldsymbol{y}$ is determined by $\mathfrak{y} = \mathfrak{x}/|\mathfrak{x}|^2$. Now consider the reflections F_i in the coordinate hyperplanes $x^i = 0$ of the pseudo-Euclidean vector space $\boldsymbol{V}$, as usual they are defined with respect to a pseudo-orthonormal basis by

$$F_i(\mathfrak{e}_i) = -\mathfrak{e}_i,\ F_i(\mathfrak{e}_j) = \mathfrak{e}_j \text{ für } j \neq i;\ i, j = 0, 1, \ldots, n+1.$$

Obviously, each of these reflections is pseudo-orthogonal, hence it generates a Möbius transformation. By means of the stereographic projection already used in Exercise 1 one can show that $s = st \circ F_{n+1} \circ st^{-1}$. Conversely, each reflection in a pseudo-Euclidean $(n+1)$-dimensional subspace $\boldsymbol{W}^{n+1} \subset \boldsymbol{V}$ obviously generates a reflection in a hypersphere or hyperplane of $\boldsymbol{E}^n$; in the sense of Möbius geometry these reflections do not differ. It is easy to see that they may be even defined as the Möbius transformations, different from the identity, whose fixed point set is a hypersphere. The next proposition justifies the name inversive geometry. □

Proposition 4. *Each Möbius transformation can be represented as the product of finitely many inversions.*[1]

[1] J. B. Wilker [111] proves that each Möbius transformation $g \in G_{n+1}$ can even be represented as the product of at most $n+2$ inversions.

Proof. According to Proposition 2.3.5 each transformation $g \in G_{n+1}$ can be represented as the product of at most $n+2$ reflections in non-isotropic, $(n+1)$-dimensional subspaces. The reflections in pseudo-Euclidean subspaces are just the inversions considered in the previous example. If the subspace of fixed vectors is Euclidean, then we may choose the basis $(\mathfrak{e}_i)$ for $\boldsymbol{V}$ so that this reflection coincides with the map F_0 defined in Example 1; hence it has the coordinate representation

$$y^0 = -x^0,\ y^j = x^j \text{ for } j = 1, \dots, n+1.$$

Since now F_0 and $-F_0$ lead to the same Möbius transformation, we may well insert $-F_0$ instead of F_0 into the representation for g. Again this map is a product of finitely many inversions:

$$-F_0 = F_1 \circ \ldots \circ F_{n+1}.$$

□

This proposition is frequently chosen to be the starting point for a definition of Möbius geometry: Accordingly, Möbius geometry is the totality of geometric notions and properties which are invariant under arbitrary inversions. Note, moreover, that Möbius space can be equipped with an *orientation*. To each pseudo-orthonormal frame $(\mathfrak{e}_i)$, $i = 0, \dots, n+1$, we assign the *South Pole* $\boldsymbol{s} := [\mathfrak{e}_0 - \mathfrak{e}_{n+1}]$ as well as the *coordinate simplex* $(\boldsymbol{a}_i) := ([\mathfrak{e}_i + \mathfrak{e}_0])$, $i = 1, \dots, n+1$. The normalized representatives occurring in the definition lie in the hyperplane $x_0 = -\langle \mathfrak{e}_0, \mathfrak{x} \rangle = 1$ and have the positive determinant 2. We may express each oriented n-simplex $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n+1})$ in this frame by normalized representatives

$$\mathfrak{x}_i = \mathfrak{e}_0 + \sum_{j=1}^{n+1} \mathfrak{e}_j \xi_i^j;$$

for the determinants we then obtain

$$D := [\mathfrak{e}_0, \mathfrak{x}_1, \dots, \mathfrak{x}_{n+1}] = \det(\xi_i^j).$$

For $D > 0$ the n-simplex is ***positively oriented*** with respect to the basis $(\mathfrak{e}_i)$. Obviously, there are two classes of pseudo-orthonormal frames: Those whose coordinate simplex with respect to a fixed frame is positively oriented, and those for which it is negatively oriented. If one of these classes is distinguished in the vector space $\boldsymbol{V}$ defining S^n, then we speak of an *oriented Möbius space*; in this situation we can distinguish between positively and negatively oriented n-simplices. As already mentioned in the case of hyperbolic spaces, a map $g \in G_{n+1}$ preserves the orientation, if its determinant is equal to one, i.e. if g belongs to the special pseudo-orthogonal group $\boldsymbol{SO}(1, n+1)$. Since this group also acts transitively on S^n, the oriented Möbius space is described as the quotient space

$$S^n \cong \boldsymbol{SO}(1, n+1)/SCE(n),$$

where $SCE(n)$ denotes the group of similarities with positive determinant; it is called the group of orientation preserving similarities of the Euclidean space.

2.7.2 Pairs of Subspheres

The definition of the Möbius space S^n and the equivariance of the action of G_{n+1} over the hyperbolic $(m+1)$-planes as well as the m-spheres discussed in the preceding chapter immediately allows to conclude properties of the invariants for pairs of subspheres from those of the invariants for pairs of hyperbolic subspaces described in Proposition 6.24 in a purely formal way:

Proposition 5. *Let F be the eqivariant isomorphism described in* (3). *Consider the hyperbolic subspaces $\boldsymbol{A}^{l+1}$ and $\boldsymbol{B}^{m+1}$ corresponding to the subspheres $S(\boldsymbol{U}) \in S_{n,l}$, $S(\boldsymbol{W}) \in S_{n,m}$, i.e.*

$$\boldsymbol{A}^{l+1} = F^{-1}(S(\boldsymbol{U})), \boldsymbol{B}^{m+1} = F^{-1}(S(\boldsymbol{W})),$$

with $0 \leq l \leq m < n$. Then the stationary invariants and the deficit of these subspaces are Möbius invariants for the pair of spheres as well; this deficit and the eigenvalues (6.73) *of the operator $A \in \mathrm{End}(\boldsymbol{W}^{\perp})$ associated with the pair of subspaces are a complete system of invariants for the action of the Möbius group on $S_{n,l} \times S_{n,m}$.* □

Note that for the dimensions of the subspaces we now have

$$\dim \boldsymbol{U} = l+2, \quad \dim \boldsymbol{U}^{\perp} = n-l, \quad \dim \boldsymbol{W} = m+2, \quad \dim \boldsymbol{W}^{\perp} = n-m,$$

so that, obviously, despite the difference in the dimensions of the spaces $\boldsymbol{U}, \boldsymbol{W}, \boldsymbol{V}$, the number of invariants nevertheless coincides.

Now we want to interpret these invariants geometrically within the framework of Möbius geometry and, if possible, express them in terms of Euclidean invariants for the pairs of spheres; in fact, Möbius invariants are all the more invariants for the subgroup of Euclidean transformations. Hence they have to be functions of the invariants in a complete Euclidean system of invariants for the pairs. Before discussing particular dimensions, we consider the Cases 1–3 described in Lemma 6.23 in general. First we again suppose that Condition A from Section 6 is satisfied. In Möbius geometry this condition amounts to the following:

Condition A. *For the concerned pair $S_1, S_2 \subset S^n$ of subspheres there is no hypersphere $\Sigma \subset S^n$ such that $S_1 \cup S_2 \subset \Sigma$.*

As is common for projective spaces we denote the subsphere generated by a set $M \subset S^n$ by $[M]$; this is the intersection of all subspheres containing M. Finally, set

$$S_1 \vee S_2 := [S_1 \cup S_2]$$

As before, the number d satisfying $\dim S_1 \vee S_2 = n - d$ is called the *deficit* of the pair of subspheres, $df(S_1, S_2) = d$.

Example 2. The hyperspheres bijectively correspond to hyperbolic hyperplanes, i.e. to $(n+1)$-dimensional pseudo-Euclidean subspaces $\boldsymbol{W} \subset \boldsymbol{V}$. These again are bijectively determined by their one-dimensional orthogonal complements $\boldsymbol{W}^\perp = [\mathfrak{u}]$, $\langle \mathfrak{u}, \mathfrak{u} \rangle > 0$. The hypersphere corresponding to $\mathfrak{u}$ is denoted by $\Sigma(\mathfrak{u})$. Thus we have

$$\Sigma(\mathfrak{u}) := \pi([\mathfrak{u}]^\perp) \cap S^n. \tag{5}$$

As for the space of hyperbolic hyperplanes, the one-sheeted hyperboloid $\langle \mathfrak{u}, \mathfrak{u} \rangle = 1$ of space-like unit vectors is a double cover for the space $S_{n,n-1}$ of hyperspheres, cf. Example 6.1, Figure 2.18. Introducing *oriented hyperspheres* by means of the orientations for hyperplanes, the set of oriented hyperplanes is bijectively represented by the space-like unit vectors. Two hyperspheres $\Sigma_1 = \Sigma(\mathfrak{u}), \Sigma_2 = \Sigma(\mathfrak{w})$ have a unique invariant:

$$|I|(\Sigma_1, \Sigma_2) := \frac{|\langle \mathfrak{u}, \mathfrak{w} \rangle|}{|\mathfrak{u}||\mathfrak{w}|}, \tag{6}$$

cf. Example 6.4. For oriented hyperspheres the sign of the scalar product has to be taken into account as well:

$$I(\Sigma_1, \Sigma_2) := \frac{\langle \mathfrak{u}, \mathfrak{w} \rangle}{|\mathfrak{u}||\mathfrak{w}|}.$$

If $|I|(\Sigma_1, \Sigma_2) < 1$, then the intersection $\boldsymbol{U} \cap \boldsymbol{W}$ of the vector spaces determining them is pseudo-Euclidean and hence defines an $(n-2)$-sphere. By definition the *angle β between the intersecting hyperspheres* is equal to the angle between the hyperbolic hyperplanes describing them: $\cos\beta := I(\Sigma_1, \Sigma_2)$, $0 < \beta \leq \pi$, for oriented hyperspheres, or $\cos\beta := |I|(\Sigma_1, \Sigma_2)$, $0 < \beta \leq \pi/2$, for non-oriented ones. If Condition A is satisfied, then $|I|(\Sigma_1, \Sigma_2) = 1$ if and only if the hyperspheres are tangent to each other; then the intersection $\boldsymbol{U} \cap \boldsymbol{W}$ is isotropic and hence defines a unique point in the Möbius space S^n. Condition A is not satisfied if and only if $\Sigma_1 = \Sigma_2$. But even in this case $|I|(\Sigma_1, \Sigma_2) = 1$. In contrast to the case of different but meeting hyperspheres the deficit is now equal to 1. In the case $|I|(\Sigma_1, \Sigma_2) > 1$ the intersection $\boldsymbol{U} \cap \boldsymbol{W}$ is Euclidean, the hyperspheres do not intersect. Further below we will find a geometric interpretation for the invariant $|I|(\Sigma_1, \Sigma_2)$ also in this case. □

We now return to the general situation described in Proposition 5. This proposition refers to Proposition 6.24 and the eigenvalues (6.73); we keep the notations used there. The extremal problem with constraints for the function (6.67) now has the following geometric interpretation: *Find the relative* *extrema of the invariant $I(\Sigma_1, \Sigma_2)$ for all the pairs of hyperspheres (Σ_1, Σ_2) containing the subspheres, i.e. satisfying $S_1 \subset \Sigma_1$ and $S_2 \subset \Sigma_2$.* We discuss the cases one by one.

Case 1. $\alpha_1 > 1$. In this case $\boldsymbol{U} \cap \boldsymbol{W} = (\boldsymbol{U}^\perp + \boldsymbol{W}^\perp)^\perp$ is Euclidean or equal to $\{\mathfrak{o}\}$, the projective subspaces intersect in the outer region $A(S^n)$, and the subspheres are disjoint. In addition, the following more precise statement can be proved:

Lemma 6. *If the maximal invariant α_1 of the subspheres $S_1^l = S(\boldsymbol{U})$, $S_2^m = S(\boldsymbol{W})$ is larger than 1, then these subspheres are disjoint, and there are hyperspheres Σ_1, Σ_2 separating S_1^l, S_2^m, i.e. satisfying*

$$S_1^l \subset \Sigma_1, \ \ S_2^m \subset \Sigma_2 \text{ and } \Sigma_1 \cap \Sigma_2 = \emptyset. \tag{7}$$

Proof. The dimension of the pseudo-Euclidean subspace $\boldsymbol{U}^\perp + \boldsymbol{W}^\perp$ is not less than 2. Hence we may find a time-like vector $\mathfrak{x} = \mathfrak{u} + \mathfrak{w} \in \boldsymbol{U}^\perp + \boldsymbol{W}^\perp$ such that $\mathfrak{u} \in \boldsymbol{U}^\perp$, $\mathfrak{w} \in \boldsymbol{W}^\perp$. Consequently, for the intersection of the hyperplanes corresponding to the space-like vectors $\mathfrak{u}, \mathfrak{w}$ the subspace

$$[\mathfrak{u}]^\perp \cap [\mathfrak{w}]^\perp = ([\mathfrak{u}] + [\mathfrak{w}])^\perp$$

is Euclidean; hence the hyperplanes intersect in the outer region of S^n, and the corresponding hyperspheres Σ_1, Σ_2 satisfy (7). □

Case 2. $\alpha_1 = 1$. Assuming Condition A the subspace $\boldsymbol{U} \cap \boldsymbol{W}$ is isotropic, cf. Lemma 6.23; *thus the intersection $S_1 \cap S_2$ consists of a single point. This point is the point of contact with the n-sphere for the tangent subspace $\boldsymbol{A}^{l+1} \cap \boldsymbol{B}^{m+1}$.*

Case 3. $\alpha_1 < 1$. In this case the subspace $\boldsymbol{A}^{l+1} \cap \boldsymbol{B}^{m+1}$ is hyperbolic. Now there are, however, two situations we have to distinguish:

Case 3.a. $\alpha_1 < 1$, and $\operatorname{Dim} \boldsymbol{A}^{l+1} \cap \boldsymbol{B}^{m+1} = 0$.

Lemma 7. *Let S_1^l, $S_2^m \subset S^n$ be two subspheres with the maximal eigenvalue $\alpha_1 < 1$ and suppose that $l + m = n - 1$ for the dimensions. Then the subspheres S_1^l, S_2^m are disjoint and* linked, *i.e. $S_1 \cap S_2 = \emptyset$, and any two hyperspheres Σ_1, Σ_2 with $S_1^l \subset \Sigma_1$, $S_2^m \subset \Sigma_2$ always intersect: $\Sigma_1 \cap \Sigma_2 \neq \emptyset$.*

Proof. Applying the dimension formula shows that by Condition A we have Case 3.a. Since in this case the intersection $\boldsymbol{A}^{l+1} \cap \boldsymbol{B}^{m+1}$ is a hyperbolic point, the corresponding subspheres are disjoint. The one-dimensional vector space corresponding to the point is time-like. Consequently, the subspace $\boldsymbol{U}^\perp + \boldsymbol{W}^\perp = (\boldsymbol{U} \cap \boldsymbol{W})^\perp$ is Euclidean. Thus, the subspace generated by two unit vectors $\mathfrak{u} \in \boldsymbol{U}^\perp$, $\mathfrak{w} \in \boldsymbol{W}^\perp$ is Euclidean, i.e., its orthogonal complement is an n-dimensional pseudo-Euclidean subspace $[\mathfrak{u}, \mathfrak{w}]^\perp = [\mathfrak{u}]^\perp \cap [\mathfrak{w}]^\perp$, which determines the intersection of hyperspheres $\Sigma_1(\mathfrak{u}) \cap \Sigma_2(\mathfrak{w})$. Hence the latter is an $(n-2)$-sphere. Since $\mathfrak{u} \in \boldsymbol{U}^\perp$, $\mathfrak{w} \in \boldsymbol{W}^\perp$ are arbitrary unit vectors, this holds for arbitrary hyperspheres with $S_1^l \subset \Sigma_1$, $S_2^m \subset \Sigma_2$. □

Case 3.b. $\alpha_1 < 1$, and $k := \operatorname{Dim} \boldsymbol{A}^{l+1} \cap \boldsymbol{B}^{m+1} > 0$. *In this case the subspheres S_1^l, S_2^m intersect in a $(k-1)$-sphere.*

In each of the cases considered above the relations

$$\sqrt{\alpha_i} = \cos\beta_i, \text{ if } \alpha_i < 1, \tag{8}$$

determine by definition the stationary values β_i, $0 < \beta_i \leq \pi/2$, of the angle, under which the hyperspheres Σ_1, Σ_2 with $S_1^l \subset \Sigma_1, S_2^m \subset \Sigma_2$ intersect.

Example 3. Already in Example 2 we discussed Condition A and the Cases 1, 2, 3.b for two hyperspheres $\Sigma_1, \Sigma_2 \subset S^n$. The dimension condition in Case 3.a can only hold on the circle, i.e. for n=1; this case will be the topic of Exercise 6 and the next example. For $n > 1$ and $I(\Sigma_1, \Sigma_2) < 1$ the hyperspheres intersect; its angle β was defined by means of the invariant I in Example 2. To interpret the invariant I geometrically also in the case of non-intersecting hyperspheres, we want to express it in terms of the Euclidean invariants for the hyperspheres. Since the Möbius invariants obviously are Euclidean invariants as well — the Euclidean group is a subgroup of the Möbius group — this has to be possible. It is easy to see that the radii r, R of the hyperspheres together with the distance $d = d(\mathfrak{z}_1, \mathfrak{z}_2)$ of their centers $\mathfrak{z}_1, \mathfrak{z}_2$ form a complete system of Euclidean invariants for the pairs of hyperspheres. As already in (6.2) let $(\mathfrak{e}_i)$, $i = 0, 1, \ldots, n+1$, (recall that the dimension was increased by 1) denote a pseudo-orthonormal basis for the vector space $\boldsymbol{V}^{n+2}$. With respect to this the equation

$$-\langle \mathfrak{e}_0, \mathfrak{x} \rangle = x^0 = 1 \tag{9}$$

defines a linear subspace, which we may interpret as the point space $\boldsymbol{E}^{n+1}$ of an $(n+1)$-dimensional Euclidean geometry by means of the scalar product on the associated vector space. Then the Möbius space S^n is represented as the unit hypersphere with center $\mathfrak{e}_0$ once we normalize the representatives by (9):

$$\mathfrak{x} = \mathfrak{e}_0 + \mathfrak{x}_1, \ \mathfrak{x}_1 = \sum_{i=1}^{n+1} \mathfrak{e}_i x^i \text{ with} \langle \mathfrak{x}_1, \mathfrak{x}_1 \rangle = 1. \tag{10}$$

We identify S^n with this unit hypersphere and project it stereographically from the South Pole $\mathfrak{e}_0 - \mathfrak{e}_{n+1}$ onto its equatorial hyperplane $\boldsymbol{E}^n \subset \boldsymbol{E}^{n+1}$ spanned by the vectors $\mathfrak{e}_1, \ldots, \mathfrak{e}_n$, cf. Exercise 1, Lemma 6.3, and Exercise 6.5. Let $S(\mathfrak{z}, r) \subset \boldsymbol{E}^n$ denote the hypersphere of the Euclidean space $\boldsymbol{E}^n$ with center $\mathfrak{z}$ and radius r, and let $\Sigma(\mathfrak{z}, r) := st^{-1}(S(\mathfrak{z}, r))$ be its inverse image under the stereographic projection. By Lemma 6.3 this is a hypersphere in S^n. We show:

Lemma 8. *In the notations introduced above, according to* (5) *the hypersphere* $\Sigma(\mathfrak{z}, r) := st^{-1}(S(\mathfrak{z}, r))$ *corresponds to the space-like unit vector*

$$\mathfrak{u}(\mathfrak{z}, r) = (\mathfrak{e}_0(1 - r^2 + \langle \mathfrak{z}, \mathfrak{z} \rangle) + 2\mathfrak{z} + \mathfrak{e}_{n+1}(1 + r^2 - \langle \mathfrak{z}, \mathfrak{z} \rangle))/2r, \tag{11}$$

i.e., $\Sigma(\mathfrak{z}, r) = \Sigma(\mathfrak{u}(\mathfrak{z}, r))$.

In order to see this we denote the coordinates of the point $\mathfrak{y} \in S(\mathfrak{z}, r)$ by (y^i) and the coordinates of the center $\mathfrak{z} \in \boldsymbol{E}^n$ by (z^i), $i = 1, \dots, n$. They are related by

$$\langle \mathfrak{y} - \mathfrak{z}, \mathfrak{y} - \mathfrak{z} \rangle = \sum_{i=1}^{n} (y^i - z^i)^2 = r^2. \tag{12}$$

>From Formula (6.20) for the inverse of the stereographic projection, in which now the index $n+1$ corresponds to the South Pole $-\mathfrak{e}_{n+1}$, we compute using (10) the normalized representative of the point $\boldsymbol{x} = [\mathfrak{x}] \in \Sigma(\mathfrak{z}, r)$ to be

$$\mathfrak{x} = \mathfrak{e}_0 + \frac{2\mathfrak{y}}{1 + \langle \mathfrak{y}, \mathfrak{y} \rangle} + \mathfrak{e}_{n+1} \frac{1 - \langle \mathfrak{y}, \mathfrak{y} \rangle}{1 + \langle \mathfrak{y}, \mathfrak{y} \rangle}. \tag{13}$$

Taking into account (12) the scalar product of the vectors (11) and (13) shows that these vectors are mutually orthogonal, which implies the assertion. □

An elementary calculation together with this lemma shows

Corollary 9. *The Möbius invariant of two oriented hyperspheres $S_1(\mathfrak{z}_1, r_1)$, $S_2(\mathfrak{z}_2, r_2)$ in the n-dimensional Euclidean space is equal to*

$$I(S_1(\mathfrak{z}_1, r_1), S_2(\mathfrak{z}_2, r_2)) = \frac{r_1^2 + r_2^2 - d^2}{2 r_1 r_2}. \tag{14}$$

Here $d = |\mathfrak{z}_1 - \mathfrak{z}_2|$ denotes the distance of the centers.

The expression (14) is frequently called the Coxeter distance of the hyperspheres, even though it, obviously, does not have the usual properties of a metric; therefore, we prefer to call it the *Coxeter invariant* for the hyperspheres. An elementary geometric consideration shows that, in the case of intersecting hyperspheres, the Coxeter invariant again determines the cosine of the intersection angle. □

Exercise 4. Derive formula (11) for $\mathfrak{u}(\mathfrak{z}, r)$ by taking the vector product of $n+1$ isotropic vectors in the pseudo-Euclidean vector space $\boldsymbol{V}^{n+2}$ which represent points in general position on the hypersphere $\Sigma^{n-1} \subset S^n$.

Exercise 5. Let $\langle \mathfrak{b}, \mathfrak{y} \rangle = p$ be the equation for a hyperplane $\boldsymbol{H}^{n-1}$ in the Euclidean space $\boldsymbol{E}^n, n > 1$. By Lemma 6.3 its inverse image under the stereographic projection is a hypersphere through the South Pole. Prove that this hypersphere is the one defined in (5) as

$$\Sigma(\mathfrak{v}) \text{ with } \mathfrak{v} = \frac{\mathfrak{b} + (\mathfrak{e}_0 - \mathfrak{e}_{n+1})p}{|\mathfrak{b}|}. \tag{15}$$

Use this to verify that the invariant $|I|(st^{-1}(\boldsymbol{H}_1), st^{-1}(\boldsymbol{H}_2))$ for two intersecting hyperplanes is equal to the cosine of their intersection angle; for parallel hyperplanes it is equal to one. An analogous fact can be proved for a pair consisting of a hyperplane $\boldsymbol{H}^{n-1}$ and a hypersphere $S^{n-1}(\boldsymbol{z}, r)$; in this case the invariant is $|I| = z/r$, where z denotes the distance from the center $\mathfrak{z}$ to the hyperplane.

Exercise 6. Show by examples that the Cases 1, 2, 3.a may occur for two 0-spheres in S^1; of course, Case 3.b is impossible there. In Case 3.a the invariant for the point pairs is equal to the cosine of the intersection angle between their generating hyperbolic lines. Figure 2.46 shows two orthogonal point pairs on a circle embedded into the Euclidean space E^3.

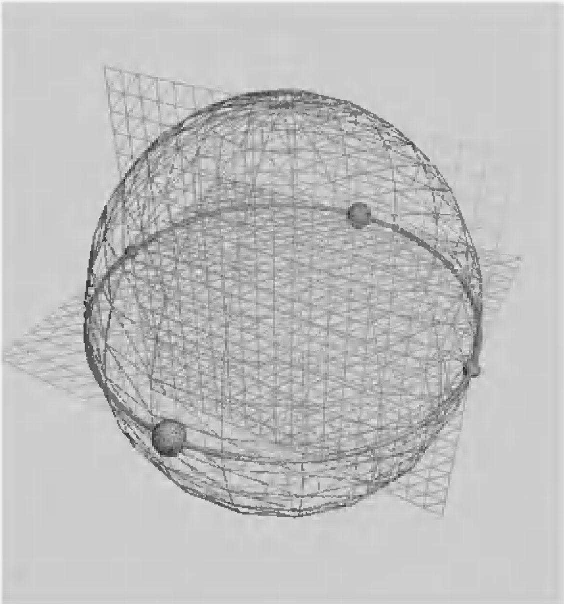


Fig. 2.46. Eigenspheres of orthogonal point pairs in $S^1 \subset S^3$.

Exercise 7. Find an explicit example for the configuration in Figure 2.46, i.e. two orthogonal 0-spheres in a circle contained in the three-dimensional sphere S^3. (Note that the the ambient Euclidean space E^3 in the figure now has to be interpreted as the sphere S^3.) For this configuration the deficit is equal to 2; the maximal eigenvalue is $\alpha_1 = 1$, it has multiplicity 2; the *eigenspheres* for the eigenvalue 1 form the pencil of spheres through the circle S^1 (cf. Figure 2.47), whereas the eigensphere for the eigenvalue 0 intersects all spheres in this pencil orthogonally. (Hint. Eigenvalues refer to the operator A defined in (6.70), which in the case considered here acts in a three-dimensional Euclidean vector space; the eigenspheres are defined by $\Sigma(\mathfrak{u})$, $\mathfrak{u}$ an eigenvector of A.)

In general, we call the hyperspheres $\Sigma(\mathfrak{w})$ for the eigenvectors $\mathfrak{w}$ of the operator A defined in Proposition 4 the *eigenhyperspheres* of the pair of spheres (S_1, S_2). Each eigenhypersphere $\Sigma_2 := \Sigma(\mathfrak{w})$ contains S_2; if the corresponding eigenvalue is $\alpha > 0$, then the projection $\mathfrak{u} := p(\mathfrak{w}) \in U^\perp$ defines a hypersphere $\Sigma_1 := \Sigma(\mathfrak{u}) \supset S_1$ for which the stationary invariant $\sqrt{\alpha} = |I|(\Sigma_1, \Sigma_2)$

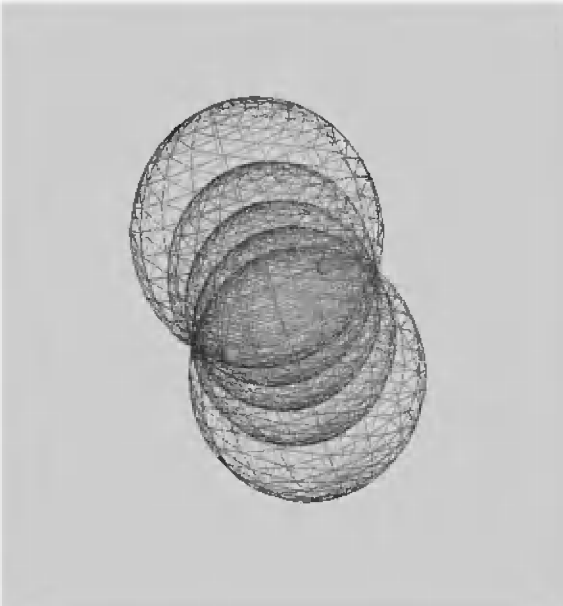


Fig. 2.47. The sphere pencil of a circle

is realized. If, however, $\alpha = 0$, then Σ_2 is orthogonal to all hyperspheres Σ containing S_1 (cf. Proposition 6.24).

Example 4. Let us again consider the one-dimensional Möbius space, i.e. the circle S^1. Here the hyperspheres are the 0-spheres, i.e. point pairs $S_1 = \{\boldsymbol{s}, \boldsymbol{n}\} \subset S^1$ with $\boldsymbol{s} \neq \boldsymbol{n}$. According to Proposition 4.7 the circle is equipped with the structure of a projective line; thus for each throw of points in S^1 the *cross ratio* CR is defined in an invariant way. Since the Möbius group $G_1 \subset \boldsymbol{PL}_2(\mathbf{R})$ consists of all projectivities of the plane leaving S^1 invariant, the CR is also invariant under the action of the Möbius group. As the points from two different 0-spheres determine a throw, it is natural to ask, how the CR of this throw is related to the Möbius invariant $|I|$. To describe this relation we use the notations and constructions from the proof of Proposition 4.7. So we represent the points $\boldsymbol{s}, \boldsymbol{n}$ by isotropic vectors $\mathfrak{s}, \mathfrak{n}$ and adapt the pseudo-orthonormal basis $(\mathfrak{c}_i)$, $i = 0, 1, 2$, to the point pair so that

$$\langle \mathfrak{s}, \mathfrak{n} \rangle = -1/2, \tag{16}$$

$$\mathfrak{c}_0 = \mathfrak{s} + \mathfrak{n}, \ \mathfrak{c}_1 = \mathfrak{s} - \mathfrak{n}, \ \mathfrak{c}_2 = \mathfrak{c}_0 \times \mathfrak{c}_1. \tag{17}$$

Because of

$$\langle \mathfrak{s}, \mathfrak{c}_2 \rangle = \langle \mathfrak{n}, \mathfrak{c}_2 \rangle = 0,$$

the tangents $\boldsymbol{T_s}, \boldsymbol{T_n}$ intersect in the outer point

$$\boldsymbol{T_s} \cap \boldsymbol{T_n} = \boldsymbol{b} := [\mathfrak{c}_2]$$

defined by $\mathfrak{c}_2$.

Consider the projective scale ξ on the tangent $\boldsymbol{T_s}$ defined by the vectors $\mathfrak{s}, \mathfrak{c}_2$ and transfer it by stereographic projection onto S^1. By (4.57) we thus obtain a parameter representation for S^1 assigning to a number, i.e. scale value, the corresponding point:

$$f : \xi \in \hat{\mathbf{R}} \longmapsto f(\xi) = [\mathfrak{x}(\xi)] := [\mathfrak{s} + \mathfrak{c}_2 \xi + \mathfrak{n} \xi^2]. \tag{18}$$

Take now a second 0-sphere:

$$S_2 = \{\boldsymbol{x}_1, \boldsymbol{x}_2\}, \boldsymbol{x}_i = [\mathfrak{x}(\xi_i)], i = 1, 2.$$

Using Formula (1.4.16) we compute the CR

$$\eta := (\boldsymbol{s}, \boldsymbol{b};\, \boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{\xi_1}{\xi_2}. \tag{19}$$

On the other hand, for the space-like unit vector determining the 0-sphere S_2 we obtain

$$\mathfrak{w} = \frac{\mathfrak{x}_1 \times \mathfrak{x}_2}{|\mathfrak{x}_1 \times \mathfrak{x}_2|} = \frac{2\mathfrak{s} + 2\mathfrak{n}\xi_1\xi_2 + \mathfrak{c}_2(\xi_1 + \xi_2)}{\xi_1 - \xi_2}. \tag{20}$$

Relation (19) for the invariant $|I|$ and its expression in terms of the CR then lead to the result we were looking for:

$$|I|(S_1, S_2) = |\langle \mathfrak{c}_2, \mathfrak{w} \rangle| = |\frac{\eta + 1}{\eta - 1}|. \tag{21}$$

Note that the CR depends on the chosen order of the points $\boldsymbol{x}_1, \boldsymbol{x}_2$; the other order, however, does not lead to a new value for $|I|$. Considering oriented 0-spheres, i.e. ordered point pairs, reversing the orientation results in a sign change for I and η, so that we might as well write

$$I(S_1, S_2) = \langle \mathfrak{c}_2, \mathfrak{w} \rangle = \frac{\eta + 1}{\eta - 1}. \tag{22}$$

Considering the inverse relation and taking into account (19) we obtain for $I = I(S_1, S_2)$ the same functional dependence

$$\eta = \frac{I + 1}{I - 1}. \tag{23}$$

These formulas will now be applied to prove that the real projective group $\boldsymbol{PL}_1$ coincides with the Möbius group. □

Proposition 10. *The projective group $\boldsymbol{PL}_1(\mathbf{R})$ for the projective structure defined on the circle S^1 according to Proposition 4.7 is equal to the Möbius group G_2.*

Proof. Let $(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4)$ be a harmonic throw in S^1; thus the CR η of this throw is equal to -1. Then by (22) the invariant I of the oriented 0-spheres $S_1 = (\boldsymbol{a}_1, \boldsymbol{a}_2)$, $S_2 = (\boldsymbol{a}_3, \boldsymbol{a}_4)$ is equal to zero. The same also holds for the images of these 0-spheres under an arbitrary Möbius transformation $g \in G_1$. So by relation (23) the images also form a harmonic throw. According to Staudt's Main Theorem, Proposition 1.4.7, g is a projectivity. Conversely, let $g \in \boldsymbol{PL}_1$ be a projectivity, and let $\boldsymbol{x}_i \in S^1$, $i = 0, 1, 2$, be points in general position, i.e. a projective frame for S^1. Let $\boldsymbol{y}_i := g(\boldsymbol{x}_i)$ denote their images under the projectivity g. We prove:

Lemma 11. *If $(\boldsymbol{x}_i)$, $(\boldsymbol{y}_i)$, $i = 0, 1, 2$, are two triples of points in general position in the Möbius space S^n, then there always is a Möbius transformation $h \in G_{n+1}$ such that $h\boldsymbol{x}_i = \boldsymbol{y}_i$. In the case $n = 1$ the transformation h is uniquely determined.*

Proof. Each of the point triples uniquely determines a circle. Since G_{n+1} acts transitively on the set of circles by Corollary 2, it suffice to consider the case $n = 1$. Let $\mathfrak{s}, \mathfrak{n}, \mathfrak{x} \in \boldsymbol{V}^3$ be isotropic vectors representing the points of the triple $(\boldsymbol{x}_i)$,

$$\boldsymbol{x}_0 = [\mathfrak{s}], \boldsymbol{x}_1 = [\mathfrak{n}], \boldsymbol{x}_2 = [\mathfrak{x}].$$

Take, moreover, a pseudo-orthonormal basis $(\mathfrak{c}_i)$ for $\boldsymbol{V}$ such that (16) and (17) hold. Obviously, the representatives of the points and the basis vectors are not uniquely determined by these conditions, yet. It is just the vector $\mathfrak{c}_2$ that is uniquely determined, if we suppose that the point pair $(\boldsymbol{x}_0, \boldsymbol{x}_1)$ is oriented; in fact,

$$\mathfrak{s} = (\mathfrak{c}_0 + \mathfrak{c}_1)/2, \quad \mathfrak{n} = (\mathfrak{c}_0 - \mathfrak{c}_1)/2, \quad \mathfrak{s} \times \mathfrak{n} = -\mathfrak{c}_2/2. \tag{24}$$

Now consider the basis representation

$$\mathfrak{x} = \mathfrak{s}\xi_1 + \mathfrak{n}\xi_2 + \mathfrak{c}_2\xi_3.$$

As $\mathfrak{x}$ is isotropic and the three points are in general position, all three components ξ_i have to be different from zero. Dividing $\mathfrak{x}$ by ξ_3 we obtain the representative for $\boldsymbol{x}_2$ that is uniquely determined by $\xi_3 = 1$. Since it is isotropic, we obtain

$$\xi_1\xi_2 = 1.$$

Next we look at the transformations still possible without violating the present conditions

$$\hat{\mathfrak{s}} = \mathfrak{s}\lambda, \quad \hat{\mathfrak{n}} = \mathfrak{n}/\lambda, \quad (\lambda \neq 0),$$

with $\lambda = \xi_1$. Because of the previous equation, we see that the additional condition

$$\mathfrak{x} = \mathfrak{s} + \mathfrak{n} + \mathfrak{c}_2 = \mathfrak{c}_0 + \mathfrak{c}_2 \tag{25}$$

uniquely determines the representatives $\mathfrak{s}, \mathfrak{n}, \mathfrak{x}$ for the points of the triple, and hence by (24) the basis $(\mathfrak{c}_i)$ is also uniquely determined by the point triple.

Analogously, the point triple $(\boldsymbol{y}_i)$ uniquely determines a pseudo-orthonormal basis $(\mathfrak{b}_i)$. The transformation $h \in G_1$ defined by the pseudo-orthogonal transformation $\gamma \in \mathbf{O}(1,2)$ with

$$\gamma(\mathfrak{c}_i) = \mathfrak{b}_i, \; i = 0, 1, 2,$$

obviously satisfies $h\boldsymbol{x}_i = \boldsymbol{y}_i$. Since, moreover, each of these transformations maps the uniquely determined bases into one another, γ and, hence also h, are uniquely determined. □

Now we conclude the proof of Proposition 10. Let $h \in G_2$ be the Möbius transformation that is uniquely determined by $h(\boldsymbol{x}_i) = \boldsymbol{y}_i$ according to Lemma 11. From what we already proved h, and thus also h^{-1}, is a projectivity of S^1. Hence the map $f := h^{-1} \circ g$ is a projectivity as well. Since this leaves the points $\boldsymbol{x}_i$ fixed, i.e. $f(\boldsymbol{x}_i) = \boldsymbol{x}_i$ for $i = 0, 1, 2$, by Proposition 1.3.14 f has to be the identity. It suffices to choose theses points as the base points and the unit point of a projective scale on S^1. Hence $g = h$ is a Möbius transformation. □

Exercise 8. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4 \in S^1$ be four different points. a) Prove the following identities for the Möbius invariants of the oriented 0-spheres determined by them: Setting

$$I((\boldsymbol{x}_1, \boldsymbol{x}_2), (\boldsymbol{x}_3, \boldsymbol{x}_4)) = \lambda$$

the following equations hold

$$\begin{aligned}
I((\boldsymbol{x}_1, \boldsymbol{x}_2), (\boldsymbol{x}_4, \boldsymbol{x}_3)) &= -\lambda, \\
I((\boldsymbol{x}_1, \boldsymbol{x}_3), (\boldsymbol{x}_2, \boldsymbol{x}_4)) &= \frac{3-\lambda}{\lambda+1}, \\
I((\boldsymbol{x}_1, \boldsymbol{x}_3), (\boldsymbol{x}_4, \boldsymbol{x}_2)) &= \frac{\lambda-3}{\lambda+1}, \\
I((\boldsymbol{x}_1, \boldsymbol{x}_4), (\boldsymbol{x}_3, \boldsymbol{x}_2)) &= \frac{\lambda+3}{\lambda-1}, \\
I((\boldsymbol{x}_1, \boldsymbol{x}_4), (\boldsymbol{x}_2, \boldsymbol{x}_3)) &= \frac{\lambda+3}{1-\lambda}.
\end{aligned}$$

These imply

$$\begin{aligned}
I((\boldsymbol{x}_1, \boldsymbol{x}_2), (\boldsymbol{x}_3, \boldsymbol{x}_4)) &= I((\boldsymbol{x}_3, \boldsymbol{x}_4), (\boldsymbol{x}_1, \boldsymbol{x}_2)) \\
= I((\boldsymbol{x}_2, \boldsymbol{x}_1), (\boldsymbol{x}_4, \boldsymbol{x}_3)) &= I((\boldsymbol{x}_4, \boldsymbol{x}_3), (\boldsymbol{x}_2, \boldsymbol{x}_1)).
\end{aligned}$$

Compare these equations with the identities proved in Section 1.4 for the CR $(\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{x}_3, \boldsymbol{x}_4)$. – b) Under these conditions none of the right-hand sides of these equations can be equal to one. – c) The throw $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4)$ is harmonic if and only if $\lambda = 0$, i.e., if the point pairs are orthogonal.

Example 5. Consider now the case $n = 2$, i.e. the sphere S^2 under the action of the Möbius group G_3. Already in Example 1.1.3 we identified the complex line $\boldsymbol{P}^1_{\mathbb{C}}$ with the *Riemann sphere* S^2. By Example 1.4.3 the continuous projectivities and anti-projectivities on $\boldsymbol{P}^1_{\mathbb{C}}$ induce the automorphism group $\operatorname{Aut} \boldsymbol{P}^1_{\mathbb{C}}$, which, in a projective scale, is represented by fractional linear and conjugate fractional linear transformations. At the same time, this group is the Möbius group G_3 of the sphere S^2, as we will soon show by interpreting the calculus of complex numbers geometrically. Historically, for the first time the Möbius group occurred in this very form. An instructive presentation of this complex form of the Möbius group G_3 can be found in Chapter 2 of the introductory textbook [27] by C. Carathéodory. □

Proposition 12. *The Möbius group G_3 of the Riemann sphere $S^2 = \boldsymbol{P}^1_{\mathbf{C}}$ as identified with the projective line is the group of continuous automorphisms* $\operatorname{Aut} \boldsymbol{P}^1_{\mathbf{C}}$.

Proof. As in Example 1.4.3 we represent the group $\operatorname{Aut} \boldsymbol{P}^1_{\mathbf{C}}$ by the *fractional linear transformations*:

$$w = f(z) = \frac{az+b}{cz+d} \quad \text{or} \tag{26}$$

$$w = f(\bar{z}) = \frac{a\bar{z}+b}{c\bar{z}+d} \tag{27}$$

with $ad - cb \neq 0$, where we have to set

$$f(-d/c) = \infty \text{ and } f(\infty) = a/c \text{ for } c \neq 0,\ f(\infty) = \infty \text{ otherwise.} \tag{28}$$

The totality of these transformations forms a group; in the proof and for brevity we will simply denote it by G. So we have to show that $G = G_3$. The transformations (26) constitute a subgroup, which we already encountered considering the transformations of projective scales on a line (cf. Example 1.2.2 and Exercise 1.2.4); as a homomorphic image of the linear group $\boldsymbol{GL}(2, \mathbf{C})$ it is isomorphic to the group $\boldsymbol{PL}_1(\mathbf{C})$ of projectivities on the complex projective line: The coefficients a, b, c, d in (26) are just the entries of the matrix $A \in \boldsymbol{GL}(2, \mathbf{C})$, cf. (1.2.10). Hence their determinant, $ad - cb$, is always different from zero. Multiplying numerator and denominator by the same number $k \in \mathbf{C}^*$ the map f defined by (26) or (27), respectively, does not change; if required, we may thus always suppose

$$ad - cb = 1, \text{ i.e. } A \in \boldsymbol{SL}(2, \mathbf{C}).$$

According to Example 1.4.5 or (1.7) this implies that the set of all transformations $f : S^2 \to S^2$ which can be represented in the form (26) is the group of projectivities $\boldsymbol{PL}_1(\mathbf{C}) \cong \boldsymbol{SL}(2, \mathbf{C})/\{\pm I_2\}$. Still adding conjugation and taking (27) into account we generate the group $\operatorname{Aut} \boldsymbol{P}_1(\mathbf{C})$. Obviously, the maps (26), (27) satisfy $f(\infty) = \infty$ if and only if $c = 0$; then we may represent them in the form

$$w = az + b \text{ or } w = a\bar{z} + b \text{ with } a \neq 0.$$

But this is the conformal Euclidean group of the complex plane, which again can be identified with the Euclidean plane, cf. Section I.2.3. Using this, multiplication by a non-zero number becomes a rotation composed with a dilation, addition corresponds to a translation, and the assignment $z \to \bar{z}$ is the reflection in the real axis. Obviously, each similarity $g \in CE(2)$ can be represented by one of these transformations f. For $c \neq 0$ the map $s : z \in S^2 \to 1/\bar{z}$ is the inversion in the unit circle with center 0:

Exercise 9. a) Verify the last assertion. – b) Prove that each transformation $f \in G$ can be represented as the composition of a similarity with the inversion s in the unit circle. (Note that s may occur several times in this composition.) – c) Prove that the map

$$h(z) := \frac{m\bar{z} - m\bar{m} + r^2}{\bar{z} - \bar{m}} = m + \frac{r^2}{\bar{z} - \bar{m}} \text{ with } r > 0 \tag{29}$$

is the inversion in the circle with center $m \in \mathbf{C}$ and radius $r > 0$ in the complex number plane. – d) Find an analogous formula for the reflection in the real line $y = at + b$, $a \neq 0$, $t \in \mathbf{R}$, in the complex plane.

With these calculations at our disposal we can now complete the proof of the proposition: Statements a) and b) together with Exercise 1 and Example 1 imply the inclusion $G \subset G_3$. Conversely, according to Proposition 4 each Möbius transformation $g \in G_3$ can be represented as a finite product of inversions, which are elements of G by Exercise 9, c), d). Since G is a group, we conclude $G_3 \subset G$. □

Exercise 10. Let $x_i, i = 1, \ldots, 4$, be four different points in $S^2 = \boldsymbol{P}^1_{\mathbf{C}}$. Prove that the four points lie on a common line or circle in $\mathbf{C}$ if and only if their cross ratio $(\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{x}_3, \boldsymbol{x}_4)$ is real.

Corollary 13. *The group of fractional linear transformations f defined by* (26) *consists of all orientation preserving Möbius transformations of the Riemann sphere S^2. This implies the following isomorphy*

$$\boldsymbol{SL}(2, \mathbf{C})/\{\pm I_2\} \cong \boldsymbol{SO}(1, 3).$$

Proof. It is not difficult to show that each transformation of the form (26) may be represented as the composition of maps from the following list:

1. Translation: $t_b : z \longmapsto z + b$.
2. Rotation with subsequent dilation: $d_a : z \longmapsto za$, $a \in \mathbf{C}^*$.
3. Taking the reciprocal: $h : z \longmapsto 1/z$.

All these maps are orientation preserving. For the translations this is obvious; in the second case, using $a = |a|\,\mathrm{e}^{\mathrm{i}\,\alpha}$, $z = |z|\,\mathrm{e}^{\mathrm{i}\,\varphi}$, we obtain:

$$d_a(z) = |a||z|\,\mathrm{e}^{\mathrm{i}(\varphi+\alpha)},$$

and this is the rotation through the angle α followed by the dilation with factor |a|. Finally, forming the reciprocal is the product of two reflections: $h(z) = \overline{1/\bar{z}}$ is the composition of the inversion in the unit circle, $s(z) = 1/\bar{z}$, with subsequent reflection in the real axis. Since each reflection has determinant -1 (cf. Example 1), this proves our assertion. Hence, cf. Proposition 4, every Möbius transformation of the form (26) can be represented as the product of an even number of reflections.[2] Now transformations of the form (27) are obtained from those of (26) by composition with the reflection in the real axis from the right; hence they can always be represented as the composition of an odd number of reflections. Thus they reverse the orientation. Therefore, the maps of the form (26) and (27) each constitute one of the components of the group G_3. According to the Examples 1.4.3 and 1.4.5 the group described by (26) is just the group $\boldsymbol{SL}(2, \mathbf{C})/\{\pm I_2\}$; on the other hand, according to Proposition 12 and what was explained at the end of Section 1 it coincides with $\boldsymbol{SO}(1,3)$. An explicit formula for the homomorphism $F : \boldsymbol{SL}(2, \mathbf{C}) \to \boldsymbol{SO}(1,3)$ in terms of a pseudo-orthogonal matrices whose entries are functions of the complex parameters a, b, c, d can be found in J. B. Wilker [111], Theorem 10. Note that in this paper the last vector in a pseudo-orthonormal frame is always time-like, i.e. its norm square is equal to -1.[3] □

2.7.3 Cross Ratios and the Riemann Sphere

In the preceding section we identified the Möbius groups of the circle S^1 and the sphere S^2 with the projective groups $\boldsymbol{PL}_1(\mathbf{R})$ and $\operatorname{Aut} \boldsymbol{P}^1_{\mathrm{C}}$, respectively. This implies (cf. Examples 4, 5 and Propositions 10, 12):

Corollary 14. *The cross ratio of four points on the circle S^1 is a Möbius invariant. On the Riemann sphere $S^2 = \boldsymbol{P}^1_{\mathrm{C}}$, the cross ratio of four points is preserved by each orientation preserving Möbius transformation, and it is transformed into its conjugate by every orientation reversing Möbius transformation.* □

For higher dimensions, $n > 2$, the projective and the Möbius groups differ substantially. nevertheless, even in the n-dimensional case the Möbius groups can be characterized as the invariance groups for suitably defined cross ratios. This method goes back to J. B. Wilker; in Section 6 of his paper [111] he actually calls this quantity the fundamental invariant of Möbius geometry. It is defined as follows: Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4$ be four different points in the n-dimensional Euclidean space. As the *absolute cross ratio* we take the expression

$$|\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{x}_3, \boldsymbol{x}_4| := \frac{|\boldsymbol{x}_1 - \boldsymbol{x}_3|}{|\boldsymbol{x}_1 - \boldsymbol{x}_4|} : \frac{|\boldsymbol{x}_2 - \boldsymbol{x}_3|}{|\boldsymbol{x}_2 - \boldsymbol{x}_4|}, \tag{30}$$

[2] In the textbook [27] C. Caratheodory proves that each transformation of the form (26) can be represented as the product of two or four inversions.

[3] The reader can find Wilker's matrices together with the necessary background form computer-algebra in the Mathematica notebook riemsph.nb at http://www-irm.mathematik.hu-berlin.de/~sulanke.

where $|\boldsymbol{x}-\boldsymbol{y}|$ denotes the norm of the position vector $\boldsymbol{x}-\boldsymbol{y}$, i.e. the Euclidean distance of the points $\boldsymbol{x}, \boldsymbol{y}$. Taking in Formula (1.4.15) the absolute value we see that Definition (30) coincides with the value of the cross ratio computed there. Obviously, the cross ratio depends continuously on its arguments; letting in (30) one of the points, e.g. $\boldsymbol{x}_4$, tend to ∞ we obtain

$$|\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{x}_3, \infty| := \lim_{\boldsymbol{x}_4 \to \infty} \frac{|\boldsymbol{x}_1 - \boldsymbol{x}_3|}{|\boldsymbol{x}_1 - \boldsymbol{x}_4|} : \frac{|\boldsymbol{x}_2 - \boldsymbol{x}_3|}{|\boldsymbol{x}_2 - \boldsymbol{x}_4|} = \frac{|\boldsymbol{x}_1 - \boldsymbol{x}_3|}{|\boldsymbol{x}_2 - \boldsymbol{x}_3|}. \tag{31}$$

Hence, taking into account the symmetry in the definition, the absolute CR is also defined if one of the points is equal to ∞. Thus

Proposition 15. *An injective map f of the n-dimensional Euclidean space extended by the point ∞ to form the n-sphere S^n is a Möbius transformation if and only if it preserves the absolute* CR *for quadruples of arbitrary points.*

Proof. We first show the invariance of the absolute cross ratio. By Proposition 10, in the case $n = 1$ the cross ratio itself remains invariant. Now each Möbius transformation can be represented as the product of finitely many inversions by Proposition 4. As reflections in hyperplanes are even isometries, it suffice to show that the absolute cross ratio is preserved by any inversion s in a hypersphere. According to Formula (4) the distance of the images of two points $\boldsymbol{x}, \boldsymbol{y}$ is computes as

$$|s(\mathfrak{x}) - s(\mathfrak{y})| = \frac{r^2}{|\mathfrak{x}||\mathfrak{y}|}|\mathfrak{x} - \mathfrak{y}|.$$

Here $\mathfrak{x}, \mathfrak{y}$ are the position vectors of the points $\boldsymbol{x}, \boldsymbol{y}$ with respect to the center of the hypersphere defining the inversion. A straightforward calculation shows that (30) implies the invariance of the absolute cross ratio for points different from ∞. Then, taking as above the limit $\boldsymbol{y} = \boldsymbol{x}_4 \to \infty$, shows that the assertion also holds, if one of the points is ∞. Conversely, let now $f : S^n \to S^n$ be an injective map preserving the absolute cross ratio and consider $\boldsymbol{y} = f(\boldsymbol{x})$. Since the Möbius group acts transitively on S^n, we find Möbius transformations $s, t \in G_{n+1}$ such that $s(\infty) = \boldsymbol{x}$, $t(\boldsymbol{y}) = \infty$. Since, by the above, Möbius transformations preserve the absolute cross ratio, the same is also true of the map $f_1 := t \circ f \circ s$. Moreover, this map has the property $f_1(\infty) = \infty$ and hence maps the Euclidean space into itself. Consider now an arbitrary triangle $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ and its image under f_1. Because of the invariance of the absolute cross ratio, by (31) the image points $\boldsymbol{y}_i = f_1(\boldsymbol{x}_i)$, $i = 1, 2, 3$ satisfy:

$$|\boldsymbol{y}_1, \boldsymbol{y}_2; \boldsymbol{y}_3, \infty| = \frac{|\boldsymbol{y}_1 - \boldsymbol{y}_3|}{|\boldsymbol{y}_2 - \boldsymbol{y}_3|} = |\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{x}_3, \infty| = \frac{|\boldsymbol{x}_1 - \boldsymbol{x}_3|}{|\boldsymbol{x}_2 - \boldsymbol{x}_3|}.$$

This implies

$$\frac{|\boldsymbol{y}_1 - \boldsymbol{y}_3|}{|\boldsymbol{x}_1 - \boldsymbol{x}_3|} = \frac{|\boldsymbol{y}_2 - \boldsymbol{y}_3|}{|\boldsymbol{x}_2 - \boldsymbol{x}_3|}.$$

Thus the image triangle $\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3$ is similar to the original triangle $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$; and since this holds for an arbitrary triangle, the map f_1 has to be a similarity; so it is a Möbius transformation. Therefore, $f = t^{-1} \circ f_1 \circ s^{-1}$ is a Möbius transformation as well. □

Remark. Note that we did not suppose f to be surjective in the proof of Proposition 15; hence the surjectivity is a consequence of injectivity together with the invariance of the absolute cross ratio for maps on the sphere. For $n = 1$, starting from the last formula a straightforward computation shows that f_1 is an affine map of the Euclidean line. We recommend that the reader proves the next corollary as an exercise; otherwise, see the paper [111] by J. B. Wilker:

Corollary 16. *Consider two families $M_1 = (\boldsymbol{x}_i)_{i\in I}$, $M_2 = (\boldsymbol{y}_i)_{i\in I}$ of points in the Möbius space S^n. Then there is a Möbius transformation $g \in G_{n+1}$ satisfying $g(\boldsymbol{x}_i) = \boldsymbol{y}_i$ for all $i \in I$ if and only if the absolute cross ratios of any four corresponding points coincide, i.e.*

$$|\boldsymbol{y}_{i_1}, \boldsymbol{y}_{i_2}; \boldsymbol{y}_{i_3}, \boldsymbol{y}_{i_4}| = |\boldsymbol{x}_{i_1}, \boldsymbol{x}_{i_2}; \boldsymbol{x}_{i_3}, \boldsymbol{x}_{i_4}|$$

for all index quadruples whose associated points in M_1 are different. If such a map g exists, then it is uniquely determined if and only if $[M_1] = S^n$, i.e. if the family M_1 is not contained in a hypersphere or hyperplane. □

To prove uniqueness we may apply Proposition 4.1 and Corollary 4.3. In the following example we refine the criterion in the corollary for the dimensions $n = 1, 2$.

Example 6. Let ξ be a projective scale on $S^1 = \boldsymbol{P}^1(\mathbf{R})$ or $S^2 = \boldsymbol{P}^1(\mathbf{C})$. Then by Definition 1.4.1, the cross ratio satisfies $\xi = (\xi, 1; 0, \infty)$, where the points are described by their parameter values. This implies the following identities for the absolute cross ratio

$$|\xi, 1; 0, \infty| = |-\xi, 1; 0, \infty| = |\bar{\xi}, 1; 0, \infty| = |-\bar{\xi}, 1; 0, \infty| = |\xi|,$$

where, obviously, the last two relations are only relevant for non-real parameters ξ. Looking at the first identity we see that, different than for the original cross ratio, it is not enough to know the absolute cross ratio for only one permutation of the four different points to ensure Möbius equivalence of the quadruples. In fact, the quadruple $(\xi, 1, 0, \infty)$ is not equivalent to the quadruple $(-\xi, 1, 0, \infty)$, since the cross ratios are different and, in the case $K = \mathbf{C}$, are not conjugate either, if $\mathrm{Re}(\xi) \neq 0$. Proposition 1.4.1, (1.4.12), and (1.4.14) imply that the cross ratios of all permutations of the quadruple $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4)$ are determined once the cross ratio of one of them is known. Note that this is false for the absolute cross ratio; it is easy to find examples showing that $|\xi| = |\eta|$ does, in general, not imply $|1-\xi| = |1-\eta|$. Nevertheless, if both the equations

$$|\boldsymbol{y}_1, \boldsymbol{y}_2; \boldsymbol{y}_3, \boldsymbol{y}_4| = |\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{x}_3, \boldsymbol{x}_4| \tag{32}$$

$$|\boldsymbol{y}_1, \boldsymbol{y}_3; \boldsymbol{y}_2, \boldsymbol{y}_4| = |\boldsymbol{x}_1, \boldsymbol{x}_3; \boldsymbol{x}_2, \boldsymbol{x}_4| \tag{33}$$

hold, then according to Proposition 1.4.1 the corresponding equations are also satisfied by the absolute cross ratios for all corresponding permutations of the quadruples. Using Corollary 16 we thus conclude that the quadruples $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4)$ and $(\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3, \boldsymbol{y}_4)$ are Möbius equivalent. Hence it suffices to verify conditions (32) and (33) for a single arrangement of any four different points in M_1. Note also the following exercise. The first case, $\xi = \eta$, occurs if both the quadruples are mapped into one another by a projective transformation (26), i.e. one preserving the cross ratio as well as the orientation, whereas the second case $\xi = \bar{\eta}$ means that both are related by a Möbius transformation of S^2 reversing the orientation, hence one of the form (27); in the following we suppose $ad - bc = 1$. □

Exercise 11. Consider $\xi, \eta \in \mathbf{C}$ and prove that the relations $|\xi| = |\eta|$ as well as $|1 - \xi| = |1 - \eta|$ both hold if and only if $\xi = \eta$ or $\xi = \bar{\eta}$.

Exercise 12. Let $\boldsymbol{x}_i \in S^2, i = 1, \ldots, 4$, be four different points on the sphere. a) Prove that there is a Möbius transformation $g \in G_3$ leaving the points $\boldsymbol{x}_1, \boldsymbol{x}_4$, fixed and interchanging the others, $g(\boldsymbol{x}_2) = \boldsymbol{x}_3$, $g(\boldsymbol{x}_3) = \boldsymbol{x}_2$, if and only if

$$|\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{x}_3, \boldsymbol{x}_4| = |\boldsymbol{x}_1, \boldsymbol{x}_3; \boldsymbol{x}_2, \boldsymbol{x}_4|. \tag{34}$$

– b) Take the four points to be in general position. Denote by $s : S^2 \to S^2$ the inversion in the circle through $\boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4$, and $\hat{\boldsymbol{x}}_1 = s(\boldsymbol{x}_1)$. If (34) holds, then the circles Σ_1 through $\boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4$ and Σ_2 through $\boldsymbol{x}_1, \hat{\boldsymbol{x}}_1, \boldsymbol{x}_4$ intersect orthogonally. The points $\boldsymbol{x}_2, \boldsymbol{x}_3$ are transformed into one another by the inversion in the circle Σ_2.

The relation between the cross ratio of four points in the Riemann sphere and the stationary invariants for the pair of 0-spheres determining the cross ratio as defined in Section 2 is described in the following proposition:

Proposition 17. *Let $S_1 = (\boldsymbol{x}_1, \boldsymbol{x}_2)$, $S_2 = (\boldsymbol{x}_3, \boldsymbol{x}_4) \subset S^2$ be two different 0-spheres in the Riemann sphere, and let $\eta = (\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{x}_3, \boldsymbol{x}_4)$ be the cross ratio they determine. Then the stationary invariants for the pair of spheres defined in Section 2 are*

$$\alpha_1(S_1, S_2) = \frac{(|\eta| + 1)^2}{|\eta - 1|^2} \geq \alpha_2(S_1, S_2) = \frac{(|\eta| - 1)^2}{|\eta - 1|^2}. \tag{35}$$

Moreover,

1. *α_1 and α_2 are independent of the orientations in the 0-spheres.*
2. *$\alpha_1 = \alpha_2$ if and only if $S_1 \cap S_2$ consists of a single point; then $\alpha_1 = \alpha_2 = 1$.*
3. *$\alpha_1 = 1 > \alpha_2$ if and only if S_1, S_2 lie in a common circle S^1 and separate each other.*

4. $\alpha_1 > \alpha_2 = 1$ *if and only if* S_1, S_2 *lie in a common circle* S^1 *and are non-separating.*
5. $\alpha_1 > 1 > \alpha_2$ *if and only if* S_1, S_2 *are in general position.*

Proof. To prove (35) we transform the Riemann sphere into the pseudo-Euclidean model for the Möbius space using the parameter representation by complex numbers as described above. So we view it as the set of generators for the isotropic cone in four-dimensional pseudo-Euclidean space. Then each 0-sphere corresponds to a two-dimensional pseudo-Euclidean subspace, and hence the orthogonal complement of that is a two-dimensional Euclidean subspace. Now we express the operator introduced in Proposition 5 in terms of the complex parameters and compute its eigenvalues; this leads to (35)[4]. Reversing the order of the points in S_1 or S_2 the cross ratio η is replaced by η^{-1} according to (1.4.7); hence the expressions (35) for α_1, α_2 do not change their value proving 1. Since the 0-spheres are not equal, at least three of the points $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4$ have to be different. By Lemma 11 we may suppose without loss of generality that the points are

$$\boldsymbol{x}_2 = 1, \boldsymbol{x}_3 = 0, \boldsymbol{x}_4 = \infty.$$

As usual, they are identified with the corresponding numbers in the Riemann sphere. Since

$$\eta = (\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{x}_3, \boldsymbol{x}_4) = (\eta, 1; 0, \infty),$$

we then have $\boldsymbol{x}_1 = \eta$. The conditions in 2–4 occur if and only if at least one of the eigenvalues is equal to 1. But this is the case if and only if the deficit is equal to 1, i.e. if the four points all lie in one circle. It is easy to verify that η is real for the special values we are considering. So we proved 5.; in fact, the four points are in general position if and only if η does not lie on the real axis (see also Exercise 10). This was Case 1 in Section 2. To prove 2 note that $\alpha_1 = \alpha_2$ if and only if $\eta = 0$ or $\eta = \infty$; because of $S_1 \neq S_2$ these are the only possibilities for a non-empty intersection $S_1 \cap S_2 = \{0\}$ or $S_1 \cap S_2 = \{\infty\}$. In both cases we have $\alpha_1 = \alpha_2 = 1$; this corresponds to Case 2 in Section 2. Statements 3 and 4 are proved by exploiting the respective conditions: That in 3 amounts to $\eta < 0$, and that in 4 is equivalent to $\eta > 0$. In 3 the condition $\alpha_1 = 1$ just means that the deficit is 1. This is Case 3.a from Section 2; the eigenvalue α_1 from Lemma 7 has to be identified with our α_2, since Condition A is actually only satisfied for the circle. Finally, statement 4 again corresponds to Case 1 from Section 2, this time referring to the circle (the real axis). □

Corollary 18. *Four different points in the Riemann sphere* S^2 *lie in a common circle if and only if their* CR *is real.* □

[4] The computations were done using the software Mathematica by Stephen Wolfram [114]. They are an interesting example showing the far-reaching capabilities of this program to perform extensive symbolic calculations. A corresponding notebook riemsph.nb: "The Riemann Sphere" can be downloaded from the homepage http://www-irm.mathematik.hu-berlin.de/~sulanke.

Example 7. We now pose the following question: Let α_1, α_2 be two non-negative real numbers. Is there a pair of 0-spheres in the Riemann sphere S^2 whose stationary invariants are these numbers? Leaving the trivial Case 2 aside, by Proposition 17 these numbers necessarily have to satisfy the following conditions:

$$\alpha_1 > \alpha_2 \geq 0, \ \alpha_1 \geq 1 \geq \alpha_2. \tag{36}$$

First we determine the possible values η of the CR in equation (35) for the given α_1, α_2. Set

$$\eta_\pm := \frac{\alpha_1 + \alpha_2 - 2}{(\sqrt{\alpha_1} \pm \sqrt{\alpha_2})^2} + \frac{2\sqrt{(\alpha_1 - 1)(1 - \alpha_2)}}{(\sqrt{\alpha_1} \pm \sqrt{\alpha_2})^2}\,\mathrm{i}\,. \tag{37}$$

Then, in general, η_+, η_- and their conjugates $\bar{\eta}_+, \bar{\eta}_-$ all are different solutions for (35); see the Mathematica notebook mentioned in Footnote 3. Since all complex numbers $\eta \neq 0, 1, \infty$ can occur as the cross ratio of four different points, corresponding pairs of 0-spheres exist. Concerning the geometry of the situation we again suppose $\boldsymbol{x}_2 = 1, \boldsymbol{x}_3 = 0$, and $\boldsymbol{x}_4 = \infty$. The unit circle with center 0 can then be characterized in a Möbius invariant way as follows: Consider the fourth harmonic point $\boldsymbol{y}$ for $\boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4$ on the circle Σ_1 determined by these points; in our special position this is the point -1 on the real axis. Then determine the circle Σ_2 through $\boldsymbol{y}, \boldsymbol{x}_2$ that is orthogonal to Σ_1. The following can be proved: Under the inversion in Σ_2 the point $\boldsymbol{x}_1 = \eta_+$ is transformed into the point η_-; on the other hand, complex conjugation is the involution in the real axis. Since these involutions map the 0-sphere $(\boldsymbol{x}_3, \boldsymbol{x}_4) = (0, \infty)$ into itself and the point $\boldsymbol{x}_2 = 1$ remains fixed, it is clear that the pairs of 0-spheres $(\eta, 1), (0, \infty)$ formed using the four solutions $\eta = \eta_+, \eta_-, \bar{\eta}_+, \bar{\eta}_-$ are Möbius equivalent. Hence they have the same stationary invariants. □

Exercise 13. Prove the assertions formulated in Corollary 18 and Example 7. Show, moreover: a) There are less than four solutions η for (35) if and only if $\alpha_1 = 1$, or $\alpha_2 = 1$, or $\alpha_2 = 0$; this again is the case if and only if x_1 belongs to $\Sigma_1 \cup \Sigma_2$. $\eta = -1$ is the only solution if and only if $\alpha_1 = 1$ and $\alpha_2 = 0$, i.e. if the throw $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4)$ is harmonic. – b) The complex numbers $\eta_+, \eta_-, \bar{\eta}_+, \bar{\eta}_-$ lie on a circle .

In the Riemann sphere hyperspheres are circles; everything said before concerning general hyperspheres holds in particular for them. The only Möbius invariant for two circles is their Coxeter invariant (14). According to (11) the space-like unit vector of a circle with center $m \in \mathbf{C}$ and radius r is

$$\mathfrak{u}(m, r) = (\mathfrak{e}_0(1 - r^2 + |m|^2) + \mathfrak{e}_1 \operatorname{Re}(m) + \mathfrak{e}_2 \operatorname{Im}(m) + \mathfrak{e}_3(1 + r^2 - |m|^2))/2r. \tag{38}$$

Correspondingly, the results formulated in Exercise 5 also hold for pairs of lines as well as pairs consisting of a line and a circle. The following exercise shows that a pair consisting of a point and a circle has no Möbius invariant at all:

Exercise 14. Prove that the Möbius group G_{n+1} acts transitively on the set of pairs

$$\{(\Sigma, \boldsymbol{a})|\boldsymbol{a} \in S^n,\ \Sigma \subset S^n \text{ a hypersphere, } \boldsymbol{a} \notin \Sigma\}$$

as well as on the set of pairs

$$\{(\Sigma, \boldsymbol{a})|\boldsymbol{a} \in S^n,\ \Sigma \subset S^n \text{ a hypersphere, } \boldsymbol{a} \in \Sigma\}.$$

Hint. Consider Σ as the absolute of an n-dimensional hyperbolic geometry.

For the sphere S^2 it finally remains to consider pairs (Σ, S^0), where Σ is a circle, and $S^0 = (z_1, z_2)$, $z_1 \neq z_2$, denotes a 0-sphere. On the set of such pairs (Σ, S^0) with $S^0 \subset \Sigma$ the Möbius group G_3 acts transitively, since the isotropy group of Σ contains the Möbius group G_2 and G_3 acts transitively on the set of circles. We may thus restrict the discussion to the case that S^0 is not contained in Σ; this assumption is the same as requiring Condition A to be satisfied. The general theory in Section 2 implies that all we have to determine is a single stationary invariant, and that can be computed as follows: Consider the four-dimensional pseudo-Euclidean vector space on which the Möbius geometry of the sphere S^2 is based. Let $\mathfrak{n} = \mathfrak{n}(\Sigma)$ denote one of the space-like unit vectors describing the points of Σ by

$$\boldsymbol{x} = [\mathfrak{x}] \in \Sigma \Longleftrightarrow \langle \mathfrak{n}(\Sigma), \mathfrak{x} \rangle = 0 \text{ and } \langle \mathfrak{x}, \mathfrak{x} \rangle = 0,$$

e.g. the one defined by (38). Note that $\mathfrak{n}$ is only determined up to sign. Below, let $\mathfrak{b}_1, \mathfrak{b}_2$ be two orthogonal space-like unit vectors, whose associated circles both contain the 0-sphere S^0; in short, we will speak of an orthonormal basis for the *pencil of circles,* i.e. the family of all circles in S^2 containing S^0. Then

$$\operatorname{invcpp}(\Sigma, S^0) = \langle \mathfrak{n}, \mathfrak{b}_1 \rangle^2 + \langle \mathfrak{n}, \mathfrak{b}_2 \rangle^2 \tag{39}$$

is independent of the choice of $\mathfrak{n}$ just described as well as of the orthonormal basis; it is the single stationary invariant of the pair (Σ, S^0) we were looking for. In fact, the matrix of the double projection described in Section 2 just reduces to the one component (39).

Exercise 15. Prove: a) All the pairs (Σ, S^0) with $\Sigma \cap S^0$ a single point are Möbius equivalent. – b) The stationary invariant of such a pair is 1, and its deficit is zero. (Hint. As in Exercise 14 one can show that each configuration with the stated property is Möbius equivalent to $(\Sigma_0, \{0, 1\})$, where Σ_0 denotes the unit circle with center 0.) – c) For $S^0 \subset \Sigma$ the stationary invariant and the deficit both are 1.

Now we want to express the Möbius invariant $\operatorname{invcpp}(\Sigma, S^0)$ in terms of the Euclidean invariants for the pair (Σ, S^0); we will derive a formula similar to (14) for the Coxeter invariant also in this case.

Proposition 19. *Let $z_1, z_2 \in \mathbf{C}$, $z_1 \neq z_2$, be two points in the complex number plane, and let $\Sigma(m, r) \subset \mathbf{C}$ be the circle with center $m \in \mathbf{C}$ and radius $r > 0$. Consider the triangle (m, z_1, z_2) and denote its side lengths as follows:*

$$a = |z_1 - z_2|,\ b = |z_2 - m|,\ c = |z_1 - m|. \tag{40}$$

Moreover, denote by α the angle of this triangle at the vertex m. Then the stationary Möbius invariant of the pair is computed by either of the formulas below:

$$\mathrm{invcpp}(\Sigma(m,r),(z_1,z_2)) = \frac{r^4 - 2r^2 bc\cos\alpha + b^2c^2}{r^2a^2}, \tag{41}$$

$$\mathrm{invcpp}(\Sigma(m,r),(z_1,z_2)) = \frac{(r^2 - bc)^2}{r^2a^2} + \frac{2bc(1-\cos\alpha)}{a^2}. \tag{42}$$

Obviously,completing the square in (41) yields (42). To prove the assertion consider the vector (11) for the circle $\Sigma(m,r)$. Find an orthonormal basis for the Euclidean vector subspace of $\boldsymbol{V}^4$ corresponding to the point pair (z_1, z_2), which is orthogonal to the isotropic vectors representing the points. Using this, the invariant can be computed by formula (39). Then the quantities occurring in this expression have to be interpreted. The details can be found in the Mathematica notebook mentioned in footnote 3. □

Now we want to derive a third expression for the invariant invcpp, making the relation to the Coxeter invariant apparent. To do so, we represent the 0-sphere $S^0 = (z_1, z_2)$ as well as the higher-dimensional spheres in center-radius form:

$$S^0(m_z, r_z, u) = (m_z + r_z\,\mathrm{e}^{\mathrm{i}\,u}, m_z - r_z\,\mathrm{e}^{\mathrm{i}\,u}),\ \ m_z \in \mathbf{C}, u \in \mathbf{R}, r_z > 0; \tag{43}$$

m_z is the center, r_z the radius, and u an angle parameter determining the direction of S^0 in the complex plane. Then

Proposition 20. *For the stationary Möbius invariant of a circle and a 0-sphere the following relation holds*

$$\mathrm{invcpp}(\Sigma(m,r), S^0(m_z,r_z,u)) = (\frac{r^2 + r_z^2 - d^2}{2r\,r_z})^2 + (\frac{d\cos\beta}{r})^2; \tag{44}$$

here $d = |m - mz|$ denotes the distance of the centers, and β is the angle between the vector $m - m_z$ and the normal to the connecting line $z_1 \vee z_2$. □

Looking at Formula (44), it is remarkable that the first term on the right-hand side is nothing but the square of the Coxeter invariant for the circles $\Sigma(m,r)$ and $\Sigma(m_z, r_z)$, whereas the second term coincides with the square of the projection of $m - m_z$ onto the normal of $S^0(m_z, r_z, u)$. In the next section we will describe a far-reaching generalization of this formula and hence omit to prove it. Of course, these formulas may be applied analogously to corresponding configurations in a Euclidean plane.

The proof of Proposition 20 can also be found in the notebook mentioned in footnote 3. This notebook contains useful tools for solving the following exercises as well as designing or modifying the accompanying illustrations.

Exercise 16. Consider the *pencil of circles* formed by the circles containing a given 0-sphere. Obviously, the Möbius group acts transitively on the set of all these pencils. Hence we may as well suppose that the pencil is just the set of all circles through the 0-sphere $S^0 = (-1, 1)$. The centers m of these circles lie on the imaginary axis; let $t := \mathrm{Im}(m)$ be chosen as the parameter for the circles in the pencil. Then the radius of the corresponding circle is $\sqrt{1+t^2}$. Denote the circle of the pencil corresponding to the parameter value t by $bc(t)$. Prove: a) The space-like vector

$$\mathfrak{bc}(t) := \{0, 0, \frac{t}{\sqrt{1+t^2}}, \frac{1}{\sqrt{1+t^2}}\},\ t \neq 0,$$

describes the circle $bc(t)$ in the pencil as the set of all points $x = [\mathfrak{x}]$, whose coordinate vector $\mathfrak{x}$ satisfies the equations

$$\langle \mathfrak{bc}(t), \mathfrak{x} \rangle = 0,\ \langle \mathfrak{x}, \mathfrak{x} \rangle = 0.$$

– b) The real axis also belongs to the pencil of circles in question. It has, however, no representation of the form explained in a). Which space-like vector determines it? – c) Through each point $z \neq \pm 1$ in the Riemann sphere there is a unique circle in the pencil. Find the corresponding parameter value t for this circle in the case $\mathrm{Im}(z) \neq 0$, and with it center and radius. For real points as well as the point ∞ the real axis is this uniquely determined circle in the pencil. – d) For each circle ca in the complex number plane with center a and radius r_a, $0 < r_a < \infty$, there is a circle $bc(t)$ in this pencil intersecting ca orthogonally. Determine such a circle. – e) If $ca = bc(s)$ itself is a circle in this pencil, then assigning the uniquely determined orthogonal circle $bc(t) = bc(s)^\perp$ to it defines a fixed-point free involution on the pencil of circles. Describe this involution in the form $t = F(s)$. Obviously, the unit circle with center 0 is in involution with the real axis; verify this by a computation.

Exercise 17. Prove: a) For each pencil of circles in the Riemann sphere there is a one-parameter family of circles orthogonally intersecting every circle in the pencil. These circles are called the *orthogonal trajectories* of the pencil. Determine the orthogonal trajectories for the pencil considered in Exercise 16. – b) Find the isotropy group H of a 0-sphere. – c) Let $M(S^0)$ denote the pencil of circles through S^0, and let $M^\perp(S^0)$ be the set of its orthogonal trajectories. Prove that H acts transitively on the product set $M(S^0) \times M^\perp(S^0)$.

Exercise 18. Let $M(S^0)$ be the pencil of circles for $S^0 = (-1, 1)$ (cf. Exercise 16), and let $S^1(m, r)$ be a test circle with center m and radius r not containing S^0. Denote by $\alpha = \mathrm{invcpp}(S^1(m, r), S^0)$ the stationary invariant (39) for this pair. Prove that there is a circle $bc(t) \in M(S^0)$ meeting the test circle $S^1(m, r)$ if and only if $\alpha \geq 1$; for $\alpha = 1$ there is a unique such circle, and for $\alpha > 1$ there exist two of these circles. Compute the parameter value t for these circles in the case $\alpha \geq 1$.

Figure 2.49 shows the situation described in Exercise 18. The test circles are blue and the circles in the pencil meeting it, if there exist any, are violet. The red circle is the one in the pencil intersecting the test circle orthogonally; the test circles are no orthogonal trajectories for the pencil. The green circle is the one in the pencil for which the Coxeter invariant (14) with respect to the test circle is extremal. Since the stationary invariant of two circles I^2 is the square of their Coxeter invariant, it

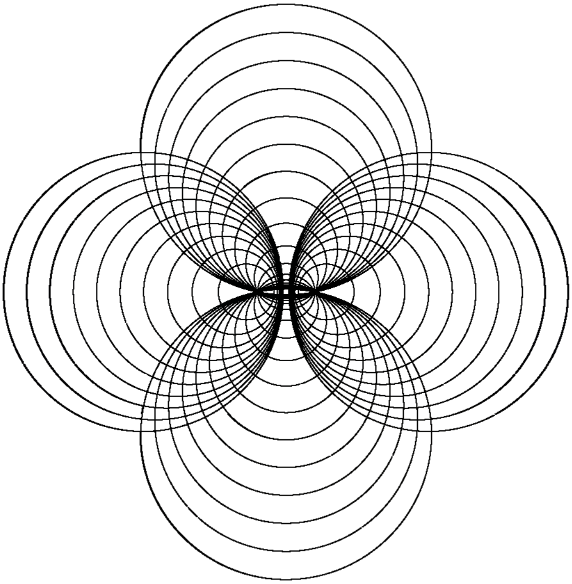


Fig. 2.48. A pencil of circles and its orthogonal trajectories.

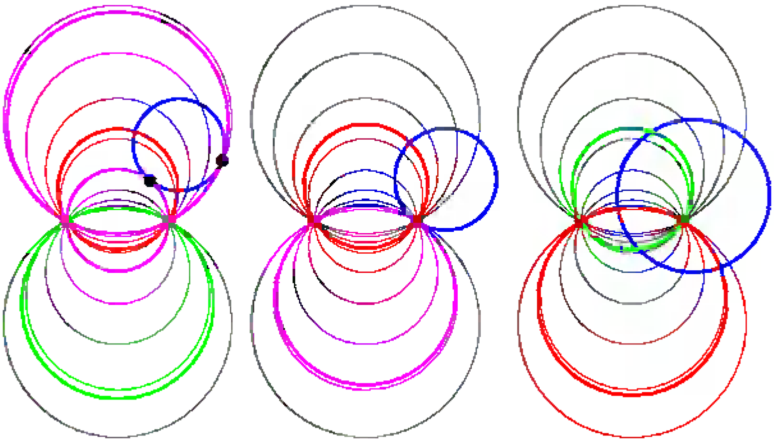


Fig. 2.49. Existence of circles in a pencil meeting a test circle.

attains its maximum for the green circle. If $\alpha > 1$, as in the figure on the left, then the test circle does not separate the points of S^0, and the pencil contains two circles meeting one another. The figure in the middle depicts the case $\alpha = 1$; the test circle contains a point of S^0, and there is a unique touching circle in the pencil, which, at the same time, realizes the maximal value $\alpha = 1$ of the stationary invariant. Finally, in the figure on the right we have $\alpha < 1$; the test circle separates the points of

S^0, each circle of the pencil meets the test circle, and the green circle realizes the smallest intersection angle occurring.

Exercise 19. Consider a pair (L, S^0) consisting of an arbitrary (real) line L and a 0-sphere S^0 in the complex plane. Proof a result analogous to Proposition 20.

Exercise 20. Replace the 0-sphere $S^0 = \{-1, 1\}$ by the 0-sphere $S^0 = \{0, \infty\}$ in Exercises 16, 17, and 18. Then the pencil of circles is replaced by the pencil of all lines through 0. Solve the exercises in this case, and draw the resulting configurations.

2.7.4 Möbius Invariants and Euclidean Invariants

In the Formulas (14) and (44) we expressed the Möbius invariant for a pair of hyperspheres and a pair (Σ^1, S^0) consisting of a circle and a 0-sphere in terms of their Euclidean invariants, this lead us to the Coxeter invariant. Now we want to derive a similar expression for pairs (Σ^{n-1}, S^l) consisting of a hypersphere and an l-sphere, $0 \le l < n$, in the n-dimensional Euclidean space $\boldsymbol{E}^n$. In fact, according to Proposition 5 these pairs are also described by a single Möbius invariant up to Möbius equivalence. Euclidian invariants of such a pair are the radii r, R of the sphere S^l and the hypersphere as well as the distance d of their centers. In the case $l = n - 1$ these already provide a complete system of Euclidean invariants. For $l < n - 1$ we still need to introduce a fourth quantity to obtain a complete system of invariants, e.g. the angle between the connecting line of the centers and the $(l+1)$-plane determined by the l-sphere, or, what amounts to the same, the distance p from the center $\boldsymbol{m}$ of the hypersphere $\Sigma = \Sigma(\boldsymbol{m}, R)$ to this $(l+1)$-plane. We choose the pseudo-orthonormal basis $(\mathfrak{e}_j)$ for the associated $(n+2)$-dimensional vector space so that the l-sphere appears as the intersection of the hypersphere $\Sigma(\boldsymbol{o}, r)$ of radius $r > 0$ with the origin $\boldsymbol{o}$ as its center and the coordinate hyperplanes through the origin determined by the basis vectors $\mathfrak{e}_\kappa$, $\kappa = l+2, \ldots, n$ as their normal vectors. Hence, from the Euclidean point of view, S^l is the l-sphere of radius r with the origin as its center lying in the coordinate-$(l+1)$-plane determined by the first $l+1$ coordinates x^α, $\alpha = 1, \ldots, l+1$. Then the vectors

$$\mathfrak{u}_1 := (\mathfrak{e}_0(1 - r^2) + \mathfrak{e}_{n+1}(1 + r^2))/2r, \ \mathfrak{u}_2 = \mathfrak{e}_{l+2}, \ldots, \mathfrak{u}_{n-l} = \mathfrak{e}_n \tag{45}$$

are an orthonormal basis for the Euclidean subspace $\boldsymbol{U}^\perp \subset \boldsymbol{V}$ determining the l-sphere S^l; by (11) the vector $\mathfrak{u}_1$ is the space-like unit vector belonging to the hypersphere considered. Likewise by (11), the vector

$$\mathfrak{w} = (\mathfrak{e}_0(1 - R^2 + d^2) + 2\mathfrak{m} + \mathfrak{e}_{n+1}(1 + R^2 - d^2))/2R$$

determines the hypersphere $\Sigma(\boldsymbol{m}, R)$; here $\mathfrak{m}$ is the position vector of the center $\boldsymbol{m}$ with respect to $\boldsymbol{o}$, and $d = |\mathfrak{m}|$ is the distance of the centers. Now we apply Proposition 5. The subspace $\boldsymbol{W}^\perp$ associated with the hypersphere $\Sigma(\boldsymbol{m}, R)$ is one-dimensional. Hence the operator A reduces to a single component, which thus is the invariant

$$\alpha(\Sigma(\boldsymbol{m}, R), S^l) := \alpha_1$$

to be computed. A straightforward computation starting from Definition (6.70) of the operator A leads to the formula

$$\alpha(\Sigma(\boldsymbol{m}, R), S^l) = \Sigma_{\kappa=1}^{n-l} \langle \mathfrak{w}, \mathfrak{u}_\kappa \rangle^2 = (\frac{r^2 + R^2 - d^2}{2rR})^2 + (\frac{p}{R})^2, \qquad (46)$$

which contains (44) as a special case; note that $p = d \cos\beta$. Hence we proved:

Proposition 21. *The only stationary Möbius invariant for pairs* (Σ^{n-1}, S^l) *consisting of a hypersphere* Σ^{n-1} *and an* l*-sphere* S^l, $0 \leq l < n$, *in the* n*-dimensional Euclidean space* $\boldsymbol{E}^n$ *is determined by* (46); *here* R *denotes the radius of* Σ^{n-1}, r *that of* S^l, d *is the distance of the centers, and* p *is the length of the projection of the vector connecting the centers onto the subspace orthogonal to the* $(l+1)$*-plane of* S^l, $p = 0$ *in the case* $l = n - 1$. □

Next we want to express the Möbius invariants of a circle pair (C_0, C_1) in the three-dimensional Euclidean space $\boldsymbol{E}^3$ in terms of their Euclidean invariants. So let $\boldsymbol{m}_i$ be the centers, let $\boldsymbol{r}_i$ be the radii, and let $\mathfrak{n}_i$ be the *position vectors* for the circles C_i, $i = 0, 1$, i.e. the unit normal vectors for the planes in which they lie. We consider the circles as non-oriented; hence the position vectors are unique only up to multiplication by ± 1. Moreover, let $\mathfrak{d} = \boldsymbol{m}_1 - \boldsymbol{m}_0$ denote the vector connecting their centers. It is easy to see that the following six quantities form a complete system of Euclidean invariants for the pair of circles (C_0, C_1):

1. The radii r_0, r_1 of the circles.
2. The angle α between the circle planes: $\cos\alpha = |\langle \mathfrak{n}_0, \mathfrak{n}_1 \rangle|$, $0 \leq \alpha \leq \pi/2$.
3. The distance $d = |\mathfrak{d}|$ of the centers.
4. The angles β, γ between the position vectors for the circles and the vector $\mathfrak{d}$ connecting their centers; for $d = 0$ these angles remain undefined.

According to Proposition 5 a circle pair has two stationary invariants $\alpha_1 \geq \alpha_2 \geq 0$, namely the eigenvectors for the matrix of their double projection A. Instead of the eigenvalues we consider the determinant and the trace of this matrix, which again are a complete system of invariants equivalent to the eigenvalues in Möbius geometry. If $\mathfrak{a}_1, \mathfrak{a}_2$ and $\mathfrak{b}_1, \mathfrak{b}_2$ are orthonormal bases for the Euclidean subspaces $\boldsymbol{W}^\perp, \boldsymbol{U}^\perp$ of the pseudo-Euclidean vector space $\boldsymbol{V}^5$ determining the circles C_0, C_1, then the coordinates of A are

$$a_{ij} = \sum_{k=1}^{n-l} \langle \mathfrak{a}_i, \mathfrak{b}_k \rangle \langle \mathfrak{a}_j, \mathfrak{b}_k \rangle, \; i, j = 1, \ldots, n - m, \qquad (47)$$

where in the case considered here we have to set $n = 3$ and $l = m = 1$. Obviously,

$$\det A = a_{11}a_{22} - a_{12}^2 = \alpha_1\alpha_2, \operatorname{tr} A = a_{11} + a_{22} = \alpha_1 + \alpha_2. \qquad (48)$$

Conversely, this provides a direct way to compute the eigenvalues. We prove

Proposition 22. *The Möbius invariants* (48) *for a circle pair* C_0, C_1 *in the Euclidean space* $\boldsymbol{E}^3$ *are expressed in terms of the Euclidean invariants for the pair according to the following formulas:*

$$\det A = \frac{((r_0^2 + r_1^2 - d^2)\cos\alpha + 2d^2\cos\beta\cos\gamma)^2}{4r_0^2 r_1^2}, \tag{49}$$

$$\operatorname{tr} A = \frac{d^4 + r_0^4 + r_1^4 + 4r_0^2 r_1^2 + 2r_0^2 r_1^2 \cos 2\alpha + 2d^2(r_0^2\cos 2\beta + r_1^2 \cos 2\gamma)}{4r_0^2 r_1^2}. \tag{50}$$

Without loss of generality, we may suppose for the proof of this proposition that the circle C_0 lies in the x, y-plane of a Euclidean coordinate system with the origin as its center. Then by (11) and (15) the vectors

$$\mathfrak{v}_1 = (\mathfrak{e}_0(1 - r_0^2) + \mathfrak{e}_4(1 + r_0^2))/2r_0, \mathfrak{v}_2 = \mathfrak{e}_3$$

form an orthonormal basis for the vector space $\boldsymbol{W}^\perp$ of the circle C_0; here the first vector defines the sphere of radius r_0 with its center at the origin, and the second one the x, y-plane. The position vector for C_0 is $\mathfrak{n}_0 = \mathfrak{e}_3$. The position vector of the second circle C_1 is an arbitrary unit vector in the Euclidean space, say

$$\mathfrak{n}_1 = \mathfrak{e}_1 \cos u \cos v + \mathfrak{e}_2 \sin u \cos v + \mathfrak{e}_3 \sin v.$$

We may choose the coordinates in the x, y-plane in such a way that the center $\boldsymbol{m}_1$ of C_1 has the position vector

$$\mathfrak{d} = \mathfrak{e}_1 x + \mathfrak{e}_3 z.$$

Again we represent the circle C_1 as the intersection of the sphere with radius r_1 and center $\boldsymbol{m}_1$ with the plane of the circle, i.e. the plane through $\boldsymbol{m}_1$ orthogonal to $\mathfrak{n}_1$. Then (11) and (15) define the vectors

$$\mathfrak{b}_1 = (\mathfrak{e}_0(1 - r_1^2 + d^2) + 2\mathfrak{d} + \mathfrak{e}_4(1 + r_1^2 - d^2))/2r_1,$$

$$\mathfrak{b}_2 = \mathfrak{n}_1 + (\mathfrak{e}_0 - \mathfrak{e}_4)\langle \mathfrak{d}, \mathfrak{n}_1 \rangle,$$

which form an orthonormal basis for the vector space $\boldsymbol{U}^\perp$ of C_1. Starting from these expressions the matrix A and its determinant as well as trace can be computed. Having calculated the Euclidean invariants,

$$\begin{aligned} \cos\alpha &= \langle \mathfrak{n}_0, \mathfrak{n}_1 \rangle = \sin v, \\ d\cos\beta &= \langle \mathfrak{n}_0, \mathfrak{d} \rangle = z, \\ d\cos\gamma &= \langle \mathfrak{n}_1, \mathfrak{d} \rangle = x \cos u \cos v + z \sin v, \end{aligned}$$

the coordinates x, z, u, v can be eliminated from $\det A$ and $\operatorname{tr} A$. This leads to the formulas (49) and (50). Since the terms involving the trigonometric functions of β and γ all contain the factor d, we can do without a definition for $d = 0$ and still have the formulas in this case. To compute the details we again applied Stephen Wolfram's Mathematica [114], cf. the notebook mcircles.nb on the homepage of R. Sulanke. □

2.7.5 Three-Dimensional Möbius Geometry

In this section we want to add some observations concerning circle geometry to the discussion of the geometry in the Möbius space S^3. Even if, as beings living in the three-dimensional, we are unable to directly image a three-dimensional sphere, we inevitably view an n-sphere as a hypersurface in an $(n+1)$-dimensional space. The dimension three, however, has the advantage that we can pass to the familiar three-dimensional Euclidean space via stereographic projection (see the hint in Exercise 1). Since this transformation is conformal, it preserves all Möbius geometric properties at least locally, and thus renders them accessible to our immediate imagination as well as its graphic representation. The geometry of spheres $\Sigma^2 \subset S^3$ and pairs (Σ^2, S^l), $l = 0, 1$, is the special case $n = 3$ for the hyperspheres considered time and again in the preceding sections, cf. in particular Proposition 21. Up to now, for circles in S^3 we only have Proposition 22, which, on the one hand, allows to compute the Möbius invariants as functions of the Euclidean configuration of the circles. It does, however, not directly yield the stationary invariants for the pair of circles. The expressions obtained for them starting from the formulas proved in Proposition 22 are rather unwieldy. Nevertheless, the stationary invariants, in particular the maximal one α_1, contain the important pieces of information concerning the position relations of the circle pair. Applying the vector product provides an alternative approach to these position relations. The series of Mathematica Notebooks called Spheres contains computational tools to study and visualize three-dimensional Möbius geometry. They can be found under the title Spheres in MathSource, or at the homepage of R. Sulanke.

In the notations of Proposition 5 let $C_0 = S(\boldsymbol{U})$, $C_1 = S(\boldsymbol{W})$ be two circles in S^3, and let $\{\mathfrak{a}_1, \mathfrak{a}_2\}$, $\{\mathfrak{b}_1, \mathfrak{b}_2\}$ be orthonormal bases for the associated Euclidean subspaces $\boldsymbol{W}^\perp$, $\boldsymbol{U}^\perp$. Under these assumptions consider the vector product in the oriented pseudo-Euclidean vector space $\boldsymbol{V}^5$ defined according to Proposition 2.2.17

$$\mathfrak{n}(C_0, C_1) := \mathfrak{a}_1 \times \mathfrak{a}_2 \times \mathfrak{b}_1 \times \mathfrak{b}_2. \tag{51}$$

We will call this the *indicator* of the circle pair. Then obviously

Lemma 23. *The indicator defined by* (51) *is uniquely determined up to the factor* ± 1 *by the circles. If the circles are oriented, then reversing the orientation in one of them changes the sign of the indicator. The indicator is assigned to the circle pair in an equivariant way; for* $g \in G_4$

$$\mathfrak{n}(gC_0, gC_1) = \pm g\mathfrak{n}(C_0, C_1).$$

The factor -1 *occurs, if* g *reverses the orientation of* S^3. □

This lemma is a straightforward consequence of Proposition 2.2.18, it immediately implies

Proposition 24. *The scalar square of the indicator is a Möbius invariant for the pair of non-oriented circles C_0, C_1. Here*

$$w(C_0, C_1) := \langle \mathfrak{n}(C_0, C_1), \mathfrak{n}(C_0, C_1) \rangle = \alpha_1 + \alpha_2 - \alpha_1\alpha_2 - 1, \tag{52}$$

where α_1, α_2 denote the stationary invariants for the pair C_0, C_1.

Proof. We use formula (2.37) with $a = -1$. By Lemma 23 we may suppose that the bases are just the canonical bases defined in deriving Proposition 6.24. For them

$$\langle \mathfrak{a}_1, \mathfrak{b}_1 \rangle = \sqrt{\alpha_1}, \ \langle \mathfrak{a}_2, \mathfrak{b}_2 \rangle = \sqrt{\alpha_2}, \ \langle \mathfrak{a}_1, \mathfrak{b}_2 \rangle = \langle \mathfrak{a}_2, \mathfrak{b}_1 \rangle = 0.$$

Inserting this into the determinant we need to compute then proves the assertion. □

Corollary 25. *Let $\mathfrak{n} = \mathfrak{n}(C_0, C_1)$ be the indicator of the circles C_0, C_1. Then there are the following mutually excluding possibilities:*

1. *If $\mathfrak{n} = \mathfrak{o}$, then the circles are not in general position, i.e., there is a sphere $\Sigma^2 \subset S^3$ containing both circles (Condition A is not satisfied).*
2. *The indicator $\mathfrak{n}$ is space-like: $w(C_0, C_1) > 0$. Then we are in Case 1; the circles are in general position, disjoint, and can be separated by spheres.*
3. *The indicator is isotropic $w(C_0, C_1) = 0$, $\mathfrak{n} \neq \mathfrak{o}$: Then we are in Case 2; the circles are in general position and intersect in a single point.*
4. *The indicator $\mathfrak{n}$ is time-like: $w(C_0, C_1) < 0$. Then we are in Case 3; the circles are disjoint and linked.*

Proof. By Proposition 2.18, 3., Situation 1 occurs if and only if the vectors $\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{b}_1, \mathfrak{b}_2$ are linearly independent. If the indicator $\mathfrak{n}$ is space-like, then $\boldsymbol{U}^\perp + \boldsymbol{W}^\perp = [\mathfrak{n}]^\perp$ is pseudo-Euclidean and has dimension 4; the assertion in 2 follows as in Lemma 6. If $\mathfrak{n}$ is isotropic, then so is $\boldsymbol{U}^\perp + \boldsymbol{W}^\perp = (\boldsymbol{U} \cap \boldsymbol{W})^\perp$, hence $\boldsymbol{U} \cap \boldsymbol{W} = [\mathfrak{n}]$; the indicator represents the single point of intersection of the circles, and 3 holds. The proof of the assertion in 4 follows from the proof of Lemma 7. □

Example 8. Let C_0, C_1 be a circle pair with maximal stationary invariant $\alpha_1 = 0$. Then, because of $0 \leq \alpha_2 \leq \alpha_1$, we also have $\alpha_2 = 0$, and the matrix A of the double projection is the zero matrix. Such circle pairs are called *orthogonal.* A particular example for this situation is the pair consisting of the z-axis, represented as the intersection of the coordinate planes, i.e. with the subspace $\boldsymbol{U}^\perp = [\mathfrak{e}_1, \mathfrak{e}_2]$, and the unit circle in the x, y-plane with its center at the origin, represented as the intersection of this plane with the unit sphere around the origin, i.e. $\boldsymbol{W}^\perp = [\mathfrak{e}_3, \mathfrak{e}_4]$. Each sphere containing the unit circle orthogonally intersects every sphere, i.e. plane containing the z-axis. This is apparently clear and can also easily be verified: The space-like unit vectors determining these spheres lie in $\boldsymbol{U}^\perp$ or $\boldsymbol{W}^\perp$, respectively, and these spaces are orthogonal. The scalar square of the indicator is $w = -1$; the circles are linked. Each pair of orthogonal circles is Möbius equivalent to the pair just described. □

Exercise 21. The notion defined in Example 8 is not to be confused with that of two perpendicular (or orthogonally intersecting) circles! Prove that two pairs of perpendicular circles in general position have equal stationary invariants $\alpha_1 = 1, \alpha_2 = 0$, and hence are Möbius equivalent. Two spheres each containing one of these circles can intersect at an arbitrary angle. Find an example for this situation and perform the corresponding calculations based on the associated Euclidean subspaces. Moreover, compute the stationary invariants for two perpendicular circles not satisfying Condition A; these again are $\alpha_1 = 1, \alpha_2 = 0$. Since now the rank condition is violated, such circle pairs obviously are not Möbius equivalent to the one considered above.

Example 9. The Mathematica notebook mcircles.nb in the series Spheres contains the description of normal forms for circle pairs. For each pair of stationary Möbius invariants a circle pair with just these invariants is specified. As described in Section 6 for pairs of hyperbolic subspaces, starting from the eigenvectors of the operator A one can adapt an orthonormal frame for the pseudo-Euclidean vector space $\boldsymbol{V}^5$ to the circle pair, in which the defining subspaces of the circles are spanned by vectors whose coefficients only depend on the invariants. Now we consider the case of *linked circles.* Since the Möbius group acts transitively on circles, we may suppose that the circle $C_0 = S(\boldsymbol{U})$ of the pair is the unit circle in the x, y-plane with center at the origin. Then the vectors $\mathfrak{e}_3, \mathfrak{e}_4$ span the defining space $\boldsymbol{U}^\perp$ for C_0. Let the second circle $C_1 = C(a, b)$ of the pair be determined by the Euclidean subspace

$$\boldsymbol{W}^\perp := [\mathfrak{a}_1, \mathfrak{a}_2],\ \mathfrak{a}_1(a) := \mathfrak{e}_1 \sin a + \mathfrak{e}_3 \cos a,\ \mathfrak{a}_2(b) := \mathfrak{e}_2 \sin b + \mathfrak{e}_4 \cos b. \quad (53)$$

Obviously, $C_1(0, 0) = C_0$, so that we have to exclude this case. It is easy to conclude from Formula (47) that the matrix A of the double projection has diagonal form in the constructed bases; its diagonal entries are the eigenvalues

$$\alpha_1 = \cos^2 a, \alpha_2 = \cos^2 b,\ 0 \le a \le b \le \pi/2,\ b \ne 0.$$

With the stated bounds we obtain a unique representative for all possible stationary invariants of linked circles. The indicator of the circle pair is the time-like vector

$$\mathfrak{n}(C_0, C(a, b)) = \mathfrak{e}_0 \sin^2[a] \sin^2[b],\ a \ne 0.$$

For $a = 0$ the circles $C(0, b)$ lie in the x, y-plane, Condition A is violated, and the indicator is the null vector. The Euclidean data of the circle $C(a, b)$ are

The center: $\mathfrak{m}((a, b) = -\mathfrak{e}_2 \sin b$,
The radius: $r(a, b) = 1/\cos b$,
The position vector: $\mathfrak{p}(a, b) = \mathfrak{e}_1 \sin a + \mathfrak{e}_3 \cos a$.

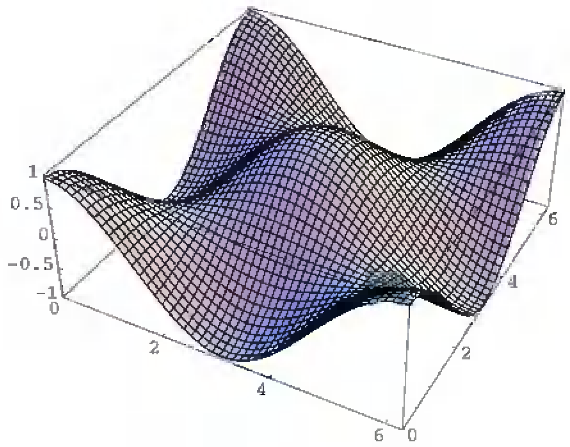


Fig. 2.50. The function $F(0.3, 1.5, s, t)$.

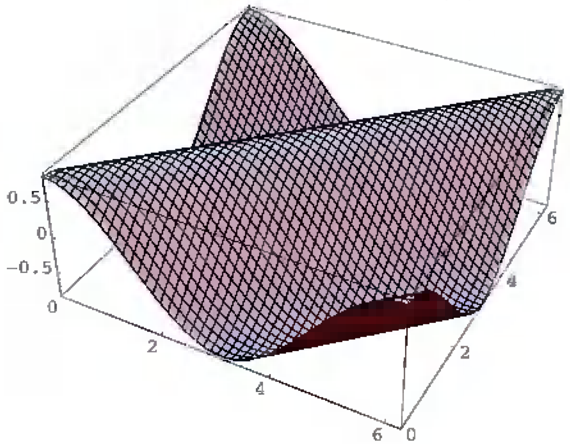


Fig. 2.51. The function $F(0.6, 0.6, s, t)$.

For $b = \pi/2$ this describes lines through the origin in the x, z-plane intersecting the x-axis at the angle a. The cosine of the intersection angle of two spheres each containing one of the circles is

$$\begin{aligned} F(a, b, s, t) &= \langle \mathfrak{e}_3 \cos t + \mathfrak{e}_4 \sin t, \mathfrak{a}_1(a) \cos s + \mathfrak{a}_2(b) \sin s \rangle \\ &= \cos a \cos s \cos t + \cos b \sin s \sin t. \end{aligned}$$

Figure 2.50 shows an example for the graph of such a function. In the case of equal eigenvalues, i.e. for $a = b$, we obtain a family of cosine functions with shifted phases,

$$F(a, a, s, t) = \cos a \cos(s - t),$$

cf. Figure 2.51. Figure 2.52 shows a part of the surface generated by the circles $C(a,a)$,
$0 \le a \le \pi/3$. We called circle pairs with equal eigenvalues *isogonal.* This is not to mean that the spheres containing the circles intersect at a constant angle. □

Fig. 2.52. The surface $\{C(a,a)|0 \le a \le \pi/3\}$.

Exercise 22. Prove that the circle pair $C(a,a), C(b,b)$ is also isogonal for $a \neq b$, and compute its stationary invariant. Thus the generating circles of the surface in Figure 2.52 are isogonally linked.

In three-dimensional Möbius geometry we still have to consider the pairs (C, S^0) consisting of a circle and a 0-sphere. Again we need two stationary Möbius invariants to characterize these pairs up to Möbius equivalence. The reader may consult the Mathematica notebook pairs.nb from the series Spheres for a discussion of these invariants and simple examples. Note also the following

Exercise 23. In the n-sphere S^n, $n \geq 2$, we consider the sets

$$M_0 := \{(C, \boldsymbol{b}) | \boldsymbol{b} \in C, C \subset S^n \text{ circle}\},$$
$$M_1 := \{(C, \boldsymbol{b}) | \boldsymbol{b} \in S^n \setminus C, C \subset S^n \text{ circle}\}.$$

Prove that the Möbius group acts transitively on each of these sets, and represent them as quotient spaces (cf. Appendix A.3).

2.7.6 Orbits, Cyclids of Dupin, and Loxodromes

In this section we want to describe orbits for subgroups of the Möbius group G_4 in the three-dimensional Möbius space S^3.

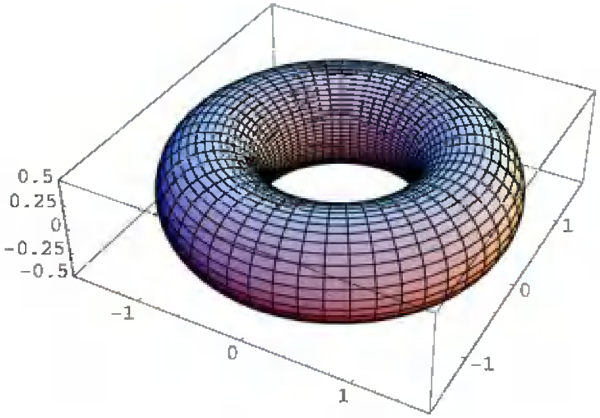


Fig. 2.53. The torus $\mathfrak{y}(s, t, \pi/7)$.

Example 10. Tori. Consider the three-dimensional Möbius space $S^3 \cong G_4/CE(3)$, cf. (1), which we view, as described in the hint in Exercise 1, as the unit hypersphere in the four-dimensional Euclidean hyperplane $\boldsymbol{E}^4$ defined by the equation $-\langle \mathfrak{e}_0, \mathfrak{x} \rangle = 1$. The orientation preserving elements of the Möbius group are the pseudo-orthogonal transformations leaving the space as well as the time orientation invariant; they form the special pseudo-orthogonal group $\boldsymbol{SO}(1, 4)$. This contains the group $\boldsymbol{SO}(4)$ of orientation preserving isometries of S^3 as the isotropy group of the vector $\mathfrak{e}_0 \in \boldsymbol{V}^5$. We decompose the four-dimensional Euclidean vector space $\boldsymbol{W}^4$ of $\boldsymbol{E}^4$ orthogonal to $\mathfrak{e}_0$ into the two two-dimensional orthogonal components

$$\boldsymbol{U}_1 = [\mathfrak{e}_1, \mathfrak{e}_2], \ \boldsymbol{U}_2 = [\mathfrak{e}_3, \mathfrak{e}_4], \ \boldsymbol{W}^4 = \boldsymbol{U}_1 \oplus \boldsymbol{U}_2.$$

In each of these components we consider the rotation group isomorphic to $\boldsymbol{SO}(2)$. In the fixed pseudo-orthonormal basis the matrices of the subgroup in G_4 generated by theses rotations have the block form

$$dd(s,t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & d(s) & 0 \\ 0 & 0 & d(t) \end{pmatrix} \text{ with } d(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}. \tag{54}$$

This, obviously Abelian group is called the ***torus group*** T^2. The orbits $T^2\mathfrak{x}$ in S^3 are circles, if $\mathfrak{x} \in \mathfrak{e}_0 + \boldsymbol{U}_1$ or $\mathfrak{x} \in \mathfrak{e}_0 + \boldsymbol{U}_2$, and surfaces, called ***tori*** , for all other unit vectors; the orbit of the point $\boldsymbol{x}(a)$ with the position vector

$$\mathfrak{x}(a) = \mathfrak{e}_0 + \mathfrak{e}_1 \cos a + \mathfrak{e}_3 \sin a$$

consists of the points with position vector

$$\mathfrak{x}(s,t,a) = \mathfrak{e}_0 + (\mathfrak{e}_1 \cos s + \mathfrak{e}_2 \sin s)\cos a + (\mathfrak{e}_3 \cos t + \mathfrak{e}_4 \sin t)\sin a. \tag{55}$$

By stereographic projection of S^3 onto the equatorial space one obtains for $0 < a < \pi/2$ surfaces of revolution

$$\mathfrak{y}(s,t,a) = ((\mathfrak{e}_1 \cos s + \mathfrak{e}_2 \sin s)\cos a + \mathfrak{e}_3 \cos t \sin a)/(1 - \sin t \sin a), \tag{56}$$

which are generated by letting a circle in the x,z-plane with its center at $(1/\cos a, 0, 0)$ and radius $\tan a$ revolve around the z-axis. They are called ***tori*** as well. Figure 2.53 shows such a torus. Note that this particular one is the orbit for a subgroup of the Möbius group, but not for a subgroup of the Euclidean group of $\boldsymbol{E}^3$. Applying inversions in arbitrary spheres leads to Möbius equivalent surfaces, which are also called *tori*. Figure 2.54 shows such a torus obtained by inversion. □

Example 11. Circular Cylinder. In the Euclidean space $\boldsymbol{E}^3$ circular cylinders are usually described as the orbits of the Abelian group $A^2 = \boldsymbol{SO}(2) \times \mathbf{R}$; here $\boldsymbol{SO}(2)$ acts as the group of rotations in the x,y-plane, and $\mathbf{R}$ acts as the group of translations in the direction of the z-axis. Of course, as a subgroup of the Euclidean group A^2 is also a subgroup of the Möbius group, so that, as the orbit of the point with position vector $\mathfrak{e}_1 r$, $r > 0$, the resulting surface in the Euclidean space $\boldsymbol{E}^3$ whose points have the position vectors

$$\mathfrak{z}(s,t,r) = (\mathfrak{e}_1 \cos s + \mathfrak{e}_2 \sin s)r + \mathfrak{e}_3 t$$

is also an orbit in the sense of Möbius geometry. Applying the inverse stereographic projection leads to corresponding orbits in the Möbius space S^3. For $r = 0$ the orbit reduces to the z-axis. It is easy to prove that all circular cylinders are Möbius equivalent. The generating lines of a circular cylinder are all parallel, from the point of view of Möbius geometry, we have to consider them as circles meeting at ∞ in the Möbius space S^3 obtained from

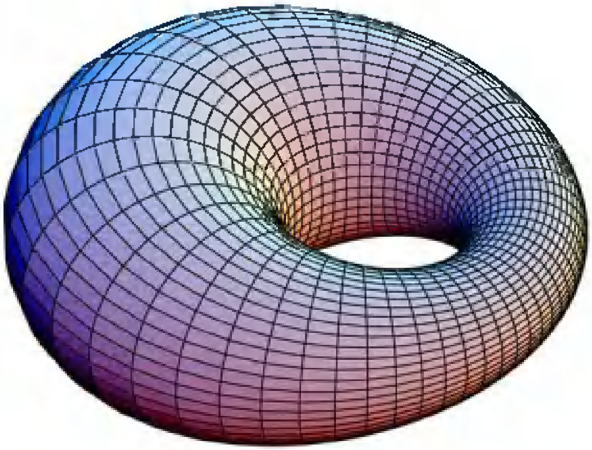


Fig. 2.54. A torus generated by inversion.

Euclidean space by adjoining just this point. ∞ is the only fixed point for all Möbius transformations from A^2. Applying the inversion in a sphere whose center does not lie on it, the circular cylinder is transformed into a surface that is generated by circles, the images of the generating lines, which meet at the center of the inversion sphere. Figure 2.55 shows such a surface. □

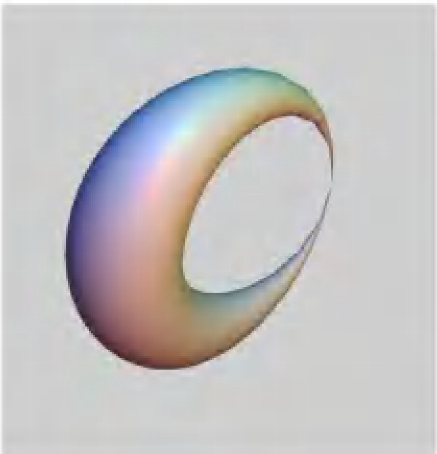


Fig. 2.55. Image of a circular cylinder under an inversion.

Example 12. Circular Cone. Analogous to the preceding example, we again start from the Euclidean space $\boldsymbol{E}^3$. In the conformal group $CE(3)$ we consider the Abelian subgroup $B^2 = \boldsymbol{SO}(2) \times \mathbf{R}^*$; here $\boldsymbol{SO}(2)$ again acts as the group of rotations in the x, y-plane, whereas the multiplicative group $\mathbf{R}^*$ of real numbers acts as the group of dilations in E^3 Note that negative values $t \in \mathbf{R}$ are not excluded. The origin $\boldsymbol{o}$ and the point ∞ are the only fixed points for all transformations from B^2. The latter is a subgroup of the isotropy group of the Möbius space S^3, cf. (1). The orbit of a point from the unit circle in the x, z-plane consists of the points with position vectors

$$\mathfrak{c}(s,t,a) = ((\mathfrak{e}_1 \cos s + \mathfrak{e}_2 \sin s)\cos a + \mathfrak{e}_3 \sin a)t \tag{57}$$

For $a = 0$ the orbit is the x, y-plane itself, for $a = \pi/2$ it is the z-axis, and for the remaining values of a we obtain the complete circular cone generated by all lines through the origin with slope angle a, where the origin, being a fixed point, has to be omitted. Again applying the inversion in a sphere whose center does not lie on the circular cone and is different from the origin, the generating lines are transformed into circles meeting at the center of the inversion sphere as well as the image of the origin under this reflection. Figure 2.56 shows such a surface. □

Fig. 2.56. Image of a circular cone under an inversion.

The surfaces discussed in the last three examples have much in common. Since, in the sense of Möbius geometry, lines are also circles, the surfaces under consideration are Möbius equivalent to surfaces of revolution generated by circles. On each of them there are two families of circles generating the surface, one consisting of the generators for the surface of revolution, and the other formed by their orthogonal trajectories. In differential geometry they were already studied in 1822 by C. Dupin, and after him they are named *cyclids of Dupin*. Depending on the type of Euclidean surface to which a cyclid of Dupin is Möbius equivalent, it is said to be of *toroidal, cylindrical,* or *conical type.* In the last two cases, the fixed points are frequently considered as singularities of the surface; note, however, that they do not belong to the surface defined as an orbit. It is clear that cyclids of Dupin with different type cannot be Möbius equivalent. The cyclids of toroidal and conical type, defined in a Euclidean context, each depend on one parameter: In the first case this is the ratio of the radius of the generating circle to the distance from the center to the axis of revolution; in the second it is the opening angle of the cone. In contrast to this, all Dupin surfaces of cylindrical type are Möbius equivalent. (Proofs as exercises!)

The cyclids of Dupin have numerous generalizations occurring in connection with classification results in differential geometry as well as *Lie sphere geometry.* The book [31] by Thomas E. Cecil, which appeared in 1992, provides an excellent introduction into this topic; it describes the historical development, contains numerous references to the literature, and new, interesting results. The first textbook on sphere geometry in dimensions 2 and 3 is the book [15] by W. Blaschke and G. Thomsen. It represents a summary of the results obtained in Möbius geometry and Lie sphere geometry up to the time of its appearance.[5] Lie sphere geometry has its origin in Lie's theory of contact transformations. The algebraic foundation of the n-dimensional Lie sphere geometry is based on the $(n+3)$-dimensional pseudo-Euclidean vector space of index 2. It contains Möbius geometry, and hence also the Euclidean, hyperbolic, and elliptic ones, as geometries defined by certain subgroups. In fact, n-dimensional hyperbolic geometry is obtained from Möbius geometry by fixing a hypersphere, and then, having chosen a scale, identifying one of the hemispheres it bounds with n-dimensional hyperbolic space. Analogously one obtains the Euclidean and the elliptic geometries as *subgeometries* of Möbius geometry.

Example 13. Loxodromes. As a preparation for the following example, the spiral surfaces, we now want to introduce a family of curves, also interesting in itself, in the Euclidean plane $\boldsymbol{E}^2$, and hence in the sphere S^2, which are themselves orbits. To this end, we consider the one-parameter subgroup $B(a) = \{\gamma_a(t) | t \in \mathbf{R}\} \subset CE(2)$ of the conformal Euclidean group of the plane, whose elements are defined in an orthonormal standard basis by the matrices

$$\gamma_a(t) := \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \mathrm{e}^{at} \quad \text{with } a = \text{const},\ a, t \in \mathbf{R}. \tag{58}$$

[5] In general, the reader should be careful with respect to historical annotations in mathematical textbooks (including the present one). Frequently, the experts don't have the necessary time and patience to completely track down the sources and study them in detail. E.g., in [31] the co-author G. Thomsen of the book [15] is neither mentioned in the text nor in the bibliography; not even his papers are quoted.

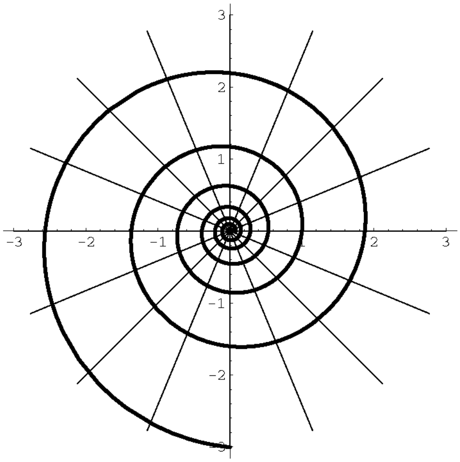


Fig. 2.57. A loxodrome in the plane.

The corresponding transformations are rotations simultaneously coupled with a dilation of the plane, where the coupling is determined by the parameter a; for $a = 0$ we obtain the rotation group $\boldsymbol{SO}(2)$ of the plane. The only fixed points of all transformations $\gamma_a(t)$ different from the identity are the origin and the point ∞. The orbits of a point $\boldsymbol{x}_0 = \mathfrak{e}_1 x_0$ on the x-axis under the group $H(a)$ are the spirals called *loxodromes* with the position vector

$$\mathfrak{x}_a(t) := (\mathfrak{e}_1 \cos t + \mathfrak{e}_2 \sin t) x_0 \, \mathrm{e}^{at} \ \text{ mit } a \neq 0,\ x_0 \neq 0, \tag{59}$$

and, in the case $a = 0$, circles, cf. Figure 2.57. The curves in the sphere resulting from applying the inverse stereographic projection to a loxodrome are called likewise; Figure 2.58 shows such a curve with the coordinate net on the sphere. A ship following a constant compass course moves along a loxodrome or, in the case $a = 0$, on a circle of latitude. In fact, the family of meridians of S^2 (with the origin as South and the point ∞ as North Pole) is transformed into itself under the transformations from $H(a)$. Since the same is true of the orbit and as these transformations are conformal, the angle at which the (tangent to the) loxodrome intersects the meridian is constant along the loxodrome. □

Example 14. The paper by R. Sulanke [101] deals with the classification of homogeneous surfaces in Möbius space S^3, i.e. with surfaces which are orbits for subgroups of the Möbius group G_4. In addition to planes, spheres, and the cyclids of Dupin this also includes the *spiral cylinders,* which we want to describe in this example. In [101] it was shown using a differential-geometric method developed by E. Cartan that this list comprises all the homogeneous surfaces in S^3. The spiral cylinders arise as surfaces in the Euclidean space

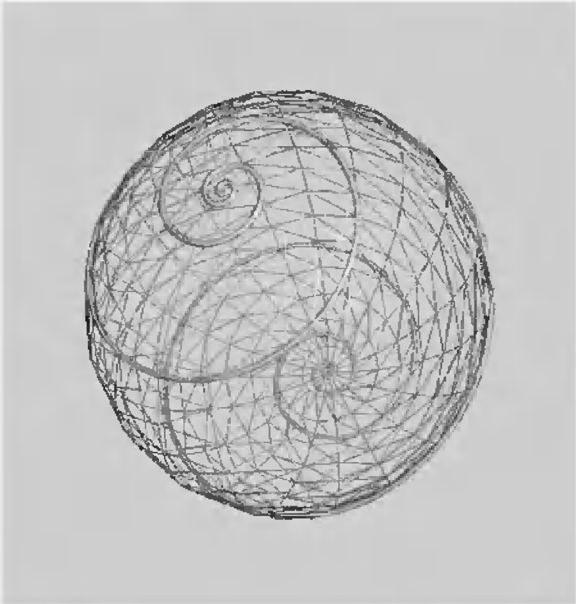


Fig. 2.58. A loxodrome on the sphere.

$\boldsymbol{E}^3$ by erecting the perpendicular cylinder on the plane of a loxodrome, cf. Figure 2.59. By (59) we thus have that

$$\mathfrak{z}_a(s,t) = \mathfrak{x}_a(t) + \mathfrak{e}_3 s$$

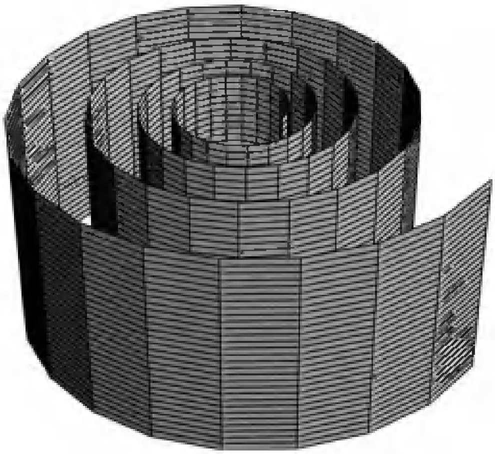


Fig. 2.59. A spiral cylinder.

is the position vector for the spiral cylinder with the parameter $a \neq 0$. The spiral cylinders are special cases of the *spiral surfaces* already introduced by S. Lie. They arise by subjecting a curve in the x, z-plane to the Möbius transformations determined by the *spiral group*

$$\gamma_a(t) := \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathrm{e}^{at} \text{ with } a = \text{const}, \ a, t \in \mathbf{R}; \tag{60}$$

for $a = 0$ we obtain the well-known *surfaces of revolution.* If the generating curve is parallel to the z-axis, then we obtain the spiral cylinder. The subgroup $M_a^2 \subset G_4$ generating the spiral cylinder as its orbit is the group depending on the parameter a generated by the transformations (60) and the parallel displacements in the direction of the z-axis; its elements are affine maps in $\boldsymbol{E}^3$ of the form

$$f_a(s,t) : \boldsymbol{x} = [\mathfrak{x}] \in \boldsymbol{E}^3 \longmapsto \boldsymbol{y} = [\gamma_a(t)\mathfrak{x} + \mathfrak{e}_3 s]. \tag{61}$$

It is easy to verify that for $a \neq 0$ the group M_a^2 is not Abelian; in fact,

$$f_a(s,t) \circ f_a(s_0,t_0) = f_a(s_0 \,\mathrm{e}^{at} + s, t_0 + t).$$

The orbit of the point $[\mathfrak{e}_1 x_0]$ on the x-axis under M_a^2 leads to a parameter representation of a spiral cylinder that is slightly different from the one above. Obviously, every point on the z-axis has the z-axis itself as its orbit. The only common fixed point of all transformations from M_a^2 is the point ∞. In Euclidean space the spiral cylinder 2.59 does actually not look too interesting; one obtains a nicer and, from the point of view of Möbius geometry, more accurate image of part of the surface by considering its behavior at infinity. More precisely, reflecting a neighborhood of ∞ by an inversion into the finite region yields Figure 2.60. There, the generating lines of the spiral cylinder appear as circles meeting at the center of the inversion sphere. □

Fig. 2.60. Inversion of a spiral cylinder.

2.8 Projective Symplectic Geometry

A *projective symplectic geometry* is understood to be the geometry of a projective space $\boldsymbol{P}^{2n-1}$ under the action of the projective symplectic group $\boldsymbol{PSp}_n$ (cf. Section 1.3) in the sense of F. Klein's Erlanger Programm. By Proposition 1.8.5 and the definition of the symplectic groups in this section we will always have to consider a vector space of even dimension, $\boldsymbol{V}^{2n}$, over a field K with a bilinear, alternating, non-degenerate scalar product $\langle , \rangle$; let F be the non-degenerate null system defined by this scalar product. The projective symplectic group $\boldsymbol{PSp}_n$ was already defined in 1.3 as the isotropy group of the null system F. According to Proposition 1.8.5 all projective symplectic geometries of equal dimension over the same field K are isomorphic. As the Examples 1.1, 1.2 show, the isotropy group $\boldsymbol{PSp}_n$ of F depends upon the algebraic properties of the field K. In general, the isotropy group of F in the group of all projectivities $\boldsymbol{PL}_{2n-1}$ is larger than the group of projectivities induced by the linear maps $g \in \boldsymbol{Sp}(\boldsymbol{V}^{2n})$. According to Example 1.2 both groups coincide in the complex projective symplectic geometry; on the other hand, Example 1.1 shows that in the real case $K = \mathbf{R}$ apart from the projectivities induced by the linear symplectic transformations $g \in \boldsymbol{Sp}(n, \mathbf{R})$ there are in addition the anti-symplectic ones, whose inducing linear transformations reverse the sign of the scalar product. The map s_o from (1.21), cf. also (1.22), is an example for this type of projectivities. To have a consistent projective point of view it is here necessary to verify, each time, the invariance with respect to the transformation s_o as well. In the following we take as the *proper projective symplectic group* $\boldsymbol{P_oSp}_n$ the group of projectivities induced by the linear transformations $g \in \boldsymbol{Sp}(\boldsymbol{V}^{2n})$; we thus have

$$\boldsymbol{P_oSp}_n = p(\boldsymbol{Sp}(\boldsymbol{V}^{2n})) \subset \boldsymbol{PSp}_n,$$

and, see Example 1.1, for the real geometry

$$\boldsymbol{PSp}_n = \boldsymbol{P_oSp}_n \cup s_o(\boldsymbol{P_oSp}_n), \quad K = \mathbf{R}. \tag{1}$$

By Example 2.3 the group $\boldsymbol{P_oSp}_n$ acts transitively on the projective space $\boldsymbol{P}^{2n-1}$, which, in that context, was called the *projective symplectic space.*

As a rule, we will use the symplectic bases defined in Section 1.3; the linear symplectic transformations are then described by matrices of the form (1.20). Choosing the point $\boldsymbol{x}_o = [\mathfrak{a}_{2n}]$ as the origin for projective symplectic space, its *isotropy group* $H \subset \boldsymbol{P_oSp}_n$ is generated by all linear transformations with matrices of the form

$$(a^i_j) \in \boldsymbol{Sp}(n, K) \;\text{ with }\; a^i_{2n} = 0 \;\text{ for }\; i < 2n,\; a^{2n}_{2n} \neq 0. \tag{2}$$

Example 1. As in Section 5 we consider the covering of the real projective space by the $(2n-1)$-sphere, cf. (5.4), which again is identified with the set of *directions*, i.e. the oriented one-dimensional subspaces of the vector space $\boldsymbol{V}^{2n}$.

The symplectic group $\boldsymbol{Sp}(n, \mathbf{R})$ acts transitively on S^{2n-1}, and the isotropy group of the direction determined by $\mathfrak{a}_{2n}$ is the subgroup of H described by matrices satisfying $a_{2n}^{2n} > 0$. The invariants of this transformation group form the subject dealt with in *spherical symplectic geometry.* □

2.8.1 Symplectic Transvections

It is remarkable fact that a statement analogous to Proposition 1.4 concerning the generation of the special linear groups by transvections also holds for the symplectic groups. Obviously, the transvections to be considered here are those preserving the symplectic scalar product; we call them *symplectic transvections.*

Lemma 1. *A transvection* $\mathfrak{x} \mapsto \mathfrak{x} + \mathfrak{b}\omega(\mathfrak{x})$ *of the symplectic vector space* $\boldsymbol{V}^{2n}$ *is symplectic if and only if* $\omega(\mathfrak{x}) = c\langle\mathfrak{b}, \mathfrak{x}\rangle$*; hence each symplectic transvection can be written in the form*

$$t_{(\mathfrak{b},c)}(\mathfrak{x}) := \mathfrak{x} + \mathfrak{b}c\langle\mathfrak{b}, \mathfrak{x}\rangle \text{ with } \mathfrak{b} \in \boldsymbol{V},\ c \in K. \tag{3}$$

Proof. For symplectic vector spaces each transformation of the form (3) is a transvection; the requirement $\omega(\mathfrak{b}) = 0$ (cf. Exercise 1.12) presents no extra condition, since $\boldsymbol{b} = [\mathfrak{b}] \in \boldsymbol{b}^{\perp}$ belongs to its own polar. If $\mathfrak{b} = \mathfrak{o}$ or $c = 0$, then (3) is the identical map, which we call the *trivial transvection.* To characterize symplecticity we consider an arbitrary transvection f of $\boldsymbol{V}$. By Exercise 1.12 it has the form $f(\mathfrak{x}) = \mathfrak{x} + \mathfrak{b}\omega(\mathfrak{x})$, where $\omega \in \boldsymbol{V}'$ belongs to the annihilator of $\mathfrak{b}$. Inserting the map f into the scalar product we see that it is a symplectic transvection if and only if for all $\mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}$

$$\langle\mathfrak{b}, \mathfrak{y}\rangle\omega(\mathfrak{x}) - \langle\mathfrak{b}, \mathfrak{x}\rangle\omega(\mathfrak{y}) = 0.$$

For each non-trivial transvection we have $\mathfrak{b} \neq \mathfrak{o}$. Since the scalar product is non-degenerate, we find a vector $\mathfrak{y}_0 \in \boldsymbol{V}$ with $\langle\mathfrak{b}, \mathfrak{y}_0\rangle \neq 0$. Hence the formula above implies the assertion

$$\omega(\mathfrak{x}) = c\langle\mathfrak{b}, \mathfrak{x}\rangle \text{ with } c = \omega(\mathfrak{y}_0)/\langle\mathfrak{b}, \mathfrak{y}_0\rangle. \square$$

For each symplectic transvection $t_{(\mathfrak{b},c)}$ the polar of $[\mathfrak{b}]$ is the hyperplane of fixed points for $t_{(\mathfrak{b},c)}$. For fixed $\mathfrak{b}$ the set comprising these transformations is an Abelian subgroup of the symplectic group isomorphic to $[K, +]$; actually, as is easily verified,

$$t_{(\mathfrak{b},a)} \circ t_{(\mathfrak{b},c)} = t_{(\mathfrak{b},a+c)}.$$

Now we prove the already announced

Proposition 2. *Each symplectic transformation* $g \in \boldsymbol{Sp}(n, K)$ *can be represented as the product of finitely many symplectic transvections.*

Proof. We will proceed by induction. The assertion for $n = 1$ is nothing but the beginning step of the induction proving Proposition 1.4; in fact, $\boldsymbol{Sp}(1, K) = \boldsymbol{SL}(2, K)$. Without loss of generality, we may suppose that g leaves a vector $\mathfrak{a} \neq \mathfrak{o}$ fixed. To see this, we have to show that the group G generated by the transvections acts transitively on the set of vectors different from the null vector. If $\mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}$ are two vectors spanning a symplectic subspace, i.e. satisfying $\langle \mathfrak{x}, \mathfrak{y} \rangle \neq 0$, then the map (3) with $\mathfrak{b} = \mathfrak{y} - \mathfrak{x}$ and $c = 1/\langle \mathfrak{y}, \mathfrak{x} \rangle$ is a transvection with

$$t_{(\mathfrak{b},c)}(\mathfrak{x}) = \mathfrak{x} + (\mathfrak{y} - \mathfrak{x})\langle \mathfrak{y} - \mathfrak{x}, \mathfrak{x} \rangle / \langle \mathfrak{y}, \mathfrak{x} \rangle = \mathfrak{y}.$$

If, however, $\langle \mathfrak{x}, \mathfrak{y} \rangle = 0$, then there is a vector $\mathfrak{c} \in \boldsymbol{V}$ with $\langle \mathfrak{x}, \mathfrak{c} \rangle \neq 0$ and $\langle \mathfrak{y}, \mathfrak{c} \rangle \neq 0$: Projectively this condition means that we have to find a point $\boldsymbol{c} = [\mathfrak{c}]$ for which $\boldsymbol{c} \notin [\mathfrak{x}]^{\perp}$ as well as $\boldsymbol{c} \notin [\mathfrak{y}]^{\perp}$, i.e. which neither belongs to the polar of $\boldsymbol{x} = [\mathfrak{x}]$ nor to that of $\boldsymbol{y} = [\mathfrak{y}]$. If the vectors $\mathfrak{x}, \mathfrak{y}$ are linearly dependent, then the polars coincide, and every point $\boldsymbol{c}$ of the complement satisfies the condition. If the vectors are linearly independent, then $n > 1$, and the polars intersect in the isotropic projective $(2n-3)$-plane $\boldsymbol{x}^{\perp} \wedge \boldsymbol{y}^{\perp} = (\boldsymbol{x} \vee \boldsymbol{y})^{\perp}$. Now, in each projective space there always is a point that does not lie in the union of two non-intersecting hyperplanes. It suffices to choose a point in each of the two hyperplanes that does not belong to their intersection; on the connecting line of these necessarily different points there is at least one more point, which then cannot belong to any of the hyperplanes. According to what we proved before we thus find symplectic transvections t_1, t_2 with $t_1(\mathfrak{y}) = \mathfrak{c}$ and $t_2(\mathfrak{c}) = \mathfrak{x}$; so, if $g(\mathfrak{x}) = \mathfrak{y} \neq \mathfrak{x}$, then $g_1 = t_2 \circ t_1 \circ g$ has the fixed point $\mathfrak{x}$.

Let now $g(\mathfrak{a}) = \mathfrak{a}$, $\mathfrak{b}$ be a vector with $\langle \mathfrak{a}, \mathfrak{b} \rangle = 1$ and $\mathfrak{b}' = g(\mathfrak{b})$. We want to show that there are transvections h_1, h_2 also leaving $\mathfrak{a}$ fixed and, moreover, satisfying $h_2 \circ h_1(\mathfrak{b}') = \mathfrak{b}$; then the transformation $g_2 := h_2 \circ h_1 \circ g$ leaves each vector in the symplectic subspace $\boldsymbol{U} = [\mathfrak{a}, \mathfrak{b}]$ spanned by $\mathfrak{a}, \mathfrak{b}$ fixed. Thus it suffices to prove that symplectic transformations with this property can be generated by finitely many symplectic transvections. If $\langle \mathfrak{b}, \mathfrak{b}' \rangle \neq 0$, then the maps

$$h_1 = t_{(\mathfrak{b}-\mathfrak{b}',c)} \text{ with } c = 1/\langle \mathfrak{b}, \mathfrak{b}' \rangle$$

and the trivial transvection h_2 each meets the requirement; to see this, note that $\langle \mathfrak{a}, \mathfrak{b} \rangle = \langle \mathfrak{a}, \mathfrak{b}' \rangle = 1$. The case $\mathfrak{b} = \mathfrak{b}'$ is trivial. If $\mathfrak{b} \neq \mathfrak{b}'$ and $\langle \mathfrak{b}, \mathfrak{b}' \rangle = 0$, then because of $\langle \mathfrak{b}', \mathfrak{a} + \mathfrak{b} \rangle \neq 0$ and using what was just proved we find a symplectic transvection h_1 mapping the pair $(\mathfrak{a}, \mathfrak{b}')$ into the pair $(\mathfrak{a}, \mathfrak{a} + \mathfrak{b})$. As $\langle \mathfrak{b}, \mathfrak{a} + \mathfrak{b} \rangle \neq 0$, in the same way we find a transvection h_2 mapping $(\mathfrak{a}, \mathfrak{a} + \mathfrak{b})$ to $(\mathfrak{a}, \mathfrak{b})$; but this is what was to be proved.

Now it is easy to complete the induction. We suppose that $g|_{\boldsymbol{U}} = \mathrm{id}_{\boldsymbol{U}}$ and consider the orthogonal decomposition into the symplectic subspaces $\boldsymbol{V} = \boldsymbol{U} + \boldsymbol{W}$ with $\boldsymbol{W} = \boldsymbol{U}^{\perp}$, both invariant under g. By the induction hypothesis we can represent the symplectic transformation $g|_{\boldsymbol{W}}$ as the product of transvections in $\boldsymbol{W}$. Let $t_{\boldsymbol{W}}$ denote such a transvection, then

$$t_{\boldsymbol{V}} = \mathrm{id}_{\boldsymbol{U}} \oplus t_{\boldsymbol{W}}$$

is a transvection in $\boldsymbol{V}$, and the product of the factors for $g|_{\boldsymbol{W}}$ extended this way yields the representation for g as the product of finitely many symplectic transvections we were looking for. □

Since each transvection has determinant 1, Proposition 2 immediately implies the following, which was already announced in (1.19), Exercise 1.3,

Corollary 3. *Every symplectic transformation $g \in \boldsymbol{Sp}(n, K)$ has determinant 1.* □

2.8.2 Subspaces

We will commence this section with the classification of subspaces $\boldsymbol{W}^k \subset \boldsymbol{V}^{2n}$ in a symplectic vectors space. At the same time, its projective interpretation will lead to the symplectic classification of k-planes. Restricting the scalar product to the subspace $\boldsymbol{W}$ we obtain an alternating bilinear form. But according to Lemma 1.8.4 these are classified by their even rank $2r$. The number $d := k - 2r$ is called the *defect* of this bilinear form; the terms rank and defect will also apply to the corresponding subspaces. Then

Proposition 4. *The rank $2r$ and the defect d of a k-dimensional subspace $\boldsymbol{W}^k$ in a symplectic vector space $\boldsymbol{V}^{2n}$ satisfy the conditions*

$$0 \leq r + d \leq n, \ 2r + d = k. \tag{4}$$

Let $M_{2n,2r,d}$ denote the set of all subspaces with rank $2r$ and defect d. Then $M_{2n,2r,d}$ is not empty if and only if (4) holds. The symplectic group $\boldsymbol{Sp}(\boldsymbol{V}^{2n})$ acts transitively on each of the sets $M_{2n,2r,d}$.

Proof. According to Lemma 1.8.4 we find a basis $\mathfrak{a}_1, \ldots, \mathfrak{a}_r, \mathfrak{b}_1, \ldots, \mathfrak{b}_{r+d}$ for $\boldsymbol{W}^k$ whose vectors have the scalar products

$$\langle \mathfrak{a}_i, \mathfrak{b}_i \rangle = -\langle \mathfrak{b}_i, \mathfrak{a}_i \rangle = 1, i = 1, \ldots, r; \langle \mathfrak{a}_i, \mathfrak{b}_j \rangle = \langle \mathfrak{a}_i, \mathfrak{a}_j \rangle = \langle \mathfrak{b}_i, \mathfrak{b}_j \rangle = 0, i \neq j; \tag{5}$$

a linearly independent vector sequence like this will be called *symplectic*. The inequality in (4) is an immediate consequence of Corollary 2.6; in fact, the subspace $[\mathfrak{b}_1, \ldots, \mathfrak{b}_{r+d}]$ is an $(r + d)$-dimensional, totally isotropic subspace. The inequality in (4) follows from $2r + d = \dim \boldsymbol{W}$. Let $(\mathfrak{a}_i), i = 1, \ldots, 2n$, be a symplectic basis; setting $\mathfrak{b}_j := \mathfrak{a}_{n+j}, j = 1, \ldots, r + d$, we obtain a symplectic vector sequence $\mathfrak{a}_1, \ldots, \mathfrak{a}_r, \mathfrak{b}_1, \ldots, \mathfrak{b}_{r+d}$, spanning a k-dimensional subspace with rank $2r$ and defect d. This again is possible if and only if (4) holds. Since subspaces of equal dimension with equal rank are isomorphic, the transitivity statement follows from Corollary 2.11 of E. Witt's Theorem 2.3. □

Exercise 1. Prove that each symplectic vector sequence in a symplectic vector space can be completed to a symplectic basis. Conclude from this the transitivity statement in Proposition 4 without applying E. Witt's Theorem.

The maximal dimension of any totally isotropic subspace in a symplectic vector space $\boldsymbol{V}^{2n}$ is n; because of their application in mechanics these subspaces, satisfying $r = 0$ as well as $d = n$, are also called *Lagrange subspaces,* see e.g. A. T. Fomenko [38]. Because of $\dim \boldsymbol{W}^{\perp} = 2n - \dim \boldsymbol{W}$, the Graßmann manifold $G_{2n,n}$ of a symplectic vector space is mapped involutively onto itself by the null system F. The Lagrange subspaces are the elements fixed by this map. Interpreting subspaces projectively leads to a corresponding statement in projective geometry. For example, the polar of a line $\boldsymbol{a} = \boldsymbol{x} \vee \boldsymbol{y}$ in the three-dimensional projective symplectic space $\boldsymbol{P}^3$ is again a line, which appears as the support of the line pencil for the polar planes corresponding to the points of $\boldsymbol{a}$:

$$F(\boldsymbol{x} \vee \boldsymbol{y}) = (\boldsymbol{x} \vee \boldsymbol{y})^{\perp} = \boldsymbol{x}^{\perp} \wedge \boldsymbol{y}^{\perp} = F(\boldsymbol{x}) \wedge F(\boldsymbol{y}).$$

Example 2. Consider the $(2n-1)$-dimensional projective space $\boldsymbol{P}^{2n-1}$ over the field K. Its points correspond to the one-dimensional subspaces of the associated symplectic vector space $\boldsymbol{V}^{2n}$. Proposition 4 implies $r = 0$, $d = 1$, and again the transitivity of the group $\boldsymbol{P_oSp}_n$ on $\boldsymbol{P}^{2n-1}$. For a hyperplane the dimension of the associated vector space is $k = 2n - 1$; then (4) implies $d = 1$, $r = n - 1$, and by Proposition 4 the group $\boldsymbol{P_oSp}_n$ acts transitively on the set of hyperplanes. The set of projective lines $\boldsymbol{b} = \boldsymbol{x} \vee \boldsymbol{y}$ splits into two orbits: The first is the linear line complex $\mathfrak{K}_F$ corresponding to the symplectic polarity F, cf. (1.8.21) consisting of the *isotropic lines*, for which $r = 0$ and $d = 2$. And the second is its complement in the Graßmann manifold $P_{2n-1,1}$ of all lines containing the *symplectic lines*, for which $r = 1$ and $d = 0$. The isotropic lines through the point $\boldsymbol{a}$ are the lines $\boldsymbol{a} \vee \boldsymbol{x}$ connecting $\boldsymbol{a}$ to the points in its polar different from $\boldsymbol{a}$: $\boldsymbol{x} \in \boldsymbol{a}^{\perp}$. All other lines through $\boldsymbol{a}$ are symplectic. The following proposition shows that instead of the null system F one may also consider the linear line complex $\mathfrak{K}_F$ of its isotropic lines as the *absolute* for the projective symplectic space. □

Proposition 5. *A projectivity $g \in \boldsymbol{PL}_{2n-1}$ of the projective symplectic space $\boldsymbol{P}^{2n-1}$ maps the absolute line complex $\mathfrak{K}_F$ to itself if and only if $g \in \boldsymbol{PSp}_n$ is projective symplectic.*

Proof. The condition is obviously necessary: each symplectic projectivity $g \in \boldsymbol{PSp}_n$ maps $\mathfrak{K}_F$ bijectively onto itself. Let $(\mathfrak{a}_i)$ be a symplectic basis, and let $\boldsymbol{a}_i = [\mathfrak{a}_i]$, $i = 1, \ldots, 2n$, be the vertices of the corresponding $(2n-1)$-simplex. Moreover, let g be a projectivity induced by the linear map $a \in \boldsymbol{GL}(\boldsymbol{V}^{2n})$ mapping $\mathfrak{K}_F$ into itself, and denote by $\mathfrak{b}_i = a(\mathfrak{a}_i)$ the images of the basis vectors. If $\langle \mathfrak{a}_i, \mathfrak{a}_j \rangle = 0$, then $\langle \mathfrak{b}_i, \mathfrak{b}_j \rangle = 0$ has to hold as well, since by assumption the image of the edge

$$g(\boldsymbol{a}_i \vee \boldsymbol{a}_j) = g(\boldsymbol{a}_i) \vee g(\boldsymbol{a}_j) = [\mathfrak{b}_i, \mathfrak{b}_j]$$

belongs to $\mathfrak{K}_F$. For the scalar products of the image vectors this yields

$$\langle \mathfrak{b}_i, \mathfrak{b}_{n+i} \rangle = -\langle \mathfrak{b}_{n+i}, \mathfrak{b}_i \rangle = \mu_i, i = 1, \ldots, n; \langle \mathfrak{b}_i, \mathfrak{b}_j \rangle = 0 \text{ else.} \tag{6}$$

Since F has rank $2n$, the determinant involving these scalar products is different from zero. This implies that all μ_i are different from zero. Setting

$$\mathfrak{c}_i := \mathfrak{b}_i \text{ and } \mathfrak{c}_{n+i} = \mathfrak{b}_{n+i}/\mu_i \text{ for } i = 1, \ldots, n,$$

the $(\mathfrak{c}_j), j = 1, \ldots, 2n$, form a symplectic basis. Let $h \in \boldsymbol{P_oSp}_n$ denote the symplectic projectivity induced by the symplectic transformation c that transforms the basis $(\mathfrak{c}_j)$ into $(\mathfrak{a}_j)$. The projectivity $h \circ g$ induced by $c \circ a$ then also transforms $\mathfrak{K}_F$ into itself. We show that $f = h \circ g$ is a symplectic projectivity; consequently, this also holds for $g = h^{-1} \circ f$. The map f is induced by $b = c \circ a$; the definition of these maps imply

$$b(\mathfrak{a}_i) = c(\mathfrak{b}_i) = c(\mathfrak{c}_i) = \mathfrak{a}_i, \tag{7}$$

$$b(\mathfrak{a}_{n+i}) = c(\mathfrak{b}_{n+i}) = c(\mathfrak{c}_{n+i})\mu_i = \mathfrak{a}_{n+i}\mu_i \tag{8}$$

for $i = 1, \ldots, n$. We assert that in (8) all μ_i have to coincide. Then the proof is complete; in fact, it is easily verified that in this case we have for arbitrary vectors $\mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}^{2n}$

$$\langle b(\mathfrak{x}), b(\mathfrak{y}) \rangle = \mu \langle \mathfrak{x}, \mathfrak{y} \rangle;$$

hence the map b is conformal symplectic, cf. Lemma 1.1. Consider, e.g., the vectors $\mathfrak{a}_1, \mathfrak{a}_2$. Using (7) and (8) we conclude from

$$\langle \mathfrak{a}_1 + \mathfrak{a}_2, \mathfrak{a}_{n+1} - \mathfrak{a}_{n+2} \rangle = 0$$

that the images under the map b satisfy

$$\langle b(\mathfrak{a}_1 + \mathfrak{a}_2), b(\mathfrak{a}_{n+1} - \mathfrak{a}_{n+2}) \rangle = \langle \mathfrak{a}_1 + \mathfrak{a}_2, \mathfrak{a}_{n+1}\mu_1 - \mathfrak{a}_{n+2}\mu_2 \rangle = \mu_1 - \mu_2 = 0;$$

this is just the assertion. □

2.8.3 Triangles

In this section we are going to classify triangles in the projective symplectic space $\boldsymbol{P}^{2n-1}$. First, let the scalar domain be an arbitrary field K; the complete classification, however, will only be carried out for $K = \mathbf{R}, \mathbf{C}$. We again use the notation for the vertices and sides familiar from school mathematics. Let the triangle $\Delta = (A, B, C)$ be non-degenerate, and suppose that its sides satisfy

$$\boldsymbol{a} = B \vee C, \boldsymbol{b} = C \vee A, \boldsymbol{c} = A \vee B.$$

The number k of isotropic sides provides a first coarse classification of triangles; the possible values are $k = 0, 1, 2, 3$. The simplest case is covered by the following exercise:

Exercise 2. Two triangles with three isotropic sides are symplectically congruent. This case can only occur in projective symplectic spaces $\boldsymbol{P}^N$ of dimensions $N \geq 5$.

If at least one of the triangle's sides is not isotropic, then for the plane $\boldsymbol{H} := A \vee B \vee C$ spanned by the triangle the parameters are $r = d = 1$, i.e. the rank of the scalar product restricted to the corresponding vector space is equal to two. Since the symplectic group acts transitively on the set of planes of this type we may suppose, without loss of generality, that the triangles in question lie in a common plane; moreover, we may assume that this plane lies in a three-dimensional projective space. So from now on we take N to be three. The pole $\boldsymbol{p} = \boldsymbol{H}^{\perp}$ belongs to $\boldsymbol{H}$; a line $\boldsymbol{h} \subset \boldsymbol{H}$ is isotropic if and only if $\boldsymbol{p} \in \boldsymbol{h}$. Consider now the case $k = 2$; let $\boldsymbol{c}$ be the only non-isotropic side of the triangle. Since both the sides $\boldsymbol{b}, \boldsymbol{a}$ are isotropic, $\boldsymbol{p} = C$ has to be their point of intersection. We may thus represent A, B, C as follows by a symplectic vector sequence

$$A = [\mathfrak{e}_1], B = [\mathfrak{e}_3], C = [\mathfrak{e}_2].$$

Then we complete this vector sequence to a symplectic basis by adding suitable vectors. In a similar way, for each triangle of this type we construct a symplectic basis with the same properties; the symplectic transformation mapping these bases into one another provides a symplectic congruence between the triangles. These considerations imply:

Proposition 6. *Two triangles in the projective symplectic space $\boldsymbol{P}^N$ each having precisely two isotropic sides are symplectically congruent. Moreover, two arbitrary pairs of isotropic lines each meeting at a point and spanning a plane that is not totally isotropic are symplectically congruent. Hence there are two classes of pairs of intersecting isotropic lines: line pairs spanning a totally isotropic plane and those spanning a plane with $r = 1$. In the three-dimensional projective symplectic space the first case cannot occur.* □

Now consider the case $k = 1$; let, say, the side $\boldsymbol{b} = A \vee C$ be isotropic and suppose that both the others are not isotropic. Then the pole $\boldsymbol{p}$ lies on the side $\boldsymbol{b}$, and we have $\boldsymbol{p} \neq C$, since otherwise the side $\boldsymbol{a}$ were also isotropic. Choose a symplectic vector sequence adapted to the triangle in the following way:

$$A = [\mathfrak{e}_1], B = [\mathfrak{e}_3], \boldsymbol{p} = [\mathfrak{e}_2]. \tag{9}$$

Since $C \in \boldsymbol{b} = A \vee \boldsymbol{p}$, $C \neq \boldsymbol{p}$, there is a representative $\mathfrak{c}$ for C with the basis representation $\mathfrak{c} = \mathfrak{e}_1 + \mathfrak{e}_2\gamma$. The relation $C \neq A$ implies $\gamma \neq 0$; replacing $\mathfrak{e}_2$ by $\mathfrak{e}_2\gamma$ and again denoting the result by $\mathfrak{e}_2$, in this symplectic vector sequence we have

$$A = [\mathfrak{e}_1], B = [\mathfrak{e}_3], C = [\mathfrak{e}_1 + \mathfrak{e}_2],$$

and as in the case $k = 2$ we obtain

Proposition 7. *Two triangles in the projective symplectic space $\boldsymbol{P}^N$ each containing precisely one isotropic side are symplectically congruent.* □

Constructing in a suitable way for the line pair in question a triangle with precisely one isotropic side one obtains

Corollary 8. *Two pairs of different intersecting lines in which corresponding lines have the same type are symplectically congruent.*

Proof. If both lines are isotropic, then the assertion is true by Proposition 6 and Exercise 2. Consider two intersecting symplectic lines $\boldsymbol{a}, \boldsymbol{c}$ intersecting in B. Then the pole $\boldsymbol{p}$ of the plane they determine lies on none of these lines. Then each line $\boldsymbol{b}$ of the plane pencil with center $\boldsymbol{p}$ not passing through B intersects both lines; let $A = \boldsymbol{b} \wedge \boldsymbol{c}, C = \boldsymbol{a} \wedge \boldsymbol{b}$ be the points of intersection. The resulting triangle has precisely one isotropic side. Proceeding similarly for the second line pair, Proposition 7 implies the assertion. The case of one isotropic and one symplectic line is treated analogously. □

Now we consider the case that all sides of the triangle are symplectic. In order to adapt the symplectic frame we first take a symplectic vector sequence satisfying (9). Let $\mathfrak{c}$ be a vector representing the point C. Its basis representation

$$\mathfrak{c} = \mathfrak{e}_1\gamma_1 + \mathfrak{e}_2\gamma_2 + \mathfrak{e}_3\gamma_3$$

has the property that all its components are different from zero; $\gamma_2 = 0$ would imply $C \in A \vee B$, and if γ_1 or γ_3 were zero, then the triangle would have an isotropic side. The representatives of A and B are only determined up to a transformation

$$\hat{\mathfrak{e}}_1 = \mathfrak{e}_1\mu,\ \hat{\mathfrak{e}}_3 = \mathfrak{e}_3/\mu.$$

The condition $\hat{\gamma}_1 = \hat{\gamma}_3$ for the transformed coefficients leads to the equation $\gamma_1/\gamma_3 = \mu^2$, which, in general, has no solution. For $K = \mathbf{C}$, however, it is always solvable, and otherwise only if γ_1/γ_3 is a square. If this is the case, then by suitably normalizing the representatives $\mathfrak{e}_2$ for $\boldsymbol{p}$ we may arrange all the coefficients in this basis representation for $\mathfrak{c}$ to have the same value $\hat{\gamma} := \hat{\gamma}_1 = \hat{\gamma}_3$; the transition to the representative $\mathfrak{c}/\hat{\gamma}$ then leads to the following canonical representation for $\mathfrak{c}$ (returning to the old notations):

$$\mathfrak{c} = \mathfrak{e}_1 + \mathfrak{e}_2 + \mathfrak{e}_3. \tag{10}$$

In real symplectic geometry the situation $\gamma_1/\gamma_3 < 0$ may well occur; in this case a transformation similar to the one above leads to the canonical representation

$$\mathfrak{c} = \mathfrak{e}_1 + \mathfrak{e}_2 - \mathfrak{e}_3. \tag{11}$$

In general, one can construct a canonical representation in which the third coefficient belongs to a complete system of representatives for the cosets in the quotient group K^*/K^{*2}. The above considerations imply:

Proposition 9. *In the complex projective symplectic space $\boldsymbol{P}^N$ all triangles without isotropic sides are symplectically congruent. In real projective symplectic geometry there are precisely two classes of symplectically congruent triangles with exclusively symplectic sides with respect to the group $\boldsymbol{P_oSp}_n$; they are distinguished by the relation* (10) *or* (11) *satisfied by the representative for C in the canonically adapted frame. With respect to the full real symplectic group all triangles with exclusively symplectic sides are congruent.*

P r o o f. The asserted congruence follows from the equality of the canonical forms by completing the symplectic vector sequence to a symplectic basis. The existence of congruent triangles which are not congruent with respect to the real proper symplectic group $\boldsymbol{P_oSp}_n$ can be shown as follows: Consider the triangles (A, B, C), $(A, B, \hat{C})$ in the real $\boldsymbol{P}^3$, where we assume that, in an arbitrarily fixed symplectic frame,

$$A = [\mathfrak{e}_1],\ B = [\mathfrak{e}_3],\ C = [\mathfrak{e}_1 + \mathfrak{e}_2 + \mathfrak{e}_3],\ \hat{C} = [\mathfrak{e}_1 + \mathfrak{e}_2 - \mathfrak{e}_3].$$

Let a be a linear transformation inducing the projectivity g satisfying $(g(A), g(B), g(C)) = (A, B, C)$. Then

$$a(\mathfrak{e}_1) = \mathfrak{e}_1\alpha, a(\mathfrak{e}_2) = \mathfrak{e}_2\beta, a(\mathfrak{e}_3) = \mathfrak{e}_3\gamma, a(\mathfrak{e}_1 + \mathfrak{e}_2 + \mathfrak{e}_3) = (\mathfrak{e}_1 + \mathfrak{e}_2 - \mathfrak{e}_3)\lambda,$$

where all occurring coefficients are different from zero. The second condition has to be satisfied, since a symplectic transformation leaves a plane together with its poles fixed. Inserting the first three equations into the last one leads to

$$\alpha = \beta = -\gamma = \lambda,\ \text{i.e.}\ \langle a(\mathfrak{e}_1), a(\mathfrak{e}_3)\rangle = -\lambda^2,$$

and hence a cannot lie in $\boldsymbol{Sp}(n, \mathbf{R})$. The linear transformation defined by

$$s(\mathfrak{e}_1) = \mathfrak{e}_1, s(\mathfrak{e}_2) = \mathfrak{e}_2, s(\mathfrak{e}_3) = -\mathfrak{e}_3, s(\mathfrak{e}_4) = -\mathfrak{e}_4$$

reverses the sign of the scalar product, and hence induces a map in the projective symplectic group $\boldsymbol{PSp}(n, \mathbf{R})$ thus proving symplectic congruence of the triangles. □

2.8.4 Skew Lines

The symplectic congruence of pairs of intersecting lines was thoroughly discussed in the previous section. This leaves the interesting task of classifying pairs of non-intersecting, i.e. skew lines. We start by considering pairs of symplectic lines. Any two skew lines always span a three-dimensional projective subspace; if the dimension of the whole space is $N \geq 5$, and if at least one of the lines is symplectic, then this subspace may be symplectic or isotropic; in the latter case its rank and defect are equal to 2.

Example 3. Consider a 5-dimensional projective space. Let $(\mathfrak{e}_i)$, $i = 1, \ldots, 6$, be a symplectic basis for the associated vector space. Then the lines

$$\boldsymbol{a} = [\mathfrak{e}_1] \vee [\mathfrak{e}_3],\ \boldsymbol{b} = [\mathfrak{e}_1 + \mathfrak{e}_2] \vee [\mathfrak{e}_3 + \mathfrak{e}_4]$$

are symplectic, and the projective subspace they span has rank and defect both equal to 2. □

The expression defining an invariant for symplectic lines presented in the following lemma can be found in the book [95] by B. A. Rosenfeld:

Lemma 10. *Let* $\boldsymbol{g}_1 = [\mathfrak{x}] \vee [\mathfrak{y}]$, $\boldsymbol{g}_2 = [\mathfrak{v}] \vee [\mathfrak{w}]$ *be two symplectic lines in the* N*-dimensional projective symplectic space* $\boldsymbol{P}^N$. *Then the number*

$$\operatorname{sym}(\boldsymbol{g}_1, \boldsymbol{g}_2) := \frac{\langle \mathfrak{x}, \mathfrak{v}\rangle\langle \mathfrak{y}, \mathfrak{w}\rangle - \langle \mathfrak{x}, \mathfrak{w}\rangle\langle \mathfrak{y}, \mathfrak{v}\rangle}{\langle \mathfrak{x}, \mathfrak{y}\rangle\langle \mathfrak{v}, \mathfrak{w}\rangle} \tag{12}$$

only depends on the lines and does not depend on the points or vectors representing them; consequently, it is a symplectic invariant for the pair of symplectic lines. Obviously,

$$\operatorname{sym}(\boldsymbol{g}_1, \boldsymbol{g}_2) = \operatorname{sym}(\boldsymbol{g}_2, \boldsymbol{g}_1). \tag{13}$$

Proof. Under a transformation of the bases $\mathfrak{x}, \mathfrak{y}$ and $\mathfrak{v}, \mathfrak{w}$ for the vector spaces associated with the lines the denominator as well as the numerator are multiplied by the determinant of the transformation matrix. The invariance follows from the invariance of the scalar products. □

Example 4. As in most cases up to now the vector representing a point $\boldsymbol{x}$ will be denoted by the corresponding Fraktur letter below, $\boldsymbol{x} = [\mathfrak{x}]$. Let, moreover, $(\mathfrak{e}_i)$, $i = 1, \ldots, 2n$, be a symplectic basis to which the coordinate simplex $(\boldsymbol{e}_i)$ corresponds. For $N = 2n - 1 = 3$ the coordinate tetrahedron has precisely two symplectic sides which are polar to one another:

$$F(\boldsymbol{e}_1 \vee \boldsymbol{e}_3) = \boldsymbol{e}_2 \vee \boldsymbol{e}_4.$$

Equation (12) immediately implies $\operatorname{sym}(\boldsymbol{e}_1 \vee \boldsymbol{e}_3, F(\boldsymbol{e}_1 \vee \boldsymbol{e}_3) = 0$. Since the symplectic group acts transitively on the set of symplectic lines, the following holds for a symplectic line $\boldsymbol{g}$ in general:

$$\operatorname{sym}(\boldsymbol{g}, F(\boldsymbol{g})) = 0. \tag{14}$$

The invariant $\operatorname{sym}(\boldsymbol{g}_1, \boldsymbol{g}_2)$ *is, however, also equal to zero if* $\boldsymbol{g}_2$ *intersects the polar of* $\boldsymbol{g}_1$. Because of the transitivity just mentioned, we may always suppose $\boldsymbol{g}_1 = \boldsymbol{e}_1 \vee \boldsymbol{e}_3$. So let $\boldsymbol{e}_2$ be the point of intersection of $\boldsymbol{g}_2 = \boldsymbol{e}_2 \vee \boldsymbol{a}$ with the polar of $\boldsymbol{g}_1$; here $\boldsymbol{a}$ is an arbitrary point not belonging to the polar plane $\boldsymbol{e}_1 \vee \boldsymbol{e}_2 \vee \boldsymbol{e}_3 = F(\boldsymbol{e}_2)$. Because of $\langle \mathfrak{e}_1, \mathfrak{e}_2\rangle = \langle \mathfrak{e}_3, \mathfrak{e}_2\rangle = 0$ the definition (12) immediately implies the assertion. Note that for pairs of skew lines not symplectically congruent the invariants sym may coincide; in fact, a pair of lines which are not polar can never be congruent to a pair of polar lines. Soon we will learn that, at least in the three-dimensional projective space, this can only happen for the values zero and one of the invariant. It is easy to prove that

$$\operatorname{sym}(\boldsymbol{g}, \boldsymbol{g}) = 1 \tag{15}$$

for each symplectic line $\boldsymbol{g}$. *The invariant* $\operatorname{sym}(\boldsymbol{g}_1, \boldsymbol{g}_2)$ *is, however, also equal to one if* $\boldsymbol{g}_2$ *intersects* $\boldsymbol{g}_1$. We may suppose that $\boldsymbol{g}_1 = \boldsymbol{e}_1 \vee \boldsymbol{e}_3$ and $\boldsymbol{g}_2 = \boldsymbol{e}_3 \vee \boldsymbol{a}$, $\boldsymbol{a} \notin F(\boldsymbol{e}_3)$; equation (12) implies the assertion. Using (11) one can show that for the line pair in 5-dimensional space spanning a plane of rank and defect 2 described in Example 3 the invariant is equal to 1. So, in summary, we constructed three mutually not congruent symplectic line pairs whose invariant is equal to one; they differ, however, in the rank or defect of the subspaces they span. □

The following exercise shows that the invariant sym is analogous to the cosine of the stationary angles for subspaces in the metric geometries.

Exercise 3. Let $\boldsymbol{g}, \boldsymbol{h}$ be two symplectic lines. The subspaces $\boldsymbol{U}, \boldsymbol{W}$ corresponding to them define orthogonal decompositions

$$\boldsymbol{V} = \boldsymbol{U} \oplus \boldsymbol{U}^{\perp}, \boldsymbol{V} = \boldsymbol{W} \oplus \boldsymbol{W}^{\perp}.$$

Consider the projections $p_1 : \boldsymbol{U} \to \boldsymbol{W}$, $p_2 : \boldsymbol{W} \to \boldsymbol{U}$ they define. Prove that the linear endomorphism $p_2 \circ p_1$ satisfies:

$$p_2 \circ p_1 = \operatorname{sym}(\boldsymbol{g}, \boldsymbol{h}) \operatorname{id}_{\boldsymbol{U}} .$$

The following proposition shows that there are non-congruent line pairs with the same invariant only for the values zero or one of the invariant.

Proposition 11. *Let $(\boldsymbol{g}_1, \boldsymbol{g}_2), (\boldsymbol{h}_1, \boldsymbol{h}_2)$ be pairs of symplectic lines in the projective symplectic space $\boldsymbol{P}^N$, $N = 2n - 1$. If their invariants are equal:*

$$\gamma := \operatorname{sym}(\boldsymbol{g}_1, \boldsymbol{g}_2) = \operatorname{sym}(\boldsymbol{h}_1, \boldsymbol{h}_2),$$

and if $\gamma \neq 0, 1$, then the line pairs are symplectically congruent.

P r o o f. According to Example 4 the line $\boldsymbol{g}_2$ can neither intersect $\boldsymbol{g}_1$ nor its polar $F(\boldsymbol{g}_1)$. Hence $\boldsymbol{g}_1, \boldsymbol{g}_2$ span a three-dimensional projective subspace. By the following lemma this subspace cannot be isotropic, hence it has rank 4:

Lemma 12. *Let $(\boldsymbol{g}_1, \boldsymbol{g}_2)$, $(\boldsymbol{h}_1, \boldsymbol{h}_2)$ be two pairs of symplectic lines each spanning a three-dimensional projective subspace $\boldsymbol{H}^3, \boldsymbol{M}^3$ with defect (and rank) 2. Then*

$$\operatorname{sym}(\boldsymbol{g}_1, \boldsymbol{g}_2) = \operatorname{sym}(\boldsymbol{h}_1, \boldsymbol{h}_2) = 1,$$

and the line pairs are symplectically congruent.

P r o o f. The assumption of the lemma imply $N \geq 5$. Once again we choose the vertices of the coordinate simplex for $\boldsymbol{P}^N$ so that $\boldsymbol{g}_1 = \boldsymbol{e}_1 \vee \boldsymbol{e}_{n+1}$. Let $\boldsymbol{k}$ be the defect line corresponding to the defect subspace of the vector space $\boldsymbol{W}^4$ associated with $\boldsymbol{H}^3$. Since the lines $\boldsymbol{g}_1, \boldsymbol{g}_2$ are symplectic, they do not intersect $\boldsymbol{k}$. The planes $\boldsymbol{k} \vee \boldsymbol{e}_1$ and $\boldsymbol{k} \vee \boldsymbol{e}_{n+1}$ intersect the line $\boldsymbol{g}_2$ at the points $\boldsymbol{a}_1$ and $\boldsymbol{a}_3$, respectively, so that $\boldsymbol{g}_2 = \boldsymbol{a}_1 \vee \boldsymbol{a}_3$; these points have to be different, since they cannot lie in $\boldsymbol{k}$: in fact, $\boldsymbol{g}_2$ is symplectic. Now we determine the points of intersection

$$\boldsymbol{b} = \boldsymbol{k} \wedge (\boldsymbol{e}_1 \vee \boldsymbol{a}_1), \boldsymbol{c} = \boldsymbol{k} \wedge (\boldsymbol{e}_{n+1} \vee \boldsymbol{a}_3). \tag{16}$$

They exist, since the lines defining them are different lines in a projective plane. In addition, they have to be different themselves; if we had $\boldsymbol{b} = \boldsymbol{c}$, then the lines $\boldsymbol{g}_1, \boldsymbol{g}_2$ would belong to the plane $\boldsymbol{e}_1 \vee \boldsymbol{e}_{n+1} \vee \boldsymbol{b}$, which contradicted the assumption. To obtain further vertices of the symplectic frame we now set $\boldsymbol{e}_2 := \boldsymbol{b}$, $\boldsymbol{e}_3 := \boldsymbol{c}$. The representatives $\mathfrak{a}_1, \mathfrak{a}_3, \mathfrak{e}_2, \mathfrak{e}_3$ are normalized to satisfy

$$\mathfrak{a}_1 = \mathfrak{e}_1 + \mathfrak{e}_2,\ \mathfrak{a}_3 = \mathfrak{e}_3 + \mathfrak{e}_{n+1}.$$

Then $\langle \mathfrak{a}_1, \mathfrak{a}_3 \rangle = 1$, and (12) implies $\operatorname{sym}(\boldsymbol{g}_1, \boldsymbol{g}_2) = 1$. The same also holds for the other line pair, for which we find vectors $\hat{\mathfrak{e}}_i, i = 1, 2, 3, 5$, with analogous position relations. Completing each of these vector sequences to a symplectic base for the vector space associated with $\boldsymbol{P}^N$ the symplectic transformation they determine maps the line pairs into one another. □

Now we return to the proof of Proposition 11. Since the subspaces spanned by the pairs are now three-dimensional as well as symplectic, and as the symplectic group acts transitively on the set of such subspaces, we may suppose that all lines in question lie in the same projective space $\boldsymbol{P}^3$, i.e. $N = 3$. As in the proof of the lemma, starting from $\boldsymbol{g}_1 = \boldsymbol{e}_1 \vee \boldsymbol{e}_3$ and the line $\boldsymbol{k} = F(\boldsymbol{g}_1)$ polar to it we construct points $\boldsymbol{a}_1, \boldsymbol{a}_3 \in \boldsymbol{g}_2$ and $\boldsymbol{b}, \boldsymbol{c} \in \boldsymbol{k}$ according to (16). Since, this time, $\boldsymbol{k}$ is symplectic and polar to $\boldsymbol{g}_1$, we now set $\boldsymbol{e}_2 := \boldsymbol{b}$ and $\boldsymbol{e}_4 := \boldsymbol{c}$. The points $(\boldsymbol{e}_i), i = 1, \ldots, 4$, are the vertices of a coordinate tetrahedron; for arbitrary representatives of the points we have by construction

$$\mathfrak{a}_1 = \mathfrak{e}_1\beta_1 + \mathfrak{e}_2\beta_2, \mathfrak{a}_3 = \mathfrak{e}_3\gamma_1 + \mathfrak{e}_4\gamma_2.$$

Here all these coefficients are different from zero, since otherwise $\boldsymbol{g}_2$ would intersect the line $\boldsymbol{g}_1$ or its polar $\boldsymbol{e}_2 \vee \boldsymbol{e}_4$; according to Example 4 this would imply $\gamma = 0$ or $\gamma = 1$. Replacing $\mathfrak{e}_2$ by $\mathfrak{e}_2\beta_2/\beta_1$ and returning to the former notation we then have

$$\mathfrak{a}_1 = (\mathfrak{e}_1 + \mathfrak{e}_2)\beta_1,$$

and dividing $\mathfrak{a}_1$ by β_1 we arrive at a representative for $\boldsymbol{a}_1$ satisfying $\mathfrak{a}_1 = \mathfrak{e}_1 + \mathfrak{e}_2$. Hence, up to a common factor, which we may set equal to one, the vectors $\mathfrak{a}_1, \mathfrak{e}_1, \mathfrak{e}_2$ are fixed. Since the basis $(\mathfrak{e}_i)$ is symplectic, the vectors $\mathfrak{e}_3$ and $\mathfrak{e}_4$ are also uniquely determined; we normalize $\mathfrak{a}_3$ so that

$$1 = \langle \mathfrak{a}_1, \mathfrak{a}_3 \rangle = \langle \mathfrak{e}_1 + \mathfrak{e}_2, \mathfrak{e}_3\gamma_1 + \mathfrak{e}_4\gamma_2 \rangle = \gamma_1 + \gamma_2.$$

Inserting this into (12) leads to:

$$\begin{gathered}\gamma = \operatorname{sym}(\boldsymbol{g}_1, \boldsymbol{g}_2) \\ = \langle \mathfrak{e}_1, \mathfrak{e}_1 + \mathfrak{e}_2 \rangle \langle \mathfrak{e}_3, \mathfrak{e}_3\gamma_1 + \mathfrak{e}_4(1 - \gamma_1) \rangle - \langle \mathfrak{e}_1, \mathfrak{e}_3\gamma_1 + \mathfrak{e}_4(1 - \gamma_1) \rangle \langle \mathfrak{e}_3, \mathfrak{e}_1 + \mathfrak{e}_2 \rangle \\ = \gamma_1.\end{gathered}$$

What we proved so far is summarized in the following lemma:

Lemma 13. *Let $(\boldsymbol{g}_1, \boldsymbol{g}_2)$ be a pair of symplectic lines with invariant $\gamma \neq 0, 1$. Then these lines span a three-dimensional symplectic subspace. There is a symplectic basis for the associated vector space $\boldsymbol{V}^{2n}$ such that*

$$\boldsymbol{g}_1 = \boldsymbol{e}_1 \vee \boldsymbol{e}_2 \text{ and } \boldsymbol{g}_2 = [\mathfrak{e}_1 + \mathfrak{e}_{n+1}] \vee [\mathfrak{e}_{n+1}\gamma + \mathfrak{e}_{n+2}(1-\gamma)]. \tag{17}$$

□

Bases of this kind will be called *canonically adapted.* Now the proof of Proposition 12 is immediate: If $(\hat{\mathfrak{e}}_i)$ is a basis canonically adapted to the line pair $\boldsymbol{h}_1, \boldsymbol{h}_2$, then, since the coordinates of the vectors determining the line pairs are equal, the transformation defined by $g(\mathfrak{e}_i) = \hat{\mathfrak{e}}_i, i = 1, \dots 2n$, maps these, and hence the line pairs as well, into one another. □

All that remains is to consider a few special cases. First we discuss pairs of symplectic lines with invariant zero.

Proposition 14. *The invariant* $\operatorname{sym}(\boldsymbol{g}_1, \boldsymbol{g}_2)$ *of two symplectic lines* $\boldsymbol{g}_1, \boldsymbol{g}_2 \subset \boldsymbol{P}^N$ *vanishes if and only if* $\boldsymbol{g}_2$ *meets the polar* $F(\boldsymbol{g}_1)$, $\boldsymbol{g}_2 \wedge F(\boldsymbol{g}_1) \neq \boldsymbol{o}$. *In this case there are two classes of congruent line pairs:* $\boldsymbol{g}_2 \subset F(\boldsymbol{g}_1)$ *and* $\boldsymbol{g}_2 \wedge F(\boldsymbol{g}_1)$ *is a single point.*

Proof. Again we start from $\boldsymbol{g}_1 = \boldsymbol{e}_1 \vee \boldsymbol{e}_3$. If $\boldsymbol{g}_2 = \boldsymbol{a}_1 \vee \boldsymbol{a}_3$, then for any of its points we have

$$[\mathfrak{a}_1 s + \mathfrak{a}_3 t] \in \boldsymbol{g}_2 \wedge F(\boldsymbol{g}_1)$$

if and only if (s, t) is a non-trivial solution of the following homogeneous system of equations:

$$\langle \mathfrak{e}_1, \mathfrak{a}_1 \rangle s + \langle \mathfrak{e}_1, \mathfrak{a}_3 \rangle t = 0, \ \langle \mathfrak{e}_3, \mathfrak{a}_1 \rangle s + \langle \mathfrak{e}_3, \mathfrak{a}_3 \rangle t = 0. \tag{18}$$

Such a solutions exists if and only if the determinant of the system vanishes, but this is just the numerator of the invariant (12). If all coefficients are equal to zero, then $\boldsymbol{g}_2 \subset F(\boldsymbol{g}_1)$; otherwise, the solution consists of a single point. Obviously, in the first case two pairs of symplectic lines are always symplectically congruent; in fact, two such lines occur as the edges of symplectic coordinate simplices. In the second case we obtain a canonically adapted coordinate simplex by first setting $\boldsymbol{e}_2 = \boldsymbol{g}_2 \wedge F(\boldsymbol{g}_1)$, then choosing one more point $\boldsymbol{e}_{n+2}$ on the line $\boldsymbol{g}_2$, and, in dimensions larger than 3, completing the basis symplectically in a suitable way. □

Now we consider the case $\operatorname{sym}(\boldsymbol{g}_1, \boldsymbol{g}_2) = 1$. It is then enough to consider the situation that the subspace spanned by the skew lines is symplectic, i.e. has defect zero; the case rank = defect = 2 was already treated in Lemma 12. In the three-dimensional space $\boldsymbol{P}^3$ the situation is particularly symmetric:

Exercise 4. Let $\boldsymbol{g}_1, \boldsymbol{g}_2 \subset \boldsymbol{P}^3$ be symplectic lines. Prove

$$\operatorname{sym}(\boldsymbol{g}_1, \boldsymbol{g}_2) + \operatorname{sym}(F(\boldsymbol{g}_1), \boldsymbol{g}_2) = 1.$$

Conclude from this that $\mathrm{sym}(\boldsymbol{g}_1, \boldsymbol{g}_2) = 1$ only if $\boldsymbol{g}_1 = \boldsymbol{g}_2$, or if the lines $\boldsymbol{g}_1, \boldsymbol{g}_2$ intersect in a point. Hence for skew lines both is impossible; this only leaves the situation described in Lemma 12.

According to Corollary 8 intersecting symplectic line pairs are always congruent. Since for higher dimensions a pair of symplectic lines always lies in a three-dimensional subspace of defect two or 0, the results formulated in Lemma 12 and Exercise 4 also hold in this case. *If the invariant of a symplectic line pair is equal to one, then the only possible cases are the three classes of congruent line pairs described in Example* 4. We leave the details to the reader; note that the restriction of the scalar product to the associated vector subspace equips each symplectic subspace of a projective symplectic space with a symplectic geometry.

In the following we consider two more types of skew line pairs in which at least one of the lines is isotropic. In this case the invariant sym is not defined. We will see that, here too, all pairs of the same type are symplectically congruent.

Example 5. The sides $\boldsymbol{e}_1 \vee \boldsymbol{e}_2$, $\boldsymbol{e}_3 \vee \boldsymbol{e}_4$ of a symplectic coordinate tetrahedron are an example for a skew pair of isotropic lines. We want to show that all such pairs, which span a three-dimensional symplectic subspace, can be described in this way: There is a symplectic frame in which they are opposite skew edges. As above, we may also suppose $N = 3$, $n = 2$ in this case. So let $\boldsymbol{g}_1, \boldsymbol{g}_2 \subset \boldsymbol{P}^3$ be a skew, isotropic line pair. On the line $\boldsymbol{g}_1$ we choose two different points $\boldsymbol{e}_1 = [\mathfrak{e}_1]$, $\boldsymbol{e}_2 = [\mathfrak{e}_2]$ as base points for the coordinate tetrahedron. Let the line $\boldsymbol{g}_2$ be spanned by the points $\boldsymbol{a} = [\mathfrak{a}]$, $\boldsymbol{b} = [\mathfrak{b}]$. We claim that there are uniquely determined points $\boldsymbol{e}_3, \boldsymbol{e}_4 \in \boldsymbol{g}_2$ for which $(\boldsymbol{e}_i), i = 1, \ldots, 4$, is a symplectic coordinate tetrahedron. The conditions $\boldsymbol{e}_3$ has to meet are expressed in the following linear system of equations for the unknown representative $\mathfrak{e}_3 = \mathfrak{a}s + \mathfrak{b}t$:

$$\begin{aligned}
\langle \mathfrak{e}_1, \mathfrak{e}_3 \rangle &= \langle \mathfrak{e}_1, \mathfrak{a} \rangle s + \langle \mathfrak{e}_1, \mathfrak{b} \rangle t = 1, \\
\langle \mathfrak{e}_2, \mathfrak{e}_3 \rangle &= \langle \mathfrak{e}_2, \mathfrak{a} \rangle s + \langle \mathfrak{e}_2, \mathfrak{b} \rangle t = 0.
\end{aligned}$$

The determinant of the scalar products for the basis $(\mathfrak{e}_1, \mathfrak{e}_2, \mathfrak{a}, \mathfrak{b})$ of the symplectic vector space $\boldsymbol{V}^4$ is the square of the determinant for this system of equations, and this is different from zero. Hence the system has a unique solution. Analogously, $\boldsymbol{e}_4 = [\mathfrak{e}_4]$ is the solution of the system of equations below:

$$\begin{aligned}
\langle \mathfrak{e}_1, \mathfrak{e}_4 \rangle &= \langle \mathfrak{e}_1, \mathfrak{a} \rangle s + \langle \mathfrak{e}_1, \mathfrak{b} \rangle t = 0, \\
\langle \mathfrak{e}_2, \mathfrak{e}_4 \rangle &= \langle \mathfrak{e}_2, \mathfrak{a} \rangle s + \langle \mathfrak{e}_2, \mathfrak{b} \rangle t = 1.
\end{aligned}$$

But this immediately shows that the line pair is symplectically congruent to the example described at the beginning. □

In general we have

Proposition 15. *Two pairs of isotropic lines in the N-dimensional projective symplectic space are symplectically congruent if and only if they span projective subspaces with equal rank and defect, respectively.* □

Exercise 5. Prove the assertion of Proposition 15 in the cases where the defect of the subspace spanned by the lines is positive, and determine the minimal dimension of a projective space in which skew line pairs spanning a subspace of defect two or four, respectively, can occur. Find examples for such line pairs.

Example 6. Now we consider the case that one of the lines is symplectic, and the other one is isotropic. We take any symplectic basis of $\boldsymbol{V}^4$ and set $\boldsymbol{g}_1 = \boldsymbol{e}_1 \vee \boldsymbol{e}_3$,

$$\mathfrak{a}_1 = \mathfrak{e}_2 + \mathfrak{e}_3,\ \mathfrak{a}_2 = \mathfrak{e}_1 + \mathfrak{e}_4, \tag{19}$$

then $\boldsymbol{g}_1$ and $\boldsymbol{g}_2 = \boldsymbol{a}_1 \vee \boldsymbol{a}_2$ form such a pair; the defect of the projective subspace spanned by $\boldsymbol{g}_1, \boldsymbol{g}_2$ is equal to zero. Conversely, we will show that for an arbitrary line pair $\boldsymbol{g}_1, \boldsymbol{g}_2$ with these properties there is a symplectic basis in which $\boldsymbol{g}_1 = \boldsymbol{e}_1 \vee \boldsymbol{e}_3$ and $\boldsymbol{g}_2 = \boldsymbol{a}_1 \vee \boldsymbol{a}_2$; here $\boldsymbol{a}_1, \boldsymbol{a}_2$ are points represented by the vectors (19). This again implies the symplectic congruence of all such pairs. □

Proposition 16. *Let $\boldsymbol{g}_1, \boldsymbol{g}_2$ be two skew lines in the N-dimensional projective symplectic space, $\boldsymbol{g}_1$ symplectic, $\boldsymbol{g}_2$ isotropic, together spanning a three-dimensional symplectic projective subspace. Then they are symplectically congruent to the pair described in Example* 6.

Proof. As earlier, in the examples with rank four and defect zero, we may suppose $N = 3$. First we choose two points $\boldsymbol{e}_1, \boldsymbol{e}_3 \in \boldsymbol{g}_1$ and represent them so that $\langle \mathfrak{e}_1, \mathfrak{e}_3 \rangle = 1$. The line $\boldsymbol{g}_2$ cannot lie in the polar of any point $\boldsymbol{x} \in \boldsymbol{g}_1$; because of $\boldsymbol{g}_2 \subset F(\boldsymbol{x})$ and, since $\boldsymbol{x} \in \boldsymbol{g}_1$ implies $F(\boldsymbol{g}_1) \subset F(\boldsymbol{x})$, we then had two different lines, the symplectic one $F(\boldsymbol{g}_1)$ and the isotropic one $\boldsymbol{g}_2$ in the plane $F(\boldsymbol{x})$. They thus determined a point of intersection $\boldsymbol{a} = F(\boldsymbol{g}_1) \wedge \boldsymbol{g}_2$. But then the polar of $\boldsymbol{a}$ were

$$F(\boldsymbol{a}) = \boldsymbol{g}_1 \vee F(\boldsymbol{g}_2) = \boldsymbol{g}_1 \vee \boldsymbol{g}_2,$$

since $\boldsymbol{g}_2$ is isotropic, i.e. $\boldsymbol{g}_2 = F(\boldsymbol{g}_2)$. As $\boldsymbol{g}_1, \boldsymbol{g}_2$ are skew, they span the whole space and not just the plane $F(\boldsymbol{a})$, so that the equalities above show a contradiction. Hence the points of intersection

$$\boldsymbol{a}_1 := \boldsymbol{g}_2 \wedge F(\boldsymbol{e}_3),\ \boldsymbol{a}_2 := \boldsymbol{g}_2 \wedge F(\boldsymbol{e}_1)$$

are well-defined. These points are different; in fact, otherwise there were a point of intersection $F(\boldsymbol{g}_1) \wedge \boldsymbol{g}_2$ with $\boldsymbol{a} = \boldsymbol{a}_1 = \boldsymbol{a}_2$, whose existence we just excluded. Take representatives $\boldsymbol{a}_i = [\mathfrak{a}_i], i = 1, 2,$. Then

$$\langle \mathfrak{a}_1, \mathfrak{e}_1 \rangle \neq 0,\ \langle \mathfrak{a}_2, \mathfrak{e}_3 \rangle \neq 0;$$

actually, otherwise we had

$$\boldsymbol{a}_1 \in F(\boldsymbol{e}_1) \wedge F(\boldsymbol{e}_3) = F(\boldsymbol{e}_1 \vee \boldsymbol{e}_3) = F(\boldsymbol{g}_1),$$

which was excluded above; analogously for $\boldsymbol{a}_2$. Thus we may normalize $\mathfrak{a}_1, \mathfrak{a}_2$ so that

$$\langle \mathfrak{a}_1, \mathfrak{e}_1 \rangle = 1, \ \langle \mathfrak{a}_2, \mathfrak{e}_3 \rangle = 1.$$

Now define

$$\mathfrak{e}_2 := \mathfrak{a}_1 - \mathfrak{e}_3, \ \mathfrak{e}_4 := \mathfrak{a}_2 - \mathfrak{e}_1.$$

It is easy to verify that $(\mathfrak{e}_j), j = 1, \ldots, 4$, is a symplectic basis obviously satisfying (19). □

Exercise 6. Find an example for a skew line pair $\boldsymbol{g}_1, \boldsymbol{g}_2 \subset \boldsymbol{P}^N$, $\boldsymbol{g}_1$ symplectic, $\boldsymbol{g}_2$ isotropic, spanning a projective subspace of rank as well as defect two. How large has N to be in order for such an example to exist? Prove that two of these line pairs are always symplectically congruent.

This example, Proposition 16, and Corollary 8 together imply

Corollary 17. *All line pairs $\boldsymbol{g}_1, \boldsymbol{g}_2 \subset \boldsymbol{P}^N$, $\boldsymbol{g}_1$ symplectic, $\boldsymbol{g}_2$ isotropic, spanning subspaces of equal rank and defect are symplectically congruent.* □

Obviously, it is an interesting task to find symplectic invariants for pairs of subspaces with arbitrary dimension and to derive corresponding classification results. Approaches to treat this problem and particular results can be found in the paper [57] by I. M. Jaglom. Since the overall situation is considerably more complicated than in the Euclidean, elliptic, or hyperbolic geometries, we only discuss the sole remaining case for dimension 3, pairs consisting of a plane and a line.

Example 7. Let $(\boldsymbol{A}, \boldsymbol{h})$ be a pair formed by a plane $\boldsymbol{A} = \boldsymbol{A}^2$ and a line $\boldsymbol{h}$ in the symplectic projective space $\boldsymbol{P}^3$. First suppose that the line $\boldsymbol{h}$ is symplectic. Then the following position relations have to be distinguished:

1. $\boldsymbol{h} \subset \boldsymbol{A}$, 2. $\boldsymbol{h}^\perp \subset \boldsymbol{A}$, 3. $\boldsymbol{h}$ and $\boldsymbol{h}^\perp$ both intersect $\boldsymbol{A}$.

Since a symplectic line $\boldsymbol{h}$ and its polar $\boldsymbol{h}^\perp$ are skew, they cannot both lie in the plane $\boldsymbol{A}$. The following is easy to show: Two pairs of the kind in question are symplectically congruent if and only if they satisfy the same position relation. Obviously, the conditions are necessary. To show that they are also sufficient we adapt, in each of the cases 1–3, a symplectic frame to the pair $(\boldsymbol{A}, \boldsymbol{h})$ in which the pair is fixed in a specific way only depending on the respective type of relation. Then the symplectic transformation mapping one of these frames to the other also transforms the pairs into one another. In Case 1 the assumption $\boldsymbol{h} \subset \boldsymbol{A}$ immediately implies $\boldsymbol{a} := \boldsymbol{A}^\perp \in \boldsymbol{h}^\perp$, and hence the pole $\boldsymbol{a}$ of $\boldsymbol{A}$ is the point of intersection $\boldsymbol{a} = \boldsymbol{h}^\perp \wedge \boldsymbol{A}$. Choosing two points $\boldsymbol{e}_1, \boldsymbol{e}_3 \in \boldsymbol{h}$, then setting $\boldsymbol{e}_2 := \boldsymbol{a}$, and finally fixing as $\boldsymbol{e}_4$ a point in $\boldsymbol{h}^\perp$ different from $\boldsymbol{e}_2$ we obtain an adapted frame. In the second case we proceed analogously by interchanging the roles of $\boldsymbol{h}, \boldsymbol{h}^\perp$. Now consider the third case and set $\boldsymbol{e}_1 := \boldsymbol{h} \wedge \boldsymbol{A}$, $\boldsymbol{e}_2 := \boldsymbol{h}^\perp \wedge \boldsymbol{A}$. Then $\boldsymbol{e}_1 \vee \boldsymbol{e}_2 \subset \boldsymbol{A}$ is an isotropic line, and hence the pole $\boldsymbol{a} = \boldsymbol{A}^\perp$ lies in $\boldsymbol{e}_1 \vee \boldsymbol{e}_2$. Here $\boldsymbol{a} \neq \boldsymbol{e}_1$, $\boldsymbol{a} \neq \boldsymbol{e}_2$, since otherwise we would be in the first or the second case: In fact, $\boldsymbol{a} = \boldsymbol{A}^\perp = \boldsymbol{e}_1 \in \boldsymbol{h}$ would imply $\boldsymbol{h}^\perp \subset \boldsymbol{a}^\perp = \boldsymbol{A}$. So if $\mathfrak{a}$ is a vector representing the pole $\boldsymbol{a}$, then we have $\mathfrak{a} = \mathfrak{e}_1 s + \mathfrak{e}_2 t$ with $s, t \neq 0$, and by a suitable normalization we can arrange $s = t = 1$. Completing $\boldsymbol{e}_1, \boldsymbol{e}_2$ by $\boldsymbol{e}_3 \in \boldsymbol{h}$ and $\boldsymbol{e}_4 \in \boldsymbol{h}^\perp$ to a symplectic frame the last equation states that it is adapted to $\boldsymbol{A}^\perp = \boldsymbol{a}, \boldsymbol{h}^\perp$ and hence also to the pair $(\boldsymbol{A}, \boldsymbol{h})$. This implies the assertion. The case of an isotropic line will be discussed in the next exercise. □

Exercise 7. Consider pairs $(\boldsymbol{A}, \boldsymbol{h})$ consisting of a plane $\boldsymbol{A}$ and an isotropic line $\boldsymbol{h}$ in the three-dimensional projective symplectic space $\boldsymbol{P}^3$. Then the only possible position relations are 1. $\boldsymbol{h} \subset \boldsymbol{A}$ or 2. $\boldsymbol{h}$ intersects $\boldsymbol{A}$. Two such pairs with equal position properties are symplectically congruent. Hint. Case 1 occurs if and only if $\boldsymbol{a} = \boldsymbol{A}^\perp \in \boldsymbol{h}$.

2.8.5 Symmetric Bilinear Forms and Quadrics

In this section we will be concerned with the classification of symmetric bilinear forms and quadrics under the action of the symplectic group. To this end we will again apply the theory of adjoint maps presented in Section 2. For higher dimensions and arbitrary bilinear forms we would have to distinguish a vast number of cases; to avoid this we will not carry out the classification in complete generality. Nevertheless, we will describe the main results and illustrate them by a series of examples in three-dimensional projective symplectic space. We start by a few considerations concerning symmetric bilinear forms. For complex vector spaces $\boldsymbol{V}^{2n}$ their symplectic classification can be found in the book by A. I. Maltzev [75], § VII.3. In the case of real vector spaces the classification is a little more involved; it is discussed in the paper [56] by I. M. Jaglom. Both sources also contain the symplectic classification of *null systems,* i.e. the linear line complexes corresponding to symmetric endomorphisms. This classification is a bit simpler than that of symmetric bilinear forms; we will not go into this. For each symmetric bilinear form b on the symplectic vector space $\boldsymbol{V}^{2n}$ there is its *associated endomorphism* a determined by

$$b(\mathfrak{x}, \mathfrak{y}) = \langle a(\mathfrak{x}), \mathfrak{y} \rangle \text{ for all } \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V},$$

which is *skew symmetric,* i.e.

$$\langle a(\mathfrak{x}), \mathfrak{y} \rangle + \langle \mathfrak{x}, a(\mathfrak{y}) \rangle = 0 \text{ for all } \mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}, \tag{20}$$

cf. Section 2.10. Recall the Gram matrix of a symplectic basis,

$$(\langle \mathfrak{e}_i, \mathfrak{e}_j \rangle) = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

and write the matrix of the endomorphism as a block matrix,

$$(a^i_j) = \begin{pmatrix} A & B \\ C & D \end{pmatrix};$$

then condition (20) expressing the skew symmetry of a is equivalent to

$$D = -A', \; B = B', \; C = C', \tag{21}$$

where A' denotes the transposed matrix of A. Lemma 2.20 and Exercise 2.17, d), immediately imply

Lemma 18. *Assigning the associated endomorphism to each symmetric bilinear form defines a linear $\boldsymbol{Sp}(\boldsymbol{V}^{2n})$-isomorphism from the space of symmetric bilinear forms onto the space of skew symmetric endomorphisms on the symplectic vector space $\boldsymbol{V}^{2n}$.* □

This reduces the classification of symmetric bilinear forms to that of skew symmetric endomorphisms. Propositions 2.30 and 2.33 then yield

Corollary 19. *If a is a splitting, skew symmetric endomorphism of the symplectic vector space $\boldsymbol{V}^{2n}$, then $\boldsymbol{V}^{2n}$ can be represented as the direct sum of orthogonal subspaces, which are invariant under a:*

$$\boldsymbol{V}^{2n} = \boldsymbol{U}_0 \oplus \bigoplus_{\lambda \neq 0} (\Delta^k(\lambda) \oplus \Delta^k(-\lambda)). \tag{22}$$

Here $\boldsymbol{U}_0$ denotes the root subspace for the eigenvalue 0, and the Jordan cells $\Delta^k(\lambda)$, $\lambda \neq 0$, are totally isotropic subspaces. According to the multiplicity of the cells $\Delta^k(\lambda)$ of a, these subspaces may occur repeatedly. Moreover, they are combined pairwise forming neutral subspaces. □

For the sake of simplicity we will mainly consider non-degenerate bilinear forms and quadrics in the sequel. Then $\boldsymbol{U}_0$ is trivial. We remark that the classification of *nilpotent* skew symmetric endomorphisms in a symplectic vector space whose only eigenvalue is $\lambda = 0$, i.e. whose characteristic polynomial is $\chi_a(x) = x^{2n}$ (cf. Definition I.5.8.2), presents quite some difficulty. We refer the reader to the literature already cited, [75] and [56]. Since the direct summands $\Delta^k(\lambda) \oplus \Delta^k(-\lambda)$ in (22) are orthogonal to one another, they again are symplectic vector spaces; hence it suffices to consider each summand separately. So let $k = n$, and take $\boldsymbol{V}^{2k} = \Delta^k(\lambda) \oplus \Delta^k(-\lambda)$ to be the Jordan decomposition for the skew symmetric operator a. Since a leaves the Jordan cell $\Delta^k(\lambda)$ invariant, and its matrix with respect to a suitable basis has the form of a Jordan box, according to Proposition 2.34 there is only one extension of a as a skew symmetric operator on $\boldsymbol{V}^{2k}$. In the basis constructed during the proof of this proposition it has the matrix

$$(a^i_j) = A^k(\lambda) := \begin{pmatrix} \Delta^k(\lambda) & 0 \\ 0 & -\Delta^k(\lambda)' \end{pmatrix}, \ \lambda \neq 0, \tag{23}$$

where now $\Delta^k(\lambda)$ also denotes the Jordan box (2.58) (cf. (21)). Now, by its very definition this basis is symplectic, and hence:

Proposition 20. *Two skew symmetric, splitting endomorphisms of a symplectic vector space whose determinants are different from zero are symplectically congruent if and only if, up to the order of the summands, they have the same Jordan decomposition; similarly, any two symmetric bilinear forms the associated endomorphisms of which have the stated property are symplectically congruent.* □

To prove this, apply Lemma 2.22 to the symplectic transformations between the summands defined as above using symplectic bases, which combine to form a symplectic transformation of the whole space. The matrix for the corresponding symmetric bilinear form on a single summand is obtained from (2.49):

$$B^k(\lambda) = \begin{pmatrix} 0 & \Delta^k(\lambda) \\ \Delta^k(\lambda)' & 0 \end{pmatrix}, \ \lambda \neq 0. \tag{24}$$

Formulas (23) and (24), applied to each direct summand of the decomposition (22) (with $\boldsymbol{U}_0 = \{\mathfrak{o}\}$) are canonical representations for the corresponding objects; in fact, they determine a complete system of invariants. In particular, as is well-known, each endomorphism of a complex vector space is splitting, so that by Proposition 20 the classification of the non-degenerate symmetric bilinear forms is complete in the case $K = \mathbf{C}$. In this case we even have more generally

Proposition 21. *Two skew symmetric or symmetric endomorphisms, or two symplectic transformations of a complex symplectic vector space are symplectically congruent if and only if, up to the order of their summands, they have the same Jordan decomposition.* □

For the proof we refer to A. I. Maltzev [75], §VII.1, Section 100. In the next example we discuss the classification of skew symmetric endomorphisms in a four-dimensional complex symplectic vector space.

Example 8. Let $a \in \mathrm{End}(\boldsymbol{V}^4)$ be a skew symmetric endomorphism of a complex symplectic vector space. Then there are only the possibilities specified below:

A. *The endomorphism a is non-degenerate.*

1. *a is diagonalizable.* Then there is a symplectic basis in which the matrix of a has the form $A^1(\lambda) \oplus A^1(\mu)$, $\lambda, \mu \neq 0$; in a suitable labelling of the basis elements this amounts to

$$(a^i_j) = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 \\ 0 & 0 & -\lambda & 0 \\ 0 & 0 & 0 & -\mu \end{pmatrix}, \quad \lambda, \mu \neq 0.$$

The matrix of the corresponding symmetric bilinear form is

$$(b_{ij}) = \begin{pmatrix} 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \mu \\ \lambda & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 \end{pmatrix}, \quad \lambda, \mu \neq 0.$$

2. *a is not diagonalizable.* Then there is a symplectic basis in which the matrix of a has the form (23) with $k = 2$:

$$(a^i_j) = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 1 & \lambda & 0 & 0 \\ 0 & 0 & -\lambda & -1 \\ 0 & 0 & 0 & -\lambda \end{pmatrix}, \quad \lambda \neq 0.$$

The matrix of the corresponding symmetric bilinear form is

$$(b_{ij}) = \begin{pmatrix} 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \\ \lambda & 0 & 0 & 0 \\ 1 & \lambda & 0 & 0 \end{pmatrix}, \quad \lambda \neq 0.$$

Hence according to Corollary (19) and Proposition 20 we described normal forms for all the non-degenerate skew symmetric endomorphisms in dimension 4.

B. *Suppose that the endomorphism a is degenerate.*

1. *a is diagonalizable.* In this case, normal forms are obtained by specializing A.1; for $\lambda = \mu = 0$ we have the zero matrix, and for $\lambda \neq 0, \mu = 0$ it is a skew symmetric endomorphism of rank 2.

2. *a is not diagonalizable and has a non-zero eigenvalue λ.* Then

$$(a^i_j) = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\lambda & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \lambda \neq 0,$$

is a normal form for endomorphisms of this kind; the Jordan box for the eigenvalue zero has order 2. Interchanging the second and third rows as well as the corresponding columns in the matrix above one obtains the Jordan normal form of a. This example shows that the Jordan cells for the eigenvalue zero are not necessarily isotropic and do not always occur in pairs. The matrix of the corresponding symmetric bilinear form is

$$(b_{ij}) = \begin{pmatrix} 0 & 0 & \lambda & 0 \\ 0 & -1 & 0 & 0 \\ \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \ \lambda \neq 0.$$

These matrices have rank three.

3. *a is not diagonalizable and does not have any non-zero eigenvalue; it has precisely one Jordan cell of order 2.* The normal form is then constructed by setting $\lambda = 0$ in Case B.2 just described.

4. *a is not diagonalizable and does not have any non-zero eigenvalue; it decomposes into two Jordan cells of order 2.*

$$(a^i_j) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

is a normal form for endomorphisms of this kind. The corresponding bilinear form is diagonal of rank 2, its defect subspace is isotropic:

$$(b_{ij}) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

5. *a is not diagonalizable and has a Jordan cell of order* 4 *with the eigenvalue zero.* The Jordan normal form of the endomorphism defined by the matrix below in a symplectic basis has the stated form:

$$(a^i_j) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

The corresponding symmetric bilinear form has the matrix

$$(b_{ij}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The following proposition (cf. A. I. Maltzev [75], §VII.3, Section 106) implies that also for the eigenvalue zero Jordan cells of dimension 3 do not occur: □

Proposition 22. *Let a be a skew symmetric, nilpotent endomorphism of the complex symplectic vector space $\boldsymbol{V}^{2n}$. Then each of its Jordan cells with odd dimension occurs with even multiplicity. Jordan cells of even dimension $k \leq 2n$, however, may occur with arbitrary multiplicity.*

Proof. Let m be the maximal dimension of a Jordan cell occurring in the Jordan decomposition of a. If $m = 1$, then a is the zero endomorphism, and the assertion is trivial. In all other cases a^{m-1} is different from the zero endomorphism; we will distinguish two possibilities:
Case A. There is a vector $\mathfrak{b} \in \boldsymbol{V}$ satisfying $a^{m-1}(\mathfrak{b}) \neq \mathfrak{o}$ as well as $\langle a^{m-1}(\mathfrak{b}), \mathfrak{b}\rangle = 0$, and *Case B*: This does not hold. If, in particular, m is even, then the endomorphism a^{m-1} is symmetric, and hence the corresponding bilinear form $b_{m-1}(\mathfrak{x}, \mathfrak{y}) := \langle a^{m-1}(\mathfrak{x}), \mathfrak{y}\rangle$ is skew symmetric of the same positive rank as a^{m-1}; so we are in Case A. In I. M. Jaglom [56] there is an ingenious proof for the following lemma:

Lemma 23. *Under the assumptions of Proposition 22 suppose, moreover, that a Jordan cell Δ_1^m of a with maximal dimension m satisfies the conditions characterizing Case A. Then there exists a further Jordan cell Δ_2^m of the same dimension such that the subspace $\Delta_1^m \oplus \Delta_2^m$ spanned by both is a symplectic subspace of dimension $2m$.* □

More precisely, the proof even shows that the Jordan cells may be chosen to be totally isotropic subspaces; the decomposition $\Delta_1^m \oplus \Delta_2^m$ has a symplectic basis in which the restriction of the endomorphism to this subspace is described by a matrix of the form (23) with $\lambda = 0$ and $k = m$. Since the subspace is symplectic and invariant under a, the same also holds for its orthogonal complement. So the argument applies here as well. After a finite number of steps we either arrive at the result, or Case B has to occur. Then $m = 2l$ is even, and there can only be a single Jordan cell of dimension m. In fact, if there were two of these Jordan cells with generating vectors $\mathfrak{c}_1, \mathfrak{c}_2$ for which by assumption

$$c_i := \langle a^{m-1}(\mathfrak{c}_i), \mathfrak{c}_i\rangle \neq 0, \ i = 1, 2,$$

then over the complex numbers the equation

$$\langle a^{m-1}(\mathfrak{c}_1 x_1 + \mathfrak{c}_2 x_2), \mathfrak{c}_1 x_1 + \mathfrak{c}_2 x_2)\rangle = c_1 x_1^2 + 2b x_1 x_2 + c_2 x_2^2 = 0$$

with

$$b := \langle a^{m-1}(\mathfrak{c}_1), \mathfrak{c}_2\rangle = \langle a^{m-1}(\mathfrak{c}_2), \mathfrak{c}_1\rangle$$

always has solutions $\mathfrak{c} = \mathfrak{c}_1 x_1 + \mathfrak{c}_2 x_2$ different from the zero vector. Moreover, as basis vectors of different Jordan cells the vectors $a^{m-1}(\mathfrak{c}_i)$ are linearly independent. Hence we also have $a^{m-1}(\mathfrak{c}) \neq \mathfrak{o}$. But this means that we are in Case A, which contradicts the assumption. Now one can prove that there are skew symmetric endomorphism in symplectic spaces of dimension m whose

Jordan decomposition consists of a single cell $\Delta^m(0)$, cf. Example 9, Case B.5. By Proposition 21 all these endomorphisms are symplectically congruent. Hence Proposition 22 is proved by further reducing dimensions. □

Example 9. Now we consider the real four-dimensional symplectic vector space $\boldsymbol{V}^4$. Apart from the splitting endomorphisms described in the previous example we still have to classify the non-splitting, skew symmetric endomorphisms. To do so we proceed as discussed in Section 1.10.4 and pass to the complex extension, which will be denoted by $\boldsymbol{V}_c$. In a real basis, i.e. one for $\boldsymbol{V}$, we represent the elements of this extension using complex coordinates, and the endomorphism a is extended $\mathbf{C}$-linearly to $\boldsymbol{V}_c$. The same holds for the scalar product (cf. Example 1.10.8), so that $\boldsymbol{V}_c$ becomes a four-dimensional complex symplectic vector space with the skew symmetric endomorphism a, to which we may apply the classification obtained in Example 8. Since the characteristic polynomial of a has real coefficients, together with each root λ of multiplicity k the complex conjugate number $\bar{\lambda}$ is also a root of the same multiplicity. We start by supposing:
C. *There is an eigenvalue $\lambda \in \mathbf{C}$ which is neither real nor purely imaginary and satisfies $\bar{\lambda} \neq -\lambda$.* Since for any skew symmetric endomorphism with λ the number $-\lambda$ also has to be an eigenvalue, $\lambda, -\lambda, \bar{\lambda}$, and $-\bar{\lambda}$ are four different eigenvalues of a; hence the complex endomorphisma has the normal form A.1 with $\mu = \bar{\lambda}$ with respect to a complex symplectic basis $(\mathfrak{c}_j)$ for $\boldsymbol{V}_c$. We decompose the basis vectors into their real and imaginary part:

$$\mathfrak{c}_j = \mathfrak{u}_j + \mathrm{i}\,\mathfrak{v}_j,\ \mathfrak{u}_j, \mathfrak{v}_j \in \boldsymbol{V},\ j = 1, \ldots, 4. \tag{25}$$

The equation $a(\mathfrak{c}_1) = \mathfrak{c}_1\lambda$ and the reality of a immediately imply $\overline{a(\mathfrak{c}_1)} = a(\overline{\mathfrak{c}_1}) = \overline{\mathfrak{c}_1}\bar{\lambda}$, so that from this, the analogous equation for $\mathfrak{c}_3$ and $-\lambda$, and using $\langle \mathfrak{c}_1, \mathfrak{c}_3 \rangle = 1$ with the definitions

$$\mathfrak{c}_2 := \overline{\mathfrak{c}_1},\ \mathfrak{c}_4 := \overline{\mathfrak{c}_3} \tag{26}$$

we obtain a symplectic basis for $\boldsymbol{V}_c$ in which a has the normal form A.1. Evaluating the scalar products $\langle \mathfrak{c}_i, \mathfrak{c}_j \rangle$ the relations (25) and (26) lead to the following: *The vectors*

$$\mathfrak{e}_1 := \mathfrak{u}_1\sqrt{2}, \mathfrak{e}_2 := \mathfrak{v}_1\sqrt{2}, \mathfrak{e}_3 := \mathfrak{u}_3\sqrt{2}, \mathfrak{e}_4 := -\mathfrak{v}_3\sqrt{2} \tag{27}$$

form a real symplectic basis $\mathfrak{e}_j \in \boldsymbol{V}$. Setting $\lambda = \alpha + \mathrm{i}\,\beta, \mu = \alpha - \mathrm{i}\,\beta$ we decompose the normal form A.1 of a into its real and imaginary parts and obtain: *In the symplectic basis* (27) *for* $\boldsymbol{V}$ *the matrix of a has the form*

$$(a^i_j) = \begin{pmatrix} \alpha & \beta & 0 & 0 \\ -\beta & \alpha & 0 & 0 \\ 0 & 0 & -\alpha & \beta \\ 0 & 0 & -\beta & -\alpha \end{pmatrix},\ \text{where } \lambda = \alpha + \mathrm{i}\,\beta, \alpha \in \mathbf{R}^*, \beta > \mathbf{0}, \tag{28}$$

is a complex eigenvalue of a. *The matrix of the associated bilinear form is*

$$(b_{ij}) = \begin{pmatrix} 0 & 0 & \alpha & -\beta \\ 0 & 0 & \beta & \alpha \\ \alpha & \beta & 0 & 0 \\ -\beta & \alpha & 0 & 0 \end{pmatrix}.$$

Thus we arrived at a real normal form; this implies that two real skew symmetric endomorphisms of type C with the same complex eigenvalues are symplectically congruent over the real numbers.

D. *There is a complex eigenvalue of the form* $\lambda = \mathrm{i}\,\alpha$ *with* $\alpha \in \mathbf{R}^*$. In this case we have four possibilities: a may be diagonalizable and non-degenerate, or it can have one or none of these properties. Each of these cases may also split into different subcases. Suppose first:
D.1. *Case* D *and a is diagonalizable.* Since with any λ the number $\bar{\lambda} = -\mathrm{i}\,\alpha$ is also an eigenvalue, we may suppose $\alpha > 0$. For the complex linear extension of a we find a complex eigenvector $\mathfrak{c} \in \boldsymbol{V}_c$ for the eigenvalue $\mathrm{i}\,\alpha$, and because of $\bar{\lambda} = -\lambda$ the conjugate vector $\bar{\mathfrak{c}}$ is an eigenvector for the eigenvalue $-\lambda$. Together with $\mathfrak{c}$ it spans a symplectic subspace. Hence we have

$$\mathfrak{c} = \mathfrak{u} + \mathrm{i}\,\mathfrak{v},\ \langle \mathfrak{c}, \bar{\mathfrak{c}} \rangle = -2\,\mathrm{i} \langle \mathfrak{u}, \mathfrak{v} \rangle \neq 0,\ \mathfrak{u}, \mathfrak{v} \in \boldsymbol{V}. \tag{29}$$

For any eigenvalue $\mathfrak{c}$ each complex multiple

$$\mathfrak{c}_1 = \mathfrak{c}\rho\,\mathrm{e}^{\mathrm{i}\,\varphi} = (\mathfrak{u}\cos\varphi - \mathfrak{v}\sin\varphi + \mathrm{i}(\mathfrak{u}\sin\varphi + \mathfrak{v}\cos\varphi))\rho$$

is also an eigenvector. Since

$$\langle \mathfrak{u}_1, \mathfrak{v}_1 \rangle = \langle \mathfrak{u}, \mathfrak{v} \rangle \rho^2,$$

we can always arrange $\langle \mathfrak{u}, \mathfrak{v} \rangle^2 = 1$. Furthermore we note that the sign of $\langle \mathfrak{u}, \mathfrak{v} \rangle$ does not depend on the choice of the eigenvector, i.e., it is an invariant for the endomorphism a. Decomposing

$$a(\mathfrak{u} + \mathrm{i}\,\mathfrak{v}) = (\mathfrak{u} + \mathrm{i}\,\mathfrak{v})\,\mathrm{i}\,\alpha$$

into real and imaginary parts we obtain

$$a(\mathfrak{u}) = -\mathfrak{v}\alpha,\ a(\mathfrak{v}) = \mathfrak{u}\alpha.$$

If now $\langle \mathfrak{u}, \mathfrak{v} \rangle = 1$, then we set $\mathfrak{e}_1 := \mathfrak{u}, \mathfrak{e}_3 := \mathfrak{v}$; in the case $\langle \mathfrak{u}, \mathfrak{v} \rangle = -1$ we correspondingly take $\mathfrak{e}_1 := \mathfrak{v}, \mathfrak{e}_3 := \mathfrak{u}$. In each case we thus obtain a symplectic subspace $\boldsymbol{U} \subset \boldsymbol{V}$ with a symplectic basis which is invariant under a, and with respect to which the restriction $a|\boldsymbol{U}$ is, due to the chosen sign, represented by a matrix of the following normal form:

$$\begin{pmatrix} 0 & \alpha \\ -\alpha & 0 \end{pmatrix} \text{ or } \begin{pmatrix} 0 & -\alpha \\ \alpha & 0 \end{pmatrix},\ \alpha > 0. \tag{30}$$

Since with $\boldsymbol{U}$ the perpendicular subspace $\boldsymbol{U}^{\perp}$ is also symplectic and invariant under a, we can represent the resulting endomorphism in this space in a canonical way taking into account the actual eigenvalues of $a|\boldsymbol{U}^{\perp}$; hence we have to distinguish the following cases:

D.1.1. *Case* D.1 *with a having further purely imaginary eigenvalues* $\mu = \pm\,\mathrm{i}\,\beta$. Then there is a real symplectic basis for $\boldsymbol{V}$ with respect to which a has the normal form

$$(a^i_j) = \begin{pmatrix} 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \beta \\ -\alpha & 0 & 0 & 0 \\ 0 & -\beta & 0 & 0 \end{pmatrix} \quad \text{with } \alpha, \beta \in \mathbf{R}^*.$$

We do not fix the signs of α, β and thus combine the four resulting possibilities into a single one. The matrix of the associated bilinear form is diagonal:

$$(b_{ij}) = \begin{pmatrix} \alpha & 0 & 0 & 0 \\ 0 & \beta & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \beta \end{pmatrix}.$$

>From this the relevance of the signs of α, β becomes obvious; they decide upon the index of the real symmetric bilinear form, which in this case may be equal to zero, two, or four.

D.1.2. *Case* D.1 *with a having the real eigenvalues* $\mu = \pm\beta$. Then there is a real symplectic basis for $\boldsymbol{V}$ with respect to which a has the normal form

$$(a^i_j) = \begin{pmatrix} 0 & 0 & \alpha & 0 \\ 0 & \beta & 0 & 0 \\ -\alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & -\beta \end{pmatrix} \quad \text{with } \alpha \in \mathbf{R}^*, \beta \geq \mathbf{0}.$$

The matrix of the associated symmetric bilinear form is

$$(b_{ij}) = \begin{pmatrix} \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta \\ 0 & 0 & \alpha & 0 \\ 0 & \beta & 0 & 0 \end{pmatrix}.$$

For $\beta = 0$ we obtain a degenerate bilinear form of rank 2, and a non-degenerate bilinear form of index one or three, otherwise. The remaining case is:

D.2. *Case* D *and a is not diagonalizable.* Then again there are two possibilities:

D.2.1. *Case* D.2 *with the complex extension of a containing a Jordan cell* $\Delta^2(\mathrm{i}\,\alpha)$. Then this extension is non-degenerate of Type A.2 from Example 8. If now a is a real skew symmetric endomorphism that, with respect to a symplectic basis, is represented by a matrix of the form

$$(a^i_j) = \begin{pmatrix} 0 & \alpha & 1 & 0 \\ -\alpha & 0 & 0 & 1 \\ 0 & 0 & 0 & \alpha \\ 0 & 0 & -\alpha & 0 \end{pmatrix} \quad \text{with } \alpha \in \mathbf{R}^*,$$

then its complex extension has the stated form. The matrix of the associated real symmetric bilinear form is

$$(b_{ij}) = \begin{pmatrix} 0 & 0 & 0 & -\alpha \\ 0 & 0 & \alpha & 0 \\ 0 & \alpha & 1 & 0 \\ -\alpha & 0 & 0 & 1 \end{pmatrix}.$$

D.2.2. *Case* D.2 *with the complex extension of* a *containing a Jordan cell* $\Delta^2(0)$. Then this extension is degenerate of Type B.2 from Example 8. If now a is a real skew symmetric endomorphism which, with respect to a symplectic basis, is represented by a matrix of the form

$$(a^i_j) = \begin{pmatrix} 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 \\ -\alpha & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \text{with } \alpha \in \mathbf{R}^*,$$

then its complex extension has the stated form. The matrix of the associated real symmetric bilinear form is even diagonal:

$$(b_{ij}) = \begin{pmatrix} \alpha & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

□

Consider now the non-degenerate skew symmetric endomorphism a of the real symplectic vector space $\boldsymbol{V}^4$. Then Proposition 20 together with the possibilities discussed in Example 9 imply that the endomorphisms A.1, A.2 from Example 8 with real eigenvalues $\lambda, \mu \in \mathbf{R}^*$ together with the following non-degenerate endomorphisms described in Example 10, C, D1.1, D1.2 with $\beta > 0$ and D2.1, form a complete system of representatives for all classes of these endomorphisms. For the degenerate endomorphisms the index of the associated bilinear form provides an additional invariant that also has to be taken into account, as the following example shows.

Example 10. We describe two degenerate skew symmetric endomorphisms, which have the same Jordan normal form and are not symplectically congruent over the real numbers. Let these endomorphisms be defined in the symplectic standard basis by the following matrices:

$$a = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

The Jordan normal forms of these endomorphisms each contain two Jordan cells of dimension two with the eigenvalue zero. Considering them over the complex numbers they are symplectically congruent according to Proposition 21; a symplectic transformation c with $b = c \circ a \circ c^{-1}$ is obtained by the equally denoted matrix

$$c = \begin{pmatrix} 1/2 & 0 & 0 & -1 \\ -\mathrm{i}/2 & 0 & 0 & -\mathrm{i} \\ 0 & 1/2 & 1 & 0 \\ 0 & -\mathrm{i}/2 & \mathrm{i} & 0 \end{pmatrix}.$$

In the case $K = \mathbf{R}$ of real numbers it is easy to show that the symmetric bilinear forms associated with a and b have the indices one and two, respectively; hence they, and by Lemma 18 also the endomorphisms themselves, cannot be symplectically congruent over the reals. □

By the last three examples we hope to have provided an insight into the topic, for the details we refer to the general study of I. M. Jaglom [56]. In the Mathematica notebook symplectic.nb on the homepage of R. Sulanke the reader can find a conceptual as well as computational introduction into elementary symplectic geometry with all the invariants and examples described in this section.

The symplectic classification of projective quadrics in real and complex geometry is an immediate consequence of the classification of symmetric bilinear forms. Multiplying the bilinear form by a non-zero constant the eigenvalues are multiplied by the same constant, and the structure of the Jordan decomposition does not change just like the quadric defined by $b(\mathfrak{x}, \mathfrak{x}) = 0$. If at least one of the eigenvalues is different from zero, then, over the complex numbers, one may always arrange one of the eigenvalues to be one, and, over the reals, that one of the eigenvalues has the absolute value one; if appropriate, one may still have to multiply by -1. This leads to a refinement of the projective geometry of quadrics. In the real case, in particular, it would be nice to have an interpretation for this refinement in the form of a geometric analysis. We hope to be able to include some more results concerning this question into the notebook mentioned above in the future.

2.9 Transformation Groups: Results and Problems

This section contains a brief overview describing some problems and results related to the *Erlanger Programm* by Felix Klein [65] and the theory of transformation groups developed at almost the same time by Sophus Lie. Right from the beginning, this theory continuously interacted with the development of geometry. Obviously, the Erlanger Programm forms only a general scheme stimulating problems, but itself providing no means to solve them. So it essentially influenced the development of contemporary geometry based on the theory of Lie groups with its algebraic and analytic methods.

Let us begin by defining a Lie group. Here we will refer to basic notions from the fields of topology and differentiable manifolds as they are collected in Appendix 4. Corresponding references to the literature can be found there.

A group G that is also a topological space and for which the group operations

$$\begin{aligned} (g,h) \in G \times G &\mapsto gh \in G, \\ g \in G &\mapsto g^{-1} \in G \end{aligned} \tag{1}$$

are continuous is called a *topological group.* If G is an analytic manifold and the operations (1) are analytic maps then G is called a *Lie group.* A deep result from the theory of topological groups states that on each such group, which, in addition, is a topological manifold, there exists a unique structure of an analytic manifold for which it is a Lie group.

Simplest examples of Lie groups are the additive groups $[\boldsymbol{V}^n, +]$ of finite-dimensional vector spaces over the field $\mathbf{R}$ of real numbers. Here, the analytic structure is already determined by a single chart,

$$\varphi : \mathfrak{x} \in \boldsymbol{V}^n \longmapsto (x^i) \in \mathbf{R}^n,$$

where (x^i) are the vector coordinates of $\mathfrak{x} = \mathfrak{a}_i x^i \in \boldsymbol{V}$ with respect to an arbitrary basis $(\mathfrak{a}_i)$. Another important example is the *linear group* $\boldsymbol{GL}(\boldsymbol{V}^n)$ of the vector space $\boldsymbol{V}^n$, which, obviously, is an open submanifold of the n^2-dimensional real vector space $\boldsymbol{L}(\boldsymbol{V}^n)$ of all linear endomorphisms in $\boldsymbol{V}^n$. Again, there is an atlas consisting of a single chart assigning to each automorphism $a \in \boldsymbol{GL}(\boldsymbol{V}^n)$ its matrix with respect to an arbitrary fixed basis for the space $\boldsymbol{V}^n$. If G is a Lie group, and $H \subset G$ is a subgroup as well as a submanifold, then H again is a Lie group with the structure of an analytic manifold canonically induced on it (see Appendix A.4.2); these subgroups are called *Lie subgroups.* Lie subgroups of $\boldsymbol{GL}(\boldsymbol{V}^n)$ are called *linear Lie groups.* For each Lie group G its Lie subgroups can be characterized by a purely topological criterion: *A subgroup $H \subset G$ is a Lie subgroup if and only if H is a closed set in G.* This criterion immediately implies that the so-called *classical groups* considered in this book (cf. Section 2.1) are linear Lie groups.

The *connected component of the unit element* $e \in G$ will be denoted by $G_\circ$; it is a Lie subgroup and also a normal subgroup of G, whose cosets are the connected components of G. The group $G_\circ$ is also called the *identity component* of G.

A *homomorphism of Lie groups* is by definition a group homomorphism that, at the same time, is an analytic map. A homomorphism of connected Lie groups $p : G \to H$ is called a *covering homomorphism* if p is a covering in the topological sense. This is equivalent to p being surjective and the kernel of $\operatorname{Ker} p$ being a *discrete* (i.e. zero-dimensional) *Lie subgroup* of G. In this case $\operatorname{Ker} p$ always belongs to the center $Z(G)$. If G is a connected Lie group, and $p : \tilde{G} \to G$ is a covering for which $\tilde{G}$ is simply connected, then there is the unique structure of a Lie group on $\tilde{G}$ such that p is a covering homomorphism (cf. C. Chevalley [32]).

In the theory of Lie groups the fact, already discovered by its founder Sophus Lie, that to each of these groups there corresponds a *Lie algebra* (cf. II.8.4) plays a fundamental part. To every Lie group G the tangent space of the manifold G at the unit element $e \in G$ is assigned as its Lie algebra $L(G) := T_e(G)$, in which a new operation called *bracket* is defined by means of the group multiplication

$$(X, Y) \in L(G) \times L(G) \longmapsto [X, Y] \in L(G).$$

It is bilinear, alternating, and satisfies the *Jacobi identity*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \text{ for all } X, Y, Z \in L(G).$$

Moreover, $L(G) = L(G_\circ)$: The Lie algebra of G only depends on the identity component of G. To each Lie subgroup $H \subset G$ there corresponds the Lie subalgebra $L(H) \subset L(G)$. If, in particular, $G = \boldsymbol{GL}(\boldsymbol{V}^n)$ is the full linear Lie group, then, as a vector space, $L(G)$ is the algebra of linear endomorphisms $\boldsymbol{L}(\boldsymbol{V}^n)$ of $\boldsymbol{V}^n$, and the bracket in $L(G)$ is the commutator determined by algebra multiplication:

$$L(\boldsymbol{GL}(\boldsymbol{V}^n)) = \boldsymbol{L}(\boldsymbol{V}^n) \text{ with } [X, Y] = X \circ Y - Y \circ X,\ X, Y \in \boldsymbol{L}(\boldsymbol{V}^n).$$

Each linear Lie group $G \subset \boldsymbol{GL}(\boldsymbol{V}^n)$ has as its Lie algebra a subalgebra $L(G) \subset \boldsymbol{L}(\boldsymbol{V}^n)$. If the subgroup $H \subset G$ of the Lie group G is connected, then it is uniquely determined by its Lie algebra $L(H) \subset L(G)$. For every homomorphism of Lie groups $f : G \to H$ its differential at the point e is canonically assigned as a homomorphism $df_e : L(H) \to L(G)$ of Lie algebras; this assignment is a covariant functor from the category of Lie groups to the category of real Lie algebras. In particular, isomorphisms of Lie groups correspond to isomorphisms of Lie algebras. If G and H are Lie groups, and if, in addition, G is simply connected, then for each homomorphism $\varphi : L(G) \to L(H)$ of their Lie algebras there always exists a homomorphism $f : G \to H$ between the Lie groups such that $\varphi = df_e$. This implies that two Lie groups G, H with isomorphic Lie algebras $L(G) \cong L(H)$ *are locally isomorphic,* i.e., there are neighborhoods of the unit elements $e \in U \subset G$, $e \in V \subset H$, as well as an analytic diffeomorphism $f : U \to V$ such that $f(xy) = f(x)f(y)$ for $x, y, xy \in U$. For each finite-dimensional real Lie algebra L there is a non-empty family of connected Lie groups G for which $L(G) \cong L$; within this family there is one simply connected Lie group, uniquely determined up to isomorphy, which is a Lie covering group for all the others.

The theory of Lie groups forms the subject of numerous textbooks and monographs. The following references represent a quite subjective selection: C. Chevalley [32], L. S. Pontryagin [88], S. Helgason [49], M. M. Postnikov [89], A. L. Onischik, E. B. Vinberg [85]. The books by H. Freudenthal, H. De Vries [40], and W. Rossmann [96] start by explaining the case most relevant for the readers of this text: linear Lie groups. The rather elementary treatment of differentiable manifolds in R. Sulanke, P. Wintgen [103] also includes an introduction into the theory of Lie groups and differentiable transformation groups.

A *complex Lie group* is understood to be a group which, at the same time, is a complex analytic manifold, and for which the group operations (1) are holomorphic maps. Examples of complex Lie groups are the additive group of a finite-dimensional vector space $\boldsymbol{V}^n$ over $\mathbf{C}$, the complex linear group $\boldsymbol{GL}(\boldsymbol{V}^n)$ corresponding to them, and all the Lie subgroups of $\boldsymbol{GL}(\boldsymbol{V}^n)$ that are complex analytic submanifolds. In particular, these include the classical groups $\boldsymbol{SL}(n, \mathbf{C})$, $\boldsymbol{Sp}(n, \mathbf{C})$, $\boldsymbol{O}(n, \mathbf{C})$, $\boldsymbol{SO}(n, \mathbf{C})$, as well as the projective groups corresponding to them. As described above in the case of real Lie groups there are analogous assignments of complex Lie algebras to the complex Lie groups, subgroups, and corresponding homomorphisms.

The modern notion of a Lie group as an analytic manifold with a group structure used today is not the approach Sophus Lie followed laying the foundations for the theory. Instead, he studies local analytic transformations between open sets of the space $\mathbf{C}^n$ which are sufficiently close to the unit element, i.e. the identical transformation. The sets of these transformations did not form a group in the algebraic sense; the operations were only defined for transformations sufficiently close to the unit element. It was not before the end of the nineteenth century that the abstract group notion took shape, and, based on it, transformation groups could be described as the action of groups on sets. Already then, however, fundamental results concerning the structure and the classification of Lie groups were obtained mainly relying on the relation between Lie groups and Lie algebras discovered by Sophus Lie. This way, W. Killing was able to classify the simple complex Lie groups already in the years 1888–1890; a connected Lie group is called simple if it is not Abelian and contains no connected normal Lie subgroups apart from the trivial one $\{e\}$ and the whole group. They correspond to the simple Lie algebras.

In the last century the structure of real and complex Lie groups as well as algebras was extensively investigated. An outline of the structure theory that arose can be found in V. V. Gorbatsevitch, A. L. Onishchik, E. B. Vinberg [86], the books S. Helgason [49] and A. L. Onishchik, E. B. Vinberg [84] contain detailed presentations. Now we want to briefly quote a few notions and results from this theory. It turned out that two quite different classes of Lie groups were to be distinguished: the *solvable* and the *semi-simple Lie groups*. By definition, a Lie group is *semi-simple* if it does not contain any non-trivial, normal, Abelian Lie subgroups. One can prove that a Lie group is semi-simple if and only if it is locally isomorphic to the direct product of simple Lie groups. Each connected Lie group G contains a unique *radical*, i.e. its largest connected, normal, solvable Lie subgroup R; G is semi-simple if and only if $R = \{e\}$. According to a result of E. Levi, each connected Lie group G can be represented in the form $G = SR$, where S is an arbitrary maximal connected, semi-simple Lie subgroup of G, and, moreover, $\dim S \cap R = 0$; such a subgroup S is called a *Levi subgroup*. A. I. Maltzev proved that all these Levi subgroups are conjugate in G. This leads to the task of classifying the simple as well as the solvable connected Lie groups. The real simple Lie groups were determined by E. Cartan, whereas the classification of solvable Lie groups is not completed yet.

Between the real and the complex Lie groups there is a natural relation, which is most easily described on the level of Lie algebras. To each real Lie algebra $\mathfrak{g}$ there corresponds its *complexification*

$$\mathfrak{g}(\mathbf{C}) := \mathfrak{g} \otimes_{\mathbf{R}} \mathbf{C},$$

which is a complex Lie algebra of the same dimension. Conversely, every complex Lie algebra $\mathfrak{g}$ can be considered as a real Lie algebra $\mathfrak{g}_{\mathbf{R}}$ of twice the dimension (restriction of the scalar domain, see Section 1.10). A real Lie subalgebra $\mathfrak{g}_0 \subset \mathfrak{g}_{\mathbf{R}}$ is called a *real form* of the complex Lie algebra $\mathfrak{g}$ if $\mathfrak{g}_0$ is a minimal real subspace spanning $\mathfrak{g}$ over $\mathbf{C}$ (cf. Definition 1.10.1). This terminology is also used for Lie groups: If G is a connected, real Lie subgroup of the connected, complex Lie group H, and if its Lie algebra $L(G) \subset L(H)$ is a real form of the Lie algebra $L(H)$, then G is called a *real form of the complex Lie group* H, and $H = G(\mathbf{C})$ a *complexification* of the Lie group G. Under this operation real solvable or semi-simple Lie groups are transformed into complex Lie groups of the same type, respectively, and the real forms of simple complex Lie groups are always simple Lie groups.

Example 1. The series of classical groups described in Section 2.1 contain many examples of simple Lie groups. Actually, as the classification results show, apart from finitely many exceptions, and up to local isomorphisms, all simple Lie groups occur among them. The groups $\boldsymbol{SL}(n, \mathbf{C})$ $(n \geq 2)$, $\boldsymbol{SO}(n, \mathbf{C})$ $(n \geq 3, n \neq 4)$, and $\boldsymbol{Sp}(n, \mathbf{C})$ $(n \geq 1)$ are connected, complex, simple Lie groups; except for these infinite series there are only five more (up to local isomorphisms) connected, complex, simple Lie groups, which are denoted by $\boldsymbol{G_2}, \boldsymbol{F_4}, \boldsymbol{E_6}, \boldsymbol{E_7}, \boldsymbol{E_8}$ and called complex simple *exceptional groups*. The connected, real, simple Lie groups are (again up to local isomorphisms) the complex simple Lie groups viewed as real ones, and the real forms of the complex simple Lie groups. In Section 2.2 we described the following connected real simple Lie groups:

$$G = \boldsymbol{U}(k, n-k),\ \boldsymbol{SL}(n, \mathbf{R});\quad G(\mathbf{C}) = \boldsymbol{SL}(n, \mathbf{C})\quad (n \geq 2);$$
$$G = \boldsymbol{SU}^*(2n) = \boldsymbol{SL}(n, \mathbf{H});\quad G(\mathbf{C}) = \boldsymbol{SL}(2n, \mathbf{C})\quad (n \geq 1);$$

$$G = \boldsymbol{SO}(l, n-l); \quad G(\mathbf{C}) = \boldsymbol{SO}(n, \mathbf{C}) \quad (n \geq 3,\ n \neq 4);$$
$$G = \boldsymbol{SO}^*(2n); \quad G(\mathbf{C}) = \boldsymbol{SO}(2n, \mathbf{C}) \quad (n \geq 3);$$
$$G = \boldsymbol{Sp}(l, n-l),\ \boldsymbol{Sp}(\boldsymbol{m}, \mathbf{R}); \quad G(\mathbf{C}) = \boldsymbol{Sp}(n, \mathbf{C}) \quad (n \geq 1).$$

Actually, the groups G above form a complete list of the real forms (up to conjugacy) of the corresponding complex simple Lie groups $G(\mathbf{C})$; in addition, there still are the real forms of the complex exceptional Lie groups, which are not considered in this book. They are called *real exceptional Lie groups.* □

In the classification of semi-simple Lie groups and in many applications of Lie groups in geometry and analysis as well, the compact Lie groups play an important part. In fact, each complex semi-simple Lie group G contains a compact real form K that is even unique up to conjugacy; moreover, G is simple if and only if K is simple. The compact real forms of the groups $\boldsymbol{SL}(n, \mathbf{C}), \boldsymbol{SO}(n, \mathbf{C}), \boldsymbol{Sp}(n, \mathbf{C})$ are their respective subgroups $\boldsymbol{SU}(n), \boldsymbol{SO}(n), \boldsymbol{Sp}(n)$. Let us also note that the radical of an arbitrary connected, compact Lie group K coincides with the identity component of its center $Z(K)$; the Levi subgroup is also uniquely determined — it is the commutator group $K_0 = (K, K)$. Moreover, we have

Proposition 1. (E. Cartan). *A maximal compact subgroup K of an arbitrary Lie group G is a Lie subgroup of G. If G has only finitely many connected components, then all these subgroups are conjugate in G, and the group G is analytically diffeomorphic to the direct product of the group K with a Euclidean space. The maximal compact subgroups of complex, connected, semi-simple Lie groups are their compact real forms.*

A further important result in the theory of Lie groups is the following

Proposition 2. (I. D. Ado). *Every finite-dimensional, real or complex Lie algebra is isomorphic to a subalgebra of the Lie algebra $\boldsymbol{L}(\boldsymbol{V})$ of endomorphisms in a finite-dimensional, real or complex vector space. Each real or complex Lie group is locally isomorphic to a real or complex linear Lie group, respectively.*

Let us now turn to the theory of Lie transformation groups, which, as already said, stood at the very beginning of the study of Lie groups. An overview presenting basic notions and results from the theory of transformation groups can be found in V. V. Gorbatsevich, A. L. Onishchik [45]. The subject of greatest relevance for geometry in this theory is the study of transitive actions on analytic manifolds. A transformation group $[G, M, \Phi]$ (cf. Appendix 4.3) is a *differentiable* or *analytic Lie transformation group* if G is a Lie group, M is a differentiable or analytic manifold, and Φ is a differentiable or analytic map, respectively. The classes of differentiable and analytic Lie transformation groups, respectively, each form a category with the *equivariant maps* as morphisms, i.e. the pairs of maps

$$(F, f) : [G, M_1] \longrightarrow [H, M_2],$$

where $F : G \to H$ is a homomorphism of Lie groups, and $f : M_1 \to M_2$ is a differentiable or analytic map, respectively, such that

$$f(gx) = F(g)f(x) \text{ for all } g \in G,\ x \in M_1.$$

If $G = H$ and $f = \mathrm{id}_G$, then the equivariant maps with this property essentially are the G-maps $f : M_1 \to M_2$ considered in Appendix 4.3, which we should now suppose to be differentiable or analytic, respectively; they form subcategories within the category of equivariant maps.

If $H \subset G$ is a Lie subgroup of the Lie group G, then the set G/H of cosets can be equipped with the unique structure of an analytic manifold such that the canonical map

$$p : g \in G \longmapsto p(g) := gH \in G/H \tag{2}$$

is open as well as continuous, and, moreover, the action of G on G/H defined in Appendix 4.3 is analytic. The proposition concerning the group models for homogeneous spaces formulated there can, with some minor changes, be transferred to Lie transformation groups. In fact, if $[G, M]$ is a Lie transformation group, then the isotropy group G_x is a Lie subgroup. Hence the analytic transformation group $[G, G/G_x]$ is defined. Moreover, the G-map

$$f : gG_x \in G/G_x \longmapsto f(gG_x) := gx \in M$$

is differentiable or analytic, respectively. If $[G, M]$ is transitive, and if the group G has an at most countable number of connected components (this condition is always fulfilled if G satisfies the second axiom of countability), then f is a diffeomorphism (or an analytic diffeomorphism), i.e. an isomorphism in the corresponding category of Lie transformation groups. If, in addition, the manifold M is connected, then the identity component $G_\circ \subset G$ acts transitively on M as well.

If G does not act transitively on M, then, by means of f, the structure of an analytic manifold that G/G_x carries may be transferred to the orbit Gx of an arbitrary point $x \in M$. In general, however, the orbit Gx is not a submanifold of M. For instance, the orbits of a differentiable action of the Lie group $[\mathbf{R}, +]$ (in other words, the trajectories of a dynamical system) may form dense subsets of M.

For an arbitrary Lie subgroup $H \subset G$ the canonical map (2) is the projection of an analytic fibration of the group G with basis G/H, whose fibres are the cosets $gH \subset G$. If $[G, M]$ is a transitive Lie transformation group, and $H = G_x$ is the isotropy group of a point $x \in M$, then, using the map f described above, the manifold M can be identified with G/G_x. This yields a fibration of G with basis M and fibres gH. This fibration allows to establish relations between the topological properties of the three manifolds G, H, and

M. If, in particular, M and H are connected, then the group G is also connected. If M is simply connected, and G is connected, then H is connected as well. For instance, using the facts that the sphere S^n is connected and satisfies $S^n \cong \boldsymbol{SO}(n+1)/\boldsymbol{SO}(n)$, cf. Exercise 5.1, one can prove by induction that the group $\boldsymbol{SO}(n)$ is connected and hence $\boldsymbol{O}(n)_\circ = \boldsymbol{SO}(n)$. Analogously one can prove that the Lie groups $\boldsymbol{U}(n), \boldsymbol{SU}(n), \boldsymbol{Sp}(n)$ are connected using their transitive actions on spheres (see Example 4 below). The connectivity properties of the non-compact classical groups can be reduced to those of their maximal compact subgroups by means of Proposition 1.

If H is a normal Lie subgroup of the group G, then the quotient group G/H is again a Lie group with respect to the structure of an analytic manifold mentioned above, and the canonical map p (cf. (2)) is an analytic homomorphism. Let us also note that there is a quite similar theory of complex Lie transformation groups, i.e. of holomorphic actions of complex Lie groups on complex manifolds.

The group models G/H are a powerful tool; they enable to reduce the study of transitive actions to the investigation of pairs (G, H) formed by a Lie group G and a Lie subgroup $H \subset G$. So, e.g., the kernel of the action of G on G/H coincides with the largest normal subgroup of G contained in H. Two transitive actions $G/H_1, G/H_2$ are G-isomorphic if and only if the subgroups H_1, H_2 are conjugate in G, i.e. if they can be transformed into one another by an inner automorphism of G; they are equivariantly isomorphic, if there is an automorphism of G mapping H_1 onto H_2. This way, the classification of the transitive actions for a specific Lie group with respect to G-isomorphisms is reduced to the determination of the conjugacy classes of its Lie subgroups, whereas its classification with respect to equivariant isomorphisms amounts to classifying these subgroups with respect to arbitrary automorphisms of G.

The classification of all the subgroups in a Lie group is a difficult and, in general, not yet finally completed task. The largest number of results are proved for semi-simple Lie groups G. Considering only connected groups and subgroups this problem can be reduced to the classification of the Lie subalgebras in the Lie algebra $L(G)$ corresponding to G. A summary of the results concerning the classification of connected Lie subgroups and subalgebras in semi-simple Lie groups or algebras, respectively, can be found in V. V. Gorbatsevich, A. L. Onishchik, E. V. Vinberg [86]. The general description of non-connected Lie subgroups is a difficult task; for discrete subgroups, i.e. Lie subgroups of dimension zero, E. B. Vinberg, V. V. Gorbatsevich, O. V. Shvartsman [107] summarize the results.

Now we want to describe a remarkable class of Lie subgroups. First note that each automorphism $s : G \to G$ determines the subgroup of its fixed elements

$$G^s := \{g \in G | s(g) = g\}.$$

A Lie subgroup $H \subset G$ is called *symmetric* if there exists an involutive automorphism $\sigma : G \to G$ such that $H \subset G^\sigma$ and $H_\circ = (G^\sigma)_\circ$. A homogeneous space M is called a *Riemannian symmetric space* if its isotropy group $H = G_b$ at an arbitrary point $b \in M$ is symmetric and compact. Every Riemannian symmetric space $M \cong G/H$ is equipped with a G-invariant Riemannian metric; for each of its points b

$$s_b(gb) := \sigma(g)b,\ g \in G, \tag{3}$$

defines an isometric diffeomorphism $s_b : M \to M$, called the *symmetry at the point* b. The symmetry s_b leaves the point b fixed; in suitable coordinates (x^j) defined on a neighborhood of B the symmetry s_b is described by the formula

$$s_b : (x^1, \ldots, x^n) \longmapsto (-x^1, \ldots, -x^n), \tag{4}$$

where the point b has coordinates $(0, \ldots, 0)$. It is easy to verify that

$$\sigma(g)x = s_b(gs_b(x)) \text{ for all } g \in G,\ x \in M. \tag{5}$$

Using this formula one can compute the automorphism σ from its symmetry s_b, if G acts effectively. The Euclidean as well as the spherical, elliptic, and hyperbolic spaces are Riemannian symmetric spaces.

Example 2. The group of motions $G = \mathfrak{E}(n)$ in the n-dimensional Euclidean space $\boldsymbol{E}^n$ acts transitively on $\boldsymbol{E}^n$. For an arbitrary point $b \in \boldsymbol{E}^n$ the isotropy group $H = G_b$ is canonically isomorphic to the group $\boldsymbol{O}(n)$ of orthogonal transformations in the associated Euclidean vector space $\boldsymbol{V}^n$. Fixing in $\boldsymbol{E}^n$ an orthonormal cartesian coordinate system $(x^1, \ldots, x^n)$ with origin b this isotropy group is identified with the group of real orthogonal matrices of order n, cf. Section 2.5.1. Hence $\mathfrak{E}(n)/\boldsymbol{O}(n)$ is the group model for the homogeneous space $[\mathfrak{E}(n), \boldsymbol{E}^n]$. In order to show that this space is Riemannian symmetric, we consider the symmetry of $\boldsymbol{E}^n$ defined by (4), and then form the corresponding inner automorphism according to (5):

$$\sigma(g) = s_b \circ g \circ s_b,\ g \in G = \mathfrak{E}(n). \tag{6}$$

Obviously, $\sigma^2 = \mathrm{id}_G$ and $H \subset G^\sigma$; moreover, since $G = H\mathfrak{T}$, where $\mathfrak{T}$ denotes the group of translations, we conclude $H = G^\sigma$. Hence H is a symmetric subgroup. The subgroup $H = \boldsymbol{O}(n)$ is compact, since it is closed in $\boldsymbol{GL}(n, \mathbf{R})$ and the entries a_{ij} of each orthogonal matrix satisfy $|a_{ij}| \leq 1$. Relation (6) immediately implies (3), so that (4) really is the symmetry at the point b.

Now we turn to hyperbolic geometry. In Section 2.6.1 we saw that the hyperbolic space $\boldsymbol{H}^n$ is a homogeneous space for the projective pseudo-orthogonal group $\boldsymbol{PO}(1, n)$. Representing the transformations by matrices with respect to a pseudo-orthonormal basis $(\mathfrak{e}_0, \ldots, \mathfrak{e}_n)$ for the associated $(n+1)$-dimensional Minkowski space this group is identified with the subgroup

$$G = \boldsymbol{O}(1, n)^+ := \{(a_{ij}) \in \boldsymbol{O}(1, n) | a_{00} > 0\}$$

of the pseudo-orthogonal group $\boldsymbol{O}(1,n)$. Here the isotropy group $H = G_b$ of the point $b = [\mathfrak{e}_0]$ is the group of matrices

$$H = \{\begin{pmatrix} 1 & \mathfrak{o}' \\ \mathfrak{o} & B \end{pmatrix} | B \in \boldsymbol{O}(n)\}, \tag{7}$$

which, obviously, is isomorphic to the orthogonal group $\boldsymbol{O}(n)$. Hence $\boldsymbol{O}(1,n)^+/\boldsymbol{O}(n)$ is a group model for the hyperbolic space $\boldsymbol{H}^n$, cf. Exercise 2.6.2. In analogy to the Euclidean case we first define the symmetry determined by the diagonal matrix

$$s_b = \operatorname{diag}(1, -1, \ldots, -1) \in G, \tag{8}$$

and then determine the involutive automorphism of the group G using (6). It is easy to verify $H = G^\sigma$ and that (8) is the symmetry at the point b.

Now we consider the spherical and the elliptic geometries. As in Section 2.5.1 denote by $S^n(r)$ the n-dimensional sphere of radius $r > 0$ whose center is the point $\mathfrak{o} \in \boldsymbol{V}^{n+1}$. Spherical geometry is then determined by the transitive action of the orthogonal group $\boldsymbol{O}(n+1)$ on $S^n(r)$; the group model is $\boldsymbol{O}(n+1)/\boldsymbol{O}(n)$. More precisely: choosing in $\boldsymbol{V}^{n+1}$ an orthonormal basis $(\mathfrak{e}_0, \ldots, \mathfrak{e}_n)$ the group G is identified with the group of orthogonal matrices of order $n+1$. The isotropy group for the point $b = \mathfrak{e}_0 r$ is again the group H determined by (7). As in hyperbolic geometry, (8) determines the symmetry at the point b and (6) the corresponding involutive automorphism σ for $G = \boldsymbol{O}(n+1)$. Now the subgroup of elements fixed by σ is

$$G^\sigma = \{\begin{pmatrix} \pm 1 & \mathfrak{o}' \\ \mathfrak{o} & B \end{pmatrix} | B \in \boldsymbol{O}(n)\} \subset \boldsymbol{O}(n+1),$$

which is isomorphic to the direct product $\boldsymbol{O}(1) \times \boldsymbol{O}(n)$. Obviously, $H \subset G^\sigma$. Since the identity component of $\boldsymbol{O}(n)$ is the special orthogonal group $\boldsymbol{SO}(n)$, we also have $H_\circ = (G^\sigma)_\circ$, i.e. H is symmetric and obviously compact; thus $S^n(r)$ is a Riemannian symmetric space.

Elliptic geometry is determined by the projective action of the orthogonal group $G = \boldsymbol{O}(n+1)$ on the real projective space $M = \mathbf{R}\boldsymbol{P}^n$, whose points are the one-dimensional subspaces of $\boldsymbol{V}^{n+1}$. Again, the symmetry at the point $b = [\mathfrak{e}_0]$ and the involutive automorphism σ are defined by formulas (8) and (6); the isotropy group G_b for the point b coincides with the group G^σ of fixed elements. The group model is $\mathbf{R}\boldsymbol{P}^n \cong G/G^\sigma$, and the space is Riemannian symmetric. Note that in this group model the action of G on M is not effective; to obtain en effective action one has to divide G by the kernel $\{\pm I_{n+1}\}$ and then define the action of $\hat{G} = G/\{\pm I_{n+1}\}$ via representatives for the cosets. The effective group model for elliptic space is thus $\hat{G}/\hat{G}^\sigma$, where now the isotropy group is the quotient group $\hat{G}^\sigma = G^\sigma/\{\pm I_{n+1}\}$; it is easy to verify the isomorphy $\hat{G}^\sigma \cong \boldsymbol{O}(n)$. □

Despite the various differing specific properties derived in Sections 2.5, 2.6, all the Riemannian symmetric spaces M^n described in Example 2 have much in common. First we note that, in all these cases, the isotropy groups of the effective group models for these spaces are isomorphic to the orthogonal group $\boldsymbol{O}(n)$. Let d denote the distance function for M^n. The *distance spheres*

$$S(\boldsymbol{m}, \rho) := \{\boldsymbol{x} \in M^n | d(\boldsymbol{m}, \boldsymbol{x}) = \rho\} \text{ for } 0 < \rho < d_M, \tag{9}$$

with center $\boldsymbol{m}$ and positive radius ρ, which is taken to be smaller than the diameter d_M of the space, are always orbits for the isotropy group of the center. As such, they themselves are homogeneous spaces; it is easy to see that they inherit a spherical geometry. In fact, all these distance spheres have a group model isomorphic to $\boldsymbol{O}(n)/\boldsymbol{O}(n-1)$. As transformation groups, all these spaces are isomorphic; note, however, that they are not isometric: Their metrics depend on the space M^n as well as the radius ρ. Since all these metrics are $\boldsymbol{O}(n)$-invariants, between any two of these orbits there is always an equivariant isomorphism $f : S_1 \to S_2$ such that

$$d_2(f(\boldsymbol{x}), f(\boldsymbol{y})) = k d_1(\boldsymbol{x}, \boldsymbol{y}), \ k \text{ a constant}, \ \boldsymbol{x}, \boldsymbol{y} \in S_1.$$

Here d_1, d_2 denote invariant metrics for the spherical geometries on the orbits (not the distances in M^n). These properties of the orbits are consequences of facts concerning the spherical geometries on the orbits discussed in Section 2.5 combined with one more that we already mentioned in 2.5, 2.6: All these spaces are *two-point homogeneous,* i.e. any two point pairs $(\boldsymbol{x}_i, \boldsymbol{y}_i) \in M^n \times M^n$, $i = 1, 2$, are mapped into one another by a transformation $g \in G$: $g\boldsymbol{x}_1 = \boldsymbol{x}_2$, $g\boldsymbol{y}_1 = \boldsymbol{y}_2$, if and only if their distances coincide, $d(\boldsymbol{x}_1, \boldsymbol{y}_1) = d(\boldsymbol{x}_2, \boldsymbol{y}_2)$. In the cases considered in Example 2 the transformation group G is, at the same time, the isometry group for the invariant metrics defined on M^n. In S. Helgason [49] the author classifies the two-point homogeneous spaces; apart from the Euclidean spaces these are the so-called *Riemannian symmetric spaces of rank* 1, whose classification can also be found there.

For the hyperbolic space $\boldsymbol{H}^n \cong \boldsymbol{O}(1, n)^+/\boldsymbol{O}(n)$ the isotropy group $\boldsymbol{O}(n)$ is a maximal compact subgroup of $\boldsymbol{O}(1, n)^+$. This follows from the fact that $\boldsymbol{H}^n$ is diffeomorphic to the Euclidean space $\boldsymbol{E}^n$ (see Lemma 6.1). In fact, the maximal compact subgroups of semi-simple Lie groups with a finite number of connected components are always symmetric Lie subgroups.

The theory of Riemannian symmetric spaces was developed by E. Cartan; it is closely related to the classification of real, simple Lie algebras. An extensive presentation of this theory as well as many geometric and analytic problems related to it can be found in the monographs by S. Helgason [49], [50], [51].

If $[G, M]$ is a Lie transformation group, and $A \subset G$ is a Lie subgroup , then restricting the action of G on M to A defines a Lie transformation group $[A, M]$. Moreover, the embedding $\iota : A \to G$ induces an equivariant morphism (ι, id_M) of this transformation group into the original one $[G, M]$, which is called an *extension of the Lie transformation group* $[A, M]$; finally, $[A, M]$ is called the *restriction of the transformation group* $[G, M]$ to the subgroup A.

According to the Erlanger Programm by F. Klein an extension of a transformation group leads to a reduction of the set of invariants and the geometric properties, whereas, conversely, restriction enlarges both these sets.

Particularly interesting is the case that the restriction of a transitive transformation group $[G, M]$ to a subgroup A is again transitive. This situation is called an *inclusion of transitive transformation groups*. The subgroup A gives such an inclusion if and only if $G = AH$, where $H = G_b \subset G$ is the isotropy group of a point $b \in M$ (cf. Exercise I.3.2.14). Then the isotropy group of the restricted action is $A_b = A \cap G_b$. Using the group models the description of the transitive restrictions for a transitive transformation group $[G, G/H]$ is thus equivalent to representing the group $G = AH$ as the product of subgroups A, H; obviously, for Lie transformation groups only Lie subgroups are to be considered. Since $G = AH$ if and only if $G = HA$ (cf. Exercise I.3.2.14), the restriction of the transformation group $[G, G/H]$ to the subgroup A is transitive if and only if the restriction of $[G, G/A]$ to the subgroup H is transitive.

The projective and Cayley-Klein geometries discussed in this book are based on transitive actions of linear groups. Among them projective geometry is the most general one: It is based on the full linear group. Since, according to the Theorem of Ado, all connected Lie groups are locally isomorphic to subgroups of the full linear group, it is quite natural that the geometries to be studied following the scheme provided by the Erlanger Programm arise by restricting the projective transformation group, where, however, transitivity is not necessarily preserved. As in hyperbolic or Möbius geometry, the action of the subgroup on its orbits has to be investigated in this case.

Example 3. Consider the natural action of the linear group $\boldsymbol{GL}(n+1, K)$ on the vector space K^{n+1}, where K is an arbitrary skew field. This action induces an action of the same group on the projective space $K\boldsymbol{P}^n$, which is transitive but not effective (cf.(1.4.22), ff.); then the kernel of this action is the center $Z(n+1)$ of the linear group, which is described by (1.2.7). The projective group is defined as the quotient group

$$\boldsymbol{PL}_n(K) := \boldsymbol{GL}(n+1, K)/Z(n+1),$$

which acts transitively and effectively on the projective space $K\boldsymbol{P}^n$, cf. Definition 1.4.3 ff. In the cases $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$ the space $K\boldsymbol{P}^n$ is an analytic manifold (cf. Appendix 4.2, Example 3), and $\boldsymbol{GL}(n+1, K)$ is a Lie group. Representing the transformations from this group by matrices with respect to inhomogeneous coordinates on $K\boldsymbol{P}^n$ we conclude from Formula (1.2.19) that in these cases $[\boldsymbol{GL}(n+1, K), K\boldsymbol{P}^n]$ is an analytic Lie transformation group. The quotient group $\boldsymbol{PL}_n(K)$, equipped with the canonical structure of a Lie group, also acts analytically on $K\boldsymbol{P}^n$. In Section 2.1.1 we proved for $K = \mathbf{C}, \mathbf{H}$ in general, and for $K = \mathbf{R}$, if n is even, that the special linear groups $\boldsymbol{SL}(n+1, K)$ induce the transformation group $\boldsymbol{PL}_n(K)$ on $K\boldsymbol{P}^n$, thus they also act transitively. For odd n the group $\boldsymbol{SL}(n+1, \mathbf{R})$ also acts transitively on $\mathbf{R}\boldsymbol{P}^n$; in fact, for any $n \geq 1$ each matrix $A \in \boldsymbol{SL}(n+1, \mathbf{R})$ can be represented as a product of the form $A = A_1 \operatorname{diag}(d, 1, \ldots, 1)$ with $d := \det(A)$ and $\det A_1 = 1$. Denoting by H the isotropy group of $[\mathfrak{e}_0]$ under the action of $\boldsymbol{GL}(n+1, \mathbf{R})$ we conclude $\boldsymbol{GL}(n+1, \mathbf{R}) = \boldsymbol{SL}(n+1, \mathbf{R})H$. Applying what we said before implies the transitivity of $\boldsymbol{SL}(n+1, \mathbf{R})$ on $\mathbf{R}\boldsymbol{P}^n$. For $K = \mathbf{C}$ the three groups $\boldsymbol{GL}(n+1, \mathbf{C})$, $\boldsymbol{SL}(n+1, \mathbf{C})$, and $\boldsymbol{PL}_n(\mathbf{C})$ are complex Lie groups acting holomorphically on the complex manifold $\mathbf{C}\boldsymbol{P}^n$.

In Sections 1.10.2 and 1.10.3 we described a representation of the group $\boldsymbol{GL}(m, \mathbf{C})$ as a subgroup of $\boldsymbol{GL}(2m, \mathbf{R})$ and a representation of the group $\boldsymbol{GL}(m, \mathbf{H})$ as a subgroup of $\boldsymbol{GL}(2m, \mathbf{C})$. These groups as well as their subgroups $\boldsymbol{SL}(m, \mathbf{C})$ and $\boldsymbol{SL}(m, \mathbf{H})$ also act transitively on $\mathbf{R}\boldsymbol{P}^{2m-1}$ or $\mathbf{C}\boldsymbol{P}^{2m-1}$, respectively; the same holds for actions of the subgroups $\boldsymbol{GL}(m, \mathbf{H})$ and $\boldsymbol{SL}(m, \mathbf{H})$ on the space $\mathbf{R}\boldsymbol{P}^{4m-1}$. These actions determine the projective geometries arising by correspondingly restricting the scalar domain. □

In Section 2.2 we introduced further classical Lie subgroups $G \subset \boldsymbol{GL}(n+1, K)$ acting transitively on $K\boldsymbol{P}^n$:

$$G = \begin{cases} \boldsymbol{O}(n+1), \boldsymbol{SO}(n+1), \boldsymbol{Sp}(m, \mathbf{R}) \ (n = 2m-1) \text{ for } K = \mathbf{R}, \\ \boldsymbol{U}(n+1), \boldsymbol{SU}(n+1), \boldsymbol{Sp}(m, \mathbf{C}), \boldsymbol{Sp}(m) \ (n = 2m-1) \text{ for } K = \mathbf{C}, \\ \boldsymbol{Sp}(n+1) \text{ for } K = \mathbf{H}. \end{cases}$$

They define Lie subgroups $\boldsymbol{PO}(n+1), \boldsymbol{PSp}(m, \mathbf{R})$ etc. of the projective groups $\boldsymbol{PL}_n(K)$ acting transitively and effectively on $K\boldsymbol{P}^n$. All these groups are associated with geometries of polarities or null systems. In particular, the group $\boldsymbol{PO}(n+1)$ corresponds to elliptic geometry, whereas the groups $\boldsymbol{PSp}(m, \mathbf{R}), \boldsymbol{PSp}(m, \mathbf{C})$ belong to the real and complex projective symplectic geometries (cf. Section 2.8).

Example 4. Closely related to the actions of the classical groups on the space $\mathbf{R}\boldsymbol{P}^n$ mentioned in Example 3 are the corresponding transitive actions on the spheres. In fact, each of these groups G acts transitively on the set $\mathbf{R}^{n+1} \setminus \{\mathfrak{o}\}$; hence it also acts transitively on the set of *rays*, i.e. oriented one-dimensional subspaces of $\mathbf{R}^{n+1}$, or their positive semi-rays. Each semi-ray intersects the unit sphere $S^n := S^n(1)$ with center $\mathfrak{o}$ in a unique point. Identifying the set of semi-rays with S^n we obtain a transitive (and, as is easy to verify, also analytic) action of G on S^n. The covering $p : S^n \to \mathbf{R}\boldsymbol{P}^n$ defined by (2.5.4) is a G-map, which transfers the action on S^n to an action on $\mathbf{R}\boldsymbol{P}^n$ by projectivities. For the transformation groups in Example 3 in which the transforming group G is the group $\boldsymbol{O}(n+1)$ or one of its subgroups it is not necessary to consider rays instead of semi-rays, since these groups map the sphere S^n into itself. If we confine ourselves to locally effective actions of connected groups, then starting from those in Example 3 and the projective transformation groups mentioned thereafter one arrives at the following list of subgroups $G \subset \boldsymbol{GL}(n+1, \mathbf{R})$ acting transitively on S^n in the way described above:

$\boldsymbol{SL}(n+1,\mathbf{R}), \boldsymbol{SO}(n+1)$ $(n \geq 1)$,
$\boldsymbol{GL}(m,\mathbf{C}), \boldsymbol{U}(m), \boldsymbol{SL}(m,\mathbf{C}), \boldsymbol{SU}(m), \boldsymbol{Sp}(m,\mathbf{R})$ $(n = 2m-1)$,
$\boldsymbol{SL}(m,\mathbf{H}), \boldsymbol{Sp}(m), \boldsymbol{Sp}(m,\mathbf{C})$ $(n = 4m-1)$.

Moreover, in Sections 2.6, 2.7 we considered the transitive action of the pseudo-orthogonal group $\boldsymbol{O}(1, n+1)$ on the quadric $Q \subset \mathbf{R}\boldsymbol{P}^n$, which, in homogeneous coordinates, is defined by equation (2.6.3). As shown there, Q is completely contained in the chart neighborhood for the inhomogeneous coordinates determined by $x^0 \neq 0$; in these it is defined by the familiar equation for a hypersphere. Hence the quadric $Q \subset \mathbf{R}\boldsymbol{P}^n$ is analytically diffeomorphic to the sphere S^n and will be identified with it. The geometry determined by the action of the subgroup $\boldsymbol{O}(1, n+1)^+ \subset \boldsymbol{O}(1, n+1)$ on it is Möbius geometry (see Section 2.7). Passing to the identity component we obtain a transitive and effective action of the connected Lie group

$$\boldsymbol{O}(1, n+1)_\circ = \boldsymbol{O}(1, n+1)^+ \cap \boldsymbol{SO}(1, n+1)$$

on the sphere S^n. Let us note that, quite analogously, one defines transitive and effective actions of the pseudo-unitary groups $\boldsymbol{SU}(1, n+1)$ and $\boldsymbol{Sp}(1, n+1)$ on the spheres $S^{2n-1} \subset \mathbf{C}\boldsymbol{P}^n$ or $S^{4n-1} \subset \mathbf{H}\boldsymbol{P}^n$, respectively. Lastly, in analogy to the projective octonion plane (which is not considered in this book, see instead H. Freudenthal [39] or M. M. Postnikov [89]), one can find a transitive action of a real form $\boldsymbol{FII}$ for the complex simple Lie group of type $\boldsymbol{F}_4$ on the sphere S^{15}. □

Examples 3 and 4 show that on the real, complex, and quaternionic projective spaces as well as on the spheres there exist numerous transitive actions of the classical groups, which give rise to rather different geometries. The obvious question is whether the ones included in the list already are all the transitive actions of Lie groups — up to equivariant isomorphisms — on these manifolds. This leads to the fundamental problem in the theory of transitive Lie transformation groups: For a given analytic manifold M, classify all transitive, analytic actions of Lie groups on M up to equivariant isomorphisms. In the terminology of group models this amounts to finding all pairs (G, H) of Lie groups G and their Lie subgroups H for which M and G/H are analytically diffeomorphic. Tackling this it appears natural to restrict the considerations to effective actions, i.e. to suppose that H does not contain any non-trivial, normal subgroup of the group G, or, at least, to *locally effective* ones; the latter means that the kernel has dimension zero, or, equivalently, that H does not contain any non-trivial, connected, normal subgroup of G. The above task to find all transitive actions on the projective and spherical spaces are special cases of this general problem.

Already the papers by S. Lie contain such a classification of locally analytic actions of complex Lie groups on the spaces $\mathbf{C}$ and $\mathbf{C}^2$, cf. V. V. Gorbatsevich, A. L. Onishchik [45]. Today, this subject is, of course, treated under its global aspect; an overview describing the results obtained so far presents [45]. Here we will introduce some general methods that are applied in these classifications, as well as conclusions drawn using them for spheres and projective spaces.

If M is connected, and if the Lie group G acts transitively on M, then the identity component $G_\circ$ also acts transitively on M. This allows to reduce the classification task to the case that G is connected. If the connected Lie group G acts transitively on the connected manifold M, then the simply connected universal covering $\tilde{G}$ acts transitively on the simply connected covering manifold $\tilde{M}$ of M. This way, the problem is further reduced to the case of simply connected manifolds M.

As mentioned in Appendix 4.2, Example 4, the spheres S^n for $n \geq 2$ as well as the complex and quaternionic projective spaces $\mathbf{C}\boldsymbol{P}^n$ and $\mathbf{H}\boldsymbol{P}^n$ for $n \geq 1$ are simply connected, whereas for $n \geq 2$ we have $\pi_1(\mathbf{R}\boldsymbol{P}^n) \cong \mathbf{Z}_2$. All these manifolds are compact. Furthermore, we will only consider the case that M is connected and compact, and that the fundamental group $\pi_1(M)$ is finite; this always holds, among others, if M is simply connected. Assuming all this we will place particular emphasis on the case that the Lie group G is compact. The importance of this case is illuminated by

Proposition 3. (D. Montgomery). *Let M be compact, and let $\pi_1(M)$ be finite. Suppose that the connected Lie group G acts transitively on M. Then the restrictions of this action to any maximal compact subgroup $K \subset G$ and to the commutator group K_0 of K also act transitively on M.*

This proposition suggests to solve the general classification problem in three steps. First, all the transitive actions of compact Lie groups on M are classified; above all, this step relies on methods from algebraic topology, cf. A. L. Onishchik [81]. The second step requires to determine the transitive actions of a connected, compact Lie group K on M which can be extended to the action of a connected, semi-simple Lie group S containing K as a maximal compact subgroup. This is achieved mainly using methods from the structure theory of real, semi-simple Lie groups and algebras. In the third step, one finally has to decide whether the actions of the connected, semi-simple Lie groups G (the compact as well as the non-compact) found so far can be extended to actions of connected Lie groups G containing S as a Levi subgroup. The methods needed in the last two steps are presented in V. V. Gorbatsevich, A. L. Onishchik [45].

Now we describe the results of the classification for the case of the spheres $M = S^n$ and begin with the transitive actions of compact Lie groups. The following proposition goes back to D. Montgomery, H. Samelson and A. Borel; a proof can also be found in [81].

Proposition 4. *An arbitrary effective, transitive action of a connected, compact Lie group on the sphere $S^n, n \geq 2$, is equivariantly isomorphic to the standard linear action of $\boldsymbol{SO}(n+1)$ or a restriction of this action to one of its following subgroups:*

$$\boldsymbol{SU}(m),\ \boldsymbol{U}(m)\ (n = 2m-1);\ \boldsymbol{Sp}(m),\ \boldsymbol{Sp}(m)\boldsymbol{U}(1),$$
$$\boldsymbol{Sp}(m)\boldsymbol{Sp}(1)(n = 4m-1);\ \boldsymbol{Spin}(9)\ (n = 15);\ \boldsymbol{Spin}(7)\ (n = 7);$$
$$\boldsymbol{G}_2\ (n = 6).$$

Here $\boldsymbol{U}(m) \subset \boldsymbol{SO}(2m)$, $\boldsymbol{Sp}(m) \subset \boldsymbol{U}(2m) \subset \boldsymbol{SO}(4m)$ are the embeddings resulting from the restriction of scalar domains (see Example 3). The subgroup $\boldsymbol{Sp}(1) \subset \boldsymbol{SO}(4m)$ is the group of dilations $\{d_q | q \in \mathbf{H}, |q| = 1\}$; each of these transformations commutes with every element of $\boldsymbol{Sp}(m)$. The group $\boldsymbol{U}(1) = \boldsymbol{Sp}(1) \cap \boldsymbol{U}(2m)$ is its subgroup defined by $\{d_q | q \in \mathbf{C}, |q| = 1\}$. By $\boldsymbol{Spin}(2l+1) \subset \boldsymbol{SO}(2^l)$ one denotes the spinor group, i.e. a simply connected, double covering of the group $\boldsymbol{SO}(2l+1)$, cf. II.8.10, or [81], Example 1.2.25. Finally, $\boldsymbol{G}_2$ is the compact real form of the corresponding complex simple Lie group considered as the automorphism group of Cayley's octonion algebra; lastly, the embedding $\boldsymbol{G}_2 \subset \boldsymbol{SO}(7)$ is obtained by using the representation of this group in the seven-dimensional space of purely imaginary octonions (cf. [39], [89]).

Determining the extensions of these actions to actions of non-compact, semi-simple Lie groups results in the following table:

S	K	M
$\boldsymbol{SO}(1,n+1)_\circ$	$\boldsymbol{SO}(n+1)$	S^n
$\boldsymbol{SU}(1,m)$	$\boldsymbol{S}(\boldsymbol{U}(1)\times\boldsymbol{U}(m))$	S^{2m-1}
$\boldsymbol{Sp}(1,m)$	$\boldsymbol{Sp}(1)\times\boldsymbol{Sp}(m))$	S^{4m-1}
$\boldsymbol{FII}$	$\boldsymbol{Spin}(9)$	S^{15}
$\boldsymbol{SL}(n+1,\mathbf{R})$	$\boldsymbol{SO}(n+1)$	S^n
$\boldsymbol{SL}(m,\mathbf{C})$	$\boldsymbol{SU}(m)$	S^{2m-1}
$\boldsymbol{Sp}(m,\mathbf{R})$	$\boldsymbol{U}(m)$	S^{2m-1}
$\boldsymbol{Sp}(m,\mathbf{C})$	$\boldsymbol{Sp}(m)$	S^{4m-1}
$\boldsymbol{SL}(m,\mathbf{H})$	$\boldsymbol{Sp}(m)$	S^{4m-1}
$\boldsymbol{SL}(m,\mathbf{H})\times\boldsymbol{Sp}(1)$	$\boldsymbol{Sp}(m)\times\boldsymbol{Sp}(1)$	S^{4m-1}
$\widetilde{\boldsymbol{SL}}(3,\mathbf{R})$	$\boldsymbol{SU}(2)$	S^3

Table 2.3. Transitive actions on spheres

Proposition 5. *Table 2.3 lists all connected, non-compact, semi-simple Lie groups S acting effectively and transitively on the sphere $M = S^n$, $n \geq 1$; furthermore, the table also displays the maximal compact subgroups K acting transitively on M as well. Each of the groups $\boldsymbol{SO}(1,n+1)_\circ, \boldsymbol{SU}(1,m)$, $\boldsymbol{Sp}(1,m), \boldsymbol{FII}$ only allows a single action on the respective sphere which is uniquely determined up to S-isomorphy; they are described in Example 4. Up to equivariance the group $\boldsymbol{SL}(n+1,\mathbf{R})$ has a unique action on S^n; it is described in Example 4. This action determines the transitive actions of its subgroups $\boldsymbol{SL}(m,\mathbf{C}), \boldsymbol{Sp}(m,\mathbf{R})$ $(n = 2m-1)$, $\boldsymbol{Sp}(m,\mathbf{C}), \boldsymbol{SL}(m,\mathbf{H})$, $\boldsymbol{SL}(m,\mathbf{H}) \times \boldsymbol{Sp}(1)$, $(n = 4m-1)$. The transitive actions of the groups $\boldsymbol{Sp}(m,\mathbf{R})$, $\boldsymbol{SL}(m,\mathbf{H})$, and $\boldsymbol{SL}(m,\mathbf{H}) \times \boldsymbol{Sp}(1)$ on S^{4m-1} resulting this way are, up to equivariant isomorphisms, the only ones, whereas for the groups $\boldsymbol{SL}(m,\mathbf{C})$, $\boldsymbol{Sp}(m,\mathbf{C})$ there exist one-parameter families of mutually non-isomorphic transitive actions on S^{2m-1} and S^{4m-1}, respectively, including the ones above. Lastly, the group $\widetilde{\boldsymbol{SL}}(3,\mathbf{R})$ (the double covering of $\boldsymbol{SL}(3, \mathbf{R})$) has a unique transitive action on $S^3 = \boldsymbol{SU}(2)$.*

Now we still have to deal with the semi-simple Lie groups S that are Levi subgroups of non-semi-simple, connected Lie groups $G = SR$ acting transitively on the spheres.

Proposition 6. *Let G be a connected Lie group acting effectively and transitively on the sphere S^n, $n \geq 2$, and let $G = SR$ be its Levi decomposition. Then the action of the Levi subgroup S on S^n is also transitive, and the radical R is Abelian. For $R \neq \{e\}$ only the following cases are possible:*

1) $S = \boldsymbol{Sp}(m,\mathbf{R})$, $n = 2m-1$, $R \cong \boldsymbol{U}(1)$,
2) $S = \boldsymbol{SL}(m,\mathbf{H})$, $n = 4m-1$, $R \cong \boldsymbol{U}(1)$,
3) $S = \boldsymbol{SU}(m)$, $n = 2m-1$, $\dim R$ *arbitrarily large,*
4) $S = \boldsymbol{Sp}(m)$, $n = 4m-1$, $\dim R$ *arbitrarily large.*

G	K	M
$\boldsymbol{SL}(n+1,\mathbf{C})$	$\boldsymbol{SU}(n+1)$	$\mathbf{C}\boldsymbol{P}^n$
$\boldsymbol{Sp}(m,\mathbf{C})$	$\boldsymbol{Sp}(m)$	$\mathbf{C}\boldsymbol{P}^{2m-1}$
$\boldsymbol{SL}(m,\mathbf{H})$	$\boldsymbol{Sp}(m)$	$\mathbf{C}\boldsymbol{P}^{2m-1}$
$\boldsymbol{SL}(n+1,\mathbf{H})$	$\boldsymbol{Sp}(n+1)$	$\mathbf{H}\boldsymbol{P}^n$

Table 2.4. Transitive actions on $\mathbf{C}\boldsymbol{P}^n$ and $\mathbf{H}\boldsymbol{P}^n$

Propositions 4–6 show that the classification of the effective, transitive actions of connected Lie groups on spheres of even dimension looks particularly simple. For $m \neq 3$ there are only the natural actions of the groups $\boldsymbol{SO}(2m+1)$, $\boldsymbol{SL}(2m+1,\mathbf{R})$, and $\boldsymbol{O}(1,2m+1)_\circ$ (up to equivariant isomorphisms). The first two of these actions can be transferred to the projective space $\mathbf{R}\boldsymbol{P}^{2m}$ by means of the canonical projection, whereas this is impossible for the third one; in fact, this action corresponds to Möbius geometry, and the Möbius group acts transitively on the set of point pairs $(\boldsymbol{x},\boldsymbol{y})$, $\boldsymbol{x} \neq \boldsymbol{y}$, in the spheres (cf. Corollary 2.7.2). From the point of view of Klein's Erlanger Programm this means that the only transitive geometries on the projective spaces $\mathbf{R}\boldsymbol{P}^{2m}$, $m \neq 3$, are the projective and the elliptic ones; passing to the spheres S^{2m}, $m \neq 3$, the Möbius geometry is added to the projective and the spherical geometries. These classical geometries were extensively discussed in the previous sections of this book. For $m = 3$ there is, in addition, the transitive action of the compact Lie group $\boldsymbol{G}_2$ on S^6, which also corresponds to an interesting geometry: the geometry of an almost-complex structure on S^6 whose automorphism group is the group $\boldsymbol{G}_2$, cf. J. A. Wolf [112]. The classification of transitive actions on spheres and real projective spaces of odd dimension leads to a considerably more involved picture.

The transitive actions of Lie groups were also classified for the complex and the quaternionic projective spaces $\mathbf{C}\boldsymbol{P}^n$ and $\mathbf{H}\boldsymbol{P}^n$, cf. A. L. Onishchik [80]. If a group $G \subset \boldsymbol{GL}(n+1,K)$ acts on the projective space $K\boldsymbol{P}^n$, then the kernel of the action Z_G is the intersection $G \cap Z(n+1)$ of G with the center $Z(n+1)$ of $\boldsymbol{GL}(n+1,K)$. The *projective action* of G is understood to be the canonical, effective action of the quotient group $PG := G/Z_G$ on KP^n. As an example we refer to the formulas (2.1.5–8) and to Exercise 2.1.1. All connected Lie groups acting transitively and effectively on the spaces $\mathbf{C}\boldsymbol{P}^n$ and $\mathbf{H}\boldsymbol{P}^n$ are simple and have trivial center.

Proposition 7. *Each effective, transitive, analytic action of a connected Lie group on one of the manifolds $M = \mathbf{C}\boldsymbol{P}^n, \mathbf{H}\boldsymbol{P}^n$, $n \geq 1$, is equivariantly isomorphic to the projective action of one of the linear groups in Table* 2.4 *described in Example 3. Apart from the Lie groups G the table also displays their maximal compact subgroups K which also act transitively on M. The group $\boldsymbol{PL}_n(\mathbf{C})$ coincides with the group of biholomorphic transformations of the complex manifold $\mathbf{C}\boldsymbol{P}^n$.*

For geometric applications inspired by F. Klein's Erlanger Programm it is interesting to determine the inclusions of transitive Lie transformation groups. The classification results formulated above contain numerous examples of such inclusions. As already mentioned, the description of all inclusions is equivalent to finding all factorizations of Lie groups into the product of two Lie subgroups. A summary of the results obtained so far concerning this topic can be found in [45]. For instance, the complete description of all factorizations of connected, compact Lie groups G into the product of two connected Lie subgroups is known. Essentially, it can be reduced to the case of a simple Lie group G. This case is completely studied in [81]; almost all the factorizations occurring there are related to inclusions of transitive actions on spheres described in Proposition 4. In [81] the extensions of transitive actions of simple, connected, compact Lie groups are described as well.

Basic Notions from Algebra and Topology

In this appendix we want to define several basic notions and fix notations which will be used frequently in this book or are needed to discuss some interesting and essential aspects. Let us start by fixing some notation. After that we will collect a few algebraic notions which, in most cases, the reader will be familiar with; this part is most of all intended to clarify the specific contents the corresponding notion will be supposed to have here. The concluding third part will be concerned with the algebra of transformation groups; there the presentation is rather more detailed, since transformation groups are essential for the geometric considerations and will frequently occur in this text. The fourth and final section will discuss some basic topological notions; they are used at several occasions in the book and are necessary for the presentation of important results concerning geometric transformation groups; without the knowledge of these notions the concluding Section 2.9 devoted to Lie transformation groups will be incomprehensible.

A.1 Notations

In this section we list — without striving for completeness — some general, frequently used notations with brief explanations. Further special symbols occurring in this book are to be found at the beginning of the index.

A.1.1 Number Domains

$\mathbf{N}$ the set of positive integers.
$\mathbf{N_0}$ the set of non-negative integers.
$\mathbf{Z}$ the ring of integers.
$\mathbf{Z}_p$ the field of residue classes of integers mod$p$, p a prime number.
$\mathbf{Q}$ the field of rational numbers.
$\mathbf{R}$ the field of real numbers.
$\mathbf{C}$ the field of complex numbers.
$\mathbf{H}$ the skew field of quaternions.

Recall that the difference between a field and a skew field is just that in a skew field the multiplication is not required to be commutative. Obviously, each field is a skew field as well, but not conversely. If $[K, +, \cdot]$ is a skew field with addition $+$ and multiplication $\cdot$, then the *opposite skew field* $[K_0, +, *]$ is defined to be the skew field with the same additive group $[K_0, +] = [K, +]$, in which, however, the multiplication $*$ is taken to be the opposite of $\cdot$:

$$\alpha * \beta := \beta \cdot \alpha \text{ for } \alpha, \beta \in K.$$

A.1.2 Foundations. Set Theory.

$A \iff B$: A holds if and only if B holds (*logical equivalence*).
$A := B$: The symbol A is *defined* by the expression B .
$A := \{x | H(x)\}$: A is the *set* containing all elements x having property $H(x)$.

$A := \{x_\iota\}_{\iota \in I}$: The *set* with the elements x_ι; here always $x_\iota \neq x_\kappa$ for $\iota \neq \kappa$; analogously for finite or countable sets: $A := \{x_1, x_2, \ldots, x_n\}$, $A := \{x_1, x_2, x_3, \ldots\}$.

$f : A \to B$, more precisely, $f : x \in A \mapsto f(x) \in B$: f is a map from the set A to the set B, assigning the element $f(x) \in B$ to the element $x \in A$.

$\mathrm{id}_M : x \in M \mapsto \mathrm{id}(x) := x \in M$: The *identity map* of the set M.

$A \subset B$: A is a *subset* of B ($A = B$ implies $A \subset B$).

$\emptyset$: The *empty set*, $\emptyset \subset M$ for all sets M.

$\mathcal{P}(\mathcal{M})$: *power set* of the set M; $\mathcal{P}(M) := \{A | A \subset M\}$.

$(X_\iota)_{\iota \in I}$: A *family* of objects $X_\iota \in Obj$, this is another notation for an assignment $\iota \in I \mapsto X_\iota \in Obj$; here Obj may be a set or a class, and $X_\iota = X_\kappa$ for $\iota \neq \kappa$ is permitted; analogously for finite or countable families: *n-tuples* $(x_1, x_2, \ldots, x_n)$, or *sequences* $(x_1, x_2, x_3, \ldots) = (x_k)_{k \in \mathbf{N}}$.

A.2 Linear Algebra

K a skew field.

$Z(K)$: The center of the skew field K.

$K^* := \{\alpha \in K | \alpha \neq 0\}$: The multiplicative group of the skew field K.

$\alpha \in K^* \mapsto \sigma_\alpha : \xi \in K \mapsto \sigma_\alpha(\xi) := \alpha\xi\alpha^{-1} \in K$: The *inner automorphism* of the skew field K corresponding to the element $\alpha \in K^*$.

$\xi, \eta \in K$ are called *conjugate* to one another if there exists $\alpha \in K^*$ such that $\eta = \sigma_\alpha(\xi)$.

$\boldsymbol{U}, \boldsymbol{V}, \ldots$ vector spaces.

$\mathfrak{x}, \mathfrak{y} \in \boldsymbol{V}$ vectors.

$n = \dim \boldsymbol{V}$: The dimension of the vector space $\boldsymbol{V} = \boldsymbol{V}^n$. Apart from a few exceptions all the spaces considered here are finite-dimensional; the dimension index n will often be omitted.

$\mathfrak{L}(M) = [M]$: The linear span of $M \subset \boldsymbol{V}$.

$(\mathfrak{a}_j)_{j=1,\ldots,n}$: A basis for $\boldsymbol{V}^n$.

$\mathfrak{x} = \mathfrak{a}_j x^j := \sum_{j=1}^n \mathfrak{a}_j x^j$: The basis representation of the vectors $\mathfrak{x} \in \boldsymbol{V}$; *the sum convention from tensor calculus*: Summations are to be taken over equally denoted upper and lower indices occurring in an expression and extend as far as the dimension prescribes. In this context, note that the j in x^j is no power but an index!

$\not\Sigma$: Termination of the sum convention.

$\boldsymbol{V}'$: The vector space dual to $\boldsymbol{V}$.

$(a_{i\alpha})$, $i = 1, \ldots, n$, $\alpha = 1, \ldots, m$: A matrix with n rows and m columns.

$(a_{i\alpha})' := (b_{\alpha i})$ with $b_{\alpha i} := a_{i\alpha}$: The *transposed matrix* of $(a_{i\alpha})$.

$\det(A)$: The determinant of the square matrix A with entries from a commutative, associative ring.

A.3 Transformation Groups

In Section I.1.4 we already introduced the first algebraic notions from the theory of transformation groups and applied them in the subsequent remarks of Section I.6.6 to formulate the principles of Klein's Erlanger Programm.

Here we recall some of these definitions and add further comments. Interesting applications and references to the literature are to be found in Section 2.9.

A *transformation group* $[G, M, \Phi]$ is taken to be a triple consisting of a group G, a set M, often called the *space* of the transformation group, and a group homomorphism $\Phi : G \to S(M)$ from G to the group of bijective maps from M onto itself. Frequently the symbol of the homomorphism Φ is omitted, and the transformation group is just denoted by $[G, M]$, where, however, the homomorphism always has to be taken as given and fixed; in geometry the homomorphism is usually clear from the context. It provides the *action* of G on M, i.e. the following map determined by Φ:

$$(g, x) \in G \times M \longmapsto gx := \Phi(g)(x) \in M.$$

Fixing g in this definition the map $x \in M \mapsto gx \in M$ is just the *transformation* $\Phi(g) \in S(M)$. Obviously, we have

$$ex = x \text{ for all } x \in M,$$
$$g(hx) = (gh)x \text{ for all } g, h \in G, x \in M,$$

which implies $g(g^{-1}x) = ex = x$. The action is called *effective* if Φ is injective; equivalently, this may be expressed by requiring that $gx = x$ for all $x \in M$ implies $g = e$, i.e., q is the unit element of the group. In this case G may be identified with the subgroup $\Phi(G) \subset S(M)$ and its elements viewed as transformations $g : M \to M$.

Two elements $x, y \in M$ are called *G-equivalent*, frequently also *congruent*, under the action in question if there is $g \in G$ such that $gx = y$. It is easy to show that this defines an equivalence relation on M. The corresponding equivalence classes are called the *orbits* for the transformation group; the orbit of $x \in M$ is the subset

$$Gx := \{gx | g \in G\}$$

consisting of all images for x under the action of G. The action is called *transitive* if it has only a single orbit, i.e., if for each pair $x, y \in M$ there is at least one $g \in G$ satisfying $gx = y$. A transitive transformation group is also called a *homogeneous space*, frequently even a *Klein space* ; already in F. Klein's [65] *Erlanger Programm* these spaces were particularly emphasized. In geometry, the homogeneous space $[G, M]$ providing the basis for the considerations is typically given, and further transformation groups are then constructed from this.

A subset $A \subset M$ is called *G-invariant* if it is mapped into itself by the transformations, i.e.

$$GA := \{gx | g \in G, x \in A\} = A.$$

Obviously, a subset is G-invariant if and only if it is a union of orbits; the orbits themselves are minimal G-invariant subsets. If A is G-invariant, then the restriction of the action to $G \times A$ defines a transformation group $[G, A]$. In particular, $[G, G_x]$ for each orbit Gx is a transitive transformation group.

Let G be a group, and let $H \subset G$ be a subgroup. Then

$$(h, g) \in H \times G \longmapsto t_h(g) := gh^{-1} \in G$$

defines an effective, but in general not transitive action of H on G (Definition I.1.4.1), whose orbits are the *left cosets of* G *by* H (Definition I.3.1.1)

$$gH := \{gh\}_{h \in H}, \ G/H := \{gH | g \in G\};$$

together they form the *quotient space* or the *quotient set of* G *by* H. The map

$$(g, aH) \in G \times G/H \longmapsto l_g(aH) := gaH \in G/H$$

defines a transitive, but in general not effective action of G on G/H. Each homogeneous space is isomorphic to such a quotient space. To make this more precise we need the notion of G-map. So let $[G, M]$ and $[G, Y]$ be two transformation groups with the same group G. A map $f : M \to Y$ is called a *G-map* if for all $g \in G$ and $x \in M$ the relation $f(gx) = gf(x)$ holds. The class of transformations groups with the same group G and the G-maps as morphisms form a category. The map inverse to a bijective G-map again is a G-map. The bijective G-maps are the isomorphisms of transformation groups with the same group G. For two arbitrary transformation groups $[G, M]$ and $[H, Y]$ the *equivariant maps* are defined as the pairs (F, f),

$$F \in \mathrm{Hom}(G, H), \ f : M \to Y \text{ satisfying } f(gx) = F(g)f(x)$$

for all $g \in G, x \in M$. The class of transformation groups is a category with the equivariant maps as its morphisms. The *equivariant isomorphisms* are the equivariant morphisms (F, f) for which F is a group isomorphism, and f is bijective; they are also called *similarities*. For fixed G, the G-maps $f : M \to Y$ form a subcategory by taking $F = \mathrm{id}_G$.

Let $[G, M]$ be a transformation group defined by the homomorphism $F : G \to S(M)$, and let $x \in M$ be an arbitrary point. The *isotropy group* or *stationary subgroup* of x is defined to be the set

$$G_x := \{h \in G | hx = x\}.$$

One can prove that the map

$$f : aG_x \in G/G_x \longmapsto f(a) := ax \in Gx$$

is well-defined, i.e., it does not depend on the element $a \in aG_x$ chosen in the left coset. The map is bijective and satisfies

$$f \circ l_g = F(g) \circ f \text{ for all } g \in G,$$

it is thus a G-isomorphism from the quotient space G/G_x onto the orbit Gx. The base point x chosen for the orbit Gx is not essential; it is easy to see that the isotropy groups of x and gx are conjugate subgroups:

$$G_{gx} = gG_xg^{-1}.$$

We will call any representation $G/H = f^{-1}(M)$ of the space M in a transitive transformation group $[G, M]$ as a quotient space using the G-isomorphism f just defined a *group model* for the homogeneous space M. By means of group models questions concerning the structure, geometric properties, and the classification of homogeneous spaces are reduced to algebraic and topological properties of the group G and its subgroups.

According to F. Klein's Erlanger Programm the G-maps reflect the geometrically relevant properties of the transformation group. In the various special geometries discussed in this book the reader finds numerous examples for G-maps. Particularly interesting are the invariants of the transformation group. Any non-empty set L can always be considered as a transformation group $[G, L]$ with the *trivial action* defined by $\Phi(g) = \mathrm{id}_L$ for all $g \in G$. In other words, $ga = a$ for all $g \in G, a \in L$. The corresponding G-maps $f : M \to L$ are called *invariants*, or, more precisely, *G-invariants* for the transformation group $[G, M]$. Invariants are nothing but the G-maps which are constant on the orbits. So they are of interest only for non-transitive transformation groups, and are used to decide whether two given elements $x, y \in M$ are G-equivalent. A *complete system of invariants* for a transformation group is understood to be a finite set $\{f_1, \dots, f_m\}$ of its invariants with the property that the equalities

$$f_1(x) = f_1(y), \dots, f_m(x) = f_m(y),\ x, y \in M,$$

imply the G-equivalence of x and y, i.e. the existence of $g \in G$ such that $gx = y$. Having found a complete system of invariants for a transformation group $[G, M]$ the classification problem for the elements of M under the action of G is solved: The intersections of the level sets for their invariant $f_\mu, \mu = 1, \dots, m$, are the orbits.

A.4 Topology

This section introduces a few basic notions from the theory of topological spaces and differentiable manifolds occasionally occurring in the text but not necessary for the understanding of the main part. For the formulation of the results concerning the theory of transformation groups summarized in Section 2.9 they are, however, indispensable. For a more detailed study the reader should consult one of the numerous textbooks, e.g. J. L. Kelley [59], W. Rinow [94], G. L. Naber [78], or G. E. Bredon [22].

A.4.1 Topological Spaces

A pair $[M, \mathcal{O}]$ consisting set M and a system $\mathcal{O} \subset \mathcal{P}(M)$ is called a *topological space* with the *system $\mathcal{O}$ of open sets* if the family $\mathcal{O}$ has the following properties:

1. $M, \emptyset \in \mathcal{O}$.
2. $U_1, U_2 \in \mathcal{O}$ implies $U_1 \cap U_2 \in \mathcal{O}$,
3. For each subsystem $\mathcal{B} \subset \mathcal{O}$ of the system of open sets the union is again open:
$$\bigcup_{U \in \mathcal{B}} U \in \mathcal{O}.$$

The system $\mathcal{O}$ of open sets is also called a *topology* on M. Each open set $U \in \mathcal{O}$ containing a given point $x \in M$ is called an *open neighborhood* of x.

Example 1. Metric Spaces A pair $[E, d]$ consisting of a non-empty set E and a real function $d : E \times E \to \mathbf{R}$ is called a *metric space* and d a *metric* or *distance function* on E if d has the following properties:

1. d is *positive definite:* $d(x, y) \geq 0,\ d(x, y) = 0 \iff x = y$,
2. d is *symmetric:* $d(x, y) = d(y, x)$,
3. d satisfies the *triangle inequality*:
$$d(x, y) \leq d(x, z) + d(z, y) \text{ for all } x, y, z \in E.$$

The subset $B(z, r) := \{x \in E | d(z, x) < r\} \subset E$ is called an *open ball* of *radius* $r > 0$ with *center* z in the metric space E. A set $U \subset E$ is called *open* if, for each of its points $z \in U$, there exists $r > 0$ such that the open ball $B(z, r) \subset U$ lies in U. It is easy to verify that every metric space with its associated system of open sets is a topological space. Moreover, using the triangle inequality, it is straightforward to show that this topology has the following property:
$\mathbf{T_2}$. *For any two different points $x, y \in E, x \neq y$, there are open sets $U_1, U_2 \subset E$ such that $x \in U_1, y \in U_2$ and $U_1 \cap U_2 = \emptyset$.*
The points x, y are said to be *separated* by the open sets U_1, U_2, and property T_2 is called the *Hausdorff separation axiom*; a topological space satisfying T_2 is called a *Hausdorff space*. The Euclidean, spherical, elliptic, or hyperbolic spaces are examples of metric, hence also Hausdorff spaces. The finite-dimensional affine point and vector spaces over $\mathbf{R}$ can always be equipped with the structure of a Euclidean space of the same dimension; the topology determined by the associated metric, however, does not depend on the chosen Euclidean structure. they are called the *natural topologies* on these spaces. Since each finite-dimensional point or vector space over $\mathbf{C}$ or $\mathbf{H}$ can also be considered as a finite-dimensional real space (cf. Section 1.10), natural topologies are defined on these spaces as well. □

A further important property of many topological spaces is expressed as the *second axiom of countability: There exists a subsystem $\mathcal{B} \subset \mathcal{O}$ of open sets* *such that each open set can be represented as the union of countably many sets* $B \in \mathcal{B}$. Such a subsystem of open sets is called a *basis* for the open sets of the topological space. The finite-dimensional point or vector space over the scalar domains **R**, **C**, **H** and many of the spaces constructed starting from them satisfy this axiom. E.g., the open cuboids with rational vertices form a basis. For certain applications the second axiom of countability has to be assumed to hold.

A map $f : M_1 \to M_2$ between topological spaces with the systems $\mathcal{O}_1$ and $\mathcal{O}_2$, respectively, of open sets is called *continuous* if the pre-image of each open set is again open, i.e. $f^{-1}(\mathcal{O}_2) \subset \mathcal{O}_1$. The map f is called *open* if the image of each open set is again open, i.e. $f(\mathcal{O}_1) \subset \mathcal{O}_2$. The map f is a *homeomorphism* if f is bijective, continuous, and open, or, in expressed equivalently, if f as well as f^{-1} are continuous.

Next we will describe three procedures to construct new topological spaces starting from given ones.

Let $[M, \mathcal{O}]$ be a topological space, and let $N \subset M$ be a non-empty subset. Define the system $\tilde{\mathcal{O}} := \{U \cap N | U \in \mathcal{O}\}$. Then the pair $[N, \tilde{\mathcal{O}}]$ satisfies the axioms of a topological space. This space is called a *topological subspace;* the topology determined by $\tilde{\mathcal{O}}$ is called the *induced topology* on N. The identical embedding $N \to M$ is a continuous map.

Let $[M, \mathcal{O}]$ be a topological space, and let $\cong$ be an equivalence relation on M. Denote by $\hat{M} = M/\cong$ the set of equivalence classes and by $p : x \in M \mapsto p(x) = \hat{x} \in \hat{M}$ the canonical map assigning its equivalence class $\hat{x} \in \hat{M}$ to each $x \in M$. Then $\hat{\mathcal{O}} := \{W \subset \hat{M} | p^{-1}(W) \in \mathcal{O}\}$ is a topology on the set $\hat{M}$, called the *quotient topology* of the topology $\mathcal{O}$ with respect to the equivalence relation $\cong$. The map p is continuous and open in the topologies $\hat{\mathcal{O}}, \mathcal{O}$.

Let $[M_i, \mathcal{O}_i]$ be two topological spaces. On the product set $M_1 \times M_2$ consider the system $\mathcal{O}$ of all sets U which can be represented as the union of arbitrary sets of the form $U_1 \times U_2$, where $U_1 \in \mathcal{O}_1$ and $U_2 \in \mathcal{O}_2$. Then $\mathcal{O}$ is a topology on the set $M_1 \times M_2$, called the *product topology* of $\mathcal{O}_1, \mathcal{O}_2$. The canonical projections $M_1 \times M_2 \to M_i$, $i = 1, 2$, are continuous and open in the corresponding topologies.

A topological space $[M, \mathcal{O}]$ is called *connected* if for any representation $M = A \cup B$ von M as the union of open sets $A, B \in \mathcal{O}$ with $A \cap B = \emptyset$ either $A = \emptyset$ or $B = \emptyset$. In other words: the space M cannot be represented as the union of two disjoint, non-empty, open sets. Each topological space can uniquely be represented as the union of mutually disjoint, non-empty, connected open sets, called its *connected components*.

A somewhat stronger notion of connectedness is that of a pathwise-connected space. A *path* in a topological space M is understood to be a continuous map $\gamma : [0, 1] \to M$ from the closed unit interval $[0, 1] \subset \mathbf{R}$ into the space M. The space M is called *pathwise-connected* if any two points can be joined by a path, i.e., for two arbitrary points $x, y \in M$ there is a path γ such that $\gamma(0) = x$ and $\gamma(1) = y$. Every pathwise-connected topological space is connected as well; the converse, however, is in general not true. Note, however, that each connected manifold (see below) is also pathwise-connected.

A path $\omega : [0,1] \to M$ is called *closed* or a *loop* if $\omega(0) = \omega(1)$. The *fundamental group* of the topological space is constructed using loops; it is important for many applications of topology. We briefly indicate the definition of this group now and refer to the textbooks on topology for the details, e.g. W. Rinow [94]. Let $x_0 \in M$ be an arbitrarily fixed point in the topological space. By $\Omega_{x_0}(M)$ we denote the set of all loops ω with $\omega(0) = \omega(1) = x_0$. Into this set we introduce an operation $*$ in the following way: $\omega := \omega_1 * \omega_2$ is the loop obtained by first moving along ω_1 and then ω_2,

$$\omega(t) := \omega_1(2t) \text{ for } 0 \leq t \leq 1/2,\ \omega(t) := \omega_2(2t-1) \text{ for } 1/2 \leq t \leq 1.$$

Two loops ω_1, ω_2 *are called homotopic* if they can be continuously deformed into one another, where the deformation in question has to keep the point x_0 fixed. In other words, the deformation is to contain only loops from $\Omega_{x_0}(M)$. This notion of homotopy is an equivalence relation, and the operation defined before can be transferred from loops to homotopy classes: products of homotopic loops are again homotopic. Hence the operation $*$ is defined in the set of homotopy classes; the latter is denoted by $\pi_1(M, x_0)$. One can prove that the operation in $\pi_1(M, x_0)$ just introduced satisfies all the axioms of a group; the resulting group, denoted $[\pi_1(M, x_0), *]$, is called the *fundamental group* of the space M at the point x_0. The unit element in this group is the *trivial loop,* $\omega(t) = x_0$ for $0 \leq t \leq 1$. If the space M is pathwise-connected, then the fundamental groups of all points $x_0 \in M$ are isomorphic; in this case, it is called the fundamental group of M and denoted by $\pi_1(M)$. A pathwise-connected topological space M is called *simply connected* if its fundamental group is trivial, $\pi_1(M) = \{e\}$; this holds if and only if each loop is homotopic to the trivial one. In other words, a topological space is simply connected if and only if it is pathwise-connected and every loop can be contracted to the point x_o.

Let M and N be pathwise-connected topological spaces. A surjective continuous map $p : M \to N$ is called a *covering* if each point $x \in N$ has an open neighborhood U such that the restriction of p to the connected component of x in the pre-image $p^{-1}(U)$ is a homeomorphism from every component onto U; the space M is then called a *covering space* for N. If the pre-image $p^{-1}(x_0)$ of any point $x_0 \in N$ consists of finitely many, say q, elements, then every point $x \in N$ has precisely q pre-images; in this case, the map p is said to be a *q-sheeted covering.* A connected topological space N is called *simply connected* if $q = 1$, i.e., each of its coverings is trivial; this holds if and only if its fundamental group is trivial. One can prove that for every pathwise-connected topological space N there exists a simply connected covering space $\tilde{N}$; if $p : \tilde{N} \to N$ and $p_1 : \tilde{N}_1 \to N$ are two such simply connected coverings, then there is a homeomorphism $\varphi : \tilde{N}_1 \to \tilde{N}$ such that $p_1 = p \circ \varphi$.

A.4.2 Differentiable Manifolds

A topological space M is called an *n-dimensional manifold* if for each point $x \in M$ there exist an open neighborhood $U \in \mathcal{O}$, $x \in U$, and a homeomorphism φ_U from U onto an open ball $B \subset \mathbf{R}^n$ in the space of real n-tuples; the homeomorphisms concerned are called the *charts* of the manifold; they define local coordinates:

$$\varphi_U : x \in U \mapsto (x^i(x)) := \varphi_U(x) \in \mathbf{R}^n.$$

Then for the points in the intersection $U \cap W \neq \emptyset$ of two intersecting chart neighborhood the *coordinate transformations* are defined:

$$\psi_{UW} : (x^i) \in \varphi_U(U \cap W) \mapsto (y^j) := \varphi_W \circ \varphi_U^{-1}((x^i)) \in \varphi_W(U \cap W),$$

these are homeomorphisms between open sets in the space $\mathbf{R}^n$. Each system of charts $\Phi = (\varphi_U)_{U \in \mathcal{B}}$, whose chart neighborhoods cover the manifold, i.e.

$$\bigcup_{U \in \mathcal{B}} U = M,$$

is called an *atlas* for the manifold and the number n the *dimension* of the manifold. If a particular atlas is distinguished on the manifold M all coordinate transformations ψ_{UW} of which are differentiable, then M equipped with this atlas is called a *differentiable manifold.* If, moreover, the functional determinants of all coordinate transformations in an atlas are positive, then the atlas is called *oriented*; the manifold is called *orientable* if there exists an oriented atlas on it. If such an atlas is distinguished, then the manifold is *oriented.* An elementary criterion for orientability can be found in R. Sulanke, P. Wintgen [103]. In order to have a greater variety of charts to choose from one usually allows for all charts whose coordinate transformations with those of the given atlas are again differentiable to be added to this atlas; we will call such charts *compatible with the atlas.*

To allow for certain techniques to be applied and to exclude too strange objects we will suppose that a manifold always satisfies the Hausdorff separation axiom T_2 as well as the second countability axiom.

Differentiable manifolds are the subject studied in differential topology. Textbooks introducing into this field are e.g. J. M. Lee [72] and [73]. Differentiable manifolds provide the right framework to apply analytic methods to solve geometric problems and forms the basis of differential geometry. As a rule, the manifolds defined in this book are even *analytic* in the sense that there is an atlas whose coordinate transformations are real analytic, i.e. locally they can be expanded into convergent power series. Usually this atlas arises quite naturally from the particular situation. For an analytic atlas a compatible chart is to be related to the charts of the atlas by analytic coordinate transformations.

The basic and simplest example of an n-dimensional analytic manifold is the real affine space $\boldsymbol{A}^n$, or, what amounts to the same with respect to the manifold structure, the Euclidean space $\boldsymbol{E}^n$. Defining cartesian coordinates $(x^i) \in \mathbf{R}^n$ with respect to a fixed frame leads to an atlas consisting of a single chart $\varphi : \boldsymbol{x} \in \boldsymbol{A}^n \mapsto (x^i(\boldsymbol{x})) \in \mathbf{R}^n$. The transformations of affine cartesian coordinates are analytic and thus determine compatible charts.

It is well-known that the space of n-tuples $\mathbf{R}^n \cong \mathbf{R}^m \times \mathbf{R}^{n-m}$ is canonically isomorphic to the product spaces on the right-hand side. Using this property it is easy to introduce product charts $\varphi_1 \times \varphi_2 : U_1 \times U_2 \to B_1 \times B_2$ and define an atlas on the product $M_1 \times M_2$ of two differentiable (or analytic) manifolds, this way equipping it with the structure of a differentiable (or analytic, respectively) manifold. Iterating this procedure we obtain arbitrary *products of finitely many differentiable (analytic) manifolds.*

Many of the manifolds occurring in this book are *submanifolds* of an affine space. The definition of this notion to follow also shows how a manifold structure arises on suitable, by no means all, subsets of a differentiable manifold. So let M be an n-dimensional differentiable (analytic) manifold, and let $N \subset M$ be a subset. For $x_0 \in N$ we call a chart $\varphi : U \to B$ for M *special* if $\varphi(U \cap N) = B \cap \mathbf{R}^k$, where $\mathbf{R}^k$ denotes the subspace of $\mathbf{R}^n$ determined by the equations $x^{k+1} = \ldots = x^n = 0$. The subset N is called a *k-dimensional submanifold* if, for each point $x \in N$, there is a special chart $\varphi : U \to B$ with $x \in U$ compatible with the atlas of M. Considering N as a topological subspace the restrictions of the special charts to N become charts on the open sets $U \cap N$, $\varphi : U \cap N \to B \cap \mathbf{R}^k$, together providing an atlas for N. With this atlas N becomes a differentiable (analytic) manifold. For example, each open subset of N is an n-dimensional submanifold. $(n-1)$-dimensional submanifolds of M are usually called *hypersurfaces* (*surfaces* in the case $k = 2$).

Example 2. Let $F : \boldsymbol{A}^n \to \mathbf{R}$ be a differentiable function defined on affine space. Its level sets $N_c := F^{-1}(c)$, $c \in \mathbf{R}$ a constant, are determined by an equation in the coordinate (x^i) of the point $\boldsymbol{x}$:

$$F(\boldsymbol{x}) = F(x^1, \ldots, x^n) = c.$$

If, at every point $\boldsymbol{x} \in N_c$ of a non-empty level set the differential does not vanish, $dF_{\boldsymbol{x}} \neq 0$, i.e. at least one of the partial derivatives of F with respect to the coordinates is non-zero, then the Implicit Function Theorem implies that N_c is a differentiable submanifold; it is even analytic, if F is analytic. For example, as the solution set for the equation

$$d(\boldsymbol{o}, \boldsymbol{x})^2 = (x^1)^2 + \ldots + (x^n)^2 = r^2$$

the hypersphere $S^{n-1}(r)$ of radius $r > 0$ with the origin $\boldsymbol{o}$ as its center is an analytic hypersurface. If $\boldsymbol{a} = (a^1, \ldots, a^n)$ is a point in the hypersphere for which $a^n > 0$, then

$$U := \{\boldsymbol{x} \in S^{n-1}(r) | x^n > 0\},\ \varphi : \boldsymbol{x} \in U \to \varphi(\boldsymbol{x}) := (x^1, \ldots, x^{n-1}) \in \mathbf{R}^{n-1},$$

is a chart for $S^{n-1}(r)$ in a neighborhood of $\boldsymbol{a}$. Similarly one defines charts for the points with $a^n < 0$. Corresponding definitions for the other coordinate axis provide an analytic atlas for the hyperspheres. □

Points in the level set at which the differential is zero are called *singularities;* the other points are the *regular* ones. It is easy to show that planes, ellipsoids, and hyperboloids are regular surfaces in three-dimensional affine space. For the circular cones defined as the solution set for an equation of the form

$$-(x^1)^2 + a^2((x^1)^2 + (x^2)^2) = 0, a \neq 0,$$

the vertex is a singularity; deleting it from the cone we obtain a surface. Example 2 and the considerations above may be generalized to differentiable functions on arbitrary differentiable manifolds.

Let $f : M_1 \to M_2$ be a continuous map between differentiable (analytic) manifolds M_1, M_2 of dimensions $\dim M_1 = n, \dim M_2 = m$. The map is called *differentiable* (*analytisch*) if, for each point $x_0 \in M_1$, there are a chart $\varphi_1 : U_1 \to \mathbf{R}^n$ defined on a neighborhood U_1 of x_0 and another one $\varphi_2 : U_2 \to \mathbf{R}^m$ defined on a neighborhood U_2 of its image point $y_0 = f(x_0)$ such that the *representation of f in the corresponding local coordinates*

$$\varphi_2 \circ f \circ \varphi_1^{-1} : (x^i) \in \varphi_1(U_1) \longmapsto (y^\alpha(x^i)) := \varphi_2 \circ f \circ \varphi_1^{-1}((x^i)) \in \varphi_2(U_2)$$

are real, differentiable (analytic) functions. This obviously means that all the real functions $y^\alpha(x^i)$, $\alpha = 1, \ldots, m$, in the n real variables $x^i, i = 1, \ldots, n$, have this property. The chain rule implies that this property is independent of the chosen charts in the given atlases for the manifolds, i.e. expresses a property of the function f itself. For example, the identical embeddding of an arbitrary submanifold $N \subset M$ is a differentiable (analytic) function. A bijective differentiable map f is called a *diffeomorphism* (*analytic diffeomorphism*) if its inverse f^{-1} is differentiable (analytic) as well.

The next example is intended to show that the projective spaces over the fields $K = \mathbf{R}, \mathbf{C}$ and those over the skew field of quaternions are real analytic manifolds. To see this we use the inhomogeneous coordinates defined in Section 1.2.3 for an arbitrary skew field K as charts.

Example 3. According to Definition 1.1.1 the projective spaces $K\boldsymbol{P}^n = \boldsymbol{P}(K^{n+1})$ are the sets of one-dimensional subspaces of the vector space K^{n+1}. We may exclude the zero vector, which, after all, only leads to the nopoint, and consider $K\boldsymbol{P}^n$ as the set of equivalence classes in $K^{n+1} \setminus \{\mathfrak{o}\}$ for the equivalence relation

$$\mathfrak{x} \equiv \mathfrak{y} :\Longleftrightarrow \mathfrak{y} = \mathfrak{x}\lambda \text{ für } \mathfrak{x}, \mathfrak{y} \in K^{n+1} \setminus \{\mathfrak{o}\},\ \lambda \in K^*;$$

these are the orbits in this set under the action of K^*: $(\mathfrak{x}, \lambda) \mapsto \mathfrak{x}\lambda^{-1}$. As already introduced in Section 1.2 we denote by $\pi : K^{n+1} \setminus \{\mathfrak{o}\} \to K\boldsymbol{P}^n$ the canonical map. In the cases $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$ each of the the vector spaces

$\mathbf{R}^{n+1}$, $\mathbf{C}^{n+1} \cong \mathbf{R}^{2n+2}$, $\mathbf{H}^{n+1} \cong \mathbf{R}^{4(n+1)}$ is equipped with the topology of a Euclidean vector space. It determines the corresponding quotient topology on the associated projective point spaces, for which the canonical map is open and continuous. Then, by Definition (1.2.8), the inhomogeneous coordinates are also defined as charts in the sense of a manifold, since these maps now are homeomorphisms as well. Lemma 1.2.6 and the coordinate transformations (1.2.9) show that this determines an analytic atlas on the projective spaces; note that, in the cases $K = \mathbf{C}, \mathbf{H}$, we have to pass to the realifications. Provided with these, the projective spaces $K\boldsymbol{P}^n$ are real analytic manifolds of dimensions $n, 2n$, and $4n$, respectively, for $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$. In the one-dimensional case the manifold topology coincides with those introduced on the projective lines in Example 1.1.3; hence, considered as real manifolds, the projective lines are analytically diffeomorphic to the circle S^1, the sphere S^2, and to the four-dimensional sphere S^4 for $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$, respectively, cf. also Example 1.2.2. □

Example 4. Consider the unit sphere $S^n \subset \mathbf{R}^{n+1}$ and the real projective space $\mathbf{R}\boldsymbol{P}^n$ as analytic manifolds (cf. Examples 2 and 3). In Example 1.10.1 we defined a two-sheeted covering $\lambda : S^n \to \mathbf{R}\boldsymbol{P}^n$ associating with each vector $\mathfrak{y} \in S^n$ the one-dimensional vector space $[\mathfrak{y}]_{\mathbf{R}} \in \mathbf{R}\boldsymbol{P}^n$ it spans, cf. Formula (1.10.7). The definitions in Examples 2, 3 easily imply that λ indeed is a two-sheeted covering in the topological sense; from the definition of the charts we conclude that λ is even analytic. It is well-known that the spheres S^n are simply connected for $n \geq 2$. In the theory of covering spaces it is then proved that the fundamental groups (see above) are

$$\pi_1(\mathbf{R}\boldsymbol{P}^n) \cong \mathbf{Z}_2 \text{ for } n \geq 2.$$

The fibrations constructed in Example 1.10.1 are classic examples of analytic fibre bundles. In the homotopy theory of fibrations one proves: *If the total space of a fibration is simply connected and the fibres are connected, then the base manifold is also simply connected* (cf. R. M. Switzer [104]). So looking at the Hopf fibration $S^{2n-1} \to \mathbf{C}\boldsymbol{P}^{n-1}$ (cf. (1.10.8)) we see that for $n \geq 1$ the complex projective spaces $\mathbf{C}\boldsymbol{P}^n$ are simply connected. Similarly, using the fibration $\mathbf{C}\boldsymbol{P}^{2n-1} \to \mathbf{H}\boldsymbol{P}^{n-1}$ defined in (1.10.14) one proves that the quaternionic projective spaces $\mathbf{H}\boldsymbol{P}^n$ are simply connected for all $n \geq 1$. □

Concluding this section we still want to briefly discuss the important notion of a *complex analytic* or, for short, *complex manifold*. Replacing in the definition of the charts on a manifold the open sets $B \subset \mathbf{R}^n$ by open sets in the complex space $B \subset \mathbf{C}^n$ determines complex coordinates on the chart neighborhoods. If all transformations of the local coordinates in a given atlas are *holomorphic*, i.e. complex analytic, then the manifold M is called an *n-dimensional complex manifold*. Each n-dimensional complex manifold is, at the same time, a $2n$-dimensional real analytic manifold in the sense defined before; on a $2n$-dimensional real manifold, however, there exists, in general, no

n-dimensional complex atlas. As in the differentiable case holomorphic maps between complex manifolds are defined as those the coordinate transformations of which all are complex analytic functions.

The complex projective spaces $\mathbf{C}\boldsymbol{P}^n$ are examples of complex manifolds; to see this, look at the charts defined in Example 3 without passing to the realifications. In the local coordinates the transformations are fractional linear maps and hence holomorphic. The only sphere on which there exists a complex analytic atlas, is $S^2 = \mathbf{C}\boldsymbol{P}^1$; in complex function theory it is called the *Riemann sphere*. One can prove that each complex manifold is orientable. Hence on the real projective spaces $\mathbf{R}\boldsymbol{P}^{2n}$, $n \geq 1$, there cannot exist any structure of a complex manifold (cf. Example 1.2.3).

References

1. I. Agricola and Th. Friedrich. *Elementargeometrie.* Vieweg, Wiesbaden, 2005.
2. A. D. Alexandrow. *Die innere Geometrie der konvexen Flächen.* Akademie-Verlag, Berlin, 1955. Translation from the Russian. Original: Moskau, Leningrad 1948.
3. F. Apéry. *Models of the Real Projective Plane.* Vieweg, Braunschweig, Wiesbaden, 1987.
4. E. Artin. *Geometric Algebra.* Interscience Publishers, New York, London, 1957.
5. R. Baer. *Linear Algebra and Projective Geometry.* Academic Press, New York, 1952.
6. L. A. Batten and A. Beutelspacher. *The Theory of Finite Linear Spaces.* Cambridge University Press, Cambridge, 1993.
7. Walter Benz. *Geometrische Transformationen unter besonderer Berücksichtigung der Lorentztransformationen.* BI Wissenschaftsverlag, Mannheim, Leipzig, Wien, Zürich, 1992.
8. Walter Benz. *Real Geometries.* BI Wissenschaftsverlag, Mannheim, Leipzig, Wien, Zürich, 1994.
9. Marcel Berger. *Géométrie II.* Nathan, Paris, 1978.
10. A. Beutelspacher. *Einführung in die endliche Geometrie I.* Bibliographisches Institut, Mannheim, Wien, Zürich, 1982.
11. A. Beutelspacher. *Einführung in die endliche Geometrie II.* Bibliographisches Institut, Mannheim, Wien, Zürich, 1983.
12. A. Beutelspacher and U. Rosenbaum. *Projektive Geometrie.* Vieweg, Braunschweig, Wiesbaden, 1992.
13. H.-G. Bigalke. *Kugelgeometrie.* Salle und Sauerländer, Frankfurt a. M., Aarau, 1984.
14. W. Blaschke. *Projektive Geometrie.* Wolfenbütteler Verlagsanstalt K. G., Wolfenbüttel, Hannover, 1948.
15. W. Blaschke and G. Thomsen. *Vorlesungen über Differentialgeometrie III, Differentialgeometrie der Kreise und Kugeln.* Grundlehren der mat. Wiss. Bd. 29. Springer-Verlag, Berlin, 1929.
16. J. Böhm and E. Hertel. *Polyedergeometrie in n-dimensionalen Räumen konstanter Krümmung.* VEB Deutscher Verlag der Wissenschaften, or Birkhäuser, Berlin bzw. Basel, 1980.

17. J. Böhm and H. Reichardt (Ed.). *Gaußsche Flächentheorie, Riemannsche Räume und Minkowski-Welt.* Teubner-Archiv zur Mathematik, Bd. 1. BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1984.
18. J. et al. Böhm. *Geometrie I, II.* Studienbücherei, Mathematik für Lehrer, Bd. 6. VEB Deutscher Verlag der Wissenschaften, Berlin, 1974, 1975.
19. N. Bourbaki. *Livre II, Algèbre.* Hermann, Paris, 1956-1960.
20. W. Boy. Über die Abbildung der projektiven Ebene auf eine im endlichen geschlossene singularitätenfreie Fläche. *Göttinger Nachr.*, pages 20–33, 1901.
21. W. Boy. Über die curvatura integra und die Topologie geschlossener Flächen. *Math. Ann.*, 57:151–184, 1903.
22. Glen E. Bredon. *Topology and Geometry.* Graduate Texts in Mathematics, V. 139. Springer, New York, Heidelberg, Berlin, 1993.
23. I. N. Bronstein and K. A. Semendjajew. *Taschenbuch der Mathematik.* BSB B. G. Teubner, Leipzig, 1985.
24. W. Burau. *Mehrdimensionale projektive und höhere Geometrie.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1961.
25. H. Busemann and P. J. Kelly. *Projective Geometry and Projective Metrics.* Academic Press Inc., Publishers, New York, 1953.
26. Peter J. Cameron. *Notes on Classical Groups.* Author's homepage, http://www.maths.qmw.ac.uk/~pjc/class_gps/, 2000.
27. C. Carathéodory. *Funktionentheorie I.* Verlag Birkhäuser, Basel, 1950.
28. E. Cartan. *La théorie des groupes fini et la géométrie différentielle, traitées par la méthode du repère mobile.* Gauthier-Villars, Paris, 1937.
29. A. Cayley. A sixth memoir on quantics. *Phil. Trans.*, 149:61–90, 1859. See also Coll. Math. Papers vol. II, Notes and References 158.
30. A. Cayley. On the non-euclidian geometry. *Math. Ann.*, 5:630–634, 1872.
31. Thomas E. Cecil. *Lie Sphere Geometry.* Universitext. Springer-Verlag, New York, Berlin, Heidelberg, ..., 1992.
32. C. Chevalley. *Theory of Lie Groups*, Volume I. Princeton University Press, Princeton, 1946.
33. J. L. Coolidge. *The elements of non-Euclidean geometry.* Clarendon Press, Oxford, 1909.
34. H. S. M. Coxeter. *Reelle projektive Geometrie der Ebene.* R. Oldenbourg, München, 1955.
35. J. Dieudonné. Les déterminants sur un corps non commutatif. *Bull. Soc. Math. France*, 71:27–45, 1943.
36. J. Dieudonné. *La Géométrie des Groupes Classiques.* Springer-Verlag, Berlin, Heidelberg, New York, troisième édition, 1971. Russian edition: Mir, Moskau, 1974.
37. Gerd Fischer (ed.). *Mathematische Modelle.* Vieweg, Akademie-Verlag, Braunschweig, Berlin, 1986.
38. A. T. Fomenko. *Symplectic Geometry.* Gordon and Breach Science Publishers, London, 1988.
39. H. Freudenthal. *Oktaven, Ausnahmegruppen, Oktavengeometrie.* Math. Inst. Rijksuniv., Utrecht, 1951.
40. H. Freudenthal and H. De Vries. *Linear Lie Groups.* Academic Press, New York, 1969.
41. F. R. Gantmacher. *Theorie der Matrizen. (Russian).* Gostechisdat, Moskau, 1953. There exist German and English translations, 1986, 1998.

42. F. R. Gantmacher. *Matrizentheorie.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1986. Translation from the Russian.
43. F. R. Gantmacher. *The Theory of Matrices. Vol. 2.* AMS Chelsea Publishing, Providence, R.I., 1998. Translation from the Russian.
44. Oswald Giering. *Vorlesungen über höhere Geometrie, unter Mitwirkung von Johann Hartl.* Vieweg, Braunschweig, Wiesbaden, 1982.
45. V. V. Gorbatsevich and A. L. Onishchik. *Lie Transformation Groups. Lie Groups and Lie Algebras I*, Volume 20 of *Encyclopaedia of Mathematical Sciences.* Springer-Verlag, New York, Heidelberg, Berlin, 1993.
46. Alfred Gray. *Differentialgeometrie. Klassische Theorie in moderner Darstellung.* Spektrum Akademischer Verlag, Heidelberg, Berlin, Oxford, 1994.
47. Larry C. Grove. *Classical groups and geometric algebra.* Graduate Studies in Mathematics. 39. AMS, Providence RI, 2001.
48. L. Heffter. *Grundlagen und analytischer Aufbau der projektiven, euklidischen, nichteuklidischen Geometrie.* B. G. Teubner, Leipzig, 1950.
49. S. Helgason. *Differential Geometry, Lie Groups, and Symmetric Spaces.* Academic Press, New York, San Francisco, London, 1978.
50. S. Helgason. *Groups and Geometric Analysis.* Academic Press, New York, San Francisco, London, 1984.
51. S. Helgason. *Geometric Analysis on Symmetric Spaces.* Mathematical Surveys and Monographs, Vol. 39. American Mathematical Society, Providence, Rhode Island, 1994.
52. D. Hilbert. Über Flächen constanter Gaußscher Krümmung. *Trans. Amer. Math. Soc.*, 2:87–99, 1901.
53. D. Hilbert. *Grundlagen der Geometrie.* B. G. Teubner, Stuttgart, 12th edition, 1968.
54. J. W. P. Hirschfeld. *Projective Geometries over Finite Fields.* Oxford University Press, Oxford, 2nd edition, 1998.
55. H. Hopf. Über die Abbildungen von Sphären auf Sphären niedrigerer Dimension. *Fund. Math.*, 25:427–440, 1935.
56. I. M. Jaglom. Quadratic and skew-symmetric bilinear forms in a real symplectic space. (Russian). *Trudi Sem. Vect.-Tens. Analisu*, 8:364–381, 1950.
57. I. M. Jaglom. On the linear subspaces of a symplectic space. (Russian). *Trudi Sem. Vect.-Tens. Analisu*, 9:309–318, 1952.
58. I. L. Kantor and A. S. Solodownikow. *Hyperkomplexe Zahlen.* Mathematische Schülerbücherei Nr. 95. BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1978. Russian edition: Nauka, Moskau, 1973.
59. J. L. Kelley. *Introduction to Smooth Manifolds.* D. Van Nostrand Company, Inc,, Princeton. N. J., 1957.
60. F. Klein. Ueber die sogenannte nicht-euklidische Geometrie. *Math. Ann.*, 4:573–625, 1871.
61. F. Klein. Ueber die sogenannte nicht-euklidische Geometrie. (Zweiter Aufsatz.). *Math. Ann.*, 6:112–145, 1873.
62. F. Klein. *Gesammelte mathematische Abhnadlungen, Bd. 1-3.* Springer–Verlag, Berlin, Heidelberg, New York, 1968. Reprints of earlier editions.
63. F. Klein. *Vorlesungen über höhere Geometrie.* Springer–Verlag, Berlin, Heidelberg, New York, 1968. Reprint of the 3rd edition 1926, revised by W. Blaschke.
64. F. Klein. *Vorlesungen über nicht-euklidische Geometrie.* Springer–Verlag, Berlin, Heidelberg, New York, 1968. Reprint of the 1st edition 1928, revised by W. Rosemann.

65. F. Klein. *Vergleichende Betrachtungen über neuere geometrische Forschungen,* Habilitationsschrift, Erlangen 1872. (*Das Erlanger Programm,* ed. H. Wussing). Ostwalds Klassiker der exakten Wissenschaften 253. Akademische Verlagsgesellschaft Geest & Portig K.-G., Leipzig, 1974.
66. W. Klingenberg. *Grundlagen der Geometrie, im C. F. Gauß Gedenkband.* B.G.Teubner Verlagsgesellschaft, Leipzig, 1957.
67. W. Klingenberg. *Lineare Algebra und Geometrie.* Springer-Verlag, Berlin etc., 1990.
68. G. Köthe. Schiefkörper unendlichen Ranges über dem Zentrum. *Math. Ann.*, 105:15–39, 1931.
69. Linus Kramer. Buildings and classical groups. *arXiv.math.*, 2003. URL: http://front.math.ucdavis.edu/math.GR/0307117.
70. L. Kronecker. Algebraische reduction der schaaren bilinearer formen. *Sitzungsber. Akademie Berlin,* pages 763–776, 1890.
71. J. Kunze. *Der Schnittort auf konvexen Verheftungsflächen.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1969.
72. John M. Lee. *Introduction to Topological Manifolds.* Graduate Texts in Mathematics, Vol. 202. Springer, New York, Heidelberg, Berlin, 2000.
73. John M. Lee. *Introduction to Smooth Manifolds.* Graduate Texts in Mathematics, Vol. 218. Springer, New York, Heidelberg, Berlin, 2003.
74. K. Leichtweiß. Karl Strubecker zum Gedenken. *Jber. d.dt.Math.-Verein.*, 94:105–117, 1992.
75. A. I. Maltzev. *Grundlagen der linearen Algebra (Russian).* Gosisdat Technisch-Theoretischer Literatur, Moskau, 1956.
76. K. Metsch. *Linear Spaces with Few Lines.* L. N. in Mathematics 1490. Springer-Verlag, Berlin, Heidelberg, 1991.
77. H. Minkowski. Raum und Zeit. *Physikalische Zeitschrift,* 10:104–111, 1909. Ges. Abh. Bd. 2.
78. Gregory L. Naber. *Topology, Geometry and Gauge fields Foundations.* Texts in Applied Mathematics, Vol. 25. Springer, New York, Heidelberg, Berlin, 1997.
79. A. P. Norden. *Einführung in die Lobatschewskische Geometrie.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1958. Translation from the Russian, Moskau 1953.
80. A. L. Onishchik. On lie groups acting transitively on compact manifolds iii. *Mat. USSR Sbornik,* 4:233–240, 1968.
81. A. L. Onishchik. *Topology of Transitive Transformation Groups.* Johann Ambrosius Barth, Leipzig, Berlin, Heidelberg, 1994.
82. A. L. Onishchik and R. Sulanke. *Algebra und Geometrie I.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1986. 2nd edition.
83. A. L. Onishchik and R. Sulanke. *Algebra und Geometrie II.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1988.
84. A. L. Onishchik and E. B. Vinberg. *Lie Groups and Algebraic Groups.* Springer-Verlag, Berlin, Heidelberg, New York, 1990.
85. A. L. Onishchik and E. B. Vinberg. *Foundations of Lie Theory. Lie Groups and Lie Algebras I,* Volume 20 of *Encyclopaedia of Mathematical Sciences.* Springer-Verlag, New York, Heidelberg, Berlin, 1993.
86. Vinberg E. B. Onishchik A. L. and V. V. Gorbatsevich. *Lie Groups and Lie Algebras III. Structure of Lie Groups and Lie Algebras*, Volume 41 of *Encyclopaedia of Mathematical Sciences.* Springer-Verlag, New York, Heidelberg, Berlin, 1994.

87. G. Pickert. *Projektive Ebenen.* Springer-Verlag, Berlin etc., 1955.
88. I. S. Pontryagin. *Topological Groups.* Oxford Graduate Texts in Mathematics. 5. Gordon and Breach Science Publishers, London, 1966.
89. M. M. Postnikov. *Leçons de géométrie. Groupes et algèbres de Lie.* Mir, Moscou, 1985. Russian original, Nauka, Moskau, 1982.
90. P. K. Raschewski. *Riemannsche Geometrie und Tensoranalysis.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1959. German translation; original: Gostechisdat, Moskau 1953; 2. Aufl. 1964.
91. H. Reichardt. *Vorlesungen über Vektor- und Tensorrechnung.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1968.
92. H. Reichardt. *Gauß und die Anfänge der nicht-euklidischen Geometrie. Mit Originalarbeiten von J. Bolyai, N. I. Lobatschewski und F. Klein.* Teubner-Archiv zur Mathematik, Bd. 4. BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1985.
93. W. Rinow. *Die innere Geometrie der metrischen Räume.* Grundlehren d. mat. Wissenschaften. Springer-Verlag, Berlin, Heidelberg, New York, 1961.
94. W. Rinow. *Lehrbuch der Topologie.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1975.
95. B. A. Rosenfeld. *Nichteuklidische Räume.* Nauka, Moskau, 1969. (Russian).
96. W. Rossmann. *Lie Groups. An Introduction through Linear Groups.* Oxford Graduate Texts in Mathematics. 5. Oxford University Press, Oxford, 2002.
97. H. et al. Salzmann. *Compact Projective Planes.* Walter de Gruyter, Berlin, New-York, 1995.
98. E. Sperner. *Einführung in die analytische Geometrie und lineare Algebra I, II.* Vandenhoeck und Ruprecht, Göttingen, 1961.
99. M. Stary. Zum Außen- und Innengebiet von Quadriken. *Journal of Geometry*, 27:87–93, 1986.
100. N. Steenrod. *The Topology of Fibre Bundles.* Princeton University Press, 1951.
101. R. Sulanke. Möbius geometry v. homogeneous surfaces in the Möbius space. *Coll. Math. Soc. J. Bolyai 46. Topics in Differential Geometry. Debrecen 1984*, 46:1141–1154, 1984.
102. R. Sulanke. Möbius invariants for pairs of spheres (s_1^m, s_2^l) in the Möbius space s^n. *Beiträge zur Algebra und Geometrie*, 41:233–246, 2000.
103. R. Sulanke and P. Wintgen. *Differentialgeometrie und Faserbündel.* VEB Deutscher Verlag der Wissenschaften, Birkhäuser, Berlin, Basel, 1972.
104. Robert M. Switzer. *Algebraic topology - homotopy and homology.* Die Grundlehren der mathematischen Wissenschaften. Band 212. Berlin-Heidelberg-New York: Springer-Verlag, 1975.
105. J. Tits. *Tabellen zu den einfachen Lie Gruppen und ihren Darstellungen.* Springer-Verlag, Berlin, Heidelberg, New York, 1967.
106. O. Veblen and J. W. Young. *Projective Geometry, I, II.* xxxxxx, Boston - New York, 1910, 1918.
107. Gorbatsevich V. V. Vinberg E. B. and O. V. Shvartsman. *Discrete Subgroups of Lie Groups. Lie Groups and Lie Algebras II*, Volume 21 of *Encyclopaedia of Mathematical Sciences.* Springer-Verlag, New York, Heidelberg, Berlin, 2000.
108. E. B. Vinberg (Ed.). *Geometry. Vol. 2. Spaces of Constant Curvature*, Volume 29 of *Encyclopaedia of Mathematical Sciences.* Springer-Verlag, Berlin, Heidelberg, New York, 1993.
109. K. Weierstrass. Zur Theorie der bilinearen und quadratischen Formen. *Monatshefte Akademie Berlin*, pages 310–338, 1867.

110. H. Weyl. *The Classical Groups, Their Invariants, and Representations.* Princeton University Press, 1939.
111. J. B. Wilker. Inversive geometry. *The Geometric Vein, The Coxeter Festschrift*, pages 379–442, 1980.
112. J. A. Wolf. The automorphism group of a homogeneous almost complex manifold. *Trans. AMS*, 144:535–543, 1969.
113. Joseph A. Wolf. *Spaces of Constant Curvature.* University of California, Berkeley, 1972.
114. Stephen Wolfram. *Mathematica.* Addison-Wesley, 3. edition, 1997.

Index