PROJECTIVE GEOMETRY

Lecture Notes by Balázs Csikós

1. INTRODUCTION

1.1. The Invention of Perspective Drawing.

Browsing through the history of painting one recognizes that the Renaissance brought a significant change in the description of the three-dimensional world on the canvas.

Traditional icons, describing generally some scenes from the Bible, have golden background, the halos of the holy persons are circular disks, the feet, legs, hands and arms are drawn in a position parallel with the plane of the painting. When a building appears in the picture parallel edges are drawn to be parallel in the picture giving the image a feeling of distortion.

The Renaissance brought a rediscovery of art, literature and ideas of ancient Greece, in particular, it raised in artists the wish to find and describe perfect beauty and harmony. Painters of the Renaissance (Mantegna, Botticelli, Leonardo da Vinci, Dürer, Michelangelo, Raffaello, Tiziano and others) banished from the pictures the unnatural static poses and the empty space behind the figures. Instead of copying the clichés of traditional painting, they started to use living people for models and existing places for background. Therefore they had to find out how to draw a realistic image of a spatial object and had to discover the basic rules of perspective drawing. They observed that parallel straight lines of the space will be represented by lines passing through a point, and that the original ratio of lengths of segments and the angle between segments are not preserved in the drawing (unless the segments are parallel to the canvas!) but are transformed according to rigorous rules. For example, the diameters of a circle will have varying length in the picture and the circle (e.g. the halo) will be represented by an ellipse (or other conic section). With the recognition and conscious application of these rules foundations of a new branch of geometry were laid.

1.2. The Real Projective Space, Points at Infinity.

Let us find the geometric formulation of perspective drawing. Consider a point O in the Euclidean space $\mathbb{E}^3$ that represents the eye of the observer and an arbitrary body B in the field of vision C (which is mathematically a convex cone). Put a transparent plate Σ between O and B. When the observer is looking at B, from every visible point P of the surface of B a ray of light of some color arrives in his eye. This ray crosses the transparent plate at the intersection point P' of the segment OP and the plane Σ. If we paint the plate at P' with the appropriate

Typeset by $\mathcal{AMS}$-TEX

color, point P becomes invisible, but it is not observable, since the ray from P is substituted by a ray from P' that has the same effect in the observer's eye. Doing this construction for all visible points of B we get a faithful image of B on the plate.

The procedure above leads us to the study of the projection Φ of the space onto the plane Σ, that assigns to a point $P \in \mathbb{E}^3$ the intersection point of the straight line OP with the plane Σ provided that $O \neq P$ and OP is not parallel to Σ. The definition of Φ makes sense also for points not in the field of vision (e.g. "behind the observer") but Φ is not defined for every point of the space. Those points for which Φ is not defined lie in the plane Σ_0 passing through O parallel to Σ.

Let us take a straight line e which is not parallel to Σ and does not go through O. Denote by e_0 the straight line passing through O parallel to e and set $Q = e \cap \Sigma_0$ and $R = e_0 \cap \Sigma$. When the point P moves along e the projecting rays OP sweep out the plane Π spanned by O and e and thus the point $P' = \Phi(P)$ moves along the line $e' = \Pi \cap \Sigma$. Φ defines a one to one mapping between the punctured straight lines $e \setminus \{Q\}$ and $e' \setminus \{R\}$. Observe that as the point P moves to infinity (no matter in which direction) P' goes to R, and moving with P towards Q P' goes to infinity. This suggests that we should attach to e and e' an extra point at infinity, E_∞ and E'_∞ respectively. Doing so we can extend Φ to a one to one mapping between the augmented straight lines $e \cup \{E_\infty\}$ and $e' \cup \{E'_\infty\}$ by setting $\Phi(E_\infty) = R$ and $\Phi(Q) = E'_\infty$. We have already met similar augmentation of the real straight line in calculus, when we introduced the symbols $\pm\infty$. The great difference, however, is that while in calculus we put different ideal points to the ends of the straight line now we put the same point to both ends. Walking through this point at infinity we can run off the straight line at one end then disappear for a moment from the Euclidean space and then return from infinity from the other end of the straight line.

Let us attach to every straight line in the Euclidean space a point at infinity and let us try to extend Φ for points of Σ_0 and for the ideal points as described above. One proves easily that the ideal points of two different straight lines will be mapped to the same point if and only if the lines are parallel. This suggests us to assume that the ideal points of two different straight lines are the same if and only if the straight lines are parallel. This way, there is a one to one correspondence between points at infinity and equivalence classes of parallel straight lines, and we may say that an ideal point is an equivalence class of parallel straight lines.

To sum up, we declare the following definitions.

- The *real projective space* $\mathbb{R}P^3$ is the union of the Euclidean space and the set of ideal points.
- A *projective plane* in $\mathbb{R}P^3$ is one of the following sets
  - an ordinary plane together with the ideal points represented by the straight lines in the plane;
  - the set of all ideal points, called the *plane at infinity*.
- A *projective straight line* in $\mathbb{R}P^3$ is one of the following sets
  - an ordinary straight line together with its ideal point;
  - the set of ideal points of a projective plane. Such sets are called *straight lines at infinity*.

Φ can be extended to the whole projective space with the exception of the point O. This extension establishes a one to one correspondence between the points of any projective plane not going through O and the projective closure of the plane Σ.

Projective geometry is a branch of geometry concerned with those properties of planar figures that are unchanged by the projection Φ. Since distances and angles are not invariant, metric concepts, which are so important in Euclidean geometry are alien to projective geometry. This does not mean however that projective theorems can not be specialized to get metric theorems. Since Φ maps ordinary and ideal points to both ordinary and ideal points, ordinary points are not distinguished from the ideal ones. This implies another important feature of projective geometry, that the concept of parallelism vanishes. Indeed, in the projective space, two different coplanar straight lines always intersect one another at exactly one point, a straight line and a plane always has a point in common and two different planes always intersect one another along a straight line.

1.3. The Topological Structure of the Projective Straight Line and Plane.

A *topological space* is a set where the neighborhoods of points are defined in such a way that the axioms below are satisfied.

- Every neighborhood of a point p contains p.
- If U is a neighborhood of p and $U \subset V$ then V is also a neighborhood of p.
- The whole space is a neighborhood of every point.
- The intersection of two neighborhoods of p is also a neighborhood of p.
- Every neighborhood U of a point p contains a smaller neighborhood $V \subset U$ of p such that V is a neighborhood of any of its points.

Sets that are neighborhoods of any of their points are called *open*. Observe that the family of open subsets describes the neighborhood structure uniquely. Indeed, a subset U of the space is a neighborhood of the point p if and only if there exists an open subset $V \subset U$ that contains p. Thus, topology can also be introduced in terms of open subsets. In that case, the above system of axioms is replaced by the following equivalent system of axioms for the open subsets.

- The whole space and the empty set are open.
- The intersection of two open sets is open.
- The union of an arbitrary collection of open sets is open.

For example, in the *standard topology of the Euclidean space*, a subset $U \subset \mathbb{E}^3$ is a neighborhood of the point $p \in \mathbb{E}^3$ if there exists a ball (i.e. a solid sphere) centered at p that is contained in U.

A topology on a set X induces a topology on every subset $Y \subset X$. Namely, a subset $V \subset Y$ is a neighborhood of the point $p \in Y$ in the *subspace topology* if there exists a neighborhood U of p in X such that V is the trace of U in Y, i.e. $V = Y \cap U$. The usual topology on the circle, sphere and other geometrical objects lying in $\mathbb{E}^3$ is the subspace topology inherited from the standard topology of $\mathbb{E}^3$.

The concept of topological space was created with the aim to find the weakest structure on a set that enables us to define continuity. Indeed, the usual "$\epsilon - \delta$" definition of continuity given in calculus for real functions is fully compatible with the following more general definitions. A mapping $f\colon X \to Y$ between the topological spaces X and Y is said to be *continuous at* $p \in X$ if for every neighborhood $U \subset Y$ of $f(p)$ in Y one can find a neighborhood V of p in X such that $f(V) \subset U$. The mapping f is said to be *continuous* if it is continuous at every point of X.

We say that two topological spaces are *topologically equivalent*, or *homeomorphic* if there is a one to one correspondence f between their points which is continuous and has continuous inverse. In this case, f is called a *homeomorphism*. Intuitively, two subsets of the Euclidean space are homeomorphic if shrinking and stretching the rubber model of one of them we can get a model of the other. (For this reason, topology has the nickname "rubber geometry".) During the deformation we may also cut the model along some edges or other subsets, but if we do this, we must later on glue together pieces that were cut apart. For example, the perimeter of a circle and a square are homeomorphic, a disk, a solid square and a hemisphere are homeomorphic, but a circle is not homeomorphic to a disk.

Another important construction is the *identification* or *factor* topology. In many cases we want to construct a new topological space by gluing together two topological spaces, or identifying different points of a given space. In both cases, what we have is a topological space X (that can be a union of two or more topological spaces) and an equivalence relation $\sim$ on X that tells which points of X should be glued together or identified. Points of the identification space $X/\sim$ will be the equivalence classes of $\sim$. A subset $\widetilde{U} \subset X/\sim$ is a neighborhood of the "point" $\widetilde{p} \in X/\sim$ if and only if the union of the equivalence classes that belong to $\tilde{U}$ is a neighborhood of every point in $\tilde{p}$.

For example, gluing together the endpoints of a segment we get a topological circle. Identifying a pair of antipodal points on a circle we get a topological space homeomorphic to the figure 8. If two antipodal pairs are identified then the factor space is homeomorphic to the character $\mathbb{O}$. If however, every point on the circle is identified with its antipodal pair, the identification space will be homeomorphic to the circle.

Let us start from a rectangle $T = \{(x, y) \mid 0 \leq x \leq a,\, 0 \leq y \leq b\}$ with vertices $A(0,0), B(a,0), C(a,b)$ and $D(0,b)$. We can obtain interesting surfaces by gluing together sides of T. If the edge $\overrightarrow{AB}$ is glued to the edge $\overrightarrow{DC}$ with the indicated orientation we get a tube. If however, two opposite edges are glued together with a twist, say $\overrightarrow{AB}$ is glued to $\overrightarrow{CD}$ then a Möbius band will be obtained.

Exercise. Show that the factor space obtained from the rectangle T by gluing together the edges $\overrightarrow{AB}$ and $\overrightarrow{BC}$ is homeomorphic to the Möbius band. (Hint: Cut the rectangle into two triangles along the diagonal BD.)

Let us consider now a point O and a projective straight line e not passing through O. There is a natural one to one correspondence between the elements of the following sets.

- **a.** the projective straight line e;
- **b.** the set $\mathcal{O}$ of those straight lines that lie in the plane Π spanned by O and e and go through O. The set $\mathcal{O}$ is called *a pencil of lines*;
- **c.** the set of antipodal pairs of points of a circle C_1 centered at O;
- **d.** the set of points of a closed semicircle $H_1 \subset C_1$ with the endpoints identified.
- **e.** any circle C_2 that goes through O.

If $f \in \mathcal{O}$ than it corresponds $f \cap e$ in e, the antipodal pair of points $f \cap C_1$ in S_1, the point(s) $f \cap H_1$ in H_1. If f is not tangent to C_2 then it corresponds in C_2 the point $f \cap C_2 \setminus \{O\}$, otherwise it corresponds to O.

Each of these sets carries a standard topology. The circles C_1 and C_2 and the semicircle H_1 are furnished with the subspace topology, from which we can derive an identification topology on the sets **c** and **d**. We define the usual topology on the pencil of lines $\mathcal{O}$ as follows. A subset $U \subset \mathcal{O}$ of the pencil is said to be a neighborhood of the straight line f if there exists a positive angle ϵ such that any rotation of f about O with angle less than ϵ belongs to U.

It is not difficult to see that the one to one correspondences between the sets **b**, **c**, **d** and **e** are homeomorphisms so the topology of the projective straight line is uniquely defined by the requirement that it should be homeomorphic to any of these spaces. A natural consequence of this definition is that *the projective straight line with its natural topology is topologically equivalent to the circle*.

The topology of the projective plane can be treated in a similar manner. Fixing a projective plane Σ in the space and a point $O \notin \Sigma$, we have natural one to one correspondences between the points of the following sets.

a'. the projective plane Σ;
b'. the set $\mathcal{O}$ of all straight lines through O;
c'. the set of antipodal pairs of points of a sphere S centered at O;
d'. the set of points of a closed hemisphere $H \subset S$ with antipodal points of the boundary identified.

Observe that the set **e** has no analog in this case. If we furnish all these spaces with their standard topologies then they become homeomorphic. With some surgery, we can get another realization of the space **d'**. Assume that the hemisphere H is chosen in such a way that its boundary circle is parallel to the plane Σ. A hyperbola drawn in the plane Σ cuts the ordinary Euclidean plane into three parts. We get two convex parts, say $\mathcal{A}$ and $\mathcal{C}$ and a band $\mathcal{B}$ between them. Projecting this picture to the hemisphere from O we obtain a splitting of the hemisphere into three parts. Let us denote by $\widetilde{\mathcal{A}}$, $\widetilde{\mathcal{B}}$ and $\widetilde{\mathcal{C}}$ these parts. The boundaries of $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{C}}$ have one arc in common with the boundary of the hemisphere, say I_1 and I_2 respectively, while $\widetilde{\mathcal{B}}$ has two ones, say J_1 and J_2. To get the projective plane from the hemisphere, we have to identify the arc I_1 with the arc I_2 and the arc J_1 with the arc J_2. When we glue $\widetilde{\mathcal{A}}$ to $\widetilde{\mathcal{C}}$ along I_1 and I_2, the resulting space is a topological disk. When the edges J_1 and J_2 of $\widetilde{\mathcal{B}}$ are glued together, a Möbius band is obtained. Both the disk and the Möbius band has a topological circle for their common boundary. We conclude that *the projective plane is topologically equivalent to a disk and a Möbius band that are glued together along their boundaries*. It is no use to try to glue together a disk and a Möbius band together along their boundary edges, because it is impossible in the 3-dimensional Euclidean space. The precise statement behind this fact is that the *projective plane is not homeomorphic to any subspace of* $\mathbb{E}^3$.

Due to the fact that the projective plane contains a Möbius band, it inherits many of its interesting properties. For example, removing the midline of a Möbius band the band remains connected. A related property of the projective plane is that a straight line does not cut the plane into two pieces – it is possible to walk from one half plane to the other through a point at infinity.

PROJECTIVE GEOMETRY

Lecture Notes by Balázs Csikós

2. LINEAR SPACES AND THE ASSOCIATED PROJECTIVE SPACES

As we saw in the introduction, there is a natural one to one connection between points of a real projective plane and the straight lines through a given point of the Euclidean space. Thus, we may say that the latter set itself is a real projective plane. Having this picture in mind we introduce higher dimensional projective spaces.

Fixing a point in the Euclidean space, we can identify every point of the space with its position vector. These vectors have a very rich algebraic structure. They can be added, multiplied by scalars (real numbers), they have length, closely related to the dot product, and there is also a cross-product. Now only the first two operations are important for us. For the Euclidean space, the set of scalars is the set $\mathbb{R}$ of real numbers. $\mathbb{R}$ has an addition and multiplication of elements. The operations listed above satisfy many algebraic identities. Since in most constructions and theorems connected with the Euclidean space we use only the circumstance that certain algebraic identities are fulfilled and the explicit sets and the explicit form of the operations are irrelevant, we can increase the power of these constructions and theorems by generalizing them to arbitrary sets that are equipped with operations having the necessary properties by postulated axioms.

2.1. Groups, Rings, Division Rings and Fields.

2.1.1. Definition. A pair $(X, *)$ is called a *group* if X is an arbitrary set, $*\colon X \times X \to X$ is a binary operation satisfying the following axioms:

$$(x * y) * z = x * (y * z) \qquad \forall x, y, z \in X \qquad \text{(associativity)};$$
$$\exists!\, e \in X \text{ such that } x * e = e * x = x \quad \forall x \in X \qquad \text{(existence of identity)};$$
$$\forall x \in X\, \exists!\, x^{-1} \in X \text{ such that } x * x^{-1} = x^{-1} * x = e \qquad \text{(existence of inverse)}.$$

Remark. Some assumptions in the above system of axioms are redundant. For example, uniqueness of the identity and inverse can be proved from the other assumptions.

The identity element is also called the unit element of the group and is sometimes denoted by 1. When the operation is called "addition" we prefer to call the identity the zero element of the group and denote it by 0, while the inverse of an element a is called negative a or minus a and is denoted by $-a$.

Typeset by $\mathcal{AMS}$-TEX

2.1.2. Definition. A group $(X, *)$ is said to be *commutative* or *abelian* if

$$x * y = y * x \quad \forall x, y \in X \quad \text{(commutativity)}.$$

Examples.

- The sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ of integer, rational, real and complex numbers form a commutative group with respect to addition.
- The sets $\mathbb{Q}^*$, $\mathbb{R}^*$, $\mathbb{C}^*$ of non-zero rational, real and complex numbers form a commutative group with respect to multiplication.
- The set $GL(n, \mathbb{R})$ of invertible $n \times n$ matrices form a non-commutative group with respect to multiplication.
- The set S_X of all bijective (= one to one and onto) mappings of a set X onto itself forms a non-commutative group under the composition $\circ$ of mappings: $(f \circ g)(x) = f(g(x))$.

2.1.3. Definition. A triple $(R, +, \cdot)$ is said to be a *ring* if R is a set , $+$ and $\cdot$ are binary operations called addition and multiplication, and the following axioms are satisfied:

- $(R, +)$ is a commutative group (called the *additive group of the ring*);
- multiplication is associative
$$(x \cdot y) \cdot z = x \cdot (y \cdot z) \quad \forall x, y, z \in R;$$
- multiplication is distributive with respect to addition from both sides
$$x \cdot (y + z) = x \cdot y + x \cdot z \text{ and } (y + z) \cdot x = y \cdot x + z \cdot x \quad \forall x, y, z \in X.$$

2.1.4. Definiton. The ring $(R, +, \cdot)$ is a *ring with identity* or *unit element* if there exists a (unique) nonzero element $e \in R$ such that

$$e \cdot x = x \cdot e = x \quad \forall x \in X$$

Remark. Some authors include the assumption on the existence of identity in the definition of the ring and allow the identity to be equal to the zero element of the group.

2.1.5. Definition. A ring $(R, +, \cdot)$ is a *division ring* or *skew field* if the set $R^* = \{x \in R \mid x \neq 0\}$ of nonzero elements of R form a group with respect to the multiplication. The group $(R^*, \cdot)$ is called then the *multiplicative group of the division ring*.

2.1.6. Definition. A *field* is a division ring with commutative multiplicative group.

Examples.

- Let m be an integer, $m\mathbb{Z} = \{n \in \mathbb{Z} \mid m \text{ divides } n\}$ the set of integer multiples of m. $(m\mathbb{Z}, +, \cdot)$ is a ring but it is a ring with identity if and only if $m = \pm 1$.
- $(\mathbb{Z}, +, \cdot)$ is a ring with identity which is not a division ring.
- Let m be a positive integer and $\mathbb{Z}_m$ denote the modulo m residuum classes of integers with the usual addition and multiplication. $\mathbb{Z}_m$ is a ring with identity and it is a field if and only if m is a prime.
- The set $Mat(n, \mathbb{R})$ of $n \times n$ matrices is a ring with respect to the usual operations, but it is not a division ring.
- Later on we shall define the division ring $\mathbb{H}$ of quaternions. $\mathbb{H}$ serves for an example of a division ring that is not a field.
- $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ and $(\mathbb{C}, +, \cdot)$ are fields.

2.2. Vector Spaces and their Subspaces. From this point on we fix a skew field $\mathbb{F}$. For geometrical applications, the most important examples are the fields of real and complex numbers and the skew field of quaternions. Elements of $\mathbb{F}$ will be called *scalars*.

2.2.1. DEFINITION. A *linear space* or *vector space over the skew field* $\mathbb{F}$ is a set V together with rules of addition and multiplication with scalars which associates to any two elements $\mathbf{a}$, $\mathbf{b}$ in V a *sum* $\mathbf{a}+\mathbf{b}$ in V, and to any $\lambda \in \mathbb{F}$ and $\mathbf{a} \in V$ a *product* $\lambda\mathbf{a}$ in V, if the following axioms hold.

- $(V, +)$ is a commutative group;
- $\lambda(\mathbf{a}+\mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b} \quad \forall\lambda \in \mathbb{F}$, $\mathbf{a}, \mathbf{b} \in V$;
- $(\lambda + \mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a} \quad \forall\lambda, \mu \in \mathbb{F}$, $\mathbf{a} \in V$;
- $(\lambda\mu)\mathbf{a} = \lambda(\mu\mathbf{a}) \quad \forall\lambda, \mu \in \mathbb{F}$, $\mathbf{a} \in V$;
- $1\mathbf{a} = \mathbf{a} \quad \forall\mathbf{a} \in V$, where 1 is the identity element of $\mathbb{F}$.

We refer to elements of V as *vectors*. We shall use bold letters for vectors and greek letters for scalars to distinguish them visually.

EXAMPLE. Let us denote by $\mathbb{F}^n$ the set of all n-tuples $\mathbf{x} = (\xi_1, \dots, \xi_n)$ consisting of elements of $\mathbb{F}$. For $\mathbf{x} = (\xi_1, \dots, \xi_n)$, $\mathbf{y} = (\eta_1, \dots, \eta_n)$ and $\lambda \in \mathbb{F}$, we define the following rules of addition and multiplication

$$\mathbf{x} + \mathbf{y} = (\xi_1 + \eta_1, \dots, \xi_n + \eta_n);$$
$$\lambda\mathbf{x} = (\lambda\xi_1, \dots, \lambda\xi_n).$$

$\mathbb{F}^n$ is a vector space over $\mathbb{F}$ with respect to these operations.

2.2.2. DEFINITION. A nonempty subset W of a vector space V is a *subspace* of V if W is closed with respect to the addition and multiplication by scalars. In other words, W is a subspace if for any $\mathbf{a}, \mathbf{b} \in W$ and $\lambda \in \mathbb{F}$, W contains $\mathbf{a}+\mathbf{b}$ and $\lambda\mathbf{a}$.

A subspace of a vector space over $\mathbb{F}$ is itself a vector space over the same skew field. The following proposition is obvious.

2.2.3. PROPOSITION. *The intersection of any family of subspaces of V is a subspace of V.*

Let S be an arbitrary subset of the vector space V. Since V contains S and V is a subspace of itself, S is contained in at least one subspace of V. The family of all subspaces of V that contain S has a minimal element, namely the intersection of all subspaces in the family. This fact gives rise to the following definitions.

2.2.4. DEFINITION. The subspace *spanned* or *generated* by a subset S of a vector space is the smallest subspace that contains S. We shall use the notation $[S]$ for the subspace generated by S.

2.2.5. DEFINITION. The *sum* or *join* $\sum_{i\in I} W_i$ of the subspaces $\{W_i \mid i \in I\}$ of a vector space is the subspace generated by their union $\bigcup_{i\in I} W_i$. The sum of the subspaces $W_1, \dots, W_n$ will also be denoted by $W_1 + \cdots + W_n$. A sum of subspaces W_i is a *direct sum* if each subspace W_i intersects the sum of the others at $\mathbf{0}$. To indicate that this is the case, direct sums are denoted by $\bigoplus_{i\in I} W_i$ or $W_1 \oplus \cdots \oplus W_n$.

2.3. Basis, Coordinates, Dimension.

2.3.1. Definition. A *linear combination* of some vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ of a vector space is a vector of the form $\lambda_1 \mathbf{a}_1 + \cdots + \lambda_n \mathbf{a}_n$, where $\lambda_i \in \mathbb{F}$. A family of vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ is *linearly independent* if

$$\lambda_1 \mathbf{a}_1 + \cdots + \lambda_n \mathbf{a}_n = \mathbf{0}$$

implies $\lambda_1 = \cdots = \lambda_n = 0$.

The set of all linear combinations of a family of vectors S is the subspace generated by S. A family of vectors is linearly dependent if and only if one of the vectors is a linear combination of the others. This implies that a minimal set of generators of a subspace of the vector space contains linearly independent vectors.

2.3.2. Proposition. *Let $\mathcal{B} \subset V$ be a subset of the vector space V. Then the following conditions for $\mathcal{B}$ are equivalent.*

a. *$\mathcal{B}$ is a minimal set of generators of V;*
b. *$\mathcal{B}$ is a maximal system of linearly independent vectors;*
c. *$\mathcal{B}$ is a linearly independent family of vectors that generate V.*
d. *Any element of the vector space V can be written as a linear combination of the elements of $\mathcal{B}$ in a unique way.*

2.3.3. Definition. We call a subset $\mathcal{B}$ of a vector space a *basis* if it satisfies one of the equivalent conditions of the proposition.

The existence of a basis follows easily if we assume that the vector space is generated by a finite set of vectors. To prove the existence of a basis in general, one has to apply Zorn's lemma. To avoid the usage of set theoretical arguments, we assume from now on that every vector space we deal with is finitely generated.

2.3.4. Definition. Let V be a vector space with basis $\mathcal{B} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$. The *coordinates of a vector* $\mathbf{v} \in V$ *with respect to the basis* $\mathcal{B}$ are the coefficients $(\lambda_1, \dots, \lambda_n)$ in the expression of $\mathbf{v}$ as a linear combination of $\mathcal{B}$

$$\mathbf{v} = \lambda_1 \mathbf{x}_1 + \cdots + \lambda_n \mathbf{x}_n.$$

2.3.5. Theorem. *Any two bases of a finitely generated vector space have the same number of elements.*

2.3.6. Definition. The number of elements of a basis of a vector space V is the *dimension* of V.

The proof of the Theorem rests upon the Exchange Lemma.

2.3.7. Exchange Lemma. *If the vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ generate the vector space V and $\mathbf{b}_1, \dots, \mathbf{b}_m$ is a family of linearly independent vectors, then $n \geq m$, and we can replace m vectors of the set $\mathbf{a}_1, \dots, \mathbf{a}_n$ with the vectors $\mathbf{b}_1, \dots, \mathbf{b}_m$ in such a way that the set $\mathbf{b}_1, \dots, \mathbf{b}_m$ together with those $\mathbf{a}_i$'s that were not replaced span V.*

Proof. We use induction on m. The case $m = 0$ is obvious. Assume now that the assertion is proved for $m-1$. Since any subset of a family of linearly independent vectors is linearly independent, the induction hypothesis can be applied to the vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ and $\mathbf{b}_1, \dots, \mathbf{b}_{m-1}$. Without loss of generality, we may assume that $\mathbf{b}_1, \ldots, \mathbf{b}_{m-1}, \mathbf{a}_m, \ldots, \mathbf{a}_n$ span V. Then $\mathbf{b}_m$ is a linear combination of these vectors

$$\mathbf{b}_m = \lambda_1 \mathbf{b}_1 + \cdots + \lambda_{m-1}\mathbf{b}_{m-1} + \mu_m \mathbf{a}_m + \cdots + \mu_n \mathbf{a}_n.$$

At least one of the coefficients $\mu_m, \ldots \mu_n$ is different from zero, otherwise $\mathbf{b}_m$ would be a linear combination of $\mathbf{b}_1, \ldots, \mathbf{b}_{m-1}$. This implies $n \geq m$. On the other hand, assuming $\mu_m \neq 0$ for example, we can express $\mathbf{a}_m$ as a linear combination of $\mathbf{b}_1, \ldots, \mathbf{b}_m, \mathbf{a}_{m+1}, \ldots, \mathbf{a}_n$

$$\mathbf{a}_m = -\mu_m^{-1}\lambda_1 \mathbf{b}_1 - \cdots - \mu_m^{-1}\lambda_{m-1}\mathbf{b}_{m-1} + \mu_m^{-1}\mathbf{b}_m - \mu_m^{-1}\mu_{m+1}\mathbf{a}_{m+1} - \cdots - \mu_m^{-1}\mu_n \mathbf{a}_n.$$

This means that the set $\mathbf{b}_1, \ldots, \mathbf{b}_m, \mathbf{a}_{m+1}, \ldots, \mathbf{a}_n$ spans the same subspace that is spanned by $\mathbf{b}_1, \ldots, \mathbf{b}_m, \mathbf{a}_m, \ldots, \mathbf{a}_n$, but the latter set generates the whole space. This proves the lemma. $\square$

Proof of Theorem 2.3.5. If $\mathcal{B}_1$ and $\mathcal{B}_2$ are two bases of a vector space, then applying the Exchange Lemma we get $\#\mathcal{B}_1 \leq \#\mathcal{B}_2$, since $\mathcal{B}_1$ is a linearly independent set, $\mathcal{B}_2$ is a system of generators. Changing the role of $\mathcal{B}_1$ and $\mathcal{B}_2$ we obtain $\#\mathcal{B}_2 \leq \#\mathcal{B}_1$, yielding $\#\mathcal{B}_1 = \#\mathcal{B}_2$. $\square$

The following proposition is an easy consequence of the definitions.

2.3.8. Proposition. *If W_1 and W_2 are subspaces of a finite dimensional vector space, then $W_1 \subset W_2$ implies* $\dim W_1 \leq \dim W_2$. *If W_1 and W_2 have the same dimension then they are equal.*

Now we prove a fundamental formula.

2.3.9. Theorem. *Let W_1 and W_2 be two subspaces of a finite dimensional linear space. Then*

$$\dim(W_1 + W_2) + \dim(W_1 \cap W_2) = \dim W_1 + \dim W_2.$$

Proof. Let us choose a basis $\mathcal{B}_{12} = (\mathbf{x}_1, \ldots, \mathbf{x}_r)$ of the intersection $W_1 \cap W_2$. Since $\mathcal{B}_{12}$ is a linearly independent subset of W_1 and W_2, we can extend it to a basis of W_1 and W_2 respectively. Let $\mathcal{B}_1 = (\mathbf{x}_1, \ldots, \mathbf{x}_r, \mathbf{y}_1, \ldots, \mathbf{y}_s)$ be a basis of W_1, $\mathcal{B}_2 = (\mathbf{x}_1, \ldots, \mathbf{x}_r, \mathbf{z}_1, \ldots, \mathbf{z}_t)$ a basis of W_2. If we prove that $\mathcal{B}^{12} = (\mathbf{x}_1, \ldots, \mathbf{x}_r, \mathbf{y}_1, \ldots, \mathbf{y}_s, \mathbf{z}_1, \ldots, \mathbf{z}_t)$ is a basis of $W_1 + W_2$ then we are ready since in that case $\dim(W_1 \cap W_2) = r$, $\dim W_1 = r + s$, $\dim W_2 = r + t$ and $\dim(W_1 + W_2) = r + s + t$. It is clear that $\mathcal{B}^{12}$ generates $W_1 + W_2$. Let us show that it contains linearly independent vectors. Consider a linear combination of these vectors that gives $\mathbf{0}$

$$\alpha_1 \mathbf{x}_1 + \cdots + \alpha_r \mathbf{x}_r + \beta_1 \mathbf{y}_1 + \cdots + \beta_s \mathbf{y}_s + \gamma_1 \mathbf{z}_1 + \cdots + \gamma_t \mathbf{z}_t = \mathbf{0}.$$

Setting

$$\mathbf{w} = \alpha_1 \mathbf{x}_1 + \cdots + \alpha_r \mathbf{x}_r + \beta_1 \mathbf{y}_1 + \cdots + \beta_s \mathbf{y}_s = -\gamma_1 \mathbf{z}_1 - \cdots - \gamma_t \mathbf{z}_t$$

the first expression for $\mathbf{w}$ shows that $\mathbf{w} \in W_1$, while the second gives $\mathbf{w} \in W_2$, hence $\mathbf{w} \in W_1 \cap W_2$. An element of W_2 belongs to $W_1 \cap W_2$ if and only if expressing it as a linear combination of the basis vectors $\mathcal{B}_2$ the coefficients of the vectors $\mathbf{z}_1, \ldots, \mathbf{z}_t$ are equal to zero. Applying this to $\mathbf{w}$ we get that $\gamma_1 = \ldots = \gamma_t = 0$ and thus $\mathbf{w} = \mathbf{0}$. The first expression for $\mathbf{w}$ is a linear combination of the basis $\mathcal{B}_1$ that gives zero. This is possible only if $\alpha_1 = \ldots = \alpha_r = \beta_1 = \ldots = \beta_s = 0$. $\square$

2.4. The Projective Space Associated to a Linear Space.

2.4.1. Definition. Let V be a linear space over a skew field $\mathbb{F}$. Let us introduce an equivalence relation on the set $V \setminus \{\mathbf{0}\}$ as follows. We say that two non-zero vectors $\mathbf{x}, \mathbf{y} \in V$ are equivalent if there exists a non-zero scalar $\lambda \in \mathbb{F}$ such that $\mathbf{x} = \lambda \mathbf{y}$. Equivalence classes of this relation are in one to one correspondence with the one dimensional subspaces of V. The set of equivalence classes is called the *projective space associated to the linear space* V and it is denoted by $P(V)$. We shall denote the equivalence class of a vector $\mathbf{0} \neq \mathbf{v} \in V$ by $[\mathbf{v}]$ and call it the *homogeneous vector represented by* $\mathbf{v}$.

For the case, when $V = \mathbb{F}^{n+1}$ we shall use the notation $\mathbb{F}P^n$ instead of $P(\mathbb{F}^{n+1})$. A point of $\mathbb{F}P^n$ is an equivalence class of an $(n+1)$-tuple of scalars, where two $(n+1)$-tuples are equivalent if they are proportional. For $\mathbf{x} = (\xi_1, \dots, \xi_{n+1}) \in \mathbb{F}^{n+1} \setminus \{\mathbf{0}\}$, the point represented by $\mathbf{x}$ is denoted by $(\xi_1 : \cdots : \xi_{n+1})$. Thus, for example, $(1 : 1 : 1)$ and $(2 : 2 : 2)$ are identical, while $(1 : 1 : 1)$ and $(1 : 2 : 3)$ are different points of $\mathbb{R}P^2$.

2.4.2. Definition. The *dimension* of the projective space associated to a linear space is defined to be one less than the dimension of the linear space

$$\dim P(V) = \dim V - 1.$$

2.4.3. Definition. A *subspace* of the projective space $P(V)$ is a projective space associated to some linear subspace W of V.

- The projectivization of the subspace $\{\mathbf{0}\}$ is the empty set. The empty set is the only -1-dimensional subspace.
- 0-dimensional subspaces of a projective space $P(V)$ are projectivizations of the 1-dimensional subspaces of V, i.e., points of the projective space.
- 1-dimensional subspaces of a projective space are called *straight lines*, 2-dimensional subspaces are called *planes*, m-dimensional subspaces are shortly called *m- planes*. A *hyperplane* is a projective subspace the dimension of which is one less than the dimension of the total space.
  The following proposition is obvious.

2.4.4. Proposition.

- *The projective space associated to the intersection of some linear subspaces W_i of a linear space is the intersection of the projectivizations $P(W_i)$ of these subspaces.*
- *Let S be a subset of the projective space $P(V)$, S' the set of vectors in V that represent an element of S. Then the smallest projective subspace of $P(V)$ that contains S is the projectivization of the linear space spanned by S'.*

Thus, we can introduce the intersection and sum or join of projective subspaces in a natural way. We shall denote these operations with the same symbols as for the case of linear spaces. The propositions below are direct consequences of the analogous propositions for linear spaces.

2.4.5. Proposition.

- *If P_1 and P_2 are projective subspaces of a projective space, and $P_1 \subset P_2$, then* $\dim P_1 \leq \dim P_2$. *If, furthermore, P_1 is not equal to P_2, then* $\dim P_1$ *is strictly less than* $\dim P_2$.

- *If P_1 and P_2 are two arbitrary subspaces of a projective space, then*

$$\dim(P_1 + P_2) + \dim(P_1 \cap P_2) = \dim P_1 + \dim P_2.$$

The last equality implies all the basic incidence theorems of the projective space. As an example, consider two straight lines e and f in the projective space. Their intersection may have dimension $-1, 0$ or 1. In the first case, the two lines have an empty intersection and generate a 3-dimensional subspace. We call such lines *skew*. In the second case, the two lines intersect one another at exactly one point and they span a plane. In the third case, the two lines coincide. Thus, we proved that two different coplanar lines intersect one another at exactly one point.

2.5. Projective coordinate systems. A coordinate system in an $(n+1)$-dimensional linear space V is given by a basis and is nothing else than a bijection between V and $\mathbb{F}^{n+1}$. Since $\mathbb{R}P^n$ is topologically not equivalent to $\mathbb{R}^n$ we can not expect the existence of a nice bijection between $\mathbb{F}P^n$ and $\mathbb{F}^n$. There are two ways out of this difficulty: we use either homogeneous coordinates or local coordinates.

2.5.1. *Homogeneous coordinates.* Fix a basis $\mathcal{B} = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{n+1})$ of V. If the point $P \in P(V)$ is represented by the vector $\mathbf{p} \in V$ then write $\mathbf{p}$ as a linear combination of the basis vectors

$$\mathbf{p} = \sum_{i=1}^{n+1} \alpha_i \mathbf{p}_i.$$

The coefficients $(\alpha_1 : \dots : \alpha_{n+1})$ are called the *homogeneous coordinates* of P with respect to the basis $\mathcal{B}$. Since representatives of P are defined only up to a scalar multiplier so are the homogeneous coordinates. This fact is indicated by the usage of : between the coordinates. The mapping $P \mapsto (\alpha_1 : \dots : \alpha_{n+1})$ defines a bijection between $P(V)$ and $\mathbb{F}P^n$.

2.5.2. *Local coordinates.* By a *local coordinate system* we mean a bijection between a subset X of $P(V)$ and $\mathbb{F}^n$. Generally we use the following local coordinate systems. Consider the subset X_i consisting of those points, the i-th homogeneous coordinate α_i of which is different from 0. Homogeneous coordinates of a point $P \in X_i$ can be normalized in a unique way so that the i-th coordinate became equal to 1. Assigning to P the rest of the coordinates, $(\alpha_i^{-1}\alpha_1, \dots, \alpha_i^{-1}\alpha_{i-1}, \alpha_i^{-1}\alpha_{i+1}, \dots, \alpha_i^{-1}\alpha_{n+1})$, we obtain a bijection between X_i and $\mathbb{F}^n$.

Observe that the complement of X_i in $P(V)$ is a hyperplane, thus, the projective space is the union of X_i, that can be identified with $\mathbb{F}^n$, and a hyperplane. This decomposition of the projective space generalizes the decomposition of the real projective space defined in the introduction into the ordinary Euclidean part and the plane at infinity.

Homogeneous coordinates of a point are related to a basis $\mathbf{p}_1, \dots, \mathbf{p}_{n+1}$ of the underlying linear space. The points $P_i = [\mathbf{p}_i]$ are vertices of a non-degenerate simplex in the projective space, called the *reference simplex* of the coordinate system. The homogeneous coordinates of P_i are $(0 : \dots : 1 : \dots : 0)$, where the only non-zero coordinate stands at the i-th position. The point Q whose homogeneous coordinates are $(1 : 1 : \dots : 1)$ is the *unit point of the coordinate system*. We are going to prove that over a commutative field, homogeneous coordinates of a point are determined by the reference simplex and the unit point.

2.5.3. DEFINITION. We say that a system of points in an n-dimensional projective space is *in general position* if any $k+1 \leq n+1$ points of it span a k-dimensional projective subspace.

Consider a system of points $S = \{P_i \mid i \in I\}$ in the projective space $P(V)$ and a system of representatives $\mathbf{S} = \{\mathbf{p}_i \mid i \in I, [\mathbf{p}_i] = P_i\}$. The system S is in general position if and only if any $k+1 \leq \dim V$ vectors of $\mathbf{S}$ span a $(k+1)$-dimensional linear subspace in V. $k+1$ vectors span a $(k+1)$-dimensional subspace if and only if the vectors are linearly independent. Since any subsystem of a linearly independent system is also a linearly independent system, a set S of at least $\dim V$ points is in general position if and only if any $\dim V$ points of it span the whole space, in other words, if no $\dim V$ points of it can be covered by a hyperplane.

Clearly, vertices of the reference simplex and the unit point of a homogeneous coordinate system yield $n+2$ points in general position. Conversely, any $n+2$ points in general position can be derived this way.

2.5.4. PROPOSITION. *Let us fix $n+2$ points $P_1, \dots, P_{n+1}, Q$ in general position in the n-dimensional projective space. Then one can choose representatives $\mathbf{p}_1, \dots,$ $\mathbf{p}_{n+1}, \mathbf{q}$ for $P_1, \dots, P_{n+1}, Q$ such that*

$$\mathbf{q} = \mathbf{p}_1 + \cdots + \mathbf{p}_{n+1},$$

and these representatives are unique up to a common scalar multiplier.

PROOF. Let us choose arbitrary representatives for $\widetilde{\mathbf{p}}_1, \dots, \widetilde{\mathbf{p}}_{n+1}, \widetilde{\mathbf{q}}$ for $P_1, \dots,$ P_{n+1}, Q respectively. By the general position condition, the vectors $\widetilde{\mathbf{p}}_1, \dots, \widetilde{\mathbf{p}}_{n+1}$ form a basis of the underlying vector space. Let us express $\widetilde{\mathbf{q}}$ as a linear combination of them

$$\widetilde{\mathbf{q}} = \alpha_1\widetilde{\mathbf{p}}_1 + \cdots + \alpha_{n+1}\widetilde{\mathbf{p}}_{n+1}.$$

The α_i-s in this expression can not be equal to zero, since if, for example, α_1 were zero, then the vectors $\widetilde{\mathbf{q}}, \widetilde{\mathbf{p}}_2, \dots, \widetilde{\mathbf{p}}_{n+1}$ were linearly dependent and that would contradict the assumption on the general position. Consequently, the choice of the non-zero vectors

$$\mathbf{p}_1 = \alpha_1\widetilde{\mathbf{p}}_1 \ \dots, \ \mathbf{p}_{n+1} = \alpha_{n+1}\widetilde{\mathbf{p}}_{n+1}, \ \mathbf{q} = \widetilde{\mathbf{q}}$$

as representatives for $P_1, \dots, P_{n+1}, Q$ will satisfy all the requirements. Uniqueness up to constant multiplier follows from the fact that having fixed a representative of Q, the above construction leaves no more freedom in the choice of representatives for $P_1, \dots, P_{n+1}$. $\square$

By the proposition above, any $n+2$ points $P_1, \dots, P_{n+1}, Q$ in general position leads to an almost unique coordinate system with reference simplex $P_1, \dots, P_{n+1}$ and unit point Q. By the identity

$$\sum_{i=1}^{n+1} \alpha_i \mathbf{p}_i = \sum_{i=1}^{n+1} (\alpha_i\lambda^{-1})(\lambda\mathbf{p}_i),$$

if we multiply with a scalar $\lambda \in \mathbb{F}^*$ the basis vectors to which homogeneous coordinates are associated, the homogeneous coordinates are multiplied by λ^{-1} from the *right*. Since two systems $(\alpha_1 : \cdots : \alpha_{n+1})$, $(\alpha'_1 : \cdots : \alpha'_{n+1})$ of homogeneous coordinates are considered to be the same if and only if there is a scalar $\lambda \in \mathbb{F}^*$ such that multiplying α_i from the *left* gives α'_i for all i, for non-commutative division rings, different coordinate systems with the same reference simplex and unit point may give different homogeneous coordinates. However, if $\mathbb{F}$ is a field, homogeneous coordinates are uniquely determined by the reference simplex and the unit point. Thus, in this case, we may speak about the *homogeneous coordinates of a point with respect to the projective coordinate system given by the reference simplex* $P_1, \ldots, P_{n+1}$ *and unit point* Q.

Let us compare Cartesian and projective coordinates for a special choice of the coordinate systems. Consider the 3-dimensional Euclidean space $\mathbb{E}^3$ with a fixed Cartesian coordinate system. The coordinates give a one to one correspondence between $\mathbb{E}^3$ and $\mathbb{R}^3$. Consider the plane defined by the equation $\Sigma = \{P(x, y, z) \mid z = 1\}$. On this plane, we introduce the Cartesian coordinate system with origin at $(0, 0, 1)$ and x- and y-axis parallel to the x- and y-axis of the spatial coordinate system respectively. The coordinates of the point $P(x, y, 1)$ in this planar coordinate system are (x, y). Now consider the projective closure of the plane Σ with points at infinity. There is a one to one correspondence between the points of this plane and the straight lines through $(0, 0, 0)$ that form the projective plane associated to $\mathbb{R}^3$. Let us introduce a projective coordinate system on this plane choosing the point at infinity of the x-axis for P_1, the point at infinity of the y-axis for P_2, the point $(0, 0, 1)$ for P_3 and the point $(1, 1, 1)$ for Q. Obviously, we may use the representatives

$$\mathbf{p}_1 = (1, 0, 0), \quad \mathbf{p}_2 = (0, 1, 0), \quad \mathbf{p}_3 = (0, 0, 1), \quad \mathbf{q} = (1, 1, 1).$$

Using these coordinate systems, consider a point $P(x, y, 1)$ in the plane Σ. Its coordinates in the planar Cartesian coordinate system are (x, y). The vector $(x, y, 1) = x\mathbf{p}_1 + y\mathbf{p}_2 + \mathbf{p}_3$ or any multiple of it represents the point P as a point of $P(\mathbb{R}^3)$. Thus the homogeneous coordinates of P are $(x : y : 1)$. In other words, a point with homogeneous coordinates $(x : y : z)$ is an ordinary point of the plane Σ if and only if $z \neq 0$ and then the Cartesian coordinates of that point are $(x/z, y/z)$. Points with homogeneous coordinates $(x : y : 0)$ are points at infinity. $(x : y : 0)$ is the point at infinity of the straight line that goes through the points (0,0) and (x,y).

Here is another example. Let P_1, P_2, P_3 be the vertices of a triangle on the Euclidean plane, Q be the baricenter of the triangle. Homogeneous coordinates of a point with respect to the pojective coordinate system determined by the reference triangle P_1, P_2, P_3 and unit point Q are the *baricentric coordinates* of the point. Baricentric coordinates have the following meaning: if we put three small balls of mass $\alpha_1, \alpha_2, \alpha_3$ respectively to the vertices of the reference triangle, then the mass center of these three balls is at the point with baricentric coordinates $(\alpha_1 : \alpha_2 : \alpha_3)$. Points at infinity have baricentric coordinates with zero sum.

2.6. The Theorems of Desargues and Pappus.

To illustrate the way we can work with homogeneous vectors, we prove two fundamental configuration theorems. These theorems are real projective theorems. They do not involve the notions of distance, angle, area etc. All we need to formulate them is the incidence of points and straight lines.

We start with two definitions.

2.6.1. DEFINITION.

- We say that *two triangles ABC and $A'B'C'$ are in perspective from a point P* if the triads PAA', PBB' and PCC' are collinear.
- We say that *two triangles with sides abc and $a'b'c'$ are in perspective with respect to the straight line p* if the triads of straight lines paa', pbb' and pcc' are concurrent.

REMARKS.

- This definition assumes that we have fixed a correspondence between the vertices (and the sides) of the two triangles.
- The triangles are not necessarily coplanar.

2.6.2. DESARGUES' THEOREM. *Two triangles are in perspective from a point if and only if they are in perspective from a line.*

PROOF. Denote by ABC and $A'B'C'$ the vertices of the two triangles. It is easy to see that if two corresponding vertices or two corresponding sides of the two triangles coincide, then the two triangles are in perspective in both senses, so we may assume that this is not the case.

Assume first that the triangles are in perspective from a point P. Since P, A, A' are collinear, we may choose representatives $\mathbf{p}$, $\mathbf{a}$, $\mathbf{a}'$ for these points in such a way that $\mathbf{a}' = \mathbf{p} + \mathbf{a}$. Similarly, we can choose vectors $\mathbf{b}, \mathbf{b}', \mathbf{c}, \mathbf{c}'$ from the underlying vector space in such a way that $B = [\mathbf{b}], B' = [\mathbf{b}'], C = [\mathbf{c}], C' = [\mathbf{c}']$ and

$$\mathbf{b}' = \mathbf{p} + \mathbf{b}, \quad \mathbf{c}' = \mathbf{p} + \mathbf{c}.$$

Since

$$\mathbf{a} - \mathbf{b} = (\mathbf{p} + \mathbf{a}) - (\mathbf{p} + \mathbf{b}) = \mathbf{a}' - \mathbf{b}',$$

the point represented by $\mathbf{a} - \mathbf{b}$ lies both on the straight lines AB and $A'B'$, i.e., $[\mathbf{a} - \mathbf{b}] = AB \cap A'B'$. Similarly,

$$[\mathbf{b} - \mathbf{c}] = BC \cap B'C', \quad [\mathbf{c} - \mathbf{a}] = CA \cap C'A'.$$

We have to show that the intersection points $AB \cap A'B'$, $BC \cap B'C'$ and $CA \cap C'A'$ are collinear. For this purpose, it suffices to show that the representatives of these points are linearly dependent. This is obvious, however, since the sum of the representatives gives zero

$$(\mathbf{a} - \mathbf{b}) + (\mathbf{b} - \mathbf{c}) + (\mathbf{c} - \mathbf{a}) = \mathbf{0}.$$

To prove the converse, assume that the triangles ABC and $A'B'C'$ are in perspective with respect to a straight line. This means that the points

$$A" = BC \cap B'C', \quad B" = AC \cap A'C', \quad C" = AB \cap A'B'$$

are collinear. Then the triangles $AA'B"$ and $BB'A"$ are in perspective from the point $C"$. Applying the first half of Desargues' Theorem to this pair of triangles we obtain that these triangles are in perspective with respect to a line. In other words, the intersection points of the corresponding sides, that is, the points

$$P = AA' \cap BB', \quad C = AB" \cap A"B, \quad C' = A'B" \cap A"B'$$

are collinear. Then the straight lines AA', BB' and CC' meet at the point P, hence the triangles ABC and $A'B'C'$ are in perspective from the point P. $\square$

Pappus lived in the 2nd half of the third century. We know very little about him, he lived probably in Alexandria. He wrote an eight volume series of books on mathematics, mainly on geometry, in which we can find the following theorem.

2.6.3. PAPPUS' THEOREM. *Let e and f be two different straight lines of the plane, $A, B, C \in e$ and $A', B', C' \in f$ be six different points, all different from the intersection point $P = e \cap f$. Then the following points are collinear*

$$A" = BC' \cap B'C, \quad B" = AC' \cap A'C, \quad C" = AB' \cap A'B.$$

Investigations on the axiomatic foundations of geometry threw a new light on this theorem. It turned out that Pappus' Theorem characterizes projective spaces over a *commutative field*.

2.6.4. THEOREM. *Pappus' Theorem is true in a projective space associated to a linear space over a division ring $\mathbb{F}$ if and only if $\mathbb{F}$ is a field.*

PROOF. Let us choose representatives for P, A and A' such that

$$\begin{aligned} P = [\mathbf{p}], \quad A = [\mathbf{a}], \quad B &= [\mathbf{p} + \mathbf{a}], \\ A' = [\mathbf{a}'], \quad B' &= [\mathbf{p} + \mathbf{a}']. \end{aligned}$$

Then

$$C = [\mathbf{p} + \lambda\mathbf{a}] \text{ and } C' = [\mathbf{p} + \mu\mathbf{a}']$$

for some $\lambda, \mu \in \mathbb{F}$. Since

$$(\mathbf{p} + \mathbf{a}) + \mathbf{a}' = (\mathbf{p} + \mathbf{a}') + \mathbf{a},$$

the point represented by this vector lies both on the lines BA' and $B'A$ and hence $C" = AB' \cap A'B = [\mathbf{p} + \mathbf{a} + \mathbf{a}']$. Similarly, the point represented by

$$(\mathbf{p} + \lambda\mathbf{a}) + \mu\mathbf{a}' = (\mathbf{p} + \mu\mathbf{a}') + \lambda\mathbf{a},$$

lies on both CA' and $C'A$, thus $B" = [\mathbf{p} + \lambda\mathbf{a} + \mu\mathbf{a}']$.

Instead of finding a representative for $A"$, let us determine a representative for the intersection $B"C" \cap B'C$. Since now it is not so easy to guess the answer as in the previous cases, we show how one can compute representative vectors for intersection points.

A representative of $B"C" \cap B'C$ must be a linear combination of $\mathbf{p} + \mathbf{a} + \mathbf{a}'$ and $\mathbf{p} + \lambda\mathbf{a} + \mu\mathbf{a}'$ and also a linear combination of $\mathbf{p} + \mathbf{a}'$ and $\mathbf{p} + \lambda\mathbf{a}$. Thus, we have to solve the equation

$$\begin{aligned} \alpha(\mathbf{p} + \mathbf{a} + \mathbf{a}') + \beta(\mathbf{p} + \lambda\mathbf{a} + \mu\mathbf{a}') &= \gamma(\mathbf{p} + \mathbf{a}') + \delta(\mathbf{p} + \lambda\mathbf{a}) \\ (\alpha + \beta)\mathbf{p} + (\alpha + \beta\lambda)\mathbf{a} + (\alpha + \beta\mu)\mathbf{a}' &= (\gamma + \delta)\mathbf{p} + \delta\lambda\mathbf{a} + \gamma\mathbf{a}' \end{aligned}$$

for α, β, γ and δ. Since the vectors $\mathbf{p}$, $\mathbf{a}$ and $\mathbf{a}'$ are linearly independent, this equation is equivalent to the following system of equations

$$\begin{aligned} \alpha + \beta &= \gamma + \delta \\ \alpha + \beta\lambda &= \delta\lambda \\ \alpha + \beta\mu &= \gamma. \end{aligned}$$

Eliminating γ and δ using the last two equations, the first equation gives

$$\begin{aligned} \alpha+\beta &= (\alpha+\beta\mu)+(\alpha\lambda^{-1}+\beta), \\ \alpha\lambda^{-1} &= \beta\mu, \\ \alpha &= \beta\mu\lambda. \end{aligned}$$

This equation is satisfied by $\beta = 1$ and $\alpha = \mu\lambda$, thus we obtain

$$B"C" \cap B'C = [\mu\lambda(\mathbf{p}+\mathbf{a}+\mathbf{a}') + (\mathbf{p}+\lambda\mathbf{a}+\mu\mathbf{a}')].$$

Changing the role of ABC and $A'B'C'$ respectively, we get

$$B"C" \cap BC' = [\lambda\mu(\mathbf{p}+\mathbf{a}+\mathbf{a}') + (\mathbf{p}+\lambda\mathbf{a}+\mu\mathbf{a}')].$$

Introducing the vectors $\mathbf{x} = \mathbf{p}+\mathbf{a}+\mathbf{a}'$ and $\mathbf{y} = \mathbf{p}+\lambda\mathbf{a}+\mu\mathbf{a}'$, we can simply write

$$B"C" \cap B'C = [\mu\lambda\mathbf{x}+\mathbf{y}] \qquad B"C" \cap BC' = [\lambda\mu\mathbf{x}+\mathbf{y}].$$

It is clear that Pappus theorem is true for the triads ABC $A'B'C'$ if and only if $B"C" \cap B'C = B"C" \cap BC'$. These points coincide if and only if their representatives are proportional. Since $\mathbf{x}$ and $\mathbf{y}$ are not proportional ($B" \neq C"$), the vectors $\mu\lambda\mathbf{x}+y$ and $\lambda\mu\mathbf{x}+\mathbf{y}$ are proportional if and only if they are equal. However, $\mu\lambda\mathbf{x}+y = \lambda\mu\mathbf{x}+\mathbf{y} \iff \mu\lambda = \lambda\mu$.

If the division ring of scalars is a field, λ and μ commute, thus Pappus theorem holds.

On the other hand, if $\mathbb{F}$ is not commutative, and say $\mu\lambda \neq \lambda\mu$, then take three linearly independent vectors $\mathbf{p}$, $\mathbf{a}$ and $\mathbf{a}'$ from the underlying vector space and consider the points $P = [\mathbf{p}]$, $A = [\mathbf{a}]$, $B = [\mathbf{p}+\mathbf{a}]$, $C = [\mathbf{p}+\lambda\mathbf{a}]$, $A' = [\mathbf{a}']$, $B' = [\mathbf{p}+\mathbf{a}']$, $C' = [\mathbf{p}+\mu\mathbf{a}']$. As it follows from the above computation, these points yield a counterexample for Pappus' theorem. $\square$

PROJECTIVE GEOMETRY

3. EXAMPLES

3.1. Projective spaces over finite fields. The simplest possible field is the field $\mathbb{F}_2 = \{0,1\}$ of two elements. Operations are the modulo 2 addition and multiplication. The projective plane over $\mathbb{F}_2$ is called the *Fano plane*. Let us describe its structure.

The linear space $\mathbb{F}_2^3$ has 8 elements, 7 non-zero elements. Since the only non-zero scalar is 1, two non-zero vectors are proportional if and only if they are equal, hence the projective plane over $\mathbb{F}_2$ has also 7 points:

$$\begin{gathered} A = (1:0:0), B = (0:1:0),\, C = (0:0:1), \\ A' = (0:1:1), B' = (1:0:1),\, C' = (1:1:0), \\ S = (1:1:1). \end{gathered}$$

If $\mathbf{a}$ and $\mathbf{b}$ are two non-zero vectors, then there are 3 non-zero linear combinations of $\mathbf{a}$ and $\mathbf{b}$: $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{a}+\mathbf{b}$, thus there are 3 points on every straight line. There are exactly 7 straight lines:

$$A'BC,\ AB'C,\ ABC',\ ASA',\ BSB',\ CSC',\ A'B'C'$$

There are two convenient ways to describe the system of straight lines.

- If ABC are represented by the vertices of a regular triangle, $A'B'C'$ are the mid-points of the sides BC, AC, AB respectively and S is the center of the triangle, then 3 points form a straight line in the Fano plane if and only if they are on a Euclidean straight line or on the inscribed circle of the triangle ABC (see Figure 1).
- If $AB'C'CA'SB$ are represented by consecutive vertices of a regular heptagon, then straight lines of the Fano plane are rotations of the triangle $AB'C$ around the center of the heptagon (see Figure 2).

The latter cyclic representation of the Fano plane can be generalized for projective planes over an arbitrary finite division ring. Consider now this general case.

If the division ring of scalars $\mathbb{F}$ has finite number of elements, say $\#\mathbb{F} = q$, then an $(n+1)$-dimensional linear space V over $\mathbb{F}$ has $\#V = q^{n+1}$ points and $\#(V \setminus \{\mathbf{0}\}) = q^{n+1} - 1$. Since the number of non-zero scalars is $q-1$, each homogeneous vector is represented by $q-1$ non-zero vectors. Thus, the number of points of the n-dimensional projective space associated to V is

$$\#P(V) = \frac{q^{n+1}-1}{q-1} = 1 + q + \cdots + q^n.$$

Finite division rings have a well developed theory. In particular, there is a classification theorem that lists all finite division rings. First of all, we recall Wedderburn's theorem.

3.1.1. Theorem (Wedderburn). *Every finite division ring is a field.*

Thus, commutativity of multiplication is a consequence of the division ring axioms if $\mathbb{F}$ is finite.

Denoting by 1 the unit element of an arbitrary field $\mathbb{F}$, we set

$$n \cdot 1 = \operatorname{sgn} n \underbrace{(1 + \cdots + 1)}_{|n| \text{ times}},$$

where n is an arbitrary integer and

$$I = \{n \in \mathbb{Z} \mid n \cdot 1 = 0\}.$$

It is not difficult to see that there exists an integer $p \geq 0$ such that I is the set of all integer multiples of p, and p is prime or equals 0.

3.1.2. Definition. p is called *the characteristic of the field* $\mathbb{F}$ and denoted by $\operatorname{char} \mathbb{F}$.

For a prime p, denote by $\mathbb{F}_p$ the field of modulo p residuum classes of integers. If $\mathbb{F}$ is a field with $\operatorname{char} \mathbb{F} = p$, then the subfield of $\mathbb{F}$ generated by 1 consists of all elements of the form $n \cdot 1$, $n \in \mathbb{Z}$ and isomorphic to $\mathbb{F}_p$. If $\operatorname{char} \mathbb{F} = 0$, then the smallest subfield that contains 1 is isomorphic to the field $\mathbb{Q}$ of rational numbers. In particular, every finite field has prime characteristic.

The following simple observation has important consequences.

3.1.3. Proposition. *If $\mathbb{F}$ is a subfield of $\tilde{\mathbb{F}}$, then $\tilde{\mathbb{F}}$ is a linear space over $\mathbb{F}$.*

3.1.4. Corollary. *If $\mathbb{F}$ is a subfield of the finite field $\tilde{\mathbb{F}}$, and the dimension of $\tilde{\mathbb{F}}$ as a linear space over $\mathbb{F}$ is n, then*

$$\#\tilde{\mathbb{F}} = (\#\mathbb{F})^n.$$

3.1.5. Corollary. *The number of elements of a finite field of characteristic p is always a power of p. If a field of p^r elements is isomorphic to a subfield of a field of p^s elements, then $p^s = (p^r)^n = p^{r \cdot n}$ for some integer n, hence r divides s.*

Now we formulate the classification theorem of finite fields.

3.1.6. Theorem. *For every power p^r of a prime p, there exists a unique (up to isomorphism) finite field $\mathbb{F}_{p^r}$ with p^r elements. $\mathbb{F}_{p^r}$ is isomorphic to a subfield of $\mathbb{F}_{p^s}$ if and only if r divides s. In this case, $\mathbb{F}_{p^s}$ contains exactly one subfield isomorphic to $\mathbb{F}_{p^r}$.*

Fields appeared in mathematics first in the work of Galois, who proved with the help of the algebraic theory of subfields of $\mathbb{C}$, that the roots of a polynomial of degree ≥ 5 can not be expressed in terms of the coefficients using only the operations $+, -, \cdot, /$ and $\sqrt[n]{\dots}$. In his honour, finite fields are called *Galois fields* as well and denoted also by $GF(p^r)$; projective planes and spaces over a Galois field are called *Galois planes and spaces*.

Let us explicitly describe the field $\mathbb{F}_q$, $q = p^s$ and its subfields. Consider the ring of polynomials $\mathbb{F}_p[x]$ of one variable x with coefficients in $\mathbb{F}_p$. The polynomial $x^{q-1} - 1$ is not irreducible. Let

$$x^{q-1} - 1 = f_1(x) \cdot \dots \cdot f_N(x)$$

be its factorization into the product of polynomials irreducible in $\mathbb{F}_p[x]$. (Finding this factorization requires an algorithm. At this level, we do not want to discuss different factorization algorithms, it is enough for us that such algorithms exist.) One can prove that the irreducible factors are different, the degree of f_i is at most s and some of the irreducible factors have degree s. Assume that the irreducible factor f_1 has degree s.

Call two polynomials in $\mathbb{F}_p[x]$ equivalent if their difference is divisible by f_1 in the ring $\mathbb{F}_p[x]$. Clearly, every equivalence class is represented by a unique polynomial of degree less then s. The number of polynomials of degree less then s is $q = p^s$, thus we have exactly q equivalence classes. Since products and sums of equivalent polynomials are equivalent, the product and sum of equivalence classes are well defined. The set of equivalence classes together with this addition and multiplication is a field of q elements.

If k divides s, then the unique subfield of $\mathbb{F}_{p^s}$ consisting of p^k elements is formed by the roots of the polynomial $x^{p^k} - x$.

The construction of cyclic representations of projective planes (and spaces) over finite fields is based on the following

3.1.7. Theorem. *The multiplicative group of a finite field $\mathbb{F}_q$ is cyclic.*

Proof. Applying the fact that every finite commutative group is the product of some cyclic groups, one can see that if the multiplicative group were not cyclic, then we could find an integer $0 < m < q - 1$ such that $\xi^m = 1$ for every $\xi \in \mathbb{F}_q^*$. But then the polynomial $x^m - 1$ would have more roots than its degree.

3.1.8. Corollary. *If $\mathbb{F}_{p^k}$ is a subfield of $\mathbb{F}_{p^s}$, and the multiplicative group of $\mathbb{F}_{p^s}$ is generated by the element ξ, then the multiplicative group of $\mathbb{F}_{p^k}$ consists of powers of* $\xi^{\frac{p^s-1}{p^k-1}}$

Proof. The subgroup generated by $\xi^{\frac{p^s-1}{p^k-1}}$ is the only subgroup having p^k-1 elements.

Now let us describe the cyclic representation of the projective plane over the finite field $\mathbb{F}_q$. The key idea is that for a 3-dimensional vector space over $\mathbb{F}_q$ we may take the field $\mathbb{F}_{q^3}$. Non-zero vectors of this linear space form the multiplicative group of $\mathbb{F}_{q^3}$. Let ξ be a generator of $\mathbb{F}^*_{q^3}$. The vectors ξ^i and ξ^j represent the same point of the projective plane over $\mathbb{F}_q$ if and only if $\xi^{i-j} \in \mathbb{F}_q$, that is, $i \equiv j \pmod{q^2+q+1}$. Thus, points of the projective plane $\mathbb{F}_qP^2$ can be represented by the vectors

$$\xi, \xi^2, \ldots, \xi^{1+q+q^2}.$$

Let us describe these points by consecutive vertices $P_1, P_2, \ldots, P_{q^2+q+1}$ of a regular q^2+q+1-gon in the Euclidean plane. Multiplication by ξ is an $\mathbb{F}_q$-linear transformation of $\mathbb{F}_{q^3}$, hence it yields a transformation of the projective plane which maps a straight line to a straight line. The action of this transformation on the vertices $P_i = [\xi^i]$ is just a rotation about the center of the polygon by angle $\alpha = 2\pi/(q^2+q+1)$.

3.1.9. Proposition. *Let $e = \{P_{i_1}, P_{i_2}, \ldots, P_{i_{q+1}}\}$ be the points of a straight line in $\mathbb{F}_qP^2$. Then rotating the $(q+1)$-qon e about the center by integer multiples of the angle α we obtain q^2+q+1 different $(q+1)$-gons, each of which forms a straight line in $\mathbb{F}_qP^2$.*

Proof. We already know that each rotation of e forms a straight line in $\mathbb{F}_qP^2$, so we only have to show that the rotations are different. Set

$$I = \{k \in \mathbb{Z} \mid \text{ rotation of } e \text{ by angle } k\alpha \text{ equals e}\}.$$

Clearly, I consists of integer multiples of a natural number d, and we have to show $d = q^2+q+1$. Since $q^2+q+1 \in I$, d divides q^2+q+1. On the other hand, the vertex set of e can be decomposed into the disjoint union of some regular $(q^2+q+1)/d$-gons. This yields that $(q^2+q+1)/d$ divides $q+1$. However, q^2+q+1 and $q+1$ are relatively primes, thus their common divisor $(q^2+q+1)/d$ must be equal to 1.

An easy enumeration shows that the number of straight lines in $\mathbb{F}_qP^2$ is exactly q^2+q+1. Therefore, we obtain any straight line as a rotation of e. Since any two points in the plane are contained in a unique straight line, points of e must have the following property:

(*) *Each of the distances $d(P_i, P_j)$ occurs exactly once among the distances between the points of e.*

Similar cyclic representation can be given for higher dimensional projective spaces as well. The only difference is that k-dimensional subspaces are obtained generally as rotations of more than one subpoligon of the regular $(q^n+q^{n-1}+\cdots+1)$-gon.

We finish this section by formulating an unsolved problem.

Question. *Assume that among the vertices $P_1, \dots, P_{q^2+q+1}$ of a regular q^2+q+1-gon we can find a $(q+1)$-gon e having property (*). Does it follow that q is a power of a prime? Can we obtain all $(q+1)$-gons having property (*) as a straight line in a cyclic representation of $\mathbb{F}_q P^2$?*

3.2. Complex Projective Spaces and the Hopf Fibration.

According to the general scheme, the n-dimensional complex projective space $\mathbb{C}P^n$ is the set of complex 1-dimensional complex linear subspaces of the $(n+1)$-dimensional complex linear space $\mathbb{C}^{n+1}$. Since $\mathbb{R}$ is a subfield of $\mathbb{C}$ and $\mathbb{C}$ is a 2-dimensional linear space over $\mathbb{R}$, every k-dimensional complex linear subspace is automatically a linear space over $\mathbb{R}$ of dimension $2k$. Thus, $\mathbb{C}^{n+1} \cong \mathbb{R}^{2n+2}$ as real vector spaces and each 1-dimensional complex linear subspace is a real 2-dimensional plane passing through the origin. However, not every 2-dimensional real subspace of $\mathbb{R}^{2n+2}$ corresponds to a complex subspace. The real planes defined by complex 1-dimensional subspaces have the remarkable property that the intersection of any two planes in the family contains only the origin. Consider now the unit sphere

$$S^{2n+1} = \{\mathbf{x} \in \mathbb{R}^{2n+2} \mid \|\mathbf{x}\| = 1\}$$

and the traces of complex 1-dimensional subspaces of $\mathbb{C}^{n+1}$ on it. What we get is a family of disjoint great circles that cover the whole sphere. Each circle corresponds to a complex 1-dimensional subspace, i.e. a point of $\mathbb{C}P^n$. This correspondence defines a mapping $\pi\colon S^{2n+1} \to \mathbb{C}P^n$. Thus, we may think of the sphere S^{2n+1} as a *bundle* of circles and we may call these circles the *fibres* of the bundle. The space of the fibres i.e. the space obtained by shrinking each fibre to a point is the complex projective space $\mathbb{C}P^n$ the factorization map is π. Splitting S^{n+1} into circles in the above way is called the *Hopf fibration*.

The precise definition of fibre bundles is given in advanced algebraic topology or differential geometry courses. Hopf fibration is a topologically non-trivial fibre bundle and it finds applications in algebraic topology. To show the beauty of this construction, we shall draw a picture of the Hopf fibration for the lowest interesting dimension $n = 1$. For S^3 is lying in $\mathbb{R}^4$ and it is not easy to visualize a four dimensional figure, our plan is to draw the image of the Hopf fibration under a stereographic projection $S^3 \to \mathbb{R}^3$. So let us recall the definition and basic properties of the stereographic projection.

3.3. Digression: Stereographic Projection and Inversion.

Let S^n be the unit sphere in $\mathbb{R}^{n+1}$ and let us identify $\mathbb{R}^n$ with the hyperplane

$$\{(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_{n+1} = 0\}$$

by the embedding

$$(x_1, \dots, x_n) \mapsto (x_1, \dots, x_n, 0).$$

We shall write elements of $\mathbb{R}^{n+1}$ in the form $(\mathbf{x}, t)$, where $\mathbf{x}$ is the n-dimensional vector formed by the first n coordinates, t is the last coordinate.

3.3.1. Definition. Let $N \in S^n$ be the North Pole of the sphere, i.e. the point represented by the vector $(\mathbf{0}, 1)$. We define the *stereographic projection*

$$\Phi\colon S^n \setminus \{N\} \to \mathbb{R}^n$$

by the rule that for $P \in S^n \setminus \{N\}$, $\Phi(P)$ is the intersection point of the straight line NP and the hyperplane $\mathbb{R}^n$.

It is not difficult to find an explicit formula for the stereographic projection. For $P = (\mathbf{x}, t) \in S^n \setminus \{N\}$, the straight line NP contains points of the form

$$\alpha(\mathbf{x}, t) + (1 - \alpha)(\mathbf{0}, 1) = (\alpha\mathbf{x}, 1 + \alpha(t - 1)),$$

where α runs over all real numbers. The last coordinate of this point is 0 if and only if $\alpha = 1/(1 - t)$, consequently,

$$\Phi\colon (\mathbf{x}, t) \mapsto \frac{1}{1-t}\mathbf{x}.$$

Stereographic projection is closely related to another transformation, the inversion.

3.3.2. Definition. Let S be a sphere of radius r centered at O in $\mathbb{R}^{n+1}$. The *inversion* with respect to S is a transformation $\Psi\colon \mathbb{R}^{n+1} \setminus \{O\} \to \mathbb{R}^{n+1} \setminus \{O\}$ defined by the rule that the image of a point P is the unique point P' for which the following two conditions hold

- P' is on the halfline starting from O and passing through P;
- $OP \cdot OP' = r^2$.

Inversion is an involutive transformation, that is, $\Psi \circ \Psi$ is the identity.

3.3.3. Proposition. *Let S be the sphere of radius $\sqrt{2}$ centered at the North Pole of the unit sphere S^n, Ψ the inversion with respect to S. Then the stereographic projection coincides with the restriction of Ψ to $S^n \setminus \{N\}$.*

Proof. Let $P = (\mathbf{x}, t) \in S^n \setminus \{N\}$ be a point, P' its stereographic projection. N, P and P' are collinear by the definition of Φ. Thus it suffices to show that the dot product of the vectors $\overrightarrow{NP} = (\mathbf{x}, t - 1)$ and $\overrightarrow{NP'} = (\mathbf{x}/(1 - t), -1)$ equals 2. However, using the fact that $\|\mathbf{x}\|^2 + t^2 = 1$ we obtain

$$\overrightarrow{NP}\overrightarrow{NP'} = \frac{\|\mathbf{x}\|^2}{1-t} + (1 - t) = \frac{1-t^2}{1-t} + (1 - t) = 2.$$

The following theorems formulate two important properties of inversions.

3.3.4. Theorem. *The image of a sphere or hyperplane A under an inversion is a sphere if A does not contain the center of the inversion and it is a hyperplane otherwise.*

Proof. Assume that a subset A of $\mathbb{R}^{n+1}$ is defined by the equation

$$A = \{\mathbf{x} \in \mathbb{R}^{n+1} \colon F(\mathbf{x}) = 0\},$$

where $F\colon \mathbb{R}^{n+1} \to \mathbb{R}$ is an arbitrary function, then a point $\mathbf{x}$ belongs to the image of A under the inversion Ψ if and only if $\Psi(\mathbf{x})$ belongs to A, thus the equation of the image is

$$\Psi(A) = \{\mathbf{x} \in \mathbb{R}^{n+1} \colon F(\Psi(\mathbf{x})) = 0\}.$$

Choosing the origin of the coordinate system at the center of the sphere of the inversion we have

$$\Psi(\mathbf{x}) = \frac{r^2}{\|\mathbf{x}\|^2}\mathbf{x}.$$

Assume that A is a sphere centered at the point $\mathbf{p}$ and having radius R. Then the equation of A is

$$A = \{\mathbf{x} \in \mathbb{R}^{n+1} \colon \|\mathbf{x} - \mathbf{p}\|^2 - R^2 = 0\}.$$

Therefore, the Ψ-image of A has equation

$$\left\| \frac{r^2}{\|\mathbf{x}\|^2}\mathbf{x} - \mathbf{p} \right\|^2 - R^2 = 0.$$

Rearranging this equation we obtain

$$\frac{r^4}{\|\mathbf{x}\|^4}\|\mathbf{x}\|^2 - 2\frac{r^2}{\|\mathbf{x}\|^2}\langle \mathbf{x}, \mathbf{p}\rangle + \|\mathbf{p}\|^2 - R^2 = 0,$$

or equivalently,

$$r^4 - 2r^2\langle \mathbf{x}, \mathbf{p}\rangle + (\|\mathbf{p}\|^2 - R^2)\|x\|^2 = 0.$$

If $\|\mathbf{p}\|^2 = R^2$, i.e. A contains the origin, then this reduces to an inhomogeneous linear equation

$$\langle \mathbf{x}, 2\mathbf{p}\rangle = r^2.$$

Therefore, $\Psi(A)$ is a hyperplane perpendicular to $\mathbf{p}$ passing through $\Psi(2\mathbf{p})$.

On the other hand, if $\|\mathbf{p}\|^2 \neq R^2$, then the equation of $\Psi(A)$ is equivalent to

$$\left\| \mathbf{x} - \frac{r^2}{\|\mathbf{p}\|^2 - R^2}\mathbf{p} \right\|^2 = \frac{r^4 R^2}{(\|\mathbf{p}\|^2 - R^2)^2}.$$

Consequently, $\Psi(A)$ is a sphere centered at $\dfrac{r^2}{\|\mathbf{p}\|^2 - R^2}\mathbf{p}$ with radius $\dfrac{r^2 R}{|\|\mathbf{p}\|^2 - R^2|}$.

Now assume that A is a hyperplane. If $\mathbf{n}$ is a unit normal vector of A, then the equation of A has the form

$$\langle \mathbf{x}, \mathbf{n}\rangle = c,$$

where c is a constant. The equation of $\Psi(A)$ is

$$\langle \frac{r^2}{\|\mathbf{x}\|^2}\mathbf{x}, \mathbf{n}\rangle = c.$$

If $c = 0$, i.e. A goes through the center of the inversion, then this equation is equivalent to $\langle \mathbf{x}, \mathbf{n}\rangle = 0$, hence $\Psi(A) = A$ is a hyperplane. If, however, $c \neq 0$, then the equation of $\Psi(A)$ can be transformed to the form

$$\left\| \mathbf{x} - \frac{r^2}{2c}\mathbf{n} \right\|^2 = \left(\frac{r^2}{2c}\right)^2,$$

which is obviously the equation of a sphere passing through the origin.

3.3.5. Corollary. *Inversion maps k-dimensional spheres and planes to k-dimensional spheres or planes.*

Proof. Lower dimensional spheres and planes can be obtained as the intersection of some spheres and hyperplanes.

3.3.6. Theorem. *Inversion preserves angles in the sense that if two curves intersect one another at a certain angle, then their images intersect each other at the same angle. We define the angle at which two curves intersect to be the angle between the tangent lines of the curves.*

Proof. Assume the two curves intersect one another at P. Let e_1 and e_2 be the tangent lines of the curves at P. The inversion takes e_1 and e_2 into two circles k_1 and k_2 respectively. The circles k_1 and k_2 intersect one another at two points, at the inversion P' of P and at the center O of the inversion. Since k_1 is in the plane spanned by O and e_1, and the system O, e_1 is symmetric to the straight line through O perpendicular to e_1, k_1 is also symmetric to this line and consequently, the tangent of k_1 at O is parallel to e_1. Similarly, the tangent of k_2 at O is parallel to e_2. After these preliminary observations the theorem can be proved as follows.

The angle between the original curves is equal to the angle between e_1 and e_2. Since the angle between two straight lines is unchanged by parallel translation, the angle between e_1 and e_2 is equal to the intersection angle of k_1 and k_2 at O. Reflection in the bisector hyperplane of the segment OP' leaves the circles k_1 and k_2 invariant, and exchanges O and P', thus the intersection angle of k_1 and k_2 is the same at O and P'. Finally the circles k_1 and k_2 are tangent to the inversion of the original curves at P', so the inversions of the original curves meet each other at same angle at P' as the circles k_1 and k_2 do.

3.3.7. Corollary.

- *Stereographic projection maps a k-dimensional sphere $A \subset S^n$ onto a k-dimensional plane or sphere depending on wether A contains the North Pole or not.*
- *Stereographic projection preserves angles.*

3.4. The Stereographic Image of the Hopf Fibration.

Our goal now is to understand the structure of Hopf fibration by drawing its image under the stereographic projection. Let

$$S^3 = \{(z_1, z_2) \in \mathbb{C}^2 \colon |z_1|^2 + |z_2|^2 = 1\}$$

be the unit sphere,

$$S^1 = \{\epsilon \in \mathbb{C} \colon |\epsilon| = 1\}$$

be the circle of unit complex numbers. If $(z_1, z_2) \in S^3$, then there exists a unique real number $0 \le a \le \pi/2$ such that

$$|z_1| = \cos a \qquad |z_2| = \sin a.$$

Let us denote by T_a the subset of S^3 belonging to the same a

$$T_a = \{(z_1, z_2) \in \mathbb{C}^2 \colon |z_1| = \cos a,\ |z_2| = \sin a\}.$$

To get all points of T_a, we have to let z_1 and z_2 run over circles of radius $\cos a$ and $\sin a$ respectively, independently of one another. If $0 < a < \pi/2$, then these are real circles, thus T_a is homeomorhic to a direct product of two circles, i.e., T_a is a topological torus. If however $a=0$ or $a=\pi/2$, then one of the circles is schrunk to a point, so T_a becomes a circle.

Circles of the Hopf fibration, or shortly Hopf circles can be given as

$$\{(\epsilon z_1, \epsilon z_2) \colon \epsilon \in S^1\},$$

where (z_1, z_2) is an arbitrary point on the circle. It is clear from this that if a Hopf circle intersects the torus T_a, then it is contained in it. As a corollary, we obtain that T_a is a disjoint union of Hopf circles. By this observation, we may split the description of the Hopf fibration into two parts. First we describe the stereographic images of the tori T_a and then we study how a typical torus T_a is covered by Hopf circles.

With the usual identification of $\mathbb{C}^2$ and $\mathbb{R}^4$, every point $(z_1, z_2) \in T_a \subset S^3$ can be represented as

$$(\cos a\, e^{iu}, \sin a\, e^{iv}) = (\cos a\, \cos u, \cos a\, \sin u, \sin a\, \cos v, \sin a\, \sin v),$$

where u and v are the arguments of the complex numbers z_1 and z_2 respectively. The stereographic projection takes this point to

$$P(a,u,v) := \frac{1}{1-\sin a\, \sin v}(\cos a\, \cos u, \cos a\, \sin u, \sin a\, \cos v).$$

Beside Hopf circles, T_a contains two other families of circles. Namely, circles of the form

$$\{(\epsilon z_1, z_2) \colon \epsilon \in S^1\}, \text{ and } \{(z_1, \epsilon z_2) \colon \epsilon \in S^1\}.$$

A circle in the first family can be parameterized by letting u run and keeping a and v constant, to get a parameterization of a member of the second family, we should let v run, while a and u are fixed. What is the stereographic image of these circles? Looking at the explicit form of $P(a,u,v)$ we see easily that if u varies while a and v are fixed, then $P(a,u,v)$ moves along a circle whose plane is parallel to the xy-plane at height $\frac{\sin a\, \cos v}{1-\sin a\, \sin v}$ above it, the center of the circle is on the z-axis, the radius of the circle is $\frac{\cos a}{1-\sin a\, \sin v}$. From this we may conclude that the stereographic projection of T_a is a surface of revolution about the z-axis.

We can get a generator curve of this surface of revolution setting $u=0$ and letting v run. Then we obtain the following points in the xz-plane

$$P(a,0,v) := \frac{1}{1-\sin a\, \sin v}(\cos a, 0, \sin a\, \cos v).$$

3.4.1. Proposition.

- *If P_1 and P_2 are two distinct points in the Euclidean plane, $\lambda \geq 0$ is a fixed real number, then the locus of those points Q on the plane for which*

$$\frac{P_1Q}{P_2Q} = \lambda$$

is a circle if $\lambda \neq 1$ and a straight line (the bisector of the segment P_1P_2) if $\lambda = 1$. The circles obtained this way are the Apollonius circles *of the points P_1 and P_2.*

- *The points $P(a,0,v)$ run over the Apollonius circle of the points $(1,0,0)$ and $(-1,0,0)$ in the xz-plane with parameter $\lambda = \operatorname{tg} a/2$.*

Proof. To prove the statement about the Apollonius circles, let us choose a Cartesian coordinate system in the plane such that the points P_1 and P_2 have coordinates $(a,0)$ and $(-a,0)$ respectively. Then the equation of the locus is

$$(x-a)^2 + y^2 = \lambda^2((x+a)^2 + y^2).$$

If $\lambda = 1$, then this equation reduces to $x = 0$, thus the locus is the y-axis, the bisector of P_1P_2.

If, however, $\lambda \neq 1$, we obtain the equation

$$\left(x - \frac{1+\lambda^2}{1-\lambda^2}a\right)^2 + y^2 = \frac{4}{(\lambda - 1/\lambda)^2}a^2,$$

which is the equation of a circle.

The second part of the statement follows from the following computation

$$\begin{aligned}
&\frac{\left(\frac{\cos a}{1-\sin a \sin v} - 1\right)^2 + \left(\frac{\sin a \cos v}{1-\sin a \sin v}\right)^2}{\left(\frac{\cos a}{1-\sin a \sin v} + 1\right)^2 + \left(\frac{\sin a \cos v}{1-\sin a \sin v}\right)^2} \\
&= \frac{(\cos a + \sin a \sin v - 1)^2 + (\sin a \cos v)^2}{(\cos a - \sin a \sin v + 1)^2 + (\sin a \cos v)^2} \\
&= \frac{2 + 2\sin a \sin v \cos a - 2\cos a - 2\sin a \sin v}{2 - 2\sin a \sin v \cos a + 2\cos a - 2\sin a \sin v} = \frac{(1-\sin a \sin v)(1-\cos a)}{(1-\sin a \sin v)(1+\cos a)} \\
&= \frac{1-\cos a}{1+\cos a} = \operatorname{tg}^2 \frac{a}{2}.
\end{aligned}$$

3.4.2. Corollary. *The stereographic projection of the torus T_a is a real Euclidean torus, obtained by rotating about the z-axis an Apollonius circle of the pair $(1,0,0)$, $(-1,0,0)$ taken in the xz-plane.*

Having understood the structure of the tori T_a let us see how the Hopf circles fill one of the tori. A Euclidean torus being a surface of revolution generated by a circle contains obviously two families of circles. However, the stereographic projections of the Hopf circles yield a third family of Euclidean circles on the torus. (The existence of a third family of circles on a torus is a surprising by-product of our considerations.) This family is invariant under rotations about the z-axis, so it is enough to describe one such circle on each torus.

It is clear, that Hopf circles are characterized by the conditions that a and $u-v$ are constant. Setting $u - v = 0$, we obtain that the points

$$P(a,u,u) = \frac{1}{1-\sin a \sin u}(\cos a \cos u, \cos a \sin u, \sin a \cos u).$$

with a fixed run over the stereographic image of a Hopf circle on $\Phi(T_a)$. Comparing the x- and z-coordinates of $P(a,u,u)$ we see that this circle is contained in the plane

$$\Sigma := \{(x,y,z) \in \mathbb{R}^3 \mid z = \operatorname{tg} a\, x\}.$$

The plane Σ goes through the y-axis and the angle between Σ and the xy-plane is just a. Σ cuts the torus in two circles. One of them is the stereographic image C_1 of the Hopf circle, the other is the reflection C_2 of this circle in the xz-plane. C_1 and C_2 intersect one another at the points

$$P = (\cos a, 0, \sin a) \quad Q = (-\cos a, 0, -\sin a).$$

Thus, by simple differential geometrical arguments, Σ is tangent to $\Phi(T_a)$ at the points P and Q.

One can turn the above analytical description into visual pictures by loading the formulae into a computer. Looking at the pictures we can observe an important topological property of the Hopf fibration: any two Hopf circles are linked.

3.5. QUATERNIONS.

By Wedderburn's Theorem (3.1.1.), every finite division ring is a field. Thus, every division ring with non-commutative multiplication must be infinite. One of the most important examples of non-commutative division rings is the division ring of quaternions. The standard notation for this division ring is $\mathbb{H}$ in honour of Hamilton, who first introduced and applied them.

Quaternions can be obtained from complex numbers in a way similar to as complex numbers are obtained from real numbers. When we introduce complex numbers, we may follow two different approaches.

- We may say that a complex number is a formal linear combination $x+iy$, where $x, y \in \mathbb{R}$ are reals, i is just a symbol. In this case, we define the sum and product of complex numbers by the formulae

$$\begin{aligned}(x+iy)+(x'+iy') &= (x+x')+i(y+y')\\ (x+iy)(x'+iy') &= (xx'-yy')+i(xy'+x'y)\end{aligned}$$

- Another possible way is to define complex numbers as a subring of the ring of 2×2 real matrices having the form $\begin{pmatrix} x & y \\ -y & x \end{pmatrix}$.

The equivalence of the two definitions is given by the correspondence $x+iy \leftrightarrow \begin{pmatrix} x & y \\ -y & x \end{pmatrix}$.

We define quaternions using the second approach. The advantage of this approach is that the verification of algebraic identities (e.g. the ring axioms) can be reduced to algebraic identities on matrices.

Let $\mathbb{H}$ denote the set of 2×2 complex matrices of the form $\begin{pmatrix} \alpha & \beta \\ -\bar\beta & \bar\alpha \end{pmatrix}$, where α and β are arbitrary complex numbers. We shall call elements of $\mathbb{H}$ quaternions.

3.5.1. Proposition. *$\mathbb{H}$ forms a division ring with respect to the usual matrix operations.*

Proof. Computing the sum and product of two elements of $\mathbb{H}$ we see that they are also in $\mathbb{H}$, thus, $\mathbb{H}$ is a subring of the ring of matrices. Indeed,

$$\begin{pmatrix} \alpha & \beta \\ -\bar\beta & \bar\alpha \end{pmatrix} + \begin{pmatrix} \alpha' & \beta' \\ -\bar\beta' & \bar\alpha' \end{pmatrix} = \begin{pmatrix} \alpha+\alpha' & \beta+\beta' \\ -\overline{(\beta+\beta')} & \overline{(\alpha+\alpha')} \end{pmatrix}$$

$$\begin{pmatrix} \alpha & \beta \\ -\bar\beta & \bar\alpha \end{pmatrix} \begin{pmatrix} \alpha' & \beta' \\ -\bar\beta' & \bar\alpha' \end{pmatrix} = \begin{pmatrix} \alpha\alpha'-\beta\bar\beta' & \alpha\beta'+\beta\bar\alpha' \\ -\bar\alpha\bar\beta'-\bar\beta\alpha' & \bar\alpha\bar\alpha'-\bar\beta\beta' \end{pmatrix}$$

Since $\mathbb{H}$ contains the unit matrix, it is a ring with a unit element. Finally, if $q = \begin{pmatrix} \alpha & \beta \\ -\bar\beta & \bar\alpha \end{pmatrix} \neq 0$, then $\det q = |\alpha|^2 + |\beta|^2 \neq 0$, therefore q is an invertible matrix, and its inverse is

$$q^{-1} = \frac{1}{\det q}\begin{pmatrix} \bar\alpha & -\beta \\ \bar\beta & \alpha \end{pmatrix} \in \mathbb{H}.$$

Thus, $\mathbb{H}$ is a division ring. $\square$

Diagonal quaternions form a subring, which is isomorphic to the ring of complex numbers. Identifying the complex number α with the matrix $\begin{pmatrix} \alpha & 0 \\ 0 & \bar\alpha \end{pmatrix}$ we may assume that $\mathbb{H}$ contains complex and real numbers ($\mathbb{R} \subset \mathbb{C} \subset \mathbb{H}$).

$\mathbb{H}$ is not commutative, as it follows from the following more stronger proposition.

3.5.2. Proposition. *$q \in \mathbb{H}$ commutes with every quaternion if and only if $q \in \mathbb{R}$*

Proof. For $q = \begin{pmatrix} \alpha & \beta \\ -\bar\beta & \bar\alpha \end{pmatrix}$ and $q' = \begin{pmatrix} 0 & \gamma \\ -\bar\gamma & 0 \end{pmatrix}$ we have

$$qq' - q'q = \begin{pmatrix} \bar\beta\gamma - \beta\bar\gamma & (\alpha-\bar\alpha)\gamma \\ (\alpha-\bar\alpha)\bar\gamma & \beta\bar\gamma - \bar\beta\gamma \end{pmatrix}.$$

This implies that if q commutes with every quaternion, then $\bar\beta\gamma$ must be real and $(\alpha-\bar\alpha)\gamma$ must vanish for every complex number γ. This yields $\beta = 0$ and $\alpha \in \mathbb{R}$, that is $q \in \mathbb{R}$.

On the other hand, real numbers obviously commute with every quaternion, since they correspond to scalar multiples of the identity matrix. $\square$

$\mathbb{H}$ is a 4-dimensional linear space over the real numbers. The matrices

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad i = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad k = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

form a basis of this vector space. Thus, every quaternion can be written uniquely as a linear combination

$$q = x + yi + zj + wk,$$

where x, y, z, w are real numbers.

The notation for the first two matrices in the above basis and the replacement of $x1$ with x in the linear combination are in keeping with the agreement on the identification of complex numbers with diagonal quaternions.

3.5.3. Definition. The *real part* of a quaternion

$$q = x + yi + zj + wk = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$$

is the real number

$$\text{Re } q = x = \text{Re } \alpha = \frac{1}{2}\text{tr } q.$$

The *imaginary part* of q is

$$\text{Im } q = yi + zj + wk$$

We call q *pure imaginary*, if $\text{Re } q = 0$.

The *conjugate* of q is the quaternion

$$\bar{q} = x - yi - zj - wk = \begin{pmatrix} \bar{\alpha} & -\beta \\ \bar{\beta} & \alpha \end{pmatrix} = \det q \cdot q^{-1}$$

The linear space $\mathbb{R}^4 \cong \mathbb{H}$ is equipped with the usual dot product

$$\langle q, q' \rangle = \text{Re } (\bar{q}q') = xx' + yy' + zz' + ww',$$

where $q = x + yi + zj + wk$ and $q' = x' + y'i + z'j + w'k$. The length of q can be expressed from the equations

$$|q|^2 = \langle q, q \rangle = x^2 + y^2 + z^2 + w^2 = |\alpha|^2 + |\beta|^2 = \det q = q\bar{q}.$$

3.5.4. Definition. A *Euclidean vector space* is a linear space over $\mathbb{R}$ equipped with a scalar product

$$\langle , \rangle : V \times V \to \mathbb{R}, \qquad (\mathbf{v}, \mathbf{w}) \mapsto \langle \mathbf{v}, \mathbf{w} \rangle$$

satisfying the axioms

$$\begin{aligned} \langle \alpha\mathbf{v} + \mathbf{v}', \mathbf{w} \rangle &= \alpha\langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{v}', \mathbf{w} \rangle, \\ \langle \mathbf{v}, \mathbf{w} \rangle &= \langle \mathbf{w}, \mathbf{v} \rangle, \\ \langle \mathbf{v}, \mathbf{v} \rangle &> 0 \text{ if } \mathbf{v} \neq \mathbf{0}. \end{aligned}$$

The standard example of an n-dimensional Euclidean vector space is $\mathbb{R}^n$ of $n \times 1$ matrices (coloumn vectors) with the usual dot product

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^T \mathbf{w},$$

where T means transposition. As a special case, quaternions form a 4-dimensional Euclidean vector space.

3.5.5. DEFINITION. Let V be a Euclidean vector space with scalar product $\langle , \rangle$. The group of (invertible) linear transformations preserving the scalar product is the *orthogonal group of V*

$$O(V) = \{A : V \to V \mid A \text{ is linear and } \langle A\mathbf{v}, A\mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{w}\rangle \quad \forall \mathbf{v}, \mathbf{w} \in V\}$$

O(n) will denote the orthogonal group of $\mathbb{R}^n$ with the standard dot product.

For the linear space $\mathbb{R}^n$ of coloumn vectors, linear transformations can be identified with $n \times n$ matrices acting on vectors by multiplication from the left. A matrix A represents an orthogonal transformation of $\mathbb{R}^n$ if and only if $\mathbf{v}^T\mathbf{w} = (A\mathbf{v})^T A\mathbf{w} = \mathbf{v}^T(A^T A)\mathbf{w}$ for all $\mathbf{v}$ and $\mathbf{w}$, i.e. $A^T A = I$. Since the equality $A^T A = I$ implies $1 = \det I = \det{}^2 A$, we obtain that the determinant of an orthogonal transformation is 1 or -1. Orthogonal transformations with determinant 1 form a subgroup of the orthogonal group, called the *special orthogonal group*.

3.5.6. DEFINITION. The *special orthogonal group of a Euclidean vector space* V is the group

$$SO(V) = \{A \in O(V) \mid \det A = 1\}.$$

We shall denote $SO(\mathbb{R}^n)$ simply by $SO(n)$.

The set S^3 of all unit quaternions form a group under quaternion multiplication. Our goal now is to show that this group is closely related to the groups $SO(3)$ and $SO(4)$. First we describe the structure of an orthogonal transformation in general.

3.5.7. LEMMA.

- *If A is a linear transformation of an odd dimensional linear space V over $\mathbb{R}$, then V contains a 1-dimensional A invariant subspace.*
- *If A is a linear transformation of a real linear space V, then V contains an at most 2 dimensional A-invariant subspace.*

PROOF. First we show that every linear transformation A of an odd dimensional real linear space V has at least one eigenvector, that is, one can find a non-zero vector $\mathbf{v} \in V$ and a number $\lambda \in \mathbb{R}$ such that $A\mathbf{v} = \lambda\mathbf{v}$. Real eigenvalues of A, i.e. those real numbers λ, for which one can find such a $\mathbf{v}$, are roots of the characteristic polynomial $p_A(\lambda) = \det(A - \lambda I)$. The degree of $p_A(\lambda) = -\lambda^{\dim V} \dots$ is odd, hence the limit of $p_A(\lambda)$ as λ tends to $\pm\infty$ is $\mp\infty$, therefore, by Bolzano's theorem p_A must have a real root. The eigenvector corresponding to the real eigenvalue spans an A invariant 1-dimensional subspace in V.

If $\dim V$ is even, the characteristic polynomial of A may not have real eigenvalues, however, it always has complex eigenvalues. Suppose $\alpha = x + iy \in \mathbb{C} \setminus \mathbb{R}$ is a comlex root of p_A. Then, corresponding to α, there is an eigenvector $\mathbf{v} + i\mathbf{w}$ of A in the complexification

$$\mathbb{C} \otimes V = V \oplus iV = \{\mathbf{v} + i\mathbf{w} | \mathbf{v}, \mathbf{w} \in V\}.$$

Taking the real and imaginary parts of $A(\mathbf{v} + i\mathbf{w}) = \alpha(\mathbf{v} + i\mathbf{w})$ we obtain

$$A\mathbf{v} = x\mathbf{v} - y\mathbf{w}, \quad A\mathbf{w} = y\mathbf{v} + x\mathbf{w},$$

showing that the subspace spanned by $\mathbf{v}$ and $\mathbf{w}$ is mapped by A into itself. $\square$

3.5.8. LEMMA. *If A is an orthogonal transformation of the Euclidean linear space $(V, \langle , \rangle)$, $W < V$ is an A invariant subspace, then the orthogonal complement $W^\perp = \{\mathbf{w}' \in V | \langle \mathbf{w}', \mathbf{w} \rangle = 0 \, \forall \mathbf{w} \in W\}$ of W is also A invariant.*

PROOF. $\mathbf{w}' \in W^\perp \iff \langle \mathbf{w}', \mathbf{w} \rangle = 0 \, \forall \mathbf{w} \in W \iff \langle A\mathbf{w}', A\mathbf{w} \rangle = 0 \, \forall \mathbf{w} \in W \iff \langle A\mathbf{w}', \mathbf{w} \rangle = 0 \, \forall \mathbf{w} \in W \iff A\mathbf{w}' \in W^\perp$. $\square$

3.5.9. THEOREM. *If $A \in O(V)$ is an orthogonal transformation of an n-dimensional Euclidean linear space $(V, \langle , \rangle)$, then V can be decomposed into the direct sum of A invariant orthogonal subspaces of dimension at most two in such a way that the restrictions of A onto the 2-dimensional factors are rotations while restrictions of A onto the 1-dimensional factors are multiplications by +1 or -1.*

3.5.10. COROLLARY. *Elements of the group $SO(3)$ are rotations about a straight line through the origin.*

PROOF. By Theorem 3.5.9, for every $A \in SO(3)$, $\mathbb{R}^3$ has an A invariant orthogonal decomposition (a) $\mathbb{R}^3 = V_1 \oplus V_2 \oplus V_3$ with $\dim V_i$=1 or (b) $\mathbb{R}^3 = V_1 \oplus V_2$ with $\dim V_1 = 2$, $\dim V_2 = 1$, where $A|_{V_1}$ is a rotation.

(a) In the first case, the number of those V_i's on which A acts by multiplication with -1 is even, that is, 0 or 2. If A acts identically on each V_i, then A is the identity, which can be thought of as a rotation about an arbitrary axis with angle 0. If, for example, $A|_{V_1}$ and $A|_{V_2}$ are reflections in the origin, $A|_{V_3}$ is the identity, then A is a half turn about V_3.

(b) Since the determinant of a rotation in a plane is 1, in the second case, $A|_{V_2}$ must be identical, therefore, A is a rotation about the 1-dimensional subspace V_2. $\square$

3.5.11. THEOREM. *For a non-zero quaternion q, conjugation $\rho_q : \mathbb{H} \to \mathbb{H}$, $\rho_q : x \mapsto qxq^{-1}$ is special orthogonal transformation. ρ_q fixes real numbers and maps the 3-dimensional space $\mathbb{R}^3$ of pure imaginary quaternions into itself. If q is a unit quaternion of the form $\cos\alpha + a\sin\alpha$, where a is a pure imaginary unit quaternion, then the restriction $\hat{\rho}_q$ of ρ_q onto the 3-dimensional subspace of pure imaginary quaternions is a rotation about the subspace spanned by a with angle 2α.*

PROOF. ρ_q is an orthogonal transformation since

$$\begin{aligned}\langle \rho_q(x), \rho_q(y) \rangle &= \mathrm{Re}\ (\overline{qxq^{-1}}qyq^{-1}) = \mathrm{Re}\ (\overline{q^{-1}}\bar{x}\bar{q}qyq^{-1}) \\ &= \mathrm{Re}\ (\frac{q}{|q|^2}\bar{x}|q|^2yq^{-1}) = \mathrm{Re}\ (q\bar{x}yq^{-1}) = \mathrm{Re}\ (\bar{x}y) = \langle x, y \rangle\end{aligned}$$

At the end of this computation we used the fact that the real part can be expressed as half the trace of the quaternion, therefore it is invariant under conjugation. For the same reason, ρ_q maps pure imaginary quaternions to pure imaginary ones. If $x \in \mathbb{R}$, then, since real numbers commute with every quaternion, we have

$$\rho_q(x) = qxq^{-1} = qq^{-1}x = x.$$

To show that ρ_q has determinant 1, it suffices to check that its restriction $\hat{\rho}_q$ is a rotation about an axis. As $\rho_q = \rho_{(q/|q|)}$, it is enough to deal with the case of unit quaternions.

Suppose that q is of the form $\cos\alpha + a\sin\alpha$, where a is a pure imaginary unit quaternion. Let b be a pure imaginary unit quaternion, orthogonal to a and set

$c = ab$. We claim that a, b, c form an orthonormal basis of the space of pure imaginary quaternions, satisfying

$$a^2 = b^2 = c^2 = -1, \quad ab = c,\ bc = a,\ ca = b,\ ba = -c,\ cb = -a,\ ac = -b.$$

Observe first, that the square of a pure imaginary quaternion is -1. Indeed, the equations $q = -\bar{q}$ and $q\bar{q} = 1$ imply $q^2 = -1$. This yields at once $a^2 = b^2 = -1$. c is a unit quaternion since $|c| = |a||b| = 1$ and it is also pure imaginary since $\mathrm{Re}\, c = \mathrm{Re}\, ab = -\mathrm{Re}\, \bar{a}b = \langle a, b\rangle = 0$. This yields $c^2 = -1$. Using these equations,

$ac = a(ab) = a^2 b = -b$;

$cb = (ab)b = ab^2 = -a$;

$ba = (-b)(-a) = (ac)(cb) = ac^2 b = -ab = -c$;

$bc = \overline{cb} = -\bar{a} = a, \quad ca = \overline{ac} = -\bar{b} = b$;

$\langle b, c\rangle = \mathrm{Re}\, \bar{b}c = -\mathrm{Re}\, a = 0, \quad \langle a, c\rangle = \mathrm{Re}\, \bar{a}c = \mathrm{Re}\, b = 0$;

as we claimed. The action of $\hat{\rho}_q$ is uniquely determined by its action on the basis vectors a, b, c. Using $q^{-1} = \bar{q} = \cos\alpha - a\sin\alpha$ and the above formulae we obtain

$$\begin{aligned}
\hat{\rho}_q(a) &= a \\
\hat{\rho}_q(b) &= (\cos^2\alpha - \sin^2\alpha)b + (2\sin\alpha\cos\alpha)c = \cos(2\alpha)b + \sin(2\alpha)c \\
\hat{\rho}_q(c) &= -(2\sin\alpha\cos\alpha)b + (\cos^2\alpha - \sin^2\alpha)c = -\sin(2\alpha)b + \cos(2\alpha)c
\end{aligned}$$

Showing that $\hat{\rho}_q$ is indeed a rotation with angle 2α about the axis spanned by a. □

3.5.12. Corollary. *The mappings $\mathbb{H}^* \to SO(3)$ and $S^3 \to SO(3)$ assigning to a non-zero quaternion q the transformation $\hat{\rho}_q$ are group homomorphisms onto the group $SO(3)$. Two elements of $\mathbb{H}^*$ are mapped to the same transformation if and only if they are proportional over the real numbers. In particular, $SO(3)$ is isomorphic to the factor groups $\mathbb{H}^*/\mathbb{R}^* \cong S^3/\{\pm 1\}$.*

3.5.13. Corollary. *$SO(3)$ as a topological space (with subspace topology obtained from the linear space topology on the linear space of all 3×3 matrices) is homeomorphic to the real projective space.*

This equivalence recovers a hidden group structure of the real projective space.

3.5.14. Theorem. *For a pair of unit quaternions (q_1, q_2), let ρ_{q_1,q_2} denote the transformation $\mathbb{H} \to \mathbb{H}$ defined by*

$$\rho_{q_1,q_2}(x) = q_1 x q_2^{-1}.$$

Then the assignment $(q_1, q_2) \mapsto \rho_{q_1,q_2}$ is a group homomorphism from $S^3 \times S^3$ onto $SO(4)$. The kernel of this homomorphism consists of two elements $(1, 1)$ and $(-1, -1)$.

Proof. ρ_{q_1,q_2} preserves length of quaternions, consequently, it preserves also scalar product of them, therefore it is an orthogonal transformation. To show that its determinant is 1, we use topological arguments. Since $S^3 \times S^3$ is path connected, we can find a continuous mapping $\gamma : [0, 1] \to S^3 \times S^3$ such that $\gamma(0) = (1, 1)$ and $\gamma(1) = (q_1, q_2)$. The function $f(t) = \det \rho_{\gamma(t)}$ is continuous and takes values in the set $\{\pm 1\}$ therefore it is constant. As a consequence, $\det \rho_{q_1,q_2} = f(1) = f(0) = 1$.

To show that the mapping we are considering is a mapping onto, let us take an arbitrary element A of $SO(4)$, and set $q = A(1)$. The transformation $\rho_{q^{-1},1} \circ A$

fixes real numbers, and, since it is a special orthogonal transformation, it acts on pure imaginary quaternions as a special othogonal transformation. By Corollary 3.5.12, this transformation is a conjugation by a unit quaternion q_2. Denoting q_1 the product qq_2, we $A = \rho_{q,1} \circ \rho_{q_2,q_2} = \rho_{q_1,q_2}$.

If ρ_{q_1,q_2} is the identity, then from $\rho_{q_1,q_2}(1) = 1$ we obtain $q_1 = q_2$, thus ρ_{q_1,q_2} is conjugation with q_1. By Corollary 3.5.12, conjugation with q_1 is the identity if and only if $q_1 = \pm 1$. $\square$

3.5.15. Corollary. *$SO(4)$ as a topological space is homeomorphic to the factor space obtained from the product $S^3 \times S^3$ of two spheres identifying the pairs $(\mathbf{v}, \mathbf{w})$, $(-\mathbf{v}, -\mathbf{w})$.*

PROJECTIVE GEOMETRY

Lecture Notes by Balázs Csikós

4. THE AXIOMATIC TREATMENT OF PROJECTIVE SPACES

4.1. The Incidence Axioms of an n-Dimensional Projective Space. In the previous units we introduced n-dimensional projective spaces, associated to a linear space over an arbitrary division ring, subspaces of projective spaces, and proved some basic properties of the subspaces. This way, we gave a generalization of the 3-dimensional real projective space. Now we go on and define projective spaces on an even more general level.

4.1.1. Definition. A system $(X, \mathcal{S}_{-1}, \mathcal{S}_0, \dots, \mathcal{S}_n)$ is called an *n-dimensional projective space*, if X is an arbitrary set, $\mathcal{S}_i$ is a family of subsets of X the elements of which are referred to as *i-dimensional subspaces of X* and the following axioms are satisfied.

- The only -1-dimensional subspace is the empty set, that is, $\mathcal{S}_{-1} = \{\emptyset\}$.
- 0-dimensional subspaces of X are the one point subsets of X, that is, $\mathcal{S}_0 = \{\{p\} \mid p \in X\}$.
- The only n-dimensional subspace of X is X, that is, $\mathcal{S}_n = \{X\}$.
- The families $\mathcal{S}_i$ and $\mathcal{S}_j$ are disjoint for $i \neq j$. A consequence of this axiom is that if W is a subspace, then $W \in \mathcal{S}_i$ for exactly one i, thus the dimension $\dim W = i$ of W is well defined.
- The intersection of subspaces is also a subspace.

This axiom gives rise to constructions such as the subspace spanned by an arbitrary subset of X and the join of subspaces. We define the *subspace $[A]$ spanned or generated by the subset $A \subset X$* as the intersection of all subspaces that contain A. $[A]$ is the smallest subspace that contains A. The *join of some subspaces* $W_1, W_2, \dots$ is the subspace generated by their union $W_1 + W_2 + \dots = [W_1 \cup W_2 \cup \dots]$.

- For any two subspaces W_1 and W_2, the following equality holds between the dimensions of the subspaces, their intersection and their join

$$\dim(W_1) + \dim(W_2) = \dim(W_1 \cap W_2) + \dim(W_1 + W_2).$$

- There exist $n+2$ points in general position, that is $n+2$ points such that no $n+1$ of them lie on a hyperplane ($= n-1$-dimensional subspace).

We prove some useful consequences of these axioms.

Typeset by $\mathcal{AMS}$-TeX

4.1.2. PROPOSITION. *If W_1 and W_2 are two subspaces such that $W_1 \subset W_2$, then $\dim W_1 \le \dim W_2$. If furthermore, $W_1 \ne W_2$ then $\dim W_1 < \dim W_2$.*

PROOF. Let us define a sequence of subspaces $W_1 = V_0 \subset V_1 \subset \cdots \subset V_l = W_2$. We start from $V_0 = W_1$. Suppose V_k is already defined. If $V_k = W_2$, we stop. If however V_k is a proper subset of W_2, then we choose a point $P \in W_2 \setminus V_k$ and set $V_{k+1} = V_k + \{P\}$. Since W_2 contains the subspaces V_k and $\{P\}$, it contains their join V_{k+1} as well. Application of the dimension formula gives

$$\dim V_{k+1} = \dim V_k + \dim\{P\} - \dim\emptyset = \dim V_k + 1.$$

Consequently,

$$\dim V_k = \dim W_1 + k.$$

Since the dimension of a subspace can not exceed the dimension of the total space, the above procedure must stop at a certain place, $V_l = W_2$ for some l. But in that case, $\dim W_2 = \dim V_l = \dim W_1 + l \ge \dim W_1$, where equality holds if and only if $l = 0$ and $W_2 = V_0 = W_1$. $\square$

4.1.3. LEMMA. *Let W be a subspace of a projective space, $\dim W \ge 0$, P a point not in W. Then the join of W and $\{P\}$ is the union of all straight lines of the form $PQ = \{P\} + \{Q\}$, where Q runs over W.*

PROOF. It is clear that for any $Q \in W$, $W + \{P\}$ contains both P and Q, thus it contains their join PQ. This yields

$$W + \{P\} \supset \bigcup_{Q \in W} PQ.$$

Now take an arbitrary point $R \in W + \{P\}$. If $R = P$, then R is covered by any straight line of the form PQ where $Q \in W$, and such lines exist by $\dim W \ge 0$. Assume that $R \ne P$. The straight line PR is contained in $W + \{P\}$ and consequently, $W + PR = W + \{P\}$. By the dimension formula,

$$\dim(W \cap PR) = \dim W + \dim PR - \dim(W + \{P\}) = 0,$$

that is, the straight line PR intersects W at a single point Q. But then R is covered by the straight line PQ. This proves

$$W + \{P\} \subset \bigcup_{Q \in W} PQ. \quad \square$$

4.1.4. PROPOSITION. *A subset W of an n-dimensional projective space is a subspace if and only if for any two different points P Q in W, W contains the straight line $PQ = \{P\} + \{Q\}$.*

PROOF. Since the join of two subspaces is the minimal subspace among those that contain the two subspaces, if W is a subspace, it contains the join of any of its two 0-dimensional subspaces $\{P\}$ and $\{Q\}$.

Now assume that W is a subset of X that contains the straight line PQ for any two different points P, Q in W. We define a sequence of subspaces of X contained in W. We set $V_0 = \emptyset$. Suppose that the subspace $V_k \subset W$ has already been defined.

If $V_k = W$, then we stop. If $V_k \neq W$, then we choose a point $P \in W \setminus V_k$ and set $V_{k+1} = V_k + \{P\}$. By the lemma above, we obtain that V_{k+1} is a union of straight lines passing through P and a point $Q \in V_k$, thus it must be a subset of W. The dimension formula yields that

$$\dim V_{k+1} = \dim V_k + \dim\{P\} - \dim \emptyset = \dim V_k + 1,$$

consequently, $\dim V_k = k$. For the dimension of subspaces is bounded by the dimension of the space, the sequence V_k must exhaust the set W, that is, $V_k = W$ for some k. But then $W = V_k$ is a subspace of X. $\square$

4.1.5. Corollary. *The family of straight lines determines the subspace structure completely. In other words, if* $(X, \mathcal{S}_{-1}, \mathcal{S}_0, \dots, \mathcal{S}_n)$ *and* $(X, \widetilde{\mathcal{S}}_{-1}, \widetilde{\mathcal{S}}_0, \dots, \widetilde{\mathcal{S}}_n)$ *are two* n*-dimensional projective space structures on the same set* X *and* $\mathcal{S}_1 = \widetilde{\mathcal{S}}_1$*, then* $\mathcal{S}_i = \widetilde{\mathcal{S}}_i$ *for any* $-1 \leq i \leq n$*.*

Proof. The previous proposition characterizes subspaces of a projective space in terms of straight lines. On the other hand, the family of all subspaces determines the dimension of each individual subspace by the following rule. A subspace W is k-dimensional if and only if there exists an increasing chain of subspaces $V_{-1} \subset V_0 \subset V_1 \subset \dots \subset V_n$ such that $V_{i-1} \neq V_i$ for $0 \leq i \leq n$ and $V_k = W$. $\square$

Let W be a k-dimensional subspace of an n-dimensional projective space $(X, \mathcal{S}_{-1}, \mathcal{S}_0, \dots, \mathcal{S}_n)$. For $-1 \leq i \leq k$, set

$$\mathcal{S}_i^W = \{A \in \mathcal{S}_i \mid A \subset W\}.$$

4.1.6. Proposition. *The system* $(W, \mathcal{S}_{-1}^W, \mathcal{S}_0^W, \dots \mathcal{S}_k^W)$ *is a* k*-dimensional projective space, that is, every subspace of a projective space inherits a projective space strucrure.*

Proof. The only non-trivial part of the statement is the existence of $k+2$ points of W in general position. Let us take $n+2$ points of X in general position, and call them $P_1, P_2, \dots, P_{n+2}$. Since joining a point to a subspace not containing it increases the dimension of the subspace by one, we can choose $n-k$ points out of $P_1, \dots, P_{n+2}$, say $P_1, \dots, P_{n-k}$ such that

$P_i \notin W + \{P_1\} + \dots + \{P_{i-1}\}$;

$\dim(W + \{P_1\} + \dots + \{P_i\}) = k + i \quad \forall 1 \leq i \leq n-k$.

We define the sets $V, V_1, \dots, V_{k+2}$ in the following way.

$V = \{P_1\} + \dots + \{P_{n-k}\}$;

$V_i = V + \{P_{n-k+i}\}$ for $1 \leq i \leq k+2$.

Since V is spanned by $n-k$ points, $\dim V \leq n-k-1$. On the the other hand, V and W span the whole space. Considering the dimension formula, this is possible only if $\dim V = n-k-1$ and $V \cap W = \emptyset$. Then $\dim V_i = n-k$ and

$$\dim(V_i \cap W) = \dim V_i + \dim W - \dim(V_i + W) = (n-k) + k - n = 0,$$

that is, V_i meets W at a single point Q_i. We claim that the points $Q_1, Q_2, \dots, Q_{k+2} \in W$ are in general position. Assume that this is not the case. Then $k+1$ of the points, say $Q_1, \dots, Q_{k+1}$, span a $\leq k-1$-dimensional subspace H. By the dimension formula,

$$\dim(V + H) = \dim V + \dim H - \dim \emptyset \leq (n-k-1) + (k-1) + 1 = n-1.$$

On the other hand,

$$\begin{aligned} V+H &= V+\{Q_1\}+\cdots+\{Q_{k+1}\} \\ &= (V+\{Q_1\})+\cdots+(V+\{Q_{k+1}\}) \\ &= (V+\{P_{n-k+1}\})+\cdots+(V+\{P_{n+1}\}) \\ &= V+\{P_{n-k+1}\}+\cdots+\{P_{n+1}\} = \\ &= \{P_1\}+\{P_2\}+\cdots+\{P_{n+1}\} = X. \end{aligned}$$

This implies

$$\dim X = \dim(V+H) \le n-1.$$

The contradiction proves the claim. □

For the low dimensional cases, the above system of axioms can be replaced by equivalent, more traditional systems of axioms. Consider first the case $n = 2$. If X is a given set, all the essential information on a projective plane structure $(X, \mathcal{S}_{-1}, \mathcal{S}_0, \mathcal{S}_1, \mathcal{S}_2)$ on X is contained in the set of straight lines $\mathcal{S}_1$ since subspaces of other dimensions are prescribed by the axioms and given together with X. Thus, we may write shortly $(X, \mathcal{S}_1)$ or $(X, \mathcal{S})$ instead of $(X, \mathcal{S}_{-1}, \mathcal{S}_0, \mathcal{S}_1, \mathcal{S}_2)$.

4.1.7. Proposition. *The system $(X, \{\emptyset\}, \{\{p\} \mid p \in X\}, \mathcal{S}, \{X\})$ satisfies the incidence axioms of a 2-dimensional projective space if and only if the pair $(X, \mathcal{S})$ has the following properties.*

- *For any two different points P, Q in X, there exists a unique straight line $e \in \mathcal{S}$ such that $P \in e$ and $Q \in e$.*
- *For any two different lines e, f in $\mathcal{S}$, there exists a unique point $P \in X$ such that $P \in e$ and $P \in f$.*
- *There exist 4 points in X such that no 3 of them can be covered by a line from $\mathcal{S}$.*

Proof. Assume first, that the system $(X, \{\emptyset\}, \{\{p\} \mid p \in X\}, \mathcal{S}, \{X\})$ satisfies the incidence axioms of a 2-dimensional projective space. Let $P \neq Q$ be two different points of X. The intersection of the 0-dimensional subspaces $\{P\}$ and $\{Q\}$ is the emptyset, hence -1-dimensional. Consequently, the join of these subspaces is one-dimensional $\{P\}+\{Q\} = e \in \mathcal{S}$. Now assume that $f \in \mathcal{S}$ is another straight line that covers P and Q. The intersection of e and f is a subspace the dimension of which is at most 1. On the other hand, $e \cap f$ contains two different points, thus it can not be a -1- or 0-dimensional subspace. If however $\dim e \cap f = 1$ then by the above proposition $e = e \cap f = f$, that proves unicity of the straight line through P and Q.

If e and f are two different straight lines, then their join is a subspace that contains e as a proper subset, thus $\dim(e+f) \ge 2$. This is possible only if $e+f = X$, $\dim(e+f) = 2$. This gives $\dim(e \cap f) = \dim e + \dim f - \dim(e+f) = 0$, that is, the intersection of the straight lines e and f contains a single point.

The last property in the proposition is the same as the last axiom for a 2-dimensional projective space. The "only if" part of the proposition is proved.

To show the "if" part, assume now that the pair $(X, \mathcal{S})$ has the properties listed in the proposition, and let us check that the system $(X, \{\emptyset\}, \{\{p\} \mid p \in X\}, \mathcal{S}, \{X\})$ satisfies the incidence axioms for a 2-dimensional projective space.

The first three axioms are fulfilled automatically.

Since X has at least 4 points, $\mathcal{S}_{-1}$, $\mathcal{S}_0$ and $\mathcal{S}_2$ are mutually disjoint. Let A, B, C, D be four points in general position. If e is a straight line, then one of these points, say A, is not in e. The straight lines through AB, AC and AD are different because of the general position condition. These straight lines intersect e at different points, since if we had $AB \cap e = AC \cap e = P$ for example, then the straight lines AB and AC would have two different points in common, A and P. Thus, we conclude that *a straight line has at least* 3 *points*, and consequently, $\mathcal{S}_1 \cap \mathcal{S}_{-1} = \mathcal{S}_1 \cap \mathcal{S}_0 = \emptyset$. We also get that X is not a straight line, since otherwise, taking another straight line e the intersection $X \cap e = e$ would have at least 3 points instead of 1.

It is obvious that the intersection of two subspaces is also a subspace if one of the subspaces is the empty set, a one point set or X. The intersection of two straight lines e and f is also a subspace, since if $e = f$ then $e \cap f = e$ and if $e \neq f$ then $e \cap f$ is a one point set.

The dimension formula is always satisfied automatically if one of the subspaces contains the other. Indeed, for $W_1 \subset W_2$ we have $W_1 \cap W_2 = W_1$ and $W_1 + W_2 = W_2$, and

$$\dim(W_1) + \dim(W_2) = \dim(W_1 \cap W_2) + \dim(W_1 + W_2).$$

There are three cases when none of the subspaces is contained in the other. When W_1 and W_2 are two different points, their intersection is the empty set, their join is the straight line through them. When one of the subspaces is a point, the other is a straight line, not passing through the point, the intersection of the subspaces is the empty set, the join of them is X. Finally, when W_1 and W_2 are two different straight lines, their intersection is a single point, their join is X. In all three cases, the dimension formula is true.

The last axiom is the same as the last property of the proposition. □

Now we give an alternative system of axioms for a 3-dimensional projective space. If X is a set, the essential information on a 3-dimensional projective space structure $(X, \mathcal{S}_{-1}, \mathcal{S}_0, \mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3)$ is given by the families of straight lines $\mathcal{S} = \mathcal{S}_1$ and planes $\mathcal{P} = \mathcal{S}_2$. Thus, we may refer to the projective space shortly as the system $(X, \mathcal{S}, \mathcal{P})$.

4.1.8. PROPOSITION. *Let X be an arbitrary set, $\mathcal{S}$ and $\mathcal{P}$ given families of subsets of X, the elements of which are called straight lines and planes respectively. Then $(X, \{\emptyset\}, \{\{p\} \mid p \in X\}, \mathcal{S}, \mathcal{P}, \{X\})$ is a 3-dimensional projective space if and only if the system $(X, \mathcal{S}, \mathcal{P})$ satisfies the following properties.*

- *For any two different points P, Q in X, there exists a unique straight line $e \in \mathcal{S}$ such that $P, Q \in e$.*
- *The intersection of two different planes Π, Σ in $\mathcal{P}$ is a straight line $\Pi \cap \Sigma \in \mathcal{S}$.*
- *If $P \in X$ is a point, $e \in \mathcal{S}$ is a straight line and $P \notin e$ then there exists a unique plane $\Sigma \in \mathcal{P}$ that contains both P and e.*
- *If $\Sigma \in \mathcal{P}$ is a plane, $e \in \mathcal{S}$ is a straight line and $e \not\subset \Sigma$ then the intersection of Σ and e is a point.*
- *If two different straight lines lie in a plane then they intersect one another at a single point.*
- *If two different straight lines have a point in common, then they are coplanar.*
- *There exist 5 points in X such that no 4 of them are coplanar.*

The proof of this proposition is longer than, but similar to the proof of the 2-dimensional case. Details of it are left to the reader.

4.2. The Principle of Duality, the Dual Space.

Given an n-dimensional projective space $(X, \mathcal{S}_{-1}, \mathcal{S}_0, \dots, \mathcal{S}_n)$, we can construct a new projective space, the dual space, the points of which are the hyperplanes of the original space. Any theorem that can be proved using the incidence axioms of the n-dimensional projective space can be applied to the dual space. Every theorem on the dual space can be interpreted as a theorem on the original space. By this procedure, every theorem of projective geometry generates another, dual theorem, which is valid once the original theorem is proved. This duality principle is one of the most fundamental principles of projective geometry.

Let $(X, \mathcal{S}_{-1}, \mathcal{S}_0, \dots, \mathcal{S}_n)$ be an n-dimensional projective space, and denote by X^* the set of hyperplanes in X: $X^* = \mathcal{S}_{n-1}$. For $W \in \mathcal{S}_i$, let us define the "anullator" of W as follows

$$W^{\perp} = \{H \in X^* \mid H \supset W\}.$$

We shall say that a subset of X^* is an i-dimensional subspace if it is the anullator of an $n-1-i$-dimensional subspace of X. Denote by

$$\mathcal{S}_i^* = \{W^{\perp} \subset X^* \mid W \in \mathcal{S}_{n-1-i}\}$$

the set of all i-dimensional subspaces of X^*.

4.2.1. Theorem. *The system $(X^*, \mathcal{S}_{-1}^*, \mathcal{S}_0^*, \dots, \mathcal{S}_n^*)$ is an n-dimensional projective space.*

4.2.2. Definition. $(X^*, \mathcal{S}_{-1}^*, \mathcal{S}_0^*, \dots, \mathcal{S}_n^*)$ is called the *dual space of the projective space* $(X, \mathcal{S}_{-1}, \mathcal{S}_0, \dots, \mathcal{S}_n)$

We shall prove the theorem after proving some lemmas.

4.2.3. Lemma. *Let W be a subspace in a projective space X, P a point not lying in W. Then there exists a hyperplane H in X such that $H \supset W$ but $P \notin H$.*

Proof. We shall use induction on the dimension of X. If $\dim X = 1$, then the statement is obvious. Assume the statement is true for spaces of dimension $< \dim X$. Joining points to W we can always find a hyperplane H_0 that contains W. If this hyperplane does not contain P, then we are done. If, however, $P \in H_0$ then we may look at H_0 as a $\dim X - 1$ dimensional projective space that contains P and W. By the induction hypothesis, we can find a hyperplane H_1 in the space H_0 such that $W \subset H_1 \not\ni P$. (H_1 will not be a hyperplane in X!). Let us choose a point $Q \notin H_0$ and consider the hyperplane $H = H_1 + \{Q\}$. H intersects H_0 in H_1, thus it contains W but does not contain P. $\square$

4.2.4. Corollary. *If W is an arbitrary subspace in a projective space, then W is the intersection of all hyperplanes that contain W.*

$$W = \bigcap_{H \in W^{\perp}} H.$$

(We use the convention that the intersection of no sets is the whole space X).

Proof. It is clear that $W \subset \bigcap_{H \in W^{\perp}} H$. On the other hand, if $P \notin W$ then there exists a hyperplane $H \supset W$ such that $P \notin H$, thus $W \supset \bigcap_{H \in W^{\perp}} H$. $\square$

4.2.5. Corollary. *If W_1 and W_2 are two subspaces of a projective space, then $W_1 \subset W_2$ if and only if $W_1^\perp \supset W_2^\perp$.*

Proof. If $W_1 \subset W_2$, then every hyperplane that contains W_2 contains W_1 as well, so $W_2^\perp \subset W_1^\perp$. On the other hand, if $W_2^\perp \subset W_1^\perp$, then by the previous corollary,

$$W_1 = \bigcap_{H \in W_1^\perp} H \subset \bigcap_{H \in W_2^\perp} H = W_2. \quad \square$$

Proof of Theorem 4.2.1.

- X is not contained in any hyperplane, so $X^\perp = \emptyset$, $\mathcal{S}_{-1}^* = \{\emptyset\}$.
- If $H \in X^*$ is a hyperplane in X, then the only hyperplane that contains H is H, thus $H^\perp = \{H\}$, and $\mathcal{S}_0^* = \{\{H\} \mid H \in X^*\}$.
- The empty set is contained in every hyperplane, consequently, $\emptyset^\perp = X^*$, and $\mathcal{S}_n^* = \{X^*\}$.
- By Corollary 4.2.4, the set $W^\perp$ defines W, (different subspaces have different anullators), and W has a uniquely defined dimension, so $W^\perp$ can not belong to both $\mathcal{S}_i^*$ and $\mathcal{S}_j^*$ for $i \neq j$.
- The intersection of subspaces of the dual space is also a subspace by the following identity.

$$\begin{aligned} W_1^\perp \cap W_2^\perp &= \{H \in X^* \mid H \supset W_1 \text{ and } H \supset W_2\} \\ &= \{H \in X^* \mid H \supset (W_1 + W_2)\} = (W_1 + W_2)^\perp. \end{aligned}$$

- To prove the dimension formula for the dual space, let us observe, that

$$\begin{aligned} W_1^\perp + W_2^\perp &= \bigcap_{W^\perp \supset W_1^\perp, W_2^\perp} W^\perp = \bigcap_{W \subset W_1, W_2} W^\perp \\ &= \bigcap_{W \subset W_1 \cap W_2} W^\perp = (W_1 \cap W_2)^\perp. \end{aligned}$$

  Using this identity,

$$\begin{aligned} \dim W_1^\perp + \dim W_2^\perp &= (\dim X - \dim W_1 - 1) + (\dim X - \dim W_2 - 1) \\ &= (\dim X - \dim(W_1 \cap W_2) - 1) + (\dim X - \dim(W_1 + W_2) - 1) \\ &= \dim(W_1 + W_2)^\perp + \dim(W_1 \cap W_2)^\perp. \end{aligned}$$

- It remains to show that X^* has $n+2$ "points" in general position. $n+2$ hyperplanes in X define $n+2$ points in general position of the dual space if and only if no $n+1$ of the hyperplanes go through a point of X. Let us choose $n+2$ points $P_1, P_2, \dots, P_{n+2} \in X$ in general position. Any n of these points span a hyperplane. For $i \neq j$, let us denote by $H_{i,j}$ the hyperplane spanned by the points $\{P_1, P_2, \dots, P_{n+2}\} \setminus \{P_i, P_j\}$. We prove that the hyperplanes $H_{1,2}, H_{2,3}, \dots, H_{n+1,n+2}, H_{n+2,1}$ are in general position. Since the role of the points $P_1, P_2, \dots, P_{n+2}$ can be cyclically permuted, it suffices to show that the first $n+1$ of these hyperplanes do not share a point in common. Let us check by induction, that

$$H_{1,2} \cap H_{2,3} \cap \cdots \cap H_{k,k+1} = \sum_{k+2 \le i \le n+2} \{P_i\}.$$

The assertion for $k=1$ is just the definition of $H_{1,2}$. Suppose the equation is known for $k-1$. Denote by W the subspace

$$H_{1,2}\cap H_{2,3}\cap\cdots\cap H_{k-1,k}=\sum_{k+1\leq i\leq n+2}\{P_i\}.$$

Since $P_{k+1}\in W\setminus H_{k,k+1}$, H and W span X and by the dimension formula $\dim(W\cap H_{k,k+1})=\dim W-1=n-k$. The span of the $n-k-1$ points $\{P_i\mid k+2\leq i\leq n+2$ is $n-k$ dimensional and is contained in $W\cap H_{k,k+1}$, hence

$$W\cap H_{k,k+1}=H_{1,2}\cap H_{2,3}\cap\cdots\cap H_{k,k+1}=\sum_{k+2\leq i\leq n+2}\{P_i\},$$

as it was to be proved. The equation for $k=n+1$ yields that the intersection $H_{1,2}\cap H_{2,3}\cap\cdots\cap H_{n+1,n+2}$ is empty. With this, the theorem is completely proved. □

Now we can formulate the duality principle more explicitly.

4.2.6. The Duality Principle. *Suppose that a theorem on the subspaces of an n-dimensional projective space can be derived from the incidence axioms and involves the following phrases: "i-dimensional subspace", "a subspace is contained in another", "the join of some subspaces", "the intersection of some subspaces". Then the theorem obtained by replacing the above phrases with the phrases "$n-i-1$-dimensional subspace", "a subspace contains another", "the join of some subspaces", "the intersection of some subspaces" respectively, is also a consequence of the incidence axioms.*

The vocabulary of the duality principle can be extended. For any macro concept defined on the base of the above atomic concepts we can produce a dual concept, and the macro and its dual can be included in the vocabulary. For example, the set of points of a straight line is often called *a range of points*. The set of hyperplanes that go through a fixed 2-codimensional subspace is called *a pencil of hyperplanes*. Obviously, these concepts are dual to one another. This way, if the theorem to be dualized contains the phrase "a range of points", it should be substituted by the phrase "a pencil of hyperplanes".

4.3. Desargues' Theorem and the Incidence Axioms. We have already proved Desargues' Theorem for projective spaces associated to a linear space over a division ring. Incidence axioms are weaker than the assumption that the projective space is associated to a linear space, so we may pose the question whether Desargues' Theorem can be proved assuming only the incidence axioms of the space. The answer depends on the dimension of the space. Incidence axioms of a projective space of dimension ≥ 3 imply Desargues' Theorem, but incidence axioms of the plane are not enough for this.

4.3.1. Proof of Desargues' Theorem from the Incidence axioms of an $n\geq 3$-dimensional projective space. Let the triangles ABC and $A'B'C'$ be in perspective from a point P. We shall distinguish two cases.

Case A. Assume that the planes $\Sigma=\{A\}+\{B\}+\{C\}$ and $\Sigma'=\{A'\}+\{B'\}+\{C'\}$ are different. The points A, B, C and P span a 3-dimensional subspace, which contains also the points A', B', C' and the planes Σ and Σ'. Σ and Σ' span this 3-dimensional space, so their intersection is a straight line e by the dimension formula. We show that the triangles ABC and $A'B'C'$ are in perspective with respect to the straight line e. Let us denote by Σ_{AB} the plane spanned by the points A, B and P. If Σ_{AB} were equal to Σ or Σ', then the system would collapse to a plane, so this is not the case. Consequently, Σ_{AB} cuts Σ and Σ' in a straight line. These straight lines are

$$\Sigma_{AB} \cap \Sigma = AB \text{ and } \Sigma_{AB} \cap \Sigma' = A'B'.$$

Since e and Σ_{AB} span a 3-dimensional space, they intersect one another at a point P_{AB}. Then

$$AB \cap A'B' \cap e = (\Sigma_{AB} \cap \Sigma) \cap (\Sigma_{AB} \cap \Sigma') \cap (\Sigma \cap \Sigma') = \Sigma_{AB} \cap \Sigma \cap \Sigma' = \Sigma_{AB} \cap e = P_{AB},$$

that is, the straight lines AB and $A'B'$ intersect one another on e. We can prove similarly that BC cuts $B'C'$ and AC cuts $A'C'$ at points on e, so we are done.

Case B. Now we assume that the triangles ABC and $A'B'C'$ lie in the same plane Σ. Since the space is at least 3-dimensional, we can find a point $S \notin \Sigma$ and since every straight line has at least three points (in general position), we can fix a point $\tilde{P}$ on the straight line SP, which is different from S and P. The plane spanned by S, P and A contains also the points A' and $\tilde{P}$ and the straight lines SA and $\tilde{P}A'$. Hence, we can define the point $\tilde{A}$ as the intersection point of the coplanar straight lines SA and $\tilde{P}A'$. We introduce similarly the points

$$\tilde{B} = SB \cap \tilde{P}B' \text{ and } \tilde{C} = SC \cap \tilde{P}C'.$$

The triangles ABC and $\tilde{A}\tilde{B}\tilde{C}$ are in perspective from the point S, the triangles $A'B'C'$ and $\tilde{A}\tilde{B}\tilde{C}$ are in perspective from the point $\tilde{P}$. Neither of these pairs of triangles are coplanar, so Case A gives that these pairs of triangles are in perspective with respect to the common straight line e of the plane of $\tilde{A}\tilde{B}\tilde{C}$ and Σ. By this,

$$\begin{array}{ccc} e \cap AB = e \cap \tilde{A}\tilde{B} & e \cap BC = e \cap \tilde{B}\tilde{C} & e \cap CA = e \cap \tilde{C}\tilde{A} \\ e \cap A'B' = e \cap \tilde{A}\tilde{B} & e \cap B'C' = e \cap \tilde{B}\tilde{C} & e \cap C'A' = e \cap \tilde{C}\tilde{A}, \end{array}$$

from which

$$e \cap AB = e \cap A'B' \quad e \cap BC = e \cap B'C' \quad e \cap CA = e \cap C'A',$$

thus the triangles ABC and $A'B'C'$ are also in perspective from e.

In the proof of Desargues' Theorem for projective spaces over a division ring (see section 2.6.), we saw that the "if" part of the theorem follows from the "only if" part. The arguments used there are universal and can be applied here as well to finish the proof. □

4.3.2. Moulton's non-desarguesian plane. The incidence axioms of a projective plane are two poor to imply Desargues' Theorem. This statement can be proved by presenting a model of a projective plane, which satisfies the incidence axioms but not Desargues' Theorem. The first non-desarguesian planes were constructed by Hilbert by algebraic methods. His ideas were simplified by Moulton, who finally arrived at a very simple model. Now we are going to describe this model.

The set of points of Moulton's plane is the set of points of an ordinary real projective plane. We think of the projective plane as the closure of the Euclidean plane $\mathbb{E}^2$ with points at infinity, and we introduce a Cartesian coordinate system on $\mathbb{E}^2$. Straight lines in Moulton's plane are subsets of the following type

- straight lines with negative slope

$$\{(x, y) \in \mathbb{E}^2 \mid y = mx + b\}, \text{ where } m < 0, m, b \in \mathbb{R},$$

  together with their points at infinity;
- vertical lines

$$\{(x, y) \in \mathbb{E}^2 \mid y = b\}, \text{ where } b \in \mathbb{R},$$

  together with their points at infinity;
- broken lines of the form

$$\{(x, y) \in \mathbb{E}^2 \mid y = \frac{3}{4}(mx + b) - \frac{1}{4} \mid mx + b \mid\}, \text{ where } m > 0, m, b \in \mathbb{R},$$

  together with the point at infinity of the straight line $y = mx + b$;
- the straight line at infinity.

It is easy to check that the projective plane together with this modified system of straight lines satisfies the incidence axioms of a projective plane, however, if we place two triangles in the plane as shown in the figure, then the triangles will be in perspective from a line but not from a point, so Desargues' Theorem fails to be valid on Moulton's plane.

4.4. Collineations.

4.4.1. Definition. Let $(X, \mathcal{S}_{-1}, \dots, \mathcal{S}_n)$ and $(Y, \mathcal{S}'_{-1}, \dots, \mathcal{S}'_n)$ be two n-dimensional projective spaces. We say that a mapping $\phi: X \to Y$ is a *collineation* if ϕ is a bijection and maps subspaces to subspaces, i.e., if for any subset $W \subset X$ of X, $W \in \mathcal{S}_i$ if and only if $\phi(W) \in \mathcal{S}'_i$.

We say that two projective spaces are *isomorphic* if there exists a collineation between them.

Isomorphic projective spaces are not different (essentially the same) from the view point of projective geometry. One of the most fundamental problems of projective geometry is to describe projective spaces up to isomorphism. A famous theorem of Hilbert says that if a projective space is desarguesian, then it is isomorphic to a projective space associated to a linear space over a division ring. This way, the classification of projective spaces of dimension ≥ 3 (since these spaces are automatically desarguesian) can be reduced to an algebraic question, the classification of division rings. Classification of projective planes, however, seems to be a very difficult problem. Search for finite projective planes is a combinatorial problem, and one might expect a classification for non-isomorphic finite projective planes. Nevertheless, we do not even know how many points a finite projective plane can have in general.

The inverse of a collineation and the composition of collineations (if defined) are collineations as well. Thus, collineations that map a projective space to itself form a group with respect to composition. This group is called the *collineation group of the space*. The collineation group contains all the symmetries of the projective space. The description of the collineation group of a given projective space is also a fundamental problem of projective geometry. This problem is solved for desaguesian spaces. For individual non-desarguesian planes the description of collineations can be a difficult problem.

4.4.2. Exercise. Describe the collineation group of Moulton's plane.

When we want to prove that a bijection between two projective spaces is a collineation it is convenient to have a simply verifiable characterization of collineations. We give one in the following proposition.

4.4.3. PROPOSITION. *If a bijection $\phi: X \to Y$ between two n-dimensional projective spaces has the property that for any three collinear points A, B and C in X the points $\phi(A)$, $\phi(B)$ and $\phi(C)$ are collinear in Y, then ϕ is a collineation.*

PROOF. By our assumption, if e is a straight line in X, then $\phi(e)$ is contained in a straight line f in Y. Let us show first that $\phi(e) = f$. Suppose to the contrary, that $f \setminus \phi(e) \neq \emptyset$. Then since ϕ is a bijection, we can find a point $P_1 \notin e$ such that $\phi(P_1) \in f$. The plane $e + \{P_1\}$ is the union of straight lines of the form P_1Q, where Q runs over e. Each of these straight lines are mapped into the straihgt line $\{\phi(P_1)\} + \{\phi(Q)\} = f$, so $\phi(\{P_1\} + e) \subset f$. Let us choose points $P_2, P_3, \dots, P_{n-1}$ in X in such a way that

- $P_i \notin e + \{P_1\} + \dots + \{P_{i-1}\}$;
- $\dim(e + \{P_1\} + \dots + \{P_i\}) = i + 1$.

Using induction on i we can show that $\phi(e+\{P_1\}+\dots+\{P_i\}) \subset f+\{\phi(P_2)\}+\dots+\{\phi(P_{i-1})\}$. For $i = n-1$ we get then that $\phi(X)$ is contained in the set spanned by f and the points $\phi(P_2), \dots, \phi(P_{n-1})$. However, the latter space has dimension $\leq (n-1)$, so it can not be the whole space Y. This contradicts the assumption that ϕ is onto.

Now let us define on Y another projective space structure calling a subset of Y an i-dimensional subspace if and only if it is the image of an i-dimensional subspace of X under ϕ. As we have seen, straight lines are the same in the original structure and in the one just defined, hence, according to one of our previous propositions, the two subspace structures on Y are the same. This conclusion is clearly equivalent to the statement that ϕ is a collineation. $\square$

We know that a subspace of a projective space inherits a projective space structure from the envelopping space. The following proposition says that any two subspaces of the same dimension are isomorphic.

4.4.4. Proposition. *Let W_1 and W_2 be two k-dimensional subspaces of a projective space. Then W_1 and W_2 are isomorphic k-dimensional projective spaces.*

We shall construct a collineation between W_1 and W_2 explicitly. For the construction, we need some lemmas.

4.4.5. Lemma. *A projective space can not be covered by two proper subspaces.*

Proof. Let V_1 and V_2 be the two subspaces. If one of them contains the other, then they have obviously no chance to cover the whole space. If none of them lies inside the other, then we can choose points $P_1 \in V_1 \setminus V_2$ and $P_2 \in V_2 \setminus V_1$. Since every straight line has at least 3 points, we can find a point Q on the straight line P_1P_2 which is different from both P_1 and P_2. This point Q is neither in V_1 nor in V_2. $\square$

4.4.6. Lemma. *Let W_1 and W_2 be two k-dimensional subspaces of a projective space. Then there exists a subspace V with the following properties*
- $W_1 \cap V = W_2 \cap V = \emptyset$;
- $W_1 + V = W_2 + V = W_1 + W_2$.

Proof. We define a sequence of subspaces $V_{-1}, V_0, \dots$. Let $V_{-1} = \emptyset$. Assume that $V_k \subset W_1 + W_2$ has already been defined in such a way that $W_1 \cap V_k = W_2 \cap V_k = \emptyset$, $\dim V_k = k$. If $V_k + W_1 = W_1 + W_2$, then V_k is a good choice for V, since $V + W_2 = W_1 + W_2$ holds because of $\dim(V + W_1) = \dim(V + W_2)$. If $V + W_1$ and $V + W_2$ are proper subspaces in $W_1 + W_2$, then we can choose a point $Q \in W_1 + W_2$ which is in neither of them and define V_{k+1} as $V_k + \{Q\}$. By finite dimensionality of the space, this procedure must stop at some point yielding a suitable V. $\square$

Proof of Proposition 4.4.4. Let V be the subspace given in the Lemma. If $\dim(W_1 \cap W_2) = l$, then $\dim(W_1 + W_2) = 2k - l$, and $\dim V = k - l - 1$. Let $Z \subset W_1$ be an m-dimensional subspace. Z and V span a $k - l + m$-dimensional subspace in $W_1 + W_2$. The intersection of this subspace with W_2 is also an m-dimensional subspace. The correspondence $Z \leftrightarrow (Z+V) \cap W_2$ is a bijection between m-dimensional subspaces of W_1 and W_2. In particular, it establishes a one to one correspondence between the points of W_1 and W_2. This correspondence is a collineation. $\square$

4.4.7. Corollary. *In a finite projective space, any two k-dimensional subspaces have the same number of points, and the same number of i-dimensional subspaces.*

Suppose that straight lines in a finite projective space have $q + 1$ points. In this case q is called the *order of the plane*. Let us determine how many i-dimensional subspaces there are in a k-dimensional subspace. Denote this number by α_i^k. Let W be a k-dimensional subspace, H a $k-1$-dimensional subspace in it, $P \in W \setminus H$. Consider the mapping $\phi\colon W \setminus \{P\} \to H$ defined by $\phi(Q) = PQ \cap H$. ϕ is a mapping onto H, since points of H are left fixed by ϕ. If $Q \in H$, then the preimage of Q

contains the points of the straight line PQ other than P, consequently, $\#\phi^{-1}(Q) = q$. This yields the recursion formula

$$\begin{aligned} \alpha_0^1 &= q+1 \\ \alpha_0^k &= \#W = q\#H + 1 = q\alpha_0^{k-1} + 1, \end{aligned}$$

from which

$$\alpha_0^k = q^k + q^{k-1} + \cdots + q + 1 = \frac{q^{k+1}-1}{q-1}.$$

It is also clear that an i-dimensional subspace in W that goes through P cuts H in an $i-1$-dimensional subspace, and this correspondence gives a bijection between i-dimensional subspaces in W that go through P and $i-1$-dimensional subspaces in H. Thus, there are α_{i-1}^{k-1} i-dimensional subspaces through a given point in a k-dimensional subspace. For there are α_0^k points in a k-dimensional subspace, and α_0^i points in an i-dimensional subspace, counting in two different way the number of pairs (P, V) such that P is a point in V and V is an i-dimensional subspace in W, we get

$$\alpha_0^k \alpha_{i-1}^{k-1} = \alpha_0^i \alpha_i^k.$$

Rearranging,

$$\alpha_i^k = \alpha_{i-1}^{k-1} \frac{\alpha_0^k}{\alpha_0^i} = \alpha_{i-1}^{k-1} \frac{q^{k+1}-1}{q^{i+1}-1}.$$

By a repeated application of this recursion formula, we obtain the following theorem.

4.4.8. Theorem. *The number of i-dimensional subspaces in a k-dimensional finite projective space of order q is equal to*

$$\alpha_i^k = \frac{(q^{k+1}-1)(q^k-1)\dots(q^{k-i+1}-1)}{(q^{i+1}-1)(q^i-1)\dots(q-1)}.$$

4.4.9. Remark. Let us introduce the notation $(n)_q = q^{n-1} + \cdots + q + 1$. $(n)_q$ is a polynomial the value of which at $q = 1$ is n. Thus, the value of the function

$$\binom{n}{k}_q = \frac{(n)_q (n-1)_q \dots (n-i+1)_q}{(i)_q (i-1)_q \dots (1)_q}$$

is the usual binomial coefficient $\binom{n}{k}$ for $q = 1$. This way, $\binom{n}{k}_q$ is a generalization of the ordinary binomial coefficient. Most of the basic theorems on ordinary binomial coefficients have an extension to generalized binomial coefficients. For example, the "q-analogue" of the standard binomial theorem is the following.

4.4.10. Generalized Binomial Theorem. *If X and Y are two elements in an algebra over the real numbers (e.g. two matrices) such that $YX = qXY$, then*

$$(X+Y)^n = \sum_{i=0}^{n} \binom{n}{i}_q X^{n-i} Y^i.$$

As a consequence, $\binom{n}{i}_q$ *is a polynomial of* q *with positive integer coefficients.*

Generalized binomial coefficients are relevant to our topic through the equation

$$\alpha_i^k = \binom{k+1}{i+1}_q.$$

Observe that the identity $\binom{n}{i}_q = \binom{n}{n-k}_q$ expresses the fact concordant with the duality principle, that the number of i- and $n-i-1$-dimensional subspaces in an n-dimensional projective space is the same.

The following theorem has many applications.

4.4.11. THEOREM. *Let* X *and* X*' be two* $n \geq 3$*-dimensional projective spaces,* $H \subset X$ *and* $H' \subset X'$ *be hyperplanes,* $\phi: H \to H'$ *a collineation,* $A, B \in X \setminus H$ *and* $A', B' \in X' \setminus H'$ *are different points, such that* $\phi(AB \cap H) = A'B' \cap H'$*. Then there exists a unique collineation* $\tilde{\phi}: X \to X'$ *such that the restriction of* $\tilde{\phi}$ *to* H *is* ϕ*,* $\tilde{\phi}(A) = A'$ *and* $\tilde{\phi}(B) = B'$*.*

PROOF. Let $P \in X$ be a point not lying on the straight line AB. Set $M = AB \cap H$, $P_A = AP \cap H$, $P_B = BP \cap H$, $M' = \phi(M)$, $P'_A = \phi(P_A)$ and $P'_B = \phi(P_B)$. The points M, P_A and P_B lie on the intersection of the plane ABP and the hyperplane H, hence they are collinear, and so are the points M', P'_A and P'_B. This implies that the points A', B', P'_A and P'_B are coplanar, so we may take the intersection point P' of the straight lines $A'P'_A$ and $B'P'_B$. It is clear that if $\tilde{\phi}$ has the required properties, it has to map P to P'. The assignment $\tilde{\phi}: P \mapsto P'$ is a bijection between $X \setminus AB$ and $X' \setminus A'B'$.

CLAIM. *The* $\tilde{\phi}$*-image of a straight line* e *is a straight line.*

Assume first that e does not intersect AB. The planes $\{A\} + e$ and $\{B\} + e$ intersect H at the straight lines e_A and e_B not passing through M. These straight lines are mapped by the collineation ϕ onto straight lines e'_A and e'_B not passing through M'. The planes $\{A'\} + e'_A$ and $\{B'\} + e'_B$ are different and intersect one another in a line, and this line is the image of e.

Now suppose that e intersects the straight line AB. We may assume that $A \notin e$. Let us choose a point $C \in X \setminus H$ which does not lie in the plane spanned by A and e. If $P \in e \setminus AB$, then the points C, P and $P_C = PC \cap H$ are collinear and contained in a straight line not intersecting the line AB. Thus, by the first case, the images of these points are also collinear. This means, that we can use the couples A, A' and C, $C' = \tilde{\phi}(C)$ instead of A, A' and B, B' to construct the image points of e. For e does not meet AC, the first case yields the second.

To extend $\tilde{\phi}$ to the straight line AB we prove the following.

CLAIM. *Let* P *be an arbitrary point on* AB*. If two straight lines* e *and* f *cross* AB *at* P*, then* $\tilde{\phi}$*-image of these straight lines meet at the same point of* $A'B'$*.*

Since e and f intersect each other, they are coplanar. Choose further points $E_1 \neq E_2$ and $F_1 \neq F_2$ on e and f, different from P. The straight lines E_1F_1 and E_2F_2 are coplanar and intersect each other at a point R not lying on AB. The images $\tilde{\phi}(E_1E_2)$, and $\tilde{\phi}(F_1F_2)$ meet at $\tilde{\phi}(R)$, hence they are coplanar. The plane spanned by them contains $\tilde{\phi}(e)$ and $\tilde{\phi}(f)$ as well, therefore these straight lines intersect each other. Since $e\{P\}$ and $f\{P\}$ are disjoint and $\tilde{\phi}$ is a bijection between $X \setminus AB$ and $X' \setminus A'B'$, the intersection point of $\tilde{\phi}(e)$ and $\tilde{\phi}(f)$ must be on $A'B'$.

To finish the proof, we define $\tilde{\phi}(P)$ for $P \in AB$ as the intersection point of $\tilde{\phi}(e) \cap A'B'$, where $e \neq AB$ is an arbitrary straight line through P in X. $\tilde{\phi}$ is obviously a collineation and the only one with the prescribed properties. $\square$

4.4.12. Corollary. *Let X and X' be two $n \geq 3$-dimensional projective spaces, and assume that a 2-dimensional subspace W of X is isomorhic to a 2-dimensional subspace W' of X'. Then X and X' are isomorphic.*

Proof. Include W and W' into a sequence of subspaces $W = W_2 \subset W_3 \subset \cdots \subset W_n = X$ $W' = W'_2 \subset W'_3 \subset \cdots \subset W'_n = X'$ such that $\dim W_i = \dim W'_i = i$ and extend a collineation between W and W' step by step, from W_k to W_{k+1}. $\square$

4.4.13. Corollary. *Let X be an $n \geq 3$-dimensional projective space, $H \subset X$ a hyperplane, $O \in X$ an arbitrary point, A and A' are points not belonging to H such that O, A and A' are collinear. Then there exists a unique collineation $\tilde{\phi}: X \to X$ with the following properties*
- *$\tilde{\phi}(P) = P$ for every $P \in H$;*
- *$\tilde{\phi}(e) = e$ for every straight line e through O;*
- *$\tilde{\phi}(A) = A'$.*

Proof. If $O \notin H$, then we can apply the theorem for $X = X'$, $O = B = B'$ and $\phi = \mathrm{id}_H$. The second property of the extension $\tilde{\phi}$ is given by the fact that every straight line through O contains two fixed point of $\tilde{\phi}$, the points O and $e \cap H$.

If $O \in H$, the let us choose a point $B \notin OA \cup H$ and define B' as the intersection point of OB and the straight line through A' and $AB \cap H$. Applying the theorem with $X = X'$, $H = H'$ and $\phi = \mathrm{id}_H$ we get $\tilde{\phi}$. To check that the second property holds, consider an arbitrary straight line e through O. Choose a point C on e not lying on the straight line AB. The triangle ABC and its image $A'B'C'$ are in perspective with respect to a line, thus, by Desargues' Theorem, they are in perspective from a point. This point must be the intersection point of AA' and BB', that is O. Consequently, C, C' and O are collinear, and the straight line $OC = e$ is mapped to the straight line $OC' = e$. $\square$

Why can we not extend a collineation between strtaight lines to planes containing the lines? A straight line, that is a 1-dimensional projective space is nothing else but a set with no structure on it. Any bijection between two straight lines is a collineation. Obviously, there is no hope that we can extend any permutation of the points of a straight line to a collineation of the plane. The reason is the following. Although a straight line in itself has no structure, a straight line embedded into a plane is given a structure by the way it is embedded. For example, if a straight line lies in a desarguesian plane, then, as we shall see in the next chapter, after fixing three points on the line, the line is induced on a division ring structure. If a bijection between two lines in a desarguesian plane has no respect to this algebraic structure, it does not extend to a collineation. Before we make the division ring structure on a straight line clear, we can try to extend the identical transformation of a straight line to a non-trivial collineation.

4.4.14. Definition. Let ϕ be a collineation of a plane. If each point of a straight line is fixed by ϕ, then it is called an *axis of the collineation*. If all the straight lines through a point are fixed (mapped to itself) by ϕ, then the point is called the *center of the collineation*. A collineation which has both axis and center is called a *central axial collineation*.

4.4.15. Existence of Central Axial Collineations on a Desarguesian Plane. *Let e be an arbitrary straight line in a desarguesian plane X, O be an arbitrary point, P and P' two different points not lying on e, such that O, P and P' are collinear. Then there exists a unique central axial collineation ϕ that has e as axis, O as center and maps P to P'.*

Proof. Assume Q is not on the straight line OP. We can construct $\phi(Q) = Q'$ in the following way. The intersection point $M = e \cap PQ$ is a fixed point of ϕ, so the point $P'Q'M$ must be collinear. On the other hand, the straight line OQ is mapped to itself, so Q must lie on it. From this we get $Q' = P'M \cap OQ$. This way, we can define ϕ outside the straight line OP. If we choose a further point R, which is neither on OP, nor on OQ, then we can use both the pair PP' and QQ' to construct $\phi(R) = R'$. An application of Desargues Theorem to the triangles PQR and $P'Q'R'$ yields that both constructions will result in the same point. The pair QQ' can also be used to construct the images of points lying on the straight line OP. This way, we can extend ϕ to the whole plane and we can also prove that replacing the pair PP' by any pair $S\phi(S)$ the mapping $\phi\colon X \to X$ constructed starting from $S\phi(S)$ is the same we have just obtained.

ϕ is a bijection, let us show that it is a collineation. Straight lines through O and e are fixed by the construction. Assume now, that $S \neq O$ is an arbitrary point not in e, and consider a straight line f through S. Set $S' = \phi(S)$, $M = e \cap f$ and let $T \in f$. Since we can use the pair SS' to construct the image of T, $\phi(T) = S'M \cap OT$. Hence, the image of f is the straight line $S'M$. Varying S and f we get that every straight line is mapped to a straight line. $\square$

4.4.16. Corollary. *Let e be a straight line on a desarguesian plane, $A \neq B$ and $A' \neq B'$ points outside e such that $AB \cap e = A'B' \cap e$. Then there exists a unique central axial collineation that has e as axis and map A to A' and B to B'.*

I. f AB and $A'B'$ are not collinear, choose for O the intersection point $AA' \cap BB'$. The central axial collineation with axis e and center O that takes A to A' will be the one we are looking for.

If AB and $A'B'$ lie on the same line, then choose a point C not on this line and not on e. Construct C' in such a way that the triangles ABC and $A'B'C'$ be in perspective with respect to the straight line e. Denote by O the point these triangles are in perspective from. Obviously, the central axial collineation with axis e, center O, that maps A to A' is good for us and it is the only one with the required properties. $\square$

PROJECTIVE GEOMETRY

LECTURE NOTES BY BALÁZS CSIKÓS

5. DESARGUESIAN PROJECTIVE SPACES

One of the fundamental problems of the axiomatic foundations of projective geometry is the description of all possible models of the system of incidence axioms of the n-dimensional projective space. This problem is solved for $n \geq 3$ and partially solved for $n = 2$ by the following theorem to the proof of which we devote the present chapter.

5.0.1. THEOREM. *Every desarguesian projective space $(X, \mathcal{S}_i)$ is isomorphic to the projective space $\mathbb{F}P^n$ over a division ring $\mathbb{F}$.*

The theorem can be broken into two theorems depending on the dimension of the space.

5.0.2. THEOREM. *Every desarguesian projective plane $\mathcal{P}$ is isomorphic to a projective plane $\mathbb{F}P^2$ over some division ring $\mathbb{F}$.*

If the dimension of the space is at least 3, we may omit the assumption that Desargues' Theorem is valid, since it can be proved from the incidence axioms (Theorem ...), thus we obtain

5.0.3. THEOREM. *Every projective space of dimension ≥ 3 is isomorphic to the projective space $\mathbb{F}P^n$ over a division ring $\mathbb{F}$.*

It is enough to prove Theorem 5.0.2. Indeed, if $(X, \mathcal{S}_i)$ is an $n \geq 3$-dimensional desarguesian projective space, then any plane $\mathcal{P} \in \mathcal{S}_2$ in it is desarguesian as well. If Theorem 5.0.2. holds, then $\mathcal{P}$ is isomorphic to the projective plane $\mathbb{F}P^2$ over some division ring $\mathbb{F}$. But $\mathbb{F}P^2$ is isomorphic also to any plane in the projective space $\mathbb{F}P^n$, and hence, by Corollary ..., X and $\mathbb{F}P^n$ are isomorphic.

We split the proof of Theorem 5.0.2. into two steps. First we construct a suitable division ring $\mathbb{F}$, then we construct a collineation between $\mathcal{P}$ and $\mathbb{F}P^2$.

5.1. Construction of the division ring $\mathbb{F}$.

Let e be a line, and fix two different points O and I on e. We define an operation $+$ on $e \setminus \{I\}$ in the following way. (The definition is motivated by Exercise **2-3.**) If $X, Y \in e \setminus \{I\}$ are arbitrary points, then $X + Y$ is constructed according to the following instructions.

- Choose two arbitrary points $A \neq B$ outside e such that $X = AB \cap e$;

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Construct the points

- $C = AI \cap BO$;
- $D = CY \cap BI$;
- $X + Y = AD \cap e$.

5.1.1. LEMMA. *The above construction of $X + Y$ is correct, i.e.,*

- *whenever a point is defined as the intersection of two lines, the lines are different;*
- *$X + Y \neq I$;*
- *the position of $X + Y$ does not depend on the choice of A and B.*

PROOF. The first two parts of the Lemma are obvious. Let us prove independence on the choice of A and B. Choose two different couples $A \neq B$ and $\tilde{A} \neq \tilde{B}$ outside e, such that $AB \cap e = \tilde{A}\tilde{B} \cap e = X$ and construct $X + Y$ with the help of both pairs, that is, set

- $C = AI \cap BO$, $\tilde{C} = \tilde{A}I \cap \tilde{B}O$;
- $D = CY \cap BI$, $\tilde{D} = \tilde{C}Y \cap \tilde{B}I$;
- $X + Y = AD \cap e$, $X \tilde{+} Y = \tilde{A}\tilde{D} \cap e$.

By Corollary ..., we can find a central axial collineation Φ that has e for its axis and maps A to $\tilde{A}$ and B to $\tilde{B}$. Then obviously, $\Phi(C) = \tilde{C}$, $\Phi(D) = \tilde{D}$ and $\Phi(X+Y) = X\tilde{+}Y$. However, Φ leaves points of e fixed, thus $X+Y = \Phi(X+Y) = X\tilde{+}Y$ as it was to be proved. $\square$

5.1.2. LEMMA. *The set $e \setminus \{I\}$ together with the operation $+$ is a commutative group, which we denote by $\mathbb{F}_{OI}$.*

PROOF. Let us show first that the addition is *associative*.

Choose three points $X, Y, Z \in e \setminus \{I\}$ and construct $X+Y$ with the help of $A \neq B$, $A, B \notin e$, $X \in AB$, i.e., set $C = AI \cap BO$, $D = CY \cap BI$ and $X + Y = AD \cap e$.

To construct $(X+Y)+Z$, we can use the auxiliary points A and D in the role of "B": $E = AI \cap DO$, $F = EZ \cap DI$, $(X+Y)+Z = AF \cap e$.

If we construct $Y+Z$ with auxiliary points C and D in the role of A and B respectively, then we see that $Y+Z = CF \cap e$. Constructing now $X+(Y+Z)$ with the help of A and B, we get $X+(Y+Z) = AF \cap e = (X+Y)+Z$.

Next we show that $O+X = X$ for any $X \in e \setminus \{I\}$. Choose $A \neq B$ not in e, such that $O = AB \cap e$. Then $C = AI \cap BO = A$, $D = CX \cap BI = AX \cap BI$, $O+X = AD \cap e = X$.

If $X \in e \setminus \{I\}$ is an arbitrary point, then we can construct a point $-X$ such that $X+(-X) = O$ in the following way. Choose two different points A, B not in e such that $AB \cap e = X$. Set $C = AI \cap BO$, $D = AO \cap BI$, $-X = CD \cap e$. Then $X+(-X) = O$ holds obviously.

Finally, let us prove that $+$ is a *commutative* operation.

Let us construct the sum $X+Y$ of two points with the auxiliary points A, B as usual, i.e., set $C = AI \cap BO$, $D = CY \cap BI$ and $X+Y = AD \cap e$. To construct $Y+X$ we use the auxiliary points $\tilde{A} = D$, $\tilde{B} = C$. According to the general procedure, $Y+X$ is obtained by setting $\tilde{C} = \tilde{A}I \cap \tilde{B}O = DI \cap CO = B$, $\tilde{D} = \tilde{C}X \cap \tilde{B}I = BX \cap CI = A$ and $Y+X = \tilde{A}\tilde{D} \cap e = DA \cap e = X+Y$. $\square$

5.1.3. Lemma. *Let $O \neq I$ and $O' \neq I'$ be arbitrary points on the plane, Φ be a collineation such that $\Phi(O) = O'$, $\Phi(I) = I'$. Then the restriction of Φ onto $\mathbb{F}_{OI}$ is a group isomorphism between the commutative groups $\mathbb{F}_{OI}$ and $\mathbb{F}_{O'I'}$*

Proof. Obvious. $\square$

5.1.4. Definition. Let $e \neq f$ be two different coplanar straight lines, $C \notin e \cup f$ be a point in their planes. Then a *perspective transformation* or simply *perspectivity* from e to f with center C is a mapping $\phi : e \to f$, defined by the equality $\phi(P) = CP \cap f$.

5.1.5. Lemma. *If a perspectivity $\phi : e \to f$ maps the points $O, I \in e$ to $O', I' \in f$ respectively, then ϕ is a group isomorphism between $\mathbb{F}_{OI}$ and $\mathbb{F}_{O'I'}$.*

Proof. The statement follows from Lemma 5.1.3 if we show that every perspectivity extends to a collineation of the plane. Consider the perspectivity ϕ and denote the center of it by C. Let M be the intersection point of e and f, $A \in e$ be a point different from M, and set $A' = \phi(A)$. By theorem ..., there exists a central axial collineation Φ of the plane with center C and axis CM such that $Phi(A) = A'$. We claim that the restriction of Φ onto e is ϕ. Since $e = AM$, we have $\Phi(e) = \Phi(AM) = A'M = f$. If $B \in e$ is an arbitrary point, then $\Phi(B)$ must lie somewhere on the line CB, since C is the center of the central axial collineation Φ, on the other hand we also have $\Phi(B) \in f$, for $\Phi(e) = f$. Therefore, $\Phi(B) = CB \cap f = \phi(B)$. $\square$

5.1.6. Lemma. *Let f be a straight line intersecting e at I, A, C be points not on $e \cup f$ such that $C \in AI$, let $\phi : e \to f$ be a perspectivity with center C, $\psi : f \to e$ be a perspectivity with center A. Then*

$$\psi \circ \phi(X) = X + Y \text{ for all } X \in \mathbb{F}_{OI},$$

where $Y = \psi \circ \phi(O)$.

Proof. The statement is a rephrasing of the definition of $X + Y$. $\square$

Our next goal is to equip the commutative groups $\mathbb{F}_{OI}$ constructed above with a multiplicative structure. For this we have to fix a third point on the line, which will play the role of the multiplicative unit.

Let O, U, I be three different points on a straight line e, $X, Y \in e \setminus \{I\}$ be two arbitrary points. We define the point $X * Y \in e \setminus \{I\}$ by the following construction:

- choose two points $A, B \notin e$ such that $X \in AB$;

  Set
- $D = OA \cap UB$;

- $C = DY \cap BI$;
- $X * Y = AC \cap e$.

5.1.7. Lemma. *The above definition of $X * Y$ is correct, i.e.,*

- *whenever a point is defined as the intersection of two lines, the lines are different;*
- $X * Y \neq I$;
- *the position of $X * Y$ does not depend on the choice of A and B.*

Proof. The proof of this lemma is completely analogous to that of Lemma 5.1.1. $\square$

Let us denote by $\mathbb{F}_{OUI}$ the set $e \setminus I$ together with the operations $+$ and $*$ given by the points O, U, I.

5.1.8. Lemma. *$\mathbb{F}_{OUI}$ is a division ring.*

Proof.

a) We prove first that $*$ is associative. Let $X, Y, Z \in e \setminus \{I\}$ be arbitrary points. Let us construct $X * Y$ with the help of the points A, B according to the definition: $D = OA \cap UB$, $C = DY \cap BI$, $X * Y = AC \cap e$. To construct $(X * Y) * Z$ let us use the auxiliary points $\tilde{A} = A$ and $\tilde{B} = C$. Then we take the points $\tilde{D} = O\tilde{A} \cap U\tilde{B} = OA \cap UC$, $\tilde{C} = \tilde{D}Z \cap \tilde{B}I$, $(X * Y) * Z = \tilde{A}\tilde{C} \cap e$. Let us construct now $Y * Z$ with the auxiliary points $A' = D$ and $B' = \tilde{B} = C$. Then $D' = OA' \cap UB' = OA \cap UC = \tilde{D}$, $C' = D'Z \cap B'I = \tilde{D}Z \cap \tilde{B}I = \tilde{C}$, $Y * Z = A'C' \cap e = D\tilde{C} \cap e$. Finally, let us construct $X * (Y * Z)$ with the help of AB. The "D"-point of the construction is D, the "C"-point of the construction is $\tilde{C} = C'$, thus

$$X * (Y * Z) = AC' \cap e = (X * Y) * Z. \quad \square$$

b) Now we show that $*$ is *distributive* with respect to $+$, i.e.,

$$X*(Y+Z) = X*Y+X*Z \text{ and } (Y+Z)*X = Y*X+Z*X \text{ for any } X,Y,Z \in e\setminus\{I\}$$

It is obvious from the definition of the multiplication that $O * Y = Y * O = O$ for any Y, thus distributivity holds if $X = O$.

Let us fix the point $X \neq O$ and denote by L_X and R_X the left and right multiplication with X

$$L_X(Y) = X * Y \qquad R_X(Y) = Y * X.$$

Both L_X and R_X can be represented as compositions of two perspectivities.

Indeed, assume that we construct the product $X * Y$ for any Y with the help of the same pair of points A, B. Let C be the intersection of BI and AU. Now if $\phi : e \to AO$ is the perspectivity with center C and $\psi : AO \to e$ is the perspectivity with center B, then the composition $\psi \circ \phi : e \to e$ takes Y to $X * Y$, hence it is the left multiplication L_X by X.

As for the right multiplication, choose two different points C and D outside e such that $CD \cap e = X$ and consider the point $A = DO \cap CU$. If $\phi : e \to CI$ is the perspectivity with center A and $\psi : CI \to e$ is the perspectivity with center D, then the composition $\psi \circ \phi : e \to e$ takes $Y \in e \setminus \{I\}$ to $Y * X = R_X(Y)$, thus $\psi \circ \phi = R_X$

Applying Lemma 5.1.5, we obtain that R_X and L_X are isomorphisms of the commutative group $\mathbb{F}_{OI}$ into itself. In particular,

$$X * (Y + Z) = L_X(Y + Z) = L_X(Y) + L_X(Z) = X * Y + X * Z$$

and

$$(Y + Z) * X = R_X(Y + Z) = R_X(Y) + R_X(Z) = Y * X + Z * X$$

for any $Y, Z \in \mathbb{F}_{OI}$.

$\mathbb{F}_{OUI}$ contains a *multiplicative unit element*, since $U * X = X * U = X$ for every X. (This identity can be derived from the definition of $*$.)

Finally, every $X \neq O$ has a right and a left *inverse*. Indeed, $X \neq O$ implies that L_X and R_X are compositions of perspectivities, therefore invertible transformations. If we set $X_L^{-1} = R_X^{-1}(U)$ and $X_R^{-1} = L_X^{-1}(U)$, then obviously

$$X_L^{-1} * X = U \quad \text{and} \quad X * X_R^{-1} = U.$$

The left and right inverse of X must coincide, since

$$X_L^{-1} = X_L^{-1} * U = X_L^{-1} * (X * X_R^{-1}) = (X_L^{-1} * X) * X_R^{-1} = U * X_R^{-1} = X_R^{-1}.$$

We conclude that X has an inverse and this finishes the proof of the Lemma. □

5.1.9. Lemma. *Let f be a straight line intersecting e at O, B, C be points not in $e \cup f$ such that $C \in BI$, let $\phi : e \to f$ be a perspectivity with center B, $\psi : f \to e$ be a perspectivity with center C. Then*

$$\psi \circ \phi(X) = X * M \text{ for all } X \in \mathbb{F}_{OUI},$$

where $M = \psi \circ \phi(U)$

Proof. The statement is a rephrasing of the definition of $X * M$. □

5.2. Construction of the collineation between $\mathcal{P}$ and $\mathbb{F}P^2$.

We know that for a projective plane $P(V)$ associated to a linear space over a division ring $\mathbb{F}$, an isomorphism between $P(V)$ and $\mathbb{F}P^2$ is established by a homogeneous coordinate system and homogeneous coordinate systems are given by 4 points in general position. This motivates our constructions below.

Let $\mathcal{P}$ be a desarguesian plane and choose four points O, I_x, I_y and U in general position. Let I be the point $OU \cap I_x I_y$ and $\mathbb{F}$ be the division ring $\mathbb{F}_{OUI}$. Elements of $\mathbb{F}$ are points of the straight line OU different from I.

5.2.1. Lemma. *Let f be an arbitrary straight line through O different from the line OI_y. For $X \in \mathbb{F}$, let us construct the point $Y \in \mathbb{F}$ as follows. Take $P = XI_y \cap f$ and set $Y = ZI_x \cap OU$. Then the quotient $X^{-1} * Y \in \mathbb{F}$ is independent of the choice of X.*

Proof. If $f \neq OI_x$, then the mapping $X \mapsto Y$ is the composition of two perspectivities satisfying the conditions of Lemma 5.1.9. Thus there exists $M \in \mathbb{F}$, independent of X, such that $Y = X * M$, or equivalently, $X^{-1} * Y = M$.

If $f = OI_x$, then $Y = O$ for every X, thus $X^{-1} * Y = O$ for every X. □

Now we define a bijection $\Phi : \mathcal{P} \to \mathbb{F}P^2$. Let $P \in \mathcal{P}$.

- If $P \notin I_x I_y$, then let $\Phi(P) = (X : Y : U)$, where $X = I_y \cap OU$, $Y = I_x \cap OU$.
- If $P \in I_x I_y$ and $P \neq I_y$, then choose a point $Q \in PO$, $Q \neq P$, take $X = I_y Q \cap OU$, $Y = I_x Q \cap OU$ and set $\Phi(P) = (X : Y : O)$. This definition is correct by Lemma 5.2.1.
- Finally, set $\Phi(I_y) = (O : U : O)$.

It is easy to check that Φ is indeed a bijection. The following Lemma compares the coordinates of points on two lines intersecting $I_x I_y$ at the same point.

5.2.2. Lemma. *Let f and f_0 be two straight lines intersecting the line $I_x I_y$ at the same point I_f, assume $O \in f_0$. Let $P \in f$ be an arbitrary point different from I_f. Consider the points $P_0 = I_y P \cap f_0$, $Y = I_x P \cap OU (\in \mathbb{F})$, $Y_0 = I_x P_0 \cap OU (\in \mathbb{F})$. Then the difference $Y - Y_0$ does not depend on the choice of P.*

Proof. We can represent the mapping $Y_0 \mapsto Y$ as the composition of four perspectivities

$$Y_0 \xmapsto{\theta} P_0 \xmapsto{\phi} P \xmapsto{\psi} I_x P \cap f \xmapsto{\theta^{-1}} Y.$$

Figure for Lemma 5.2.1 Figure for Lemma 10

The centers of $\theta : OU \to f$, $\phi : f \to f_0$ and $\psi : f_0 \to f$ are I_x, I_y and I_x respectively. By Lemma 5.1.3, θ is a group isomorphism between $\mathbb{F}_{OI}$ and $\mathbb{F}_{OI_f}$. By Lemma 5.1.6, for the composition $\psi \circ \phi : \mathbb{F}_{OI_f} \to \mathbb{F}_{OI_f}$ we have

$$\psi \circ \phi(P_0) = P_0 + \psi \circ \phi(O),$$

where $+$ denotes here the addition in $\mathbb{F}_{OI_f}$. Combining these two results, we obtain

$$Y = Y_0 + \theta^{-1} \circ \psi \circ \phi(O). \quad \square$$

The following Lemma finishes the proof of the theorem.

5.2.3. Lemma. *The bijection $\Phi : \mathcal{P} \to \mathbb{F}P^2$ is a collineation.*

Proof. According to Proposition ..., it is enough to prove that the image of a straight line f in $\mathcal{P}$ is contained in a straight line of $\mathbb{F}P^2$.

- If $f = I_x I_y$, then the last homogeneous coordinate of $\Phi(P)$ for $P \in f$ is O, therefore $\Phi(f)$ is contained in the line spanned by the points $(U : O : O)$ and $(O : U : O)$.
- If f passes through I_y and $f \neq I_x I_y$, then

$$\Phi(f) = \{(X_0 : Y : U) \mid X_0 = f \cap OU,\, Y \in \mathbb{F}\} \cup \{(O : U : O)\},$$

  thus $\Phi(f)$ is contained in the straight line spanned by $(X_0 : O : U)$ and $(O : U : O)$.
- If f does not go through I_y, then consider the point $I_f = f \cap I_x I_y$ and the straight line $f_0 = OI_f$. By Lemma 5.2.1, there exists a constant $M \in \mathbb{F}$ such that

$$\Phi(f_0) = \{(X : X * M : Z) \mid X, Z \in \mathbb{F}\}$$

  By Lemma 5.2.2, we can find a constant $N \in \mathbb{F}$ such that

$$\Phi(f) = \{(X : X * M + N : U) \mid X \in \mathbb{F}\} \cap \{(X : X * M : O)\},$$

  therefore, $\Phi(f)$ is contained in the straight line spanned by $(U : M : O)$ and $(O : N : U)$. $\square$

We have completed the proof of the theorem.

PROJECTIVE GEOMETRY

Lecture Notes by Balázs Csikós

6. THE FUNDAMENTAL THEOREM OF PROJECTIVE GEOMETRY

When we introduced projective spaces axiomatically, we did it with the hope that the theorems we prove on the base of the incidence axioms will be applicable to a large variety of spaces more general than the projective spaces associated to a linear space over a division ring. However, according to the main theorem of the previous unit, the axiomatic approach brings more generality only for dimension 2. Dealing only with projective spaces over a division ring we loose only non-desarguesian projective planes.

From this point on we shall consider the projective space $\mathbb{F}P^n = P(\mathbb{F}^{n+1})$ over the division ring $\mathbb{F}$, where $n \geq 2$. Let us denote by $Coll(\mathbb{F}P^n)$ the group of all collineations of the space. The Fundamental Theorem of Projective Geometry gives a description of this group. Before formulating the theorem, we define two subgroups of $Coll(\mathbb{F}P^n)$ and prove their characteristic properties.

6.1. The Projective General Linear Group $PGL(n+1, \mathbb{F})$.

6.1.1. Definition. Let V and W be linear spaces over $\mathbb{F}$. We call a mapping $\Phi : V \to W$ *linear* if

$$\Phi(\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2) = \alpha_1 \Phi(\mathbf{v}_1) + \alpha_2 \Phi(\mathbf{v}_2)$$

for every $\mathbf{v}_1, \mathbf{v}_2 \in V$ and $\alpha_1, \alpha_2 \in \mathbb{F}$.

Composition of linear mappings is also linear. If a linear mapping is a bijection, then its inverse is linear as well. In particular, invertible linear transformations of a linear space V into itself form a group, the so-called *general linear group of* V, which we shall denote by $GL(V)$. We shall also use the notation $GL(n, \mathbb{F})$ for the group $GL(\mathbb{F}^n)$.

If we fix a basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ of the linear space V, then the images of the basis vectors under a linear transformation $\Phi : V \to V$ can be written as a linear combination of the basis vectors

$$\Phi(\mathbf{v}_i) = a_i^1 \mathbf{v}_1 + \dots + a_i^n \mathbf{v}_n.$$

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

The $n \times n$ matrix $A = (a_i^j)$ with a_i^j at the i-th position of the j-th row is *the matrix of* Φ with respect to the basis $\mathbf{v}_1, \dots, \mathbf{v}_n$. The assignment $\Phi \mapsto A$ is a one-to-one correspondence between linear transformations from V to V and $n \times n$ matrices. If $\mathbf{x} \in V$ has coordinates $(x^1, \dots, x^n)$ with respect to the basis $\mathbf{v}_1, \dots, \mathbf{v}_n$, then we have

$$\Phi(\mathbf{x}) = \Phi(\sum_{i=1}^{n} x^i \mathbf{v}_i) = \sum_{i=1}^{n} x^i \Phi(\mathbf{v}_i) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} x^j a_j^i \right) \mathbf{v}_i.$$

Therefore, we obtain the coordinates of $\Phi(\mathbf{x})$ if we multiply the row vector of coordinates of $\mathbf{x}$ with the matrix A from the right.

Now let $\Phi \in GL(n+1, \mathbb{F})$ be an invertible linear transformation. Then Φ maps linear subspaces of $\mathbb{F}^{n+1}$ onto linear subspaces of the same dimension. In particular, Φ induces a mapping on the 1-dimensional subspaces, that is on the points of the associated projective space. This mapping

$$[\Phi] : \mathbb{F}P^n \to \mathbb{F}P^n \qquad [\Phi]([\mathbf{x}]) = [\Phi(\mathbf{x})]$$

is obviously a collineation.

It is straightforward that if $\Phi, \Psi \in GL(n+1, \mathbb{F})$, then

$$[\Phi \circ \Psi] = [\Phi] \circ [\Psi] \text{ and } [\Phi^{-1}] = [\Phi]^{-1}.$$

Thus, collineations obtained from invertible linear transformations form a subgroup of the group of all collineations. We denote this subgroup by $PGL(n+1, \mathbb{F})$ and call it the *projective general linear group*.

The assignment $\Phi \mapsto [\Phi]$ is a surjective group homomorphism. The kernel of this homomorphism consists of linear transformations $\Phi \in GL(n+1, \mathbb{F})$ such that $[\Phi(\mathbf{x})] = [\mathbf{x}]$ for every $\mathbf{0} \neq \mathbf{x} \in \mathbb{F}^{n+1}$. Let us find more explicit description for these transformations.

6.1.2. Definition. A vector $\mathbf{x} \neq \mathbf{0}$ in a linear space V is said to be an *eigenvector* of the linear mapping $\Phi : V \to V$ if there exists a scalar $\lambda \in \mathbb{F}$ such that $\Phi(\mathbf{x}) = \lambda \mathbf{x}$. In this case λ is said to be the *eigenvalue* corresponding to the eigenvector $\mathbf{x}$. The scalar λ is an eigenvalue of Φ if it corresponds to an eigenvector.

Observe that if $\mathbf{x}$ is an eigenvector of Φ with eigenvalue λ, then for any $0 \neq \mu \in \mathbb{F}$, the vector $\mu\mathbf{x}$ is also an eigenvector with eigenvalue $\mu\lambda\mu^{-1}$.

It is clear from the definition that $[\mathbf{x}] \in \mathbb{F}P^n$ is a fixed point of the collineation $[\Phi] \in PGL(n+1, \mathbb{F})$ if and only if $\mathbf{x}$ is an eigenvector of Φ.

6.1.3. Lemma. *If the vectors of a basis $\mathbf{v}_1, \dots, \mathbf{v}_{n+1} \in \mathbb{F}^{n+1}$ and the sum $\mathbf{q}$ of the basis vectors are eigenvectors of a linear transformation Φ, then the eigenvalues corresponding to the basis vectors are equal.*

Proof. If the eigenvalue corresponding to $\mathbf{v}_i$ is λ_i and the eigenvalue corresponding to $\mathbf{q}$ is λ, then

$$\lambda \mathbf{v}_1 + \dots + \lambda \mathbf{v}_{n+1} = \Phi(\mathbf{q}) = \Phi(\mathbf{v}_1) + \dots + \Phi(\mathbf{v}_{n+1}) = \lambda_1 \mathbf{v}_1 + \dots + \lambda_{n+1} \mathbf{v}_{n+1}.$$

Since decomposition of vectors into the linear combination of basis vectors is unique, this equation implies $\lambda_1 = \dots = \lambda_{n+1} = \lambda$. $\square$

Let us denote by $Z(\mathbb{F}^*)$ the center of the multiplicative group of $\mathbb{F}$, that is the group of those non-zero elements that commute with all elements of $\mathbb{F}^*$

$$Z(\mathbb{F}^*) = \{\lambda \in \mathbb{F}^* \mid \lambda\mu = \mu\lambda \ \forall \mu \in \mathbb{F}^*\}$$

The division ring $\mathbb{F}$ is a field if and only if $Z(\mathbb{F}^*) = \mathbb{F}^*$

6.1.4. PROPOSITION. *Two linear transformations $\Phi, \Psi \in GL(n+1, \mathbb{F})$ induce the same collineation $[\Phi] = [\Psi]$ if and only if $\Phi \circ \Psi^{-1}$ is multiplication with a scalar $\lambda \in Z(\mathbb{F}^*)$.*

PROOF. $[\Phi] = [\Psi]$ if and only if $[\Phi \circ \Psi^{-1}]$ fixes every point of the projective space. This happens if and only if every non-zero vector is an eigenvector of the linear transformation $\Lambda = \Phi \circ \Psi^{-1}$.

Let us choose a basis $\mathbf{v}_1, \dots, \mathbf{v}_{n+1}$ and apply Lemma 6.1.3. According to the Lemma, there exists a scalar $\lambda \in \mathbb{F}^*$ such that $\Lambda(\mathbf{v}_i) = \lambda \mathbf{v}_i$ for each basis vector. If $\mu \in \mathbb{F}^*$ is a non-zero scalar, then $\mu\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n+1}$ is also a basis, to which we can apply the Lemma. In this case it yields that the eigenvalue corresponding to $\mu\mathbf{v}_1$, that is $\mu\lambda\mu^{-1}$ is equal to the eigenvalue corresponding to $\mathbf{v}_2$, that is λ. The equality $\mu\lambda\mu^{-1} = \lambda$ for all $\mu \in \mathbb{F}^*$ means $\lambda \in Z(\mathbb{F}^*)$. Now choose a vector $\mathbf{x}$, write it as a linear combination of the basis vectors and apply to it Λ. What we get is

$$\begin{aligned}
\Lambda(\mathbf{x}) &= \Lambda(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \dots + x_{n+1}\mathbf{v}_{n+1}) \\
&= x_1\Lambda(\mathbf{v}_1) + x_2\Lambda(\mathbf{v}_2) + \dots + x_{n+1}\Lambda(\mathbf{v}_{n+1}) \\
&= x_1\lambda\mathbf{v}_1 + x_2\lambda\mathbf{v}_2 + \dots + x_{n+1}\lambda\mathbf{v}_{n+1} \\
&= \lambda x_1\mathbf{v}_1 + \lambda x_2\mathbf{v}_2 + \dots + \lambda x_{n+1}\mathbf{v}_{n+1} = \lambda\mathbf{x},
\end{aligned}$$

as it was to be proved. □

6.1.5. COROLLARY. *The projective general group $PGL(n+1, \mathbb{F})$ is the factor group of the general linear group $GL(n+1, \mathbb{F})$ with respect to the normal subgroup consisting of multiplications by scalars in $Z(\mathbb{F}^*)$. Since the latter group is isomorphic to $Z(\mathbb{F}^*)$, we can write*

$$PGL(n+1, \mathbb{F}) \cong GL(n+1, \mathbb{F})/Z(\mathbb{F}^*).$$

6.1.6. DEFINITION. Let X be a set, G be a group. We say that the group G *acts on* X, if we are given a mapping

$$T : G \times X \to X$$

such that denoting $T(g, x)$ shortly by gx we have
- $1x = x \ \forall x \in X$, where $1 \in G$ is the unit element of the group;
- $g_1(g_2x) = (g_1g_2)x \ \forall x \in X,\ g_1, g_2 \in G$.

If a G-action on X is given by the mapping T, then we can assign to every group element $g \in G$ a transformation $T_g : X \to X$ by the rule $T_g(x) = gx$. The first axiom of group actions is equivalent to the requirement that T_1 is the is the identity of X, the second axiom says that $T_{g_1} \circ T_{g_2} = T_{g_1g_2}$. From these two axioms we can derive that T_g is a bijection (a permutation of X) and its inverse is $T_{g^{-1}}$. Thus, we can think of a group action also as a group homomorphism from the group G to the permutation group of X.

For example, the collineation group and the projective general linear group act on the projective space, the general linear group acts on the linear space in an obvious way. When a group acts on a set X its action defines an action on the k-point subsets, or more generally, on the family of all subsets of X. The projective general linear group acts also on straight lines, on k-dimensional subspaces. The isometry group of the euclidean plane act on triangles, polygons, circles, ellipses, etc.

6.1.7. Definition. We say that a group action $T : G \times X \to X$ is *transitive* if for any $x, y \in X$ there exists $g \in G$ such that $gx = y$. The action is said to be *simply transitive* if the group element g with this property is unique for any pair of points.

6.1.8. Theorem. *The group $PGL(n+1, \mathbb{F})$ acts transitively on ordered $(n+2)$-tuples in general position.*

Proof. Let $A_1, \ldots, A_{n+2}$ $B_1, \ldots, B_{n+2}$ be two ordered $(n+2)$-tuples in general position. By Proposition 2.5.4, we can choose representatives $\mathbf{a}_1, \ldots, \mathbf{a}_{n+1}$ for $A_1, \ldots, A_{n+1}$ and $\mathbf{b}_1, \ldots, \mathbf{b}_{n+1}$ for $B_1, \ldots, B_{n+1}$ such that

$$[\mathbf{a}_1 + \cdots + \mathbf{a}_{n+1}] = A_{n+2} \text{ and } [\mathbf{b}_1 + \cdots + \mathbf{b}_{n+1}] = B_{n+2}.$$

According to the assumption on general position, the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_{n+1}$ are linearly independent, so they form a basis of the linear space $\mathbb{F}^{n+1}$. Similarly, the vectors $\mathbf{b}_1, \ldots, \mathbf{b}_{n+1}$ yield a basis of $\mathbb{F}^{n+1}$ as well. Thus, there is (a unique) linear transformation $\Phi \in GL(n+1, \mathbb{F})$ such that Φ takes $\mathbf{a}_i$ to $\mathbf{b}_i$ for $1 \leq i \leq n+1$. Then by linearity of Φ, we have

$$\Phi(\mathbf{a}_1 + \cdots + \mathbf{a}_{n+1}) = \mathbf{b}_1 + \cdots + \mathbf{b}_{n+1},$$

hence the collineation $[\Phi] \in PGL(n+1, \mathbb{F})$ maps A_i to B_i for $1 \leq i \leq n+2$. □

6.1.9. Corollary. *If $\mathbb{F}$ is a field, then the group $PGL(n+1, \mathbb{F})$ acts simply transitively on ordered $(n+2)$-tuples in general position.*

Proof. If there were two different projective general linear transformations $[\Phi]$ and $[\Psi]$ that mapped A_i to B_i for $i = 1, \ldots, n+2$, then the transformation $[\Phi^{-1} \circ \Psi] \in PGL(n+1, \mathbb{F})$ would be different from identity and would fix the points $A_1, \ldots, A_{n+2}$. Therefore it suffices to show that if the transformation $[\Lambda] \in PGL(n+1, \mathbb{F})$, with $\Lambda \in GL(n+1, \mathbb{F})$ fixes $n+2$ points in general position, say $A_1, \ldots, A_{n+2}$, then $[\Lambda]$ is the identity. Choosing representatives $\mathbf{a}_1, \ldots, \mathbf{a}_{n+1}$ for $A_1, \ldots, A_{n+1}$ in such a way that the sum $\mathbf{a}_1 + \cdots + \mathbf{a}_{n+1}$ represents A_{n+2} we obtain a basis consisting of eigenvectors of Λ such that the sum of the basis vectors is also an eigenvector of Λ. But then, by Lemma 6.1.3. above, Λ is a multiplication with a scalar $\lambda \in \mathbb{F}^*$ and therefore $[\Lambda]$ is identical. □

6.2. Collineations induced by automorphisms of $\mathbb{F}$.

Recall that an *automorphism* of a division ring (or field) $\mathbb{F}$ is a bijection $\phi : \mathbb{F} \to \mathbb{F}$ such that

- $\phi(0) = 0$, $\phi(1) = 1$;
- $\phi(x + y) = \phi(x) + \phi(y)$ for all $x, y \in \mathbb{F}$;
- $\phi(xy) = \phi(x)\phi(y)$ for all $x, y \in \mathbb{F}$.

Automorphisms of $\mathbb{F}$ form a group with respect to composition. We shall denote this group by $Aut(\mathbb{F})$. If $\mathbb{F}$ is not commutative, the automorphism group contains a non-trivial subgroup $Aut_0(\mathbb{F})$ consisting of conjugations, i.e. transformations of the form $\phi_a(x) = a^{-1}xa$, where $a \in \mathbb{F}^*$. Since $\phi_a = \phi_b$ if and only if $ab^{-1} \in Z(\mathbb{F}^*)$, this subgroup is isomorphic to the factor group $\mathbb{F}^*/Z(\mathbb{F}^*)$.

Let ϕ be an automorpism of $\mathbb{F}$. Consider the transformation $\tilde{\phi} : \mathbb{F}^{n+1} \to \mathbb{F}^{n+1}$ defined by

$$\tilde{\phi} : (x_1, x_2, \ldots, x_{n+1}) \mapsto (\phi(x_1), \phi(x_2), \ldots, \phi(x_{n+1})).$$

$\tilde{\phi}$ is not a linear transformation since instead of linearity it satisfies

$$\tilde{\phi}(\lambda\mathbf{x} + \mu\mathbf{y}) = \phi(\lambda)\tilde{\phi}(\mathbf{x}) + \phi(\mu)\tilde{\phi}(\mathbf{y}),$$

nevertheless, it maps k-dimensional linear subspaces of $\mathbb{F}^{n+1}$ onto k-dimensional linear subspaces. In particular, it induces a collineation $[\tilde{\phi}]$ of $\mathbb{F}P^n$ by the formula

$$[\tilde{\phi}][\mathbf{x}] = [\tilde{\phi}(\mathbf{x})].$$

Since $[\widetilde{\phi \circ \psi}] = [\tilde{\phi}] \circ [\tilde{\psi}]$ and $[\tilde{\phi}]^{-1} = [\widetilde{\phi^{-1}}]$, collineations that are induced from an automorphism of $\mathbb{F}$ form a subgroup $\widetilde{Aut}(\mathbb{F})$ of the collineation group $Coll(\mathbb{F}P^n)$.

If an automorphism ϕ is different from the identity of $\mathbb{F}$, then $[\tilde{\phi}]$ is different from the identity of $\mathbb{F}P^n$, since if $\phi(x) \neq x$, then

$$[\tilde{\phi}](x : 1 : 1 : \cdots : 1) = (\phi(x) : 1 : 1 : \cdots : 1) \neq (x : 1 : 1 : \cdots : 1).$$

Thus, the group $\widetilde{Aut(\mathbb{F})}$ is isomorphic to the group of automorphisms of $\mathbb{F}$.

The following theorem gives a characterization of those collineations that are induced from a field automorphism.

6.2.1. Theorem. *Let $\mathbf{e}_1, \dots, \mathbf{e}_{n+1}$ be the standard basis in $\mathbb{F}^{n+1}$, where $\mathbf{e}_i$ is the $0-1$ vector whose only non-zero coordinate is at the i-th position. Let $P_i \in \mathbb{F}P^n$ be the point represented by $\mathbf{e}_i$ and let Q be the point represented by $\mathbf{e}_1 + \cdots + \mathbf{e}_{n+1}$. Then a collineation Ψ is induced from a field automorphism if and only if Ψ fixes the points $P_1, \dots, P_{n+1}$ and Q.*

Proof. Since a field automorphism fixes 0 and 1, every collineation in $\widetilde{Aut}(\mathbb{F})$ leaves the points P_i, $i = 1, \dots, n+1$ and Q fixed.

Assume now that Ψ is a collineation fixing the points P_i, $i = 1, \dots, n+1$ and Q. In this case, Ψ maps the line P_iP_{n+1} into itself, thus we can find a bijection $\phi_i : \mathbb{F} \to \mathbb{F}$ such that

$$\Psi([\lambda\mathbf{e}_i + \mathbf{e}_{n+1}]) = [\phi_i(\lambda)\mathbf{e}_i + \mathbf{e}_{n+1}] \text{ for } i = 1, \dots, n.$$

6.2.2. Lemma. *The bijections $\phi_i : \mathbb{F} \to \mathbb{F}$ are in $Aut(\mathbb{F})$.*

Proof. By exercises 2-3 and 2-4, the mapping $\theta : \mathbb{F} \to P_iP_{n+1}$ defined by $\lambda \mapsto [\lambda\mathbf{e}_i + \mathbf{e}_{n+1}]$ is an isomorphism between the division ring $\mathbb{F}$ and the division ring $\mathbb{F}_{P_{n+1}Q_iP_i}$, where $Q_i = [\mathbf{e}_i + \mathbf{e}_{n+1}]$. (For the definition of the division ring $\mathbb{F}_{P_{n+1}Q_iP_i}$ see section **5.1**.) Ψ fixes P_i, P_{n+1} and Q_i. Q_i is fixed since it is the intersection point of the line P_iP_{n+1} with the hyperplane spanned by the points $\{P_j \mid 1 \leq j \leq n, i \neq j\}$ and Q. This implies that the restriction of Ψ onto the line P_iP_{n+1} gives an automorphis of $\mathbb{F}_{P_{n+1}Q_iP_i}$. Therefore, $\phi_i = \theta^{-1} \circ \Psi \circ \theta$ is an automorphism of $\mathbb{F}$. $\square$

Let us consider the hyperplanes

$$H^i_\lambda = \{(x_1 : x_2 : \cdots : x_{n+1}) \mid x_i = x_{n+1}\lambda\}$$

where $i = 1, 2, \dots, n$, $\lambda \in \mathbb{F}$.

H^i_λ is spanned by the points $\{P_j \mid 1 \le j \le n,\ i \ne j\}$ and $P^i_\lambda = [\lambda \mathbf{e}_i + \mathbf{e}_{n+1}]$. Since the points $\{P_j \mid 1 \le j \le n\, i \ne j\}$ are fixed by Ψ, while $\Psi(P^i_\lambda) = P^i_{\phi_i(\lambda)}$,

$$\Psi(H^i_\lambda) = H^i_{\phi_i(\lambda)}.$$

This allows us to determine the Ψ-image of a point with homogeneous coordinates $(\lambda_1 : \cdots : \lambda_n : 1)$. Indeed, since this point is the intersection of the hyperplanes $H^i_{\lambda_i}$, $i = 1, \ldots, n$, we have

$$\Psi(\lambda_1 : \cdots : \lambda_n : 1) = \Psi\left(\cap_{i=1}^n H^i_{\lambda_i}\right) = \cap_{i=1}^n H^i_{\phi_i(\lambda_i)} = (\phi_1(\lambda_1) : \cdots : \phi_n(\lambda_n) : 1)$$

6.2.3. Lemma. *The automorphisms $\phi_1, \ldots, \phi_n$ are equal.*

Proof. Let $\lambda \in \mathbb{F}$ be an arbitrary scalar and consider the point $(\lambda : \cdots : \lambda : 1)$. This point is on the line $P_{n+1}Q$. Since the points P_{n+1} and Q are fixed by Ψ, the image of $(\lambda : \cdots : \lambda : 1)$, that is $(\phi_1(\lambda) : \cdots : \phi_n(\lambda) : 1)$ is also on the line $P_{n+1}Q$. However, points of the line $P_{n+1}Q$ are characterized by the property that their first n homogeneous coordinates are equal. □

Let us denote the automorphism $\phi_1 = \cdots = \phi_n$ simply by ϕ and show that $\Psi = [\tilde{\phi}]$. Let us take an arbitrary point $(\lambda_1 : \cdots : \lambda_{n+1}) \in \mathbb{F}P^n$. If $\lambda_{n+1} \ne 0$, then we already know that

$$\begin{aligned}\Psi(\lambda_1 : \cdots : \lambda_{n+1}) &= \Psi(\lambda_{n+1}^{-1}\lambda_1 : \cdots : \lambda_{n+1}^{-1}\lambda_n : 1)\\ &= (\phi(\lambda_{n+1}^{-1}\lambda_1) : \cdots : \phi(\lambda_{n+1}^{-1}\lambda_n) : 1)\\ &= (\phi(\lambda_1) : \cdots : \phi(\lambda_{n+1})) = [\tilde{\phi}](\lambda_1 : \cdots : \lambda_{n+1}).\end{aligned}$$

If $\lambda_{n+1} = 0$, then let us represent the point $(\lambda_1 : \cdots : \lambda_n : 0)$ as the intersection point of the line through P_{n+1} and $(\lambda_1 : \cdots : \lambda_n : 1)$ with the hyperplane spanned by $P_1, \ldots, P_n$. This representation shows us that $\Psi(\lambda_1 : \cdots : \lambda_n : 0)$ is the intersection point of the straight line through P_{n+1} and $(\phi(\lambda_1) : \cdots : \phi(\lambda_n) : 1)$ with the hyperplane spanned by $P_1, \ldots, P_n$, i.e.,

$$\Psi(\lambda_1 : \cdots : \lambda_n : 0) = (\phi(\lambda_1) : \cdots : \phi(\lambda_n) : 0). \quad □$$

Now we are ready to formulate

6.3. The Fundamental Theorem of Projective Geometry.

- *Any collineation of the projective space $\mathbb{F}P^n$ can be obtained as the composition of a projective general linear transformation and a collineation induced by an automorphism of $\mathbb{F}$.*
- *The intersection of the subgroups $PGL(n+1, \mathbb{F})$ and $\widetilde{Aut(\mathbb{F})}$ consists of collineations corresponding to conjugations in $Aut_0(\mathbb{F})$. In particular,*

$$PGL(n+1, \mathbb{F}) \cap \widetilde{Aut(\mathbb{F})} \cong Aut_0(\mathbb{F}) \cong \mathbb{F}^*/Z(\mathbb{F}^*).$$

- *$PGL(n+1, \mathbb{F})$ is a normal subgroup of $Coll(\mathbb{F}P^n)$*

Proof.

• Let Φ be an arbitrary collineation of $\mathbb{F}P^n$. Let $\mathbf{e}_1, \dots, \mathbf{e}_{n+1}$ be the standard basis of $\mathbb{F}^{n+1}$, $P_i = [\mathbf{e}_i]$, $Q = [\mathbf{e}_1 + \dots + \mathbf{e}_{n+1}]$. Since the group $PGL(n+1, \mathbb{F})$ acts transitively on ordered $(n+2)$-tuples in general position, one can find a projective general linear transformation Ψ_0 such that $\Phi(P_i) = \Psi_0(P_i)$ for $i = 1, \dots, n+1$ and $\Phi(Q) = \Psi_0(Q)$. But then the collineation $\Psi_1 = \Psi_0^{-1} \circ \Phi$ leaves the points $P_1, \dots, P_{n+1}$ and Q fixed, therefore $\Psi_1 \in \widetilde{Aut(\mathbb{F})}$. The decomposition $\Phi = \Psi_0 \circ \Psi_1$ is what we wanted to achieve.

• If $\Lambda \in GL(n+1, \mathbb{F})$ yields a collineation which is also in $\widetilde{Aut(\mathbb{F})}$, then $[\Lambda]$ fixes the points $P_1, \dots, P_{n+1}$ and Q. This means, that the basis vectors $\mathbf{e}_1, \dots, \mathbf{e}_{n+1}$ and their sum are eigenvectors of Λ. If λ is the common eigenvalue corresponding to these vectors, then we have

$$\begin{aligned}[\Lambda](\mu_1 : \dots : \mu_{n+1}) &= (\mu_1\lambda : \dots : \mu_{n+1}\lambda) \\ &= (\lambda^{-1}\mu_1\lambda : \dots : \lambda^{-1}\mu_{n+1}\lambda) = [\tilde{\phi}_\lambda](\mu_1 : \dots : \mu_{n+1}),\end{aligned}$$

where ϕ_λ is the conjugation with λ.

• Recall that a subgroup H of a group G is a *normal subgroup* if for any $h \in H$ and $g \in G$, the conjugate $g^{-1}hg$ is also in H.

Let $\Phi \in PGL(n+1, \mathbb{F})$, $\Psi \in Coll(\mathbb{F}P^n)$ be arbitrary collineations. Let us write Ψ as $\Psi_0 \circ \Psi_1$, where $\Psi_0 \in PGL(n+1, \mathbb{F})$ and $\Psi_1 \in \widetilde{Aut(\mathbb{F})}$. Then we have

$$\Psi^{-1} \circ \Phi \circ \Psi = \Psi_1^{-1} \circ (\Psi_0^{-1} \circ \Phi \circ \Psi_0) \circ \Psi_1.$$

Since here $(\Psi_0^{-1} \circ \Phi \circ \Psi_0)$ is in the projective general linear group, it suffices to show that the conjugation of a projective general linear transformation with a collineation induced by an automorphism of $\mathbb{F}$ is a projective general linear transformation. This follows from the analogous statement for transformations of $\mathbb{F}^{n+1}$: If Λ is a linear transformation of $\mathbb{F}^{n+1}$, $\phi \in Aut(\mathbb{F})$, then $\tilde{\phi}^{-1} \circ \Lambda \circ \tilde{\phi}$ is a linear transformation. Indeed,

$$\begin{aligned}\tilde{\phi}^{-1} \circ \Lambda \circ \tilde{\phi}(\lambda\mathbf{a} + \mu\mathbf{b}) &= \tilde{\phi}^{-1} \circ \Lambda(\phi(\lambda)\tilde{\phi}(\mathbf{a}) + \phi(\mu)\tilde{\phi}(\mathbf{b})) \\ &= \tilde{\phi}^{-1}(\phi(\lambda)\Lambda(\tilde{\phi}(\mathbf{a})) + \phi(\mu)\Lambda(\tilde{\phi}(\mathbf{b}))) \\ &= \phi^{-1}(\phi(\lambda))\tilde{\phi}^{-1}(\Lambda(\tilde{\phi}(\mathbf{a}))) + \phi^{-1}(\phi(\mu))\tilde{\phi}^{-1}(\Lambda(\tilde{\phi}(\mathbf{b}))) \\ &= \lambda\tilde{\phi}^{-1}(\Lambda(\tilde{\phi}(\mathbf{a}))) + \mu\tilde{\phi}^{-1}(\Lambda(\tilde{\phi}(\mathbf{b}))). \quad \square\end{aligned}$$

Let us recall a definition from algebra.

6.3.1. Definition. Let H_1 and H_2 be two subgroups of a group G. We say that G is the *semidirect product of the subgroups* H_1 *and* H_2 and write $G = H_1 \rtimes H_2$ or $G = H_2 \ltimes H_1$ if

- any $g \in G$ can be written as the product of an element of H_1 and an element of H_2;
- the intersection $H_1 \cap H_2$ contains only the unit element of G;
- H_1 is a normal subgroup of G.

If the division ring $\mathbb{F}$ is a field, then $\mathbb{F}^* = Z(\mathbb{F}^*)$ and the Fundamental Theorem of Projective Geometry gives the following

6.3.2. Corollary. *If $\mathbb{F}$ is a field, then*

$$Coll(\mathbb{F}P^n) \cong PGL(n+1, \mathbb{F}) \rtimes \widetilde{Aut(\mathbb{F})}.$$

When we apply the Fundamental Theorem we have to know the automorphism group of $\mathbb{F}$. Here are some examples.

If $\mathbb{F} = GF(p^r)$ is the Galois field of order p^r, then $Aut(\mathbb{F})$ is a cyclic group of order r. The generator of the automorphism group is the automorphism $\phi(x) = x^p$.

If $\mathbb{F} = \mathbb{Q}$ is the field of rational numbers, then $\mathbb{F}$ has no non-trivial automorphism. Indeed, if $\phi : \mathbb{Q} \to \mathbb{Q}$ is a field automorphism, then using induction on n we can show first that $\phi(n) = n$ for natural numbers. For initial values $n = 0, 1$ the statement follows from the definition of a field automorphism, the induction step comes from the identity $\phi(n+1) = \phi(n) + \phi(1) = \phi(n) + 1$.

If $-n$ is a negative integer, then from

$$0 = \phi(0) = \phi(n + (-n)) = \phi(n) + \phi(-n) = n + \phi(-n)$$

we can derive that $\phi(-n) = -n$, thus ϕ fixes all integers.

If q is an arbitrary rational number, then it has the form $q = m/n$, where m and n are integers. Applying ϕ to q we obtain $\phi(q) = \phi(m)/\phi(n) = m/n = q$, from which we conclude that ϕ is the identity.

6.3.3. Corollary. *The collineation group of the projective space $\mathbb{Q}P^n$ over the rational numbers is isomorphic to the group $PGL(n+1, \mathbb{Q})$.*

The most important field for (classical) geometry is the field $\mathbb{F} = \mathbb{R}$ of real numbers. Let us prove that *the field of real numbers has no non-trivial automorphisms.*

If $\phi : \mathbb{R} \to \mathbb{R}$ is a field automorphism, then we can show that $\phi(q) = q$ for all rational numbers just as for the case $\mathbb{F} = \mathbb{Q}$.

The key observation is that positive real numbers can be characterized in terms of multiplication. Namely, a real number is positive if and only if it is the square of a non-zero number. This characterization shows that ϕ maps positive numbers to positive numbers.

ϕ preserves ordering as well. Indeed, if $a < b$, then $b - a$ is positive, therefore $\phi(b - a) = \phi(b) - \phi(a)$ is positive,

If ϕ were not the identity, then we could find a real number a such that $\phi(a) \neq a$. Assume for example that $a < \phi(a)$. Since the set of rational numbers is dense in $\mathbb{R}$, the interval $(a, \phi(a))$ contains a rational number q. This leads to a contradiction since $a < q$, but $\phi(a) > q = \phi(q)$. The case $a > \phi(a)$ can be treated similarly.

6.3.4. Corollary. *The collineation group of the real projective space $\mathbb{R}P^n$ is isomorphic to the group $PGL(n+1, \mathbb{R})$.*

The field $\mathbb{F} = \mathbb{C}$ of complex numbers is less rigid. It has infinite $(2^{\aleph_0})$ automorphisms. These automorphisms fix the rational numbers but they can thoroughly mix up other real numbers. The only non-trivial continuous automorphism of $\mathbb{C}$ is the complex conjugation $z \mapsto \bar{z}$.

PROJECTIVE GEOMETRY

.

7. CROSS-RATIO PRESERVING TRANSFORMATIONS BETWEEN LINES

From this point on we shall consider only projective spaces $\mathbb{F}P^n$ over a *field* $\mathbb{F}$.

7.1. Cross-ratio.

7.1.1. Definition. Let A, B, C, D be four different points on a straight line, and let $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ be arbitrary representatives of these points. Since the points are collinear, we can split $\mathbf{c}$ and $\mathbf{d}$ as the linear combination of $\mathbf{a}$ and $\mathbf{b}$

$$\begin{aligned}\mathbf{c} &= \gamma_1\mathbf{a} + \gamma_2\mathbf{b}\\ \mathbf{d} &= \delta_1\mathbf{a} + \delta_2\mathbf{b}.\end{aligned}$$

Then the number

$$(A, B, C, D) = \frac{\gamma_2}{\gamma_1} : \frac{\delta_2}{\delta_1}$$

is called the *cross-ratio of the points* A, B, C, D.

7.1.2. Proposition. *The cross-ratio depends only on the points A, B, C, D but not on the choice of the representatives.*

Proof. If we choose another system of representatives

$$\mathbf{a}' = x\mathbf{a}, \quad \mathbf{b}' = y\mathbf{b}, \quad \mathbf{c}' = z\mathbf{c}, \quad \mathbf{d}' = w\mathbf{d},$$

then we have

$$\begin{aligned}\mathbf{c}' &= \frac{z}{x}\gamma_1\mathbf{a}' + \frac{z}{y}\gamma_2\mathbf{b}'\\ \mathbf{d}' &= \frac{w}{x}\delta_1\mathbf{a}' + \frac{w}{y}\delta_2\mathbf{b}'.\end{aligned}$$

If we compute the cross-ratio with the help of the new system of representatives, we obtain

$$\frac{\frac{z}{y}\gamma_2}{\frac{z}{x}\gamma_1} : \frac{\frac{w}{y}\delta_2}{\frac{w}{x}\delta_1} = \frac{\gamma_2}{\gamma_1} : \frac{\delta_2}{\delta_1}. \quad \square$$

We list some important properties of the cross-ratio.

7.1.3. Proposition.

(i) *The cross-ratio depends on the order of A, B, C, D in the following way*

$$(A, B, C, D) = \frac{1}{(B, A, C, D)} = \frac{1}{(A, B, D, C)} = 1 - (A, C, B, D).$$

(Iterating these transpositions, we can compute the cross-ratio of any permutation.)

(ii) *If $\mathbf{a}$ and $\mathbf{b}$ are linearly independent vectors, and for $x \in \mathbb{F}$ we denote by P_x the point $[\mathbf{a} + x\mathbf{b}]$, then*

$$(P_x, P_y, P_z, P_w) = \frac{x - z}{z - y} : \frac{x - w}{w - y}.$$

This formula follows at once from the decompositions

$$\mathbf{a} + z\mathbf{b} = \frac{z-y}{x-y}(\mathbf{a} + x\mathbf{b}) + \frac{x-z}{x-y}(\mathbf{a} + y\mathbf{b}),$$

$$\mathbf{a} + w\mathbf{b} = \frac{w-y}{x-y}(\mathbf{a} + x\mathbf{b}) + \frac{x-w}{x-y}(\mathbf{a} + y\mathbf{b}).$$

(iii) *In particular,if $\mathbb{F} = \mathbb{R}$ and we represent the projective space as the union of the euclidean space and the hyperplane at infinity,* $\mathbf{b}$ *is a unit vector that represents a point at infinity, then the correspondence $x \leftrightarrow P_x = [\mathbf{a} + x\mathbf{b}]$ is a euclidean coordinate system (sometimes called a "ruler") on the line. This coordinate system gives the line an orientation and the difference $x - y$ is just the signed distance of P_x and P_y, which we denote by $\overrightarrow{P_yP_x}$. Thus we obtain that if A, B, C, D are four different points on a directed straight line in the Euclidean space, then*

$$(A, B, C, D) = \frac{\overrightarrow{AC}}{\overrightarrow{CB}} \frac{\overrightarrow{AD}}{\overrightarrow{DB}}.$$

(iv) *If A, B, C are collinear points in the euclidean space, D is the point at infinity of the line AB, then the cross-ratio (A, B, C, D) can also be expressed in terms of distances as*

$$(A, B, C, D) = -\frac{\overrightarrow{AC}}{\overrightarrow{CB}}.$$

(v) $(A, B, C, D) < 0$ *if and only if A, B separates C, D on the projective line.*

(vi) *The cross-ratio of four different points is never equal to* 0 *or* 1. *If we are given three different points A, B, C on a straight line e and a number $x \in \mathbb{F}\backslash\{0, 1\}$, then there exists a unique point $D \in e$ different from A, B, C such that $(A, B, C, D) = x$.*

Proof. The proof is left to the reader. □

It is not difficult to follow what happens with the cross-ratio when we apply a collineation to the points A, B, C, D.

7.1.4. Proposition. *Let A, B, C, D be four different points on a line in $\mathbb{F}P^n$.*

- *If $\Phi \in PGL(n+1, \mathbb{F})$, then*

$$(\Phi(A), \Phi(B), \Phi(C), \Phi(D)) = (A, B, C, D).$$

- *If $\Phi = [\tilde{\phi}]$ is a collineation induced by the field automorphism ϕ, then*

$$(\Phi(A), \Phi(B), \Phi(C), \Phi(D)) = \phi((A, B, C, D)).$$

Proof. • Assume first that $\Phi = [\Lambda]$, where Λ is a linear transformation. Then the points $\Phi(A), \Phi(B), \Phi(C), \Phi(D)$ are represented by the vectors $\Lambda(\mathbf{a}), \Lambda(\mathbf{b}), \Lambda(\mathbf{c}), \Lambda(\mathbf{d})$ respectively, where $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$ are representatives of A, B, C, D respectively. Now if

$$\mathbf{c} = \gamma_1\mathbf{a} + \gamma_2\mathbf{b}$$
$$\mathbf{d} = \delta_1\mathbf{a} + \delta_2\mathbf{b},$$

then

$$\Lambda(\mathbf{c}) = \gamma_1\Lambda(\mathbf{a}) + \gamma_2\Lambda(\mathbf{b})$$
$$\Lambda(\mathbf{d}) = \delta_1\Lambda(\mathbf{a}) + \delta_2\Lambda(\mathbf{b}),$$

therefore

$$(\Phi(A), \Phi(B), \Phi(C), \Phi(D)) = \frac{\gamma_2}{\gamma_1} : \frac{\delta_2}{\delta_1} = (A, B, C, D).$$

• If $\Phi = [\tilde{\phi}]$ is a collineation induced by the field automorphism ϕ, then the points $\Phi(A), \Phi(B), \Phi(C), \Phi(D)$ are represented by the vectors $\tilde{\phi}(\mathbf{a}), \tilde{\phi}(\mathbf{b}), \tilde{\phi}(\mathbf{c}), \tilde{\phi}(\mathbf{d})$, for which we have

$$\tilde{\phi}(\mathbf{c}) = \phi(\gamma_1)\tilde{\phi}(\mathbf{a}) + \phi(\gamma_2)\tilde{\phi}(\mathbf{b})$$
$$\tilde{\phi}(\mathbf{d}) = \phi(\delta_1)\tilde{\phi}(\mathbf{a}) + \phi(\delta_2)\tilde{\phi}(\mathbf{b}),$$

thus

$$(\Phi(A), \Phi(B), \Phi(C), \Phi(D)) = \frac{\phi(\gamma_2)}{\phi(\gamma_1)} : \frac{\phi(\delta_2)}{\phi(\delta_1)} = \phi((A, B, C, D)). \quad \square$$

7.1.5. Corollary. *A collineation of the projective space $\mathbb{F}P^n$ preserves cross-ratio if and only if it is a projective general linear transformation.*

7.2. Characterizations of cross-ratio preserving transformations between straight lines.

7.2.1. Theorem. *Let $\phi : e \to f$ be a bijection between two straight lines. Then the following three conditions on ϕ are equivalent:*
(i) ϕ preserves cross-ratio;
(ii) ϕ extends to a projective general linear transformation of the whole space;
(iii) ϕ is the composition of a finite number of perspectivities. (For the definition of perspectivity see 5.1.4.)

Proof. The implication $(ii) \Longrightarrow (i)$ has already been proved in 7.1.4.

To prove $(iii) \Longrightarrow (ii)$ it suffices to show that every perspectivity extends to a projective general linear transformation. Consider the perspectivity $\psi : e \to f$ with center O. Set $M = e \cup f$ and let $P \in e$ be an arbitrary point different from M, $P' = \psi(P)$. Choose representative vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ for M, P, O respectively in such a way that $\mathbf{b} + \mathbf{c}$ represents P'. The vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ are linearly independent, therefore they can be extended to a basis $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{e}_1, \dots, \mathbf{e}_s$. Let L be the linear transformation which maps $\mathbf{b}$ to $\mathbf{b} + \mathbf{c}$ and fixes all the other vectors of this basis.

We claim that the collineation associated to L is an extension of ψ. Any point $Q \in e$ can be represented by a linear combination $\lambda\mathbf{a} + \mu\mathbf{b}$. $[L]$ maps such a point Q to the point $Q' = [\lambda\mathbf{a} + \mu(\mathbf{b} + \mathbf{c})] = [(\lambda\mathbf{a} + \mu\mathbf{b}) + \mu\mathbf{c}]$. It is clear that Q' is on the lines $f = MP'$ and also on the line OP, thus, $Q' = \psi(Q)$.

Before proving the implication (i)$\Longrightarrow$(iii), let us formulate a corollary of the implication (iii)$\Longrightarrow$(i).

7.2.2. Pappus' Theorem. *Perspectivities preserve cross-ratio.*

It remains to show that (i)$\Longrightarrow$(iii). Let A, B, C be three different points on e, A', B', C' be their images. First we show that we can map A, B, C to A', B', C' respectively by the composition of some perspectivities. We distinguish three cases.

Case a: *e and f are different and coplanar.* In this case, $e \cap f = M$ is a point. M can be equal to at most one of A, B, C and at most one of A', B', C', thus we may assume without loss of generality that $A \neq M \neq A'$. Let g be the line through the intersection points $A" = AB' \cap A'B$ and $C" = AC' \cap A'C$, and set $A" = g \cap AA'$. The perspectivity from e to g with center A' takes the points A, B, C to $A", B", C"$ respectively, while the perspectivity from g to f with center A takes the points $A", B", C"$ to A', B', C' respectively, so the composition of these two perspectivities maps A, B, C to A', B', C'.

Case b: $e = f$. Choose a straight line g coplanar with e but different from it, and an arbitrary perspectivity from e to g that takes A, B, C to A^*, B^*, C^* respectively. Applying Case a, we can construct two perspectivities, the composition of which takes A^*, B^*, C^* to A', B', C' respectively, therefore, we can map A, B, C to A', B', C' with the help of three perspectivities.

Case c: *e and f are skew.* This case can be handled similarly to Case b. Choose a point E in e and F in f and let g be the line EF. Suppose that an arbitrary perspectivity takes A, B, C to A^*, B^*, C^*. Applying Case a, we can construct two perspectivities, the composition of which takes A^*, B^*, C^* to A', B', C' respectively, therefore, we can map A, B, C to A', B', C' with the help of three perspectivities.

Denote by $\tilde{\phi}$ the composition of perspectivities that takes A, B, C to A', B', C'. We show that $\phi = \tilde{\phi}$. Since both ϕ and $\tilde{\phi}$ preserve cross-ratio and they coincide on the points A, B, C we need only the following proposition.

7.2.3. Proposition. *If ϕ and $\tilde{\phi}$ are arbitrary cross-ratio preserving transformations between the lines e and f and they coincide at three different points $A, B, C \in e$, then they are equal.*

Proof. Let $D \in e$ be an arbitrary point different from A, B and C. Then we have

$$\begin{aligned}(\phi(A),&\phi(B), \phi(C), \phi(D)) = (A, B, C, D) \\ &= (\tilde{\phi}(A), \tilde{\phi}(B), \tilde{\phi}(C), \tilde{\phi}(D)) = (\phi(A), \phi(B), \phi(C), \tilde{\phi}(D)),\end{aligned}$$

from which we derive by 7.1.3.(vi) that $\phi(D) = \tilde{\phi}(D)$.

This finishes the proof of Theorem 7.2.1. $\square$

Pappus theorem gives rise to the following definition

7.2.4. Definition. Let a, b, c, d be four different coplanar lines, passing through a point O. Suppose that a straight not passing through P meets a, b, c, d at the points A, B, C, D respectively. Then we define the *cross ratio of the lines a, b, c, d* by

$$(a, b, c, d) = (A, B, C, D).$$

7.3. Cross-ratio preserving transformations between coplanar lines.

7.3.1. Theorem. *Let e and f be two different coplanar straight lines, $C = e \cap f$. A cross-ratio preserving transformation $\phi : e \to f$ is a perspectivity if and only if $\phi(C) = C$.*

Proof. It is clear that if ϕ is a perspectivity, then $\phi(C) = C$.

Suppose now that ϕ is a cross ratio preserving transformation with $\phi(C) = C$. Choose two points A and B in e different from C and define O as the intersection point of $A\phi(A)$ and $B\phi(B)$. Then ϕ coincides with the perspectivity from e to f with center O at three different points (A, B, C), thus, by Proposition 7.2.3, ϕ coincides with this perspectivity.

7.3.2. Definition. We say that four collinear points A, B, C, D *form a harmonic range* if the cross-ratio (A, B, C, D) is equal to -1.

Remarks. Since the cross-ratio of four different points is never equal to 0 and 1, there are no harmonic ranges in projective spaces over fields of characteristic 2.

Since cross-ratio depends on the order of points, not every permutation of a harmonic range yields a harmonic range. A permutation of the harmonic range $ABCD$ is a harmonic range if and only if the permutation fixes the set $\{\{A, B\}, \{C, D\}\}$.

As an application of Pappus' Theorem (7.2.2.), we prove the so called Harmonic Construction Theorem, which gives a method to construct the fourth point of a harmonic range from three points of it, using only a ruler.

7.3.3. Harmonic Construction Theorem. . *Let $XYZW$ be four points in a plane such that no three of them are collinear, $A = XY \cap ZW$, $B = YZ \cap WX$ be the intersection points of the opposite sides of the quadrangle $XYZW$, $C = AB \cap XZ$, $D = AB \cap YW$ be the intersection points of the diagonals of the quadrangle $XYZW$ with the line AB. Then*

- *$C = D$ if and only if $\mathrm{char}\,\mathbb{F} = 2$;*
- *if $\mathrm{char}\,\mathbb{F} \neq 2$, then $ABCD$ are different and form a harmonic range.*

Proof. Incidence axioms of the plane imply easily, that the points X, Y, Z, W, A, B, C are all different and D may coincide only with C. In Chapter 5 we turned the set $AB \setminus \{B\}$ into a group $\mathbb{F}_{AB}$ isomorphic to the additive group of $\mathbb{F}$. Constructing the sum $C + D$ with the help of the auxiliary points $XYZW$, we see that $C + D = A$ is the 0 element of the group. Hence, $C = D$ happens if and only if $1 + 1 = 0$ in $\mathbb{F}$.

Now suppose that $\mathrm{char}\,\mathbb{F} \neq 2$. Denote by E the intersection point $XZ \cap YW$. A perspectivity with center Y takes the points $ABCD$ to $XZCE$ respectively, so by Pappus' Theorem (7.2.2.)

$$(A, B, C, D) = (X, Z, C, E).$$

Projecting back (X, Z, C, E) into the line AB from center W, we obtain

$$(X, Z, C, E) = (B, A, C, D).$$

Applying 7.1.3.(i), we get

$$(A, B, C, D) = (B, A, C, D) = \frac{1}{(A, B, C, D)},$$

from which $(A, B, C, D)^2 = 1$ or equivalently $(A, B, C, D) = \pm 1$. The cross-ratio of four different points never takes the values $0, 1$, so the cross-ratio (A, B, C, D) must be equal to -1. $\square$

7.3.4. Steiner Axis Theorem. *Let e and f be different coplanar lines, $\phi : e \to f$ a cross-ratio preserving transformation between them. Then the set of intersection points*

$$\{X\phi(Y) \cap Y\phi(X) \mid X, Y \in e,\ \phi(Y) \neq X \neq Y \neq \phi(X)\}$$

is contained in a (unique) straight line.

7.3.5. Definition. The straight line swept by the intersection points $X\phi(Y)\cap Y\phi(X)$ is called the *Steiner axis of* ϕ.

Proof. We prove the theorem first for perspectivities. Let $\phi : e \to f$ be a perspectivity with center O, $N = e \cap f$, $X \neq Y \in e$, $X' = \phi(X)$, $Y' = \phi(Y)$, $Z = XY' \cap YX'$, $g = NO$. We prove, that Z is located on a straight line passing through N, the position of which does not depend on X and Y. If X or Y coincides with N, then $Z = N$ is contained in any line through N, so it is enough to deal with the case when X and Y are different from N. In that case $Z \neq N$ and we may consider the line $h = NZ$ and the point $W = h \cap OY$. Applying the Harmonic Construction Theorem (7.3.3.) to the quadrangle $NX'ZX$, we obtain that if $\operatorname{char}\mathbb{F} = 2$ then $W = O$ and $h = NO = g$, and if $\operatorname{char}\mathbb{F} \neq 2$, then

$$(e, f, g, h) = (Y, Y', W, O) = -1.$$

In both cases, h is uniquely determined by the lines e, f, g, so it does not depend on X and Y. We conclude that h is the Steiner axis.

Suppose now that ϕ is not a perspectivity. Consider the points $N = e \cap f$, $K = \phi^{-1}(N)$, $L = \phi(N)$. K, L, N are different points by Theorem 7.3.1. We show that in this case, the line KL is the Steiner axis. Let us take two different points $X, Y \in e$ and set $X' = \phi(X)$, $Y' = \phi(Y)$, $Z = XY' \cap X'Y$. If one of X and Y coincides with K or N then Z coincides with K or L respectively, hence Z is in the line KL. Suppose now that X, Y, K, N are different. Since ϕ is cross-ratio preserving,

$$(X, Y, K, N) = (X', Y', N, L).$$

Applying Proposition 7.1.3.(i), we get

$$(X', Y', N, L) = (Y', X', L, N),$$

therefore, there is a cross-ratio preserving transformation between e and f that sends X, Y, K, N to Y', X', L, N respectively. This transformation fixes $N = e \cap f$, so that it must be a perspectivity by Theorem 7.3.1. This means, that the lines XY', YX' and KL must go through one point (the center of perspectivity), or equivalently, $Z = XY' \cap YX'$ must lie on KL. □

Remark. Theorem 2.6.3 of Pappus follows easily from the Steiner Axis Theorem, since if $A, B, C \in e$ and $A', B', C' \in f$ are different points, then there is a unique cross-ratio preserving transformation from e to f sending A, B, C to A', B', C' respectively, and then the points

$$A" = BC' \cap CB' \quad B" = CA' \cap AC' \quad C" = AB' \cap BA'$$

are located on the Steiner axis of this transformation.

7.3.6. Corollary. *A cross-ratio preserving transformation between two coplanar straight lines is a perspectivity if and only if its Steiner axis goes through the intersection points of the lines.*

Proof. Follows from the explicit description of the Steiner axis given in the proof of Theorem 7.3.4.

7.4. Cross-ratio preserving transfomations of a line, involutions.

By Theorem 7.2.1, the group of cross-ratio preserving transformations of a line into itself is isomorphic to the group $PGL(2,\mathbb{F})$. We know from Corollary 6.1.9 that this group acts simply transitively on 3-tuples in general position, where 3 points are said to be in general position if they are different. In particular, if a cross-ratio preserving transformation of a line fixes 3 points, then it must be the identity, or equivalently, a cross-ratio preserving transformation different from the identity has at most two fixed points.

7.4.1. Definition. A cross-ratio preserving transformation $\phi : e \to e$ is said to be a *hyperbolic, parabolic* or *elliptic* transformation if it has exactly 2, 1 or 0 fixed points respectively.

Fixed points of a cross-ratio preserving transformation associated to the linear transformation L correspond to eigenvectors of L. To find eigenvectors, first we have to find the eigenvalues of L, which are the roots of the characteristic polynomial $p(\lambda) = \det(L - \lambda I)$, where I is the unit matrix. When L acts on a 2-dimensional linear space, the characteristic polynomial is a second degree polynomial of the form

$$p(\lambda) = \lambda^2 - \operatorname{tr} L\lambda + \det L.$$

The number of eigenvalues of L and their multiplicities inform us on the possible numbers of fixed points of $[L]$.

It is clear that if L has two different eigenvalues in $\mathbb{F}$, then $[L]$ is a hyperbolic transformation. If L has no eigenvalues in $\mathbb{F}$ then $[L]$ is elliptic. When a second degree polynomial with coefficients in $\mathbb{F}$ has a root in $\mathbb{F}$, then its second root is also in $\mathbb{F}$, therefore, if $[L]$ is a parabolic transformation, then L must have an eigenvalue of multiplicity 2 in $\mathbb{F}$. Conversely, if L has an eigenvalue of multiplicity 2 in $\mathbb{F}$, then $[L]$ is either a parabolic transformation or the identity. Indeed, $[L]$ must have at least one fixed point, on the other hand, if $[L]$ has two different fixed points, then we can find a basis $\mathbf{a}, \mathbf{b}$ of the 2-dimensional linear space consisting of eigenvectors of L with the same eigenvalue. In that case however, L must be a constant multiple of the identity.

If $\operatorname{char} \mathbb{F} \neq 2$, then we can use the usual formula

$$\lambda_{1,2} = \frac{\operatorname{tr} L \pm \sqrt{(\operatorname{tr} L)^2 - 4 \det L}}{2}$$

to find the roots of the characteristic polynomial of L. Thus we have proved the following proposition.

7.4.2. Proposition. *Suppose* $\operatorname{char}\mathbb{F} \neq 2$. *Let* $\phi = [L]$ *be a cross-ratio preserving transformation of a straight line, different from the identity, set* $D = (\operatorname{tr} L)^2 - 4 \det L$. *Then*

- ϕ *is hyperbolic if and only if* D *is non-zero and has a square root in* $\mathbb{F}$*;*
- ϕ *is parabolic if and only if* $D = 0$*;*
- ϕ *is elliptic, if and only if* D *has no square root in* $\mathbb{F}$.

7.4.3. Definition. If G is a group with unit element e, $a \in G$, then there is a unique natural number n such that the set

$$\{m \in \mathbb{Z} | a^m = e\}$$

consists of integer multiples of n. This n is called the *order* of a. Elements of order 2 are called *involutions*. In other words, $a \in G$ is an involution if and only if $a^2 = e \neq a$.

In the following let us study involutions of the group $PGL(2, \mathbb{F})$.

7.4.4. THEOREM. *(a) Suppose* $\mathrm{char}\,\mathbb{F} \neq 2$ *Let* $L : V \to V$ *be a linear transformation of a 2-dimensional linear space,* $\phi = [L] : P(V) \to P(V)$ *the associated cross-ratio preserving transformation of the straight line* $P(V)$*. Then* ϕ *is an involution if and only if* $\mathrm{tr}\, L = 0$*.*

(b) If $\mathrm{char}\,(\mathbb{F}) = 2$*, then* ϕ *is an involution if and only if* $\mathrm{tr}\, L = 0$ *and* L *is not a scalar multiple of the identity.*

PROOF. The characteristic polynomial p of L has the form

$$p(\lambda) = \lambda^2 - (L)\lambda + \det L.$$

We shall use the Cayley-Hamilton theorem, saying that every linear transformation is a "root" of its characteristic polynomial, that is $p(L) = 0$.

Suppose that ϕ is an involution. This means that L^2 is a scalar multiple of the identity: $L^2 = \mu I$. Comparing this to the equation $p(L) = 0$ we obtain the equation

$$(\det L + \mu)I = (L)L.$$

Since $\phi \neq id_{P(V)}$, L can not be a scalar multiple of the identity, the last equation implies $L = (\det L + \mu) = 0$.

Conversely, suppose $\mathrm{tr}\, L = 0$. Then the Cayley-Hamilton theorem gives $L^2 = (-\det L)I$, therefore, $\phi^2 = [L^2] = [(-\det L)I] = id_{P(V)}$. On the other hand, ϕ is not the identity, otherwise L would be a scalar multiple of the identity matrix, say $L = \nu I$, and then $L = 2\nu = 0$ would imply $L = 0$, contradicting the assumption that L is invertible. $\square$

7.4.5. THEOREM. *(a) If* $\mathrm{char}\,\mathbb{F} \neq 2$*, then there are no parabolic inversions.*

(b) If $\mathrm{char}\,\mathbb{F} = 2$*, then there are no hyperbolic involutions.*

PROOF. (a) Suppose that the inversion ϕ is associated to the linear transformation L. Then, $L = 0$, and the characteristic polynomial of L has the form $\lambda^2 + \det L$ with $\det L \neq 0$. Therefore, there are either no eigenvalues of L in $\mathbb{F}$ or there are two different ones. Namely, if λ is an eigenvalue, then $-\lambda$ is another one, and $\lambda \neq -\lambda$. In the first case L has no eigenvector and ϕ has no fixed points, in the second L has an eigenvector to both eigenvalues, thus ϕ has two fixed points.

(b) Since $\lambda = -\lambda$ in a field of characteristic 2, L has either no eigenvalue, or one eigenvalue of multiplicity 2. In the first case $[L]$ has no fixed points, in the second one $[L]$ is elliptic. Indeed, if L had two linearly independent eigenvectors with the same eigenvalue, then L would be a scalar multiple of the identity. $\square$

7.4.6. THEOREM. *Suppose* $\mathrm{char}\,\mathbb{F} \neq 2$*. If* ϕ *is a hyperbolic inversion with fixed points* F_1, F_2*, and* P *is an arbitrary point different from the fixed points, then*

$$(F_1, F_2, P, \phi(P)) = -1.$$

Proof. Let $\phi = [L]$, $F_1 = [\mathbf{a}]$, $F_2 = [\mathbf{b}]$, $P = [\alpha\mathbf{a} + \beta\mathbf{b}]$. L has eigenvalues $\pm\lambda$ and its eigenvectors are $\mathbf{a}$, $\mathbf{b}$, therefore we may assume that $L\mathbf{a} = \lambda\mathbf{a}$ and $L\mathbf{b} = -\lambda\mathbf{b}$. Then we have

$$\phi(P) = [L(\alpha\mathbf{a} + \beta\mathbf{b})] = \lambda\alpha\mathbf{a} - \lambda\beta\mathbf{b},$$

from which, by the definition of cross-ratio,

$$(F_1, F_2, P, \phi(P)) = \frac{\beta}{\alpha} : \frac{-\lambda\beta}{\lambda\alpha} = -1. \quad \square$$

7.4.7. Theorem. *Let A, A', B, B' four not necessarily different points on a line e, such that $\{A, A'\} \cap \{B, B'\} = \emptyset$. Then there is a unique involution $\phi : e \to e$ such that $\phi(A) = A'$ and $\phi(B) = B'$.*

Proof. *Existence.* Choose representatives $\mathbf{a}$ and $\mathbf{b}$ for A and B. Since $A \neq B$, $\mathbf{a}$ and $\mathbf{b}$ are linearly independent. As $A' \neq B$, A' can be represented by a vector of the form $\mathbf{a} + \alpha\mathbf{b}$. Similarly, B' can be represented by a vector of the form $\beta\mathbf{a} - \mathbf{b}$. Since $A' \neq B'$, their representatives are linearly independent, thus there is a unique linear transformation L, which maps the basis $\mathbf{a}, \mathbf{b}$ to the basis $\mathbf{a} + \alpha\mathbf{b}, \beta\mathbf{a} - \mathbf{b}$. Setting $\phi = [L]$, we have $\phi(A) = A'$, $\phi(B) = B'$ and ϕ is an involution since $L = 0$.

Unicity. If $A = A'$ and $B = B'$ are the fixed points of an involution then the image of a point $P \notin \{A, B\}$ is uniquely determined by

$$(A, B, P, \phi(P)) = -1.$$

(See Theorem 7.4.6.). If, for example, $A \neq A'$, then ϕ is uniquely determined by the condition that it must map the points A, A', B to the points A', A, B' respectively. $\square$

7.4.8. Theorem. *If $\operatorname{char}\mathbb{F} \neq 2$, then any cross-ratio preserving transformation of a line into itself can be obtained as the composition of two involutions. In particular, involutions generate the group $PGL(\mathbb{F}P^1)$.*

Proof. Since $PGL(\mathbb{F}P^1)$ acts simply transitively on ordered 3-tuples of points, it suffices to show that if A, B, C and A', B', C' are two 3-tuples, then there are 2 involutions the composition of which maps A, B, C to A', B', C' respectively. The identity is the square of any involution, so it is enough to deal with transformations different from the identity. Such a transformation has at most two fixed points, while the straight lines have $|\mathbb{F}| + 1 \geq 4$ points, so we may assume that A and B are not fixed points of the transformation. Then $\{A, B'\} \cap \{A', B\} = \emptyset$, thus, by Theorem 7.4.7, there is an involution ϕ_1 such that $\phi_1(A) = B'$ and $\phi_1(B) = A'$. Let C^* be the image of C under ϕ_1. Clearly, $\{A', B'\} \cap \{C^*, C'\} = \emptyset$, so applying again Theorem 7.4.7, we get an involution ϕ_2 such that $\phi_2(A') = B'$ and $\phi_2(C^*) = C'$. By the construction, the composition $\phi_2 \circ \phi_1$ takes A, B, C to A', B', C'. $\square$