PROJECTIVE GEOMETRY

b3 course 2003

Nigel Hitchin

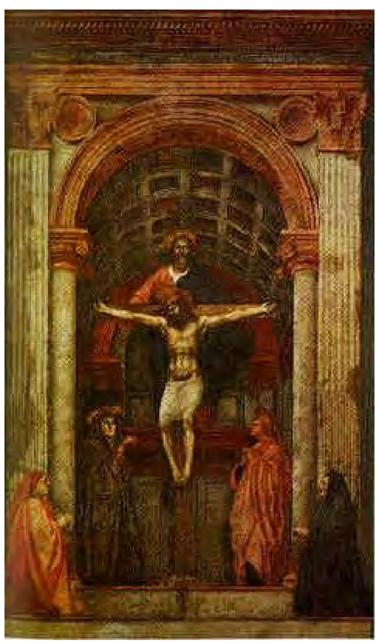
hitchin@maths.ox.ac.uk

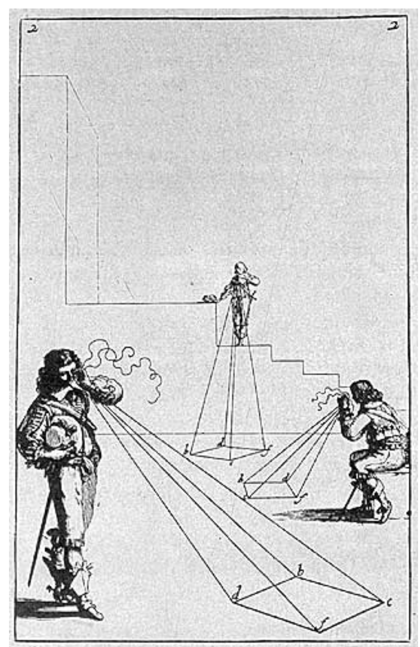
1 Introduction

This is a course on projective geometry. Probably your idea of geometry in the past has been based on triangles in the plane, Pythagoras' Theorem, or something more analytic like three-dimensional geometry using dot products and vector products. In either scenario this is usually called *Euclidean geometry* and it involves notions like distance, length, angles, areas and so forth. So what's wrong with it? Why do we need something different?

Here are a few reasons:

- Projective geometry started life over 500 years ago in the study of perspective drawing: the distance between two points on the artist's canvas does not represent the true distance between the objects they represent so that Euclidean distance is not the right concept.

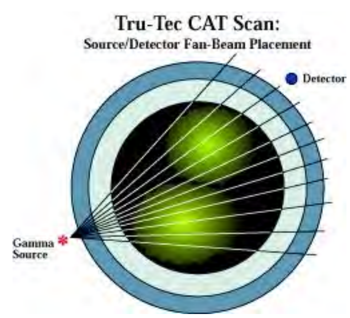
  

  The techniques of projective geometry, in particular homogeneous coordinates, provide the technical underpinning for perspective drawing and in particular for the modern version of the Renaissance artist, who produces the computer graphics we see every day on the web.

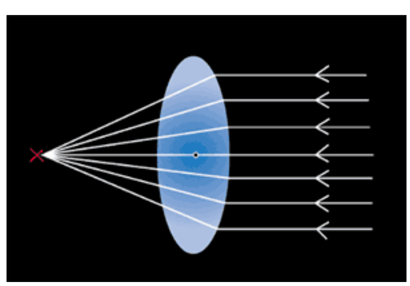
- Even in Euclidean geometry, not all questions are best attacked by using distances and angles. Problems about intersections of lines and planes, for example are not really metric. Centuries ago, projective geometry used to be called "descriptive geometry" and this imparts some of the flavour of the subject. This doesn't mean it is any less quantitative though, as we shall see.

- The Euclidean space of two or three dimensions in which we usually envisage geometry taking place has some failings. In some respects it is incomplete and asymmetric, and projective geometry can counteract that. For example, we know that through any two points in the plane there passes a unique straight line. But we can't say that any two straight lines in the plane intersect in a unique point, because we have to deal with parallel lines. Projective geometry evens things out – it adds to the Euclidean plane extra points at infinity, where parallel lines intersect. With these new points incorporated, a lot of geometrical objects become more unified. The different types of conic sections – ellipses, hyperbolas and parabolas – all become the same when we throw in the extra points.

- It may be that we are only interested in the points of good old $\mathbf{R}^2$ and $\mathbf{R}^3$ but there are always other spaces related to these which don't have the structure of a vector space – the space of lines for example. We need to have a geometrical and analytical approach to these. In the real world, it is necessary to deal with such spaces. The CT scanners used in hospitals essentially convert a series of readings from a subset of the space of straight lines in $\mathbf{R}^3$ into a density distribution.



At a simpler level, an optical device maps incoming light rays (oriented lines) to outgoing ones, so how it operates is determined by a map from the space of straight lines to itself.



> Projective geometry provides the means to describe analytically these auxiliary spaces of lines.

In a sense, the basic mathematics you will need for projective geometry is something you have already been exposed to from your linear algebra courses. Projective geometry is essentially a geometric realization of linear algebra, and its study can also help to make you understand basic concepts there better. The difference between the points of a vector space and those of its dual is less apparent than the difference between a point and a line in the plane, for example. When it comes to describing the space of lines in three-space, however, we shall need some additional linear algebra called *exterior algebra* which is essential anyway for other subjects such as differential geometry in higher dimensions and in general relativity. At this level, then, you will need to recall the basic properties of :

- vector spaces, subspaces, sums and intersections
- linear transformations
- dual spaces

After we have seen the essential features of projective geometry we shall step back and ask the question "What is geometry?" One answer given many years ago by Felix Klein was the rather abstract but highly influential statement: "Geometry is the study of invariants under the action of a group of transformations". With this point of view both Euclidean geometry and projective geometry come under one roof. But more than that, non-Euclidean geometries such as spherical or hyperbolic geometry can be treated in the same way and we finish these lectures with what was historically a driving force for the study of new types of geometry — Euclid's axioms and the parallel postulate.

2 Projective spaces

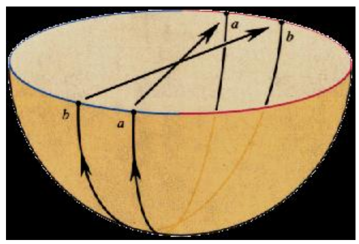
2.1 Basic definitions

Definition 1 *Let V be a vector space. The projective space $P(V)$ of V is the set of 1-dimensional vector subspaces of V.*

Definition 2 *If the vector space V has dimension $n+1$, then $P(V)$ is a projective space of dimension n. A 1-dimensional projective space is called a projective line, and a 2-dimensional one a projective plane.*

For most of the course, the field F of scalars for our vector spaces will be either the real numbers $\mathbf{R}$ or complex numbers $\mathbf{C}$. Our intuition is best served by thinking of the real case. So the projective space of $\mathbf{R}^{n+1}$ is the set of lines through the origin. Each such line intersects the unit n-sphere $S^n = \{x \in \mathbf{R}^{n+1} : \sum_i x_i^2 = 1\}$ in two points $\pm u$, so from this point of view $P(\mathbf{R}^{n+1})$ is S^n with antipodal points identified. Since each line intersects the lower hemisphere, we could equally remove the upper hemisphere and then identify opposite points on the equatorial sphere.

When $n = 1$ this is just identifying the end points of a semicircle which gives a circle, but when $n = 2$ it becomes more difficult to visualize:



If this were a course on topology, this would be a useful starting point for looking at some exotic topological spaces, but it is less so for a geometry course. Still, it does explain why we should think of $P(\mathbf{R}^{n+1})$ as n-dimensional, and so we shall write it as $P^n(\mathbf{R})$ to make this more plain.

A better approach for our purposes is the notion of a representative vector for a point of $P(V)$. Any 1-dimensional subspace of V is the set of multiples of a non-zero vector $v \in V$. We then say that v is a representative vector for the point $[v] \in P(V)$. Clearly if $\lambda \neq 0$ then λv is another representative vector so

$$[\lambda v] = [v].$$

Now suppose we choose a basis $\{v_0, \ldots, v_n\}$ for V. The vector v can be written

$$v = \sum_{i=0}^{n} x_i v_i$$

and the $n+1$-tuple $(x_0, \ldots, x_n)$ provides the coordinates of $v \in V$. If $v \neq 0$ we write the corresponding point $[v] \in P(V)$ as $[v] = [x_0, x_1, \ldots, x_n]$ and these are known as homogeneous coordinates for a point in $P(V)$. Again, for $\lambda \neq 0$

$$[\lambda x_0, \lambda x_1, \ldots, \lambda x_n] = [x_0, x_1, \ldots, x_n].$$

Homogeneous coordinates give us another point of view of projective space. Let $U_0 \subset P(V)$ be the subset of points with homogeneous coordinates $[x_0, x_1, \ldots, x_n]$

such that $x_0 \neq 0$. (Since if $\lambda \neq 0$, $x_0 \neq 0$ if and only if $\lambda x_0 \neq 0$, so this is a well-defined subset, independent of the choice of $(x_0, \ldots, x_n)$). Then, in U_0,

$$[x_0, x_1, \ldots, x_n] = [x_0, x_0(x_1/x_0), \ldots, x_0(x_n/x_0)] = [1, (x_1/x_0), \ldots, (x_n/x_0)].$$

Thus we can uniquely represent any point in U_0 by one of the form $[1, y_1, \ldots, y_n]$, so

$$U_0 \cong F^n.$$

The points we have missed out are those for which $x_0 = 0$, but these are the 1-dimensional subspaces of the n-dimensional vector subspace spanned by $v_1, \ldots, v_n$, which is a projective space of one lower dimension. So, when $F = \mathbf{R}$, instead of thinking of $P^n(\mathbf{R})$ as S^n with opposite points identified, we can write

$$P^n(\mathbf{R}) = \mathbf{R}^n \cup P^{n-1}(\mathbf{R}).$$

A large chunk of real projective n-space is thus our familiar $\mathbf{R}^n$.

Example: The simplest example of this is the case $n = 1$. Since a one-dimensional projective space is a single point (if $\dim V = 1$, V is the only 1-dimensional subspace) the projective line $P^1(F) = F \cup pt$. Since $[x_0, x_1]$ maps to $x_1/x_0 \in F$ we usually call this extra point $[0, 1]$ the point ∞. When $F = \mathbf{C}$, the complex numbers, the projective line is what is called in complex analysis the *extended complex plane* $\mathbf{C} \cup \{\infty\}$.

Having said that, there are many different copies of F^n inside $P^n(F)$, for we could have chosen x_i instead of x_0, or coordinates with respect to a totally different basis. Projective space should normally be thought of as a homogeneous object, without any distinguished copy of F^n inside.

2.2 Linear subspaces

Definition 3 *A linear subspace of the projective space $P(V)$ is the set of 1-dimensional vector subspaces of a vector subspace $U \subseteq V$.*

Note that a linear subspace is a projective space in its own right, the projective space $P(U)$.

Recall that a 1-dimensional projective space is called a projective line. We have the following two propositions which show that projective lines behave nicely:

Proposition 1 *Through any two distinct points in a projective space there passes a unique projective line.*

Proof: Let $P(V)$ be the projective space and $x, y \in P(V)$ distinct points. Let u, v be representative vectors. Then u, v are linearly independent for otherwise $u = \lambda v$ and

$$x = [u] = [\lambda v] = [v] = y.$$

Let $U \subseteq V$ be the 2-dimensional vector space spanned by u and v, then $P(U) \subset P(V)$ is a line containing x and y.

Suppose $P(U')$ is another such line, then $u \in U'$ and $v \in U'$ and so the space spanned by u, v (namely U) is a subspace of U'. But U and U' are 2-dimensional so $U = U'$ and the line is thus unique. □

Proposition 2 *In a projective plane, two distinct projective lines intersect in a unique point.*

Proof: Let the projective plane be $P(V)$ where $\dim V = 3$. Two lines are defined by $P(U_1), P(U_2)$ where U_1, U_2 are distinct 2-dimensional subspaces of V. Now from elementary linear algebra

$$\dim V \geq \dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$$

so that

$$3 \geq 2 + 2 - \dim(U_1 \cap U_2)$$

and

$$\dim(U_1 \cap U_2) \geq 1.$$

But since U_1 and U_2 are 2-dimensional,

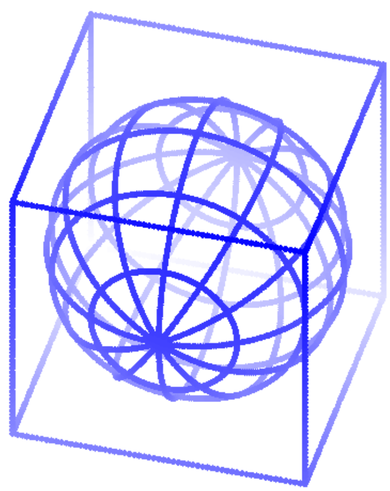
$$\dim(U_1 \cap U_2) \leq 2$$

with equality if and only if $U_1 = U_2$. As the lines are distinct, equality doesn't occur and so we have the 1-dimensional vector subspace

$$U_1 \cap U_2 \subset V$$

which is the required point of intersection in $P(V)$. □

Remark: The model of projective space as the sphere with opposite points identified illustrates this proposition, for a projective line in $P^2(\mathbf{R})$ is defines by a 2-dimensional subspace of $\mathbf{R}^3$, which intersects the unit sphere in a great circle. Two great circles intersect in two antipodal points. When we identify opposite points, we just get one intersection.



Instead of the spherical picture, let's consider instead the link between projective lines and ordinary lines in the plane, using the decomposition

$$P^2(\mathbf{R}) = \mathbf{R}^2 \cup P^1(\mathbf{R}).$$

Here we see that the real projective plane is the union of $\mathbf{R}^2$ with a projective line $P^1(\mathbf{R})$. Recall that this line is given in homogeneous coordinates by $x_0 = 0$, so it corresponds to the 2-dimensional space spanned by $(0, 1, 0)$ and $(0, 0, 1)$. Any 2-dimensional subspace of $\mathbf{R}^3$ is defined by a single equation

$$a_0x_0 + a_1x_1 + a_2x_2 = 0$$

and if a_1 and a_2 are not both zero, this intersects $U_0 \cong \mathbf{R}^2$ (the points where $x_0 \neq 0$) where

$$0 = a_0 + a_1(x_1/x_0) + a_2(x_2/x_0) = a_0 + a_1y_1 + a_2y_2$$

which is an ordinary straight line in $\mathbf{R}^2$ with coordinates y_1, y_2. The projective line has one extra point on it, where $x_0 = 0$, i.e. the point $[0, a_2, -a_1]$. Conversely, any straight line in $\mathbf{R}^2$ extends uniquely to a projective line in $P^2(\mathbf{R})$.

Two lines in $\mathbf{R}^2$ are parallel if they are of the form

$$a_0 + a_1y_1 + a_2y_2 = 0, \qquad b_0 + a_1y_1 + a_2y_2 = 0$$

but then the added point to make them projective lines is the same one: $[0, a_2, -a_1]$, so the two lines meet at a single point on the "line at infinity" $P^1(\mathbf{R})$.

2.3 Projective transformations

If V, W are vector spaces and $T : V \to W$ is a linear transformation, then a vector subspace $U \subseteq V$ gets mapped to a vector subspace $T(U) \subseteq W$. If T has a non-zero kernel, $T(U)$ may have dimension less than that of U, but if $\ker T = 0$ then $\dim T(U) = \dim U$. In particular, if U is one-dimensional, so is $T(U)$ and so T gives a well-defined map

$$\tau : P(V) \to P(W).$$

Definition 4 *A projective transformation from $P(V)$ to $P(W)$ is the map τ defined by an invertible linear transformation $T : V \to W$.*

Note that if $\lambda \neq 0$, then λT and T define the same linear transformation since

$$[(\lambda T)(v)] = [\lambda(T(v))] = [T(v)].$$

The converse is also true: suppose T and T' define the same projective transformation τ. Take a basis $\{v_0, \ldots, v_n\}$ for V, then since

$$\tau([v_i]) = [T'(v_i)] = [T(v_i)]$$

we have

$$T'(v_i) = \lambda_i T(v_i)$$

for some non-zero scalars λ_i and also

$$T'(\sum_{i=0}^{n} v_i) = \lambda T(\sum_{i=0}^{n} v_i)$$

for some non-zero λ. But then

$$\sum_{i=0}^{n} \lambda T(v_i) = \lambda T(\sum_{i=0}^{n} v_i) = T'(\sum_{i=0}^{n} v_i) = \sum_{i=0}^{n} \lambda_i T(v_i).$$

Since T is invertible, $T(v_i)$ are linearly independent, so this implies $\lambda_i = \lambda$. Then $T'(v_i) = \lambda T(v_i)$ for all basis vectors and hence for all vectors and so

$$T' = \lambda T.$$

Example: You are, in fact, already familiar with one class of projective transformations – Möbius transformations of the extended complex plane. These are just projective transformations of the complex projective line $P^1(\mathbf{C})$ to itself. We describe points in $P^1(\mathbf{C})$ by homogeneous coordinates $[z_0, z_1]$, and then a projective transformation τ is given by

$$\tau([z_0, z_1]) = ([az_0 + bz_1, cz_0 + dz_1])$$

where $ad - bc \neq 0$. This corresponds to the invertible linear transformation

$$T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

It is convenient to write $P^1(\mathbf{C}) = \mathbf{C} \cup \{\infty\}$ where the point ∞ is now the 1-dimensional space $z_1 = 0$. Then if $z_1 \neq 0$, $[z_0, z_1] = [z, 1]$ and

$$\tau([z, 1]) = [az + b, cz + d]$$

and if $cz + d \neq 0$ we can write

$$\tau([z, 1]) = [\frac{az + b}{cz + d}, 1]$$

which is the usual form of a Möbius transformation, i.e.

$$z \mapsto \frac{az + b}{cz + d}.$$

The advantage of projective geometry is that the point $\infty = [1, 0]$ plays no special role. If $cz + d = 0$ we can still write

$$\tau([z, 1]) = [az + b, cz + d] = [az + b, 0] = [1, 0]$$

and if $z = \infty$ (i.e. $[z_0, z_1] = [1, 0]$) then we have

$$\tau([1, 0]) = [a, c].$$

Example: If we view the real projective plane $P^2(\mathbf{R})$ in the same way, we get some less familiar transformations. Write $P^2(\mathbf{R}) = \mathbf{R}^2 \cup P^1(\mathbf{R})$ where the projective line at infinity is $x_0 = 0$. A linear transformation $T : \mathbf{R}^3 \to \mathbf{R}^3$ can then be written as the matrix

$$T = \begin{pmatrix} d & b_1 & b_2 \\ c_1 & a_{11} & a_{12} \\ c_2 & a_{21} & a_{22} \end{pmatrix}$$

and its action on $[1, x, y]$ can be expressed, with $\mathbf{v} = (x, y) \in \mathbf{R}^2$, as

$$\mathbf{v} \mapsto \frac{1}{\mathbf{b} \cdot \mathbf{v} + d}(A\mathbf{v} + \mathbf{c})$$

where A is the 2×2 matrix a_{ij} and $\mathbf{b}, \mathbf{c}$ the vectors $(b_1, b_2), (c_2, c_2)$. These are the 2-dimensional versions of Möbius transformations. Each one can be considered as a composition of

- an invertible linear transformation $\mathbf{v} \mapsto A\mathbf{v}$
- a translation $\mathbf{v} \mapsto \mathbf{v} + \mathbf{c}$
- an inversion $\mathbf{v} \mapsto \mathbf{v}/(\mathbf{b} \cdot \mathbf{v} + d)$

Clearly it is easier here to consider projective transformations defined by 3×3 matrices, just ordinary linear algebra.

Example: A more geometric example of a projective transformation is to take two lines $P(U), P(U')$ in a projective plane $P(V)$ and let $K \in P(V)$ be a point disjoint from both. For each point $x \in P(U)$, the unique line joining K to x intersects $P(U')$ in a unique point $X = \tau(x)$. Then

$$\tau : P(U) \to P(U')$$

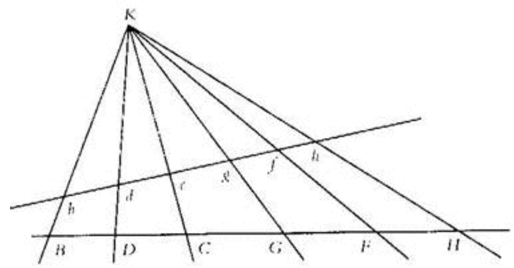
is a projective transformation.

To see why, let W be the 1-dimensional subspace of V defined by $K \in P(V)$. Then since K does not lie in $P(U')$, $W \cap U' = 0$. This means that

$$V = W \oplus U'.$$

Now take $a \in U$ as a representative vector for x. It can be expressed uniquely as $a = w + a'$, with $w \in W$ and $a' \in U'$. The projective line joining K to x is defined by the 2-dimensional vector subspace of V spanned by w and a and so $a' = a - w$ is a representative vector for $\tau(x)$. In linear algebra terms the map $a \mapsto a'$ is just the linear projection map $P : V \to U'$ restricted to U. It has zero kernel since K does not lie in $P(U)$, and hence $W \cap U = 0$. Thus $T : U \to U'$ is an isomorphism and τ is a projective transformation.

If we restrict to the points in $\mathbf{R}^2$, then this is what this *projection from K* looks like:



A linear transformation of a vector space of dimension n is determined by its value on n linearly independent vectors. A similar statement holds in projective space. The analogue of linear independence is the following

Definition 5 *Let $P(V)$ be an n-dimensional projective space, then $n+2$ points in $P(V)$ are said to be in general position if each subset of $n+1$ points has representative vectors in V which are linearly independent.*

Example: Any two distinct points in a projective line are represented by linearly independent vectors, so any three distinct points are in general position.

Theorem 3 *If $X_1, \ldots, X_{n+2}$ are in general position in $P(V)$ and $Y_1, \ldots, Y_{n+2}$ are in general position in $P(W)$, then there is a unique projective transformation $\tau : P(V) \to P(W)$ such that $\tau(X_i) = Y_i$, $1 \leq i \leq n+2$.*

Proof: First choose representative vectors $v_1, \ldots, v_{n+2} \in V$ for the points $X_1, \ldots, X_{n+2}$ in $P(V)$. By general position the first $n+1$ vectors are linearly independent, so they form a basis for V and there are scalars λ_i such that

$$v_{n+2} = \sum_{i=1}^{n+1} \lambda_i v_i \tag{1}$$

If $\lambda_i = 0$ for some i, then (1) provides a linear relation amongst a subset of $n+1$ vectors, which is not possible by the definition of general position, so we deduce that $\lambda_i \neq 0$ for all i. This means that each $\lambda_i v_i$ is also a representative vector for x_i, so (1) tells us that we could have chosen representative vectors v_i such that

$$v_{n+2} = \sum_{i=1}^{n+1} v_i \tag{2}$$

Moreover, given v_{n+2}, these v_i are unique for

$$\sum_{i=1}^{n+1} v_i = \sum_{i=1}^{n+1} \mu_i v_i$$

implies $\mu_i = 1$ since $v_1, \ldots, v_{n+1}$ are linearly independent.

[*Note: This is a very useful idea which can simplify the solution of many problems*].

Now do the same for the points $Y_1, \ldots Y_{n+2}$ in $P(W)$ and choose representative vectors such that

$$w_{n+2} = \sum_{i=1}^{n+1} w_i \tag{3}$$

Since $v_1, \ldots, v_{n+1}$ are linearly independent, they form a basis for V so there is a unique linear transformation $T : V \to W$ such that $Tv_i = w_i$ for $1 \leq i \leq n+1$. Since $w_1, \ldots, w_{n+1}$ are linearly independent, T is invertible. Furthermore, from (2) and (3)

$$Tv_{n+2} = \sum_{i=1}^{n+1} Tv_i = \sum_{i=1}^{n+1} w_i = w_{n+2}$$

and so T defines a projective transformation τ such that $\tau(X_i) = Y_i$ for all $n+2$ vectors v_i.

To show uniqueness, suppose T' defines another projective transformation τ' with the same property. Then $T'v_i = \mu_i w_i$ and

$$\mu_{n+2} w_{n+2} = T'v_{n+2} = \sum_{i=1}^{n+1} T'v_i = \sum_{i=1}^{n+1} \mu_i w_i.$$

But by the uniqueness of the representation (3), we must have $\mu_i/\mu_{n+2} = 1$, so that $T'v_i = \mu_{n+2} Tv_i$ and $\tau' = \tau$. □

Examples:

1. In $P^1(\mathbf{C})$ take the three distinct points $[0,1],[1,1],[1,0]$ and any other three distinct points X_1, X_2, X_3. Then there is a unique projective transformation taking X_1, X_2, X_3 to $[0,1],[1,1],[1,0]$. In the language of complex analysis, we can say that there is a unique Möbius transformation taking any three distinct points to $0, 1, \infty$.

2. In any projective line we could take the three points $[0,1],[1,1],[1,0]$ and then for X_1, X_2, X_3 any permutation of these. Now projective transformations of a space

to itself form a group under composition, so we see that the group of projective transformations of a line to itself always contains a copy of the symmetric group S_3. In fact if we take the scalars to be the field $\mathbf{Z}_2$ with two elements 0 and 1, the *only* points on the projective line are $[0, 1], [1, 1], [1, 0]$, and S_3 is the full group of projective transformations.

As an example of the use of the notion of general position, here is a classical theorem called Desargues' theorem. In fact, Desargues (1591-1661) is generally regarded as the founder of projective geometry. The proof we give here uses the method of choosing representative vectors above.

Theorem 4 *(Desargues) Let A, B, C, A', B', C' be distinct points in a projective space $P(V)$ such that the lines AA', BB' CC' are distinct and concurrent. Then the three points of intersection $AB \cap A'B'$, $BC \cap B'C'$, $CA \cap C'A'$ are collinear.*

Proof: Let P be the common point of intersection of the three lines AA', BB', CC'. Since P, A, A' lie on a projective line and are distinct, they are in general position, so as in (2) we choose representative vectors p, a, a' such that

$$p = a + a'.$$

These are vectors in a 2-dimensional subspace of V. Similarly we have representative vectors b, b' for B, B' and c, c' for C, C' with

$$p = b + b' \qquad p = c + c'.$$

It follows that $a + a' = b + b'$ and so

$$a - b = b' - a' = c''$$

and similarly

$$b - c = c' - b' = a'' \qquad c - a = a' - c' = b''.$$

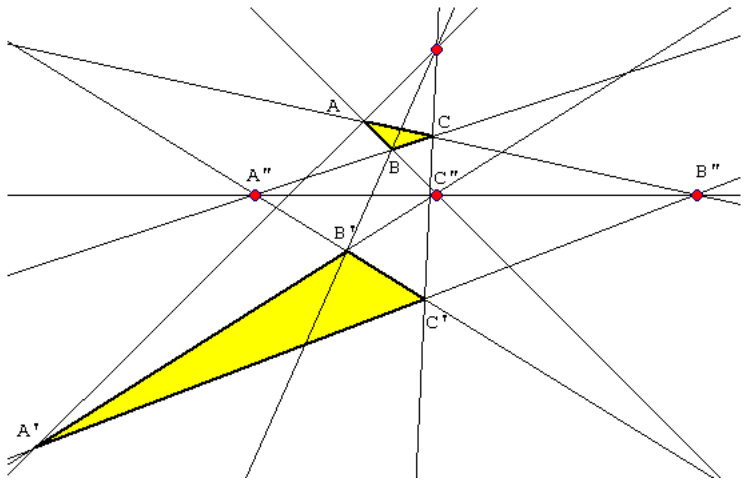
But then

$$c'' + a'' + b'' = a - b + b - c + c - a = 0$$

and a'', b'', c'' are linearly independent and lie in a 2-dimensional subspace of V. Hence the points A'', B'', C'' in $P(V)$ represented by a'', b'', c'' are collinear.

Now since $c'' = a - b$, c'' lies in the 2-dimensional space spanned by a and b, so C'' lies on the line AB. Since c'' also equals $b' - a'$, C'' lies on the line $A'B'$ and so c'' represents the point $AB \cap A'B'$. Repeating for B'' and A'' we see that these are the three required collinear points. □

Desargues' theorem is a theorem in projective space which we just proved by linear algebra – linear independence of vectors. However, if we take the projective space $P(V)$ to be the real projective plane $P^2(\mathbf{R})$ and then just look at that part of the data which lives in $\mathbf{R}^2$, we get a theorem about perspective triangles in the plane:



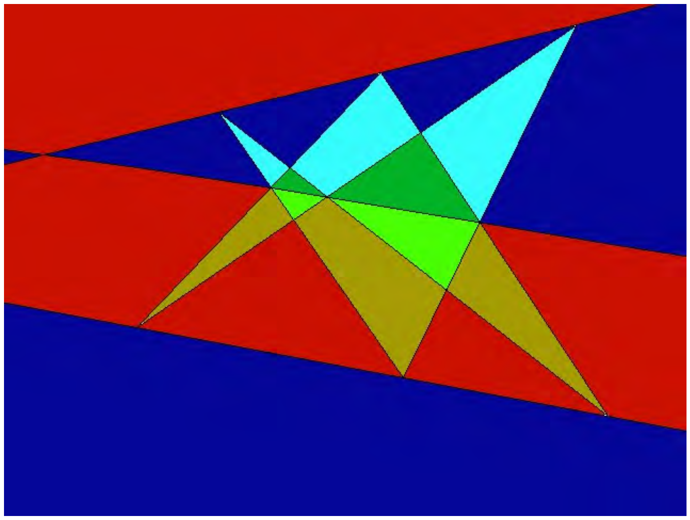
Here is an example of the use of projective geometry – a "higher form of geometry" to prove simply a theorem in $\mathbf{R}^2$ which is less accessible by other means. Another theorem in the plane for which these methods give a simple proof is Pappus' theorem. Pappus of Alexandria (290-350) was thinking again of plane Euclidean geometry, but his theorem makes sense in the projective plane since it only discusses collinearity and not questions about angles and lengths. It means that we can transform the given configuration by a projective transformation to a form which reduces the proof to simple linear algebra calculation:

Theorem 5 *(Pappus) Let A, B, C and A', B', C' be two pairs of collinear triples of distinct points in a projective plane. Then the three points $BC' \cap B'C, CA' \cap C'A, AB' \cap A'B$ are collinear.*

Proof: Without loss of generality, we can assume that A, B, C', B' are in general position. If not, then two of the three required points coincide, so the conclusion is trivial. By Theorem 3, we can then assume that

$$A = [1, 0, 0], \quad B = [0, 1, 0], \quad C' = [0, 0, 1], \quad B' = [1, 1, 1].$$

The line AB is defined by the 2-dimensional subspace $\{(x_0, x_1, x_2) \in F^3 : x_2 = 0\}$, so the point C, which lies on this line, is of the form $C = [1, c, 0]$ and $c \neq 0$ since $A \neq C$. Similarly the line $B'C'$ is $x_0 = x_1$, so $A' = [1, 1, a]$ with $a \neq 1$.



The line BC' is defined by $x_0 = 0$ and $B'C$ is defined by the span of $(1,1,1)$ and $(1,c,0)$, so the point $BC' \cap B'C$ is represented by the linear combination of $(1,1,1)$ and $(1,c,0)$ for which $x_0 = 0$, i.e.

$$(1,1,1) - (1,c,0) = (0,1-c,1).$$

The line $C'A$ is given by $x_1 = 0$, so similarly $CA' \cap C'A$ is represented by

$$(1,c,0) - c(1,1,a) = (1-c,0,-ca).$$

Finally AB' is given by $x_1 = x_2$, so $AB' \cap A'B$ is

$$(1,1,a) + (a-1)(0,1,0) = (1,a,a).$$

But then

$$(c-1)(1,a,a) + (1-c,0,-ca) + a(0,1-c,1) = 0.$$

Thus the three vectors span a 2-dimensional subspace and so the three points lie on a projective line. □

2.4 Duality

Projective geometry gives, as we shall see, a more concrete realization of the linear algebra notion of duality. But first let's recall what dual spaces are all about. Here are the essential points:

- Given a finite-dimensional vector space V over a field F, the dual space V' is the vector space of linear transformations $f : V \to F$.

- If $v_1, \dots, v_n$ is a basis for V, there is a *dual basis* $f_1, \dots f_n$ of V' characterized by the property $f_i(v_j) = 1$ if $i = j$ and $f_i(v_j) = 0$ otherwise.
- If $T : V \to W$ is a linear transformation, there is a natural linear transformation $T' : W' \to V'$ defined by $T'f(v) = f(Tv)$.

Although a vector space V and its dual V' have the same dimension there is no natural way of associating a point in one with a point in the other. We can do so however with vector subspaces:

Definition 6 *Let $U \subseteq V$ be a vector subspace. The annihilator $U^o \subset V'$ is defined by $U^o = \{f \in V' : f(u) = 0$ for all $u \in U\}$.*

The annihilator is clearly a vector subspace of V' since $f(u) = 0$ implies $\lambda f(u) = 0$ and if also $g(u) = 0$ then $(f + g)(u) = f(u) + g(u) = 0$. Furthermore, if $U_1 \subseteq U_2$ and $f(u) = 0$ for all $u \in U_2$, then in particular $f(u) = 0$ for all $u \in U_1$, so that

$$U_2^o \subseteq U_1^o.$$

We also have:

Proposition 6 $$\dim U + \dim U^o = \dim V.$$

Proof: Let $u_1, \dots, u_m$ be a basis for U and extend to a basis $u_1, \dots, u_m, v_1, \dots, v_{n-m}$ of V. Let $f_1, \dots, f_n$ be the dual basis. Then for $m+1 \le i \le n$, $f_i(u_j) = 0$ so $f_i \in U^o$. Conversely if $f \in U^o$, write

$$f = \sum_{i=1}^{n} c_i f_i$$

Then $0 = f(u_i) = c_i$, and so f is a linear combination of f_i for $m + 1 \le i \le n$. Thus $f_{m+1}, \dots, f_n$ is a basis for U^o and

$$\dim U + \dim U^o = m + n - m = n = \dim V.$$

□

If we take the dual of the dual we get a vector space V'', but this is naturally isomorphic to V itself. To see this, define $S : V \to V''$ by

$$Sv(f) = f(v).$$

This is clearly linear in v, and $\ker S$ is the set of vectors such that $f(v) = 0$ for all f, which is zero, since we could extend $v = v_1$ to a basis, and $f_1(v_1) \neq 0$. Since $\dim V = \dim V'$, S is an isomorphism. Under this transformation, for each vector subspace $U \subseteq V$, $S(U) = U^{oo}$. This follows since if $u \in U$, and $f \in U^0$

$$Su(f) = f(u) = 0$$

so $S(U) \subseteq U^{oo}$. But from (6) the dimensions are the same, so we have equality.

Thus to any vector space V we can naturally associate another vector space of the same dimension V', and to any projective space $P(V)$ we can associate another one $P(V')$. Our first task is to understand what a point of $P(V')$ means in terms of the original projective space $P(V)$.

From the linear algebra definition of dual, a point of $P(V')$ has a non-zero representative vector $f \in V'$. Since $f \neq 0$, it defines a surjective linear map

$$f : V \to F$$

and so

$$\dim \ker f = \dim V - \dim F = \dim V - 1.$$

If $\lambda \neq 0$, then $\dim \ker \lambda f = \dim \ker f$ so the point $[f] \in P(V')$ defines unambiguously a vector subspace $U \subset V$ of dimension one less than that of V, and a corresponding linear subspace $P(U)$ of $P(V)$.

Definition 7 *A hyperplane in a projective space $P(V)$ is a linear subspace $P(U)$ of dimension $\dim P(V) - 1$ (or* codimension *one).*

Conversely, a hyperplane defines a vector subspace $U \subset V$ of dimension $\dim V - 1$, and so we have a 1-dimensional quotient space V/U and a surjective linear map

$$\pi : V \to V/U$$

defined by $\pi(v) = v + U$. If $\nu \in V/U$ is a non-zero vector then

$$\pi(v) = f(v)\nu$$

for some linear map $f : V \to F$, and then $U = \ker f$. A different choice of ν changes f to λf, so the hyperplane $P(U)$ naturally defines a point $[f] \in P(V')$. Hence,

Proposition 7 *The points of the dual projective space $P(V')$ of a projective space $P(V)$ are in natural one-to-one correspondence with the hyperplanes in $P(V)$.*

The surprise here is that the space of hyperplanes should have the structure of a projective space. In particular there are linear subspaces of $P(V')$ and they demand an interpretation. From the point of view of linear algebra, this is straightforward: to each $m+1$-dimensional vector subspace $U \subseteq V$ of the $n+1$-dimensional vector space V we associate the $n-m$-dimensional annihilator $U^o \subseteq V'$. Conversely, given $W \subseteq V'$, take $W^o \subset V''$ then $W^o = S(U)$ for some U and since $S(U) = U^{oo}$, it follows that

$$W = U^o.$$

Thus taking the annihilator defines a one-to-one correspondence between vector subspaces of V and vector subspaces of V'. We just need to give this a geometrical interpretation.

Proposition 8 *A linear subspace $P(W) \subseteq P(V')$ of dimension m in a dual projective space $P(V')$ of dimension n consists of the hyperplanes in $P(V)$ which contain a fixed linear subspace $P(U) \subseteq P(V)$ of dimension $n-m-1$.*

Proof: As we saw above, $W = U^o$ for some vector subspace $U \subseteq V$, so $f \in W$ is a linear map $f : V \to F$ such that $f(U) = 0$. This means that $U \subset \ker f$ so the hyperplane defined by f contains $P(U)$. □

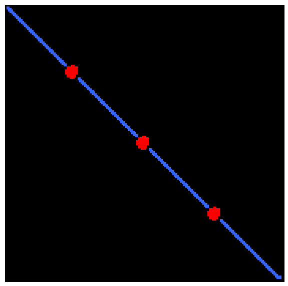
A special case of this is a hyperplane in $P(V')$. This consists of the hyperplanes in $P(V)$ which pass through a fixed point $X \in P(V)$, and this describes geometrically the projective transformation defined by S

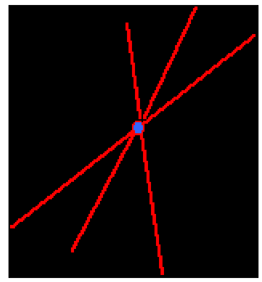
$$P(V) \cong P(V'').$$

All these features are somewhat clearer in low dimensions. A hyperplane in a projective line is just a point, so there is a natural isomorphism $P(V) \cong P(V')$ here and duality gives nothing new. In a projective plane however, a hyperplane is a line, so $P(V')$ is the space of lines in $P(V)$. The space of lines passing through a point $X \in P(V)$ constitutes a line X^o in $P(V')$. Given two points X, Y there is a unique line joining them. So there must be a unique point in $P(V')$ which lies on the two lines X^o, Y^o. Duality therefore shows that Proposition 2 is just the same as Proposition 1, if we apply the latter to the dual projective plane $P(V')$.

Here is another example of dual configurations:

three collinear points *three concurrent lines*





In general, any result of projective geometry when applied to the dual plane $P(V')$ can be reinterpreted in $P(V)$ in a different form. In principle then, we get two theorems for the price of one. As an example take Desargues' Theorem, at least in the way we formulated it in (4). Instead of applying it to the projective plane $P(V)$, apply it to $P(V')$. The theorem is still true, but it says something different in $P(V)$. For example, our starting point in $P(V')$ consists of seven points, which now become seven lines in $P(V)$. So here is the dual of Desargues' theorem:

Theorem 9 *(Desargues) Let $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$ be distinct lines in a projective plane $P(V)$ such that the points $\alpha \cap \alpha'$, $\beta \cap \beta'$, $\gamma \cap \gamma'$ are distinct and collinear. Then the lines joining $\alpha \cap \beta$, $\alpha' \cap \beta'$ and $\beta \cap \gamma$, $\beta' \cap \gamma'$ and $\gamma \cap \alpha, \gamma' \cap \alpha'$ are concurrent.*

Here the dual theorem starts with three points lying on a line and ends with three lines meeting in a point – looked at the right way, we have the *converse* of Desargues' Theorem.

Now look at Pappus' theorem. Instead of two triples of collinear points, the dual statement of the theorem gives two triples of concurrent lines α, β, γ passing through A and α', β', γ' passing through A'. Define B on α to be $\alpha \cap \gamma'$ and C to be $\alpha \cap \beta'$. Define B' on α' to be $\alpha' \cap \beta$ and C' to be $\alpha' \cap \gamma$.

The dual of Pappus says that the lines joining $\{\beta \cap \gamma', \beta' \cap \gamma\}, \{\gamma \cap \alpha', \gamma' \cap \alpha\}, \{\alpha \cap \beta', \alpha' \cap \beta\}$ are concurrent at a point P. By definition of B, B', C, C', the last two are $\{BC', B'C\}$, which therefore intersect in P. Now A lies on β and by definition so does B' so β is the line AB'. Similarly $A'B$ is the line γ'. Likewise A lies on γ and by definition so does C' so AC' is γ and $A'C$ is β'.

Thus the intersection of $\{AB', A'B\}$ is $\beta \cap \gamma'$ and the intersection of $\{AC', A'C\}$ is $\beta' \cap \gamma$. The dual of Pappus' theorem says that the line joining these passes through

P, which is the intersection of $\{BC', B'C\}$. These three points are thus collinear and this is precisely Pappus' theorem itself.

Finally, we can use duality to understand something very down-to-earth – the space of straight lines in $\mathbf{R}^2$. When we viewed the projective plane $P^2(\mathbf{R})$ as $\mathbf{R}^2 \cup P^1(\mathbf{R})$ we saw that a projective line not equal to the line at infinity $P^1(\mathbf{R})$ intersected $\mathbf{R}^2$ in an ordinary straight line. Since we now know that the lines in $P^2(\mathbf{R})$ are in one-to-one correspondence with another projective plane – the dual plane – we see that we only have to remove a single point from the dual plane, the point giving the line at infinity, to obtain the space of lines in $\mathbf{R}^2$. So in the sphere model, we remove the north and south poles and identify antipodal points.

Concretely parametrize the sphere in the usual way:

$$x_1 = \sin\theta \sin\phi, \qquad x_2 = \sin\theta\cos\phi, \qquad x_3 = \cos\theta$$

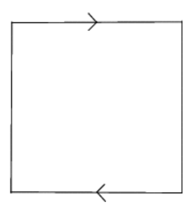
then with the poles removed the range of values is $0 < \theta < \pi$, $0 \leq \phi < 2\pi$. The antipodal map is

$$\theta \mapsto \pi - \theta, \qquad \phi \mapsto \phi + \pi.$$

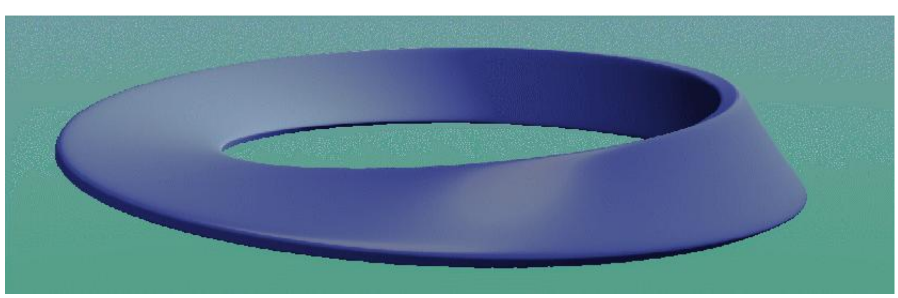
We can therefore identify the space of lines in $\mathbf{R}^2$ as the pairs

$$(\theta, \phi) \in (0, \pi) \times [0, \pi]$$

where we identify $(\theta, 0)$ with $(\pi - \theta, \pi)$:



and this is the *Möbius band.*



2.5 Exercises

1. Let U_1, U_2 and U_3 be the 2-dimensional vector subspaces of $\mathbf{R}^3$ defined by

$$x_0 = 0, \qquad x_0 + x_1 + x_2 = 0, \quad 3x_0 - 4x_1 + 5x_2 = 0$$

respectively. Find the vertices of the "triangle" in $P^2(\mathbf{R})$ whose sides are the projective lines $P(U_1), P(U_2), P(U_3)$.

2. Let U_1, U_2 be vector subspaces of V. Show that the linear subspace

$$P(U_1 + U_2) \subseteq P(V)$$

is the set of points obtained by joining each $X \in P(U_1)$ and $Y \in P(U_2)$ by a projective line.

3. Prove that three skew (i.e. non-intersecting) lines in $P^3(\mathbf{R})$ have an infinite number of transversals (i.e. lines meeting all three).

4. Find the projective transformation $\tau : P^1(\mathbf{R}) \to P^1(\mathbf{R})$ for which

$$\tau[0,1] = [1,0], \qquad \tau[1,0] = [1,1], \qquad \tau[1,1] = [0,1]$$

and show that $\tau^3 = \mathrm{i}d$.

5. Let $T : V \to V$ be an invertible transformation. Show that if $v \in V$ is an eigenvector of T, then $[v] \in P(V)$ is a fixed point of the projective transformation τ defined by T. Prove that any projective transformation of $P^2(\mathbf{R})$ has a fixed point.

6. Let V be a 3-dimensional vector space with basis v_1, v_2, v_3 and let $A, B, C \in P(V)$ be expressed in homogeneous coordinates relative to this basis by

$$A = [2,1,0], \qquad B = [0,1,1], \qquad C = [-1,1,2].$$

Find the coordinates with respect to the dual basis of the three points in the dual space $P(V')$ which represent the lines AB, BC and CA.

3 Quadrics

3.1 Quadratic forms

The projective geometry of quadrics is the geometrical version of the part of linear algebra which deals with symmetric bilinear forms – the generalization of the dot product $\mathbf{a} \cdot \mathbf{b}$ of vectors in $\mathbf{R}^3$. We recall:

Definition 8 *A symmetric bilinear form on a vector space V is a map $B : V \times V \to F$ such that*

- $B(v, w) = B(w, v)$
- $B(\lambda_1 v_1 + \lambda_2 v_2, w) = \lambda_1 B(v_1, w) + \lambda_2 B(v_2, w)$

The form is said to be nondegenerate if $B(v, w) = 0$ for all $w \in V$ implies $v = 0$.

If we take a basis $v_1, \ldots, v_n$ of V, then $v = \sum_i x_i v_i$ and $w = \sum_i y_i v_i$ so that

$$B(v, w) = \sum_{i,j} B(v_i, v_j) x_i y_j$$

and so is uniquely determined by the symmetric matrix $\beta_{ij} = B(v_i, v_j)$. The bilinear form is nondegenerate if and only if β_{ij} is nonsingular.

We can add symmetric bilinear forms: $(B+C)(v, w) = B(v, w) + C(v, w)$ and multiply by a scalar $(\lambda B)(v, w) = \lambda B(v, w)$ so they form a vector space isomorphic to the space of symmetric $n \times n$ matrices which has dimension $n(n+1)/2$. If we take a different basis

$$w_i = \sum_j P_{ji} v_j$$

then

$$B(w_i, w_j) = B(\sum_k P_{ki} v_k, \sum_\ell P_{\ell j} v_\ell) = \sum_{k,\ell} P_{ki} B(v_k, v_\ell) P_{\ell j}$$

so that the matrix $\beta_{ij} = B(v_i, v_j)$ changes under a change of basis to

$$\beta' = P^T \beta P.$$

Most of the time we shall be working over the real or complex numbers where we can divide by 2 and then we often speak of the quadratic form $B(v, v)$ which determines the bilinear form since

$$B(v + w, v + w) = B(v, v) + B(w, w) + 2B(v, w)$$

Here we have the basic result:

Theorem 10 *Let B be a quadratic form on a vector space V of dimension n over a field F. Then*

- *if $F = \mathbf{C}$, there is a basis such that if $v = \sum_i z_i v_i$*

$$B(v, v) = \sum_{i=1}^{m} z_i^2$$

- *if $F = \mathbf{R}$, there is a basis such that*

$$B(v, v) = \sum_{i=1}^{p} z_i^2 - \sum_{i=j}^{q} z_j^2.$$

If B is nondegenerate then $m = n = p + q$.

Proof: The proof is elementary – just *completing the square*. We note that changing the basis is equivalent to changing the coefficients x_i of v by an invertible linear transformation.

First we write down the form in one basis, so that

$$B(v, v) = \sum_{i,j} \beta_{ij} x_i x_j$$

and ask: *is there a term $\beta_{ii} \neq 0$?*. If not, then we create one. If the coefficient of $x_i x_j$ is non-zero, then putting $y_i = (x_i + x_j)/2, y_j = (x_i - x_j)/2$ we have

$$x_i x_j = y_i^2 - y_j^2$$

and so we get a term $\beta'_{ii} \neq 0$.

If there is a term $\beta_{ii} \neq 0$, then we note that

$$\frac{1}{\beta_{ii}}(\beta_{i1}x_1 + \ldots + \beta_{in}x_n)^2 = \beta_{ii}x_i^2 + 2\sum_{k \neq i} \beta_{ik}x_k x_i + R$$

where R involves the x_k with $k \neq i$. So if

$$y_i = \beta_{i1}x_1 + \ldots + \beta_{in}x_n$$

then

$$B(v,v) = \frac{1}{\beta_{ii}}y_i^2 + B_1$$

where B_1 is a quadratic form in the $n-1$ variables x_k, $k \neq i$.

We now repeat the procedure to find a basis such that if v has coefficients $y_1, \ldots, y_n$, then

$$B(v,v) = \sum_{i=1}^{m} c_i y_i^2.$$

Over $\mathbf{C}$ we can write $z_i = \sqrt{c_i}y_i$ and get a sum of squares and over $\mathbf{R}$ we put $z_i = \sqrt{|c_i|}y_i$ to get the required expression. □

Example: Consider the quadratic form in $\mathbf{R}^3$:

$$B(v,v) = x_1x_2 + x_2x_3 + x_3x_1.$$

We put

$$y_1 = (x_1 + x_2)/2, \qquad y_2 = (x_1 - x_2)/2$$

to get

$$B(v,v) = y_1^2 - y_2^2 + x_3(2y_1).$$

Now complete the square:

$$B(v,v) = (y_1 + x_3)^2 - y_2^2 - x_3^2$$

so that with $z_1 = y_1 + x_3, z_2 = y_2, y_3 = x_3$ we have $p = 1, q = 2$.

3.2 Quadrics and conics

Definition 9 *A quadric in a projective space $P(V)$ is the set of points whose representative vectors satisfy $B(v,v) = 0$ where B is a symmetric bilinear form on V. The quadric is said to be nonsingular if B is nondegenerate. The dimension of the quadric is* $\dim P(V) - 1$.

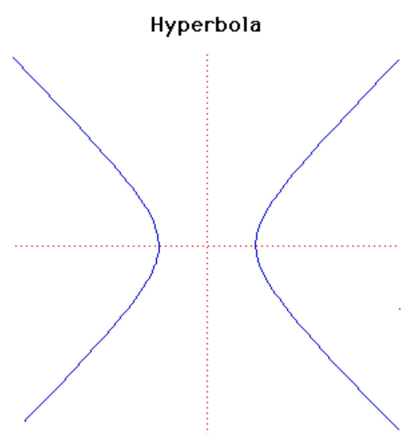
A quadric in a projective line is either empty or a pair of points. A quadric in a projective plane, a one-dimensional quadric, is called a conic.

Note that bilinearity of B means that $B(\lambda v, \lambda v) = \lambda^2 B(v, v)$ so that the set of points $[v] \in P(V)$ such that $B(v, v) = 0$ is well-defined. Also, clearly B and λB define the same quadric. The converse is not true in general, because if $F = \mathbf{R}$ and B is positive definite, then $B(v, v) = 0$ implies $v = 0$ so the quadric defined by B is the empty set. A little later we shall work over the complex numbers in general, as it makes life easier. But for the moment, to get some intuition, let us consider conics in $P^2(\mathbf{R})$ which are non-empty, and consider the intersection with $\mathbf{R}^2 \subset P^2(\mathbf{R})$ defined by the points $[x_0, x_1, x_2]$ such that $x_0 \neq 0$. Using coordinates $x = x_1/x_0, y = x_2/x_0$ this has the equation

$$\beta_{11}x^2 + 2\beta_{12}xy + \beta_{22}y^2 + 2\beta_{01}x + 2\beta_{02}y + \beta_{00} = 0.$$

Examples:

1. Consider the hyperbola $xy = 1$:



In $P^2(\mathbf{R})$ it is defined by the equation

$$x_1x_2 - x_0^2 = 0$$

and the line at infinity $x_0 = 0$ meets it where $x_1x_2 = 0$ i.e. at the two points $[0, 1, 0], [0, 0, 1]$.

Now look at it a different way: as in Theorem 10 we rewrite the equation as

$$(\frac{1}{2}(x_1 + x_2))^2 - (\frac{1}{2}(x_1 - x_2))^2 - x_0^2 = 0$$

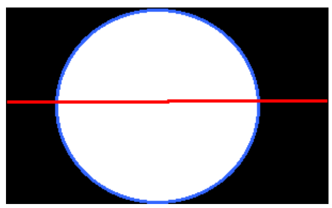
and then if $x_1 + x_2 \neq 0$, we put

$$y_1 = \frac{x_1 - x_2}{x_1 + x_2}, \quad y_2 = \frac{2x_0}{x_1 + x_2}$$

and the conic intersects the copy of $\mathbf{R}^2 \subset P^2(\mathbf{R})$ (the complement of the line $x_1+x_2 = 0$) in the circle

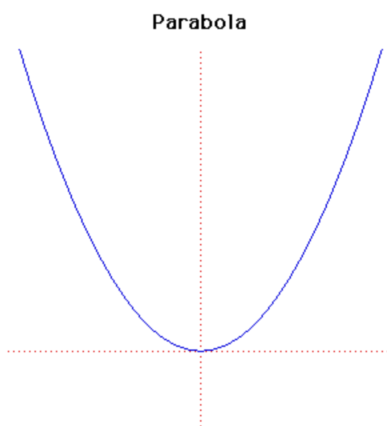
$$y_1^2 + y_2^2 = 1.$$

The original line at infinity $x_0 = 0$ meets this in $y_2 = 0$:



So a projective transformation allows us to view the two branches of the hyperbola as the two semicircles on each side of the line at infinity.

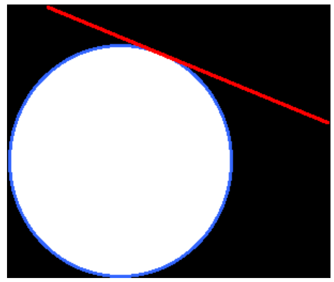
2. Now look at the parabola $y = x^2$.



The equation in homogeneous coordinates is

$$x_2 x_0 = x_1^2$$

and the line $x_0 = 0$ meets it in one point $[0, 0, 1]$. In projective space it still looks like a circle, but the single branch of the parabola is the complement of the point where the line at infinity meets the circle tangentially:



Thus the three different types of conics in $\mathbf{R}^2$ – ellipses, hyperbolas and parabolas – all become circles when we add in the points at infinity to make them sit in $P^2(\mathbf{R})$.

3.3 Rational parametrization of the conic

Topologically, we have just seen that the projective line $P^1(\mathbf{R})$ and a conic in $P^2(\mathbf{R})$ are both homeomorphic to a circle. In fact a much stronger result holds over any field.

Theorem 11 *Let C be a nonsingular conic in a projective plane $P(V)$ over the field F, and let A be a point on C. Let $P(U) \subset P(V)$ be a projective line not containing A. Then there is a bijection*

$$\alpha : P(U) \to C$$

such that, for $X \in P(U)$,the points $A, X, \alpha(X)$ are collinear.

Proof: Suppose the conic is defined by the nondegenerate symmetric bilinear form B. Let $a \in V$ be a representative vector for A, then $B(a, a) = 0$ since A lies on the conic. Let $x \in P(U)$ be a representative vector for $X \in P(U)$. Then a and x are linearly independent since X does not lie on the line $P(U)$. Extend a, x to a basis a, x, y of V.

Now B restricted to the space spanned by a, x is not identically zero, because if it were, the matrix of B with respect to this basis would be of the form

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

which is singular. So at least one of $B(x, x)$ and $B(a, x)$ is non-zero.

Any point on the line AX is represented by a vector of the form $\lambda a + \mu x$ and this lies on the conic C if

$$0 = B(\lambda a + \mu x, \lambda a + \mu x) = 2\lambda\mu B(a, x) + \mu^2 B(x, x).$$

When $\mu = 0$ we get the point X. The other solution is $2\lambda B(a, x) + \mu B(x, x) = 0$ i.e. the point with representative vector

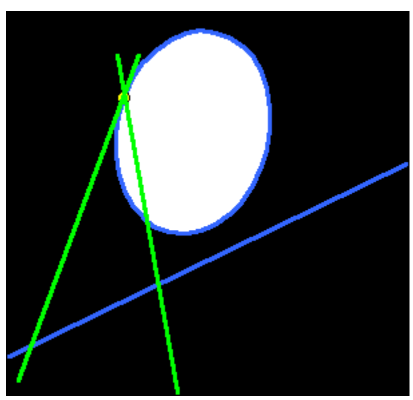
$$w = B(x, x)a - 2B(a, x)x \tag{4}$$

which is non-zero since the coefficients are not both zero.

We define the map $\alpha : P(U) \to C$ by

$$\alpha(X) = [w]$$

which has the collinearity property of the statement of the Theorem. If $Y \in C$ is distinct from A, then the line AY meets the line $P(U)$ in a unique point, so α^{-1} is well-defined on this subset. By the definition of α in (4), $\alpha(X) = A$ if and only if $B(a, x) = 0$. Since B is nonsingular $f(x) = B(a, x)$ is a non-zero linear map from V to F and so defines a line, which meets $P(U)$ in one point. Thus α has a well-defined inverse and is therefore a bijection. $\square$



Remark: There is a more invariant way of seeing this map by using duality. The line A^o in $P(V')$ is dual to the point A. Each point $Y \in A^0$ defines a line Y^o in $P(V)$ through A which intersects the conic C in a second point $\alpha(Y)$. What we do in the more concrete approach of the theorem is to compose this natural bijection with the projective transformation $A^0 \to P(U)$ defined by $Y \mapsto Y^0 \cap P(U)$.

Example: Consider the case of the conic

$$x_0^2 + x_1^2 - x_2^2 = 0.$$

Take $A = [1, 0, 1]$ and the line $P(U)$ defined by $x_0 = 0$. Note that this conic and the point and line are defined over any field since the coefficients are 0 or 1.

A point $X \in P(U)$ is of the form $X = [0, 1, t]$ or $[0, 0, 1]$ and the map α is

$$\begin{aligned} \alpha([0,1,t]) &= [B((0,1,t),(0,1,t))(1,0,1) - 2B((1,0,1),(0,1,t))(0,1,t)] \\ &= [1 - t^2, 2t, 1 + t^2] \end{aligned}$$

or $\alpha([0, 0, 1]) = [-1, 0, 1]$.

This has an interesting application if we use the field of rational numbers $F = \mathbf{Q}$. Suppose we want to find all right-angled triangles whose sides are of integer length. By Pythagoras, we want to find positive integer solutions to

$$x^2 + y^2 = z^2.$$

But then $[x, y, z]$ is a point on the conic. Conversely, if $[x_0, x_1, x_2]$ lies on the conic, then multiplying by the least common multiple of the denominators of the rational numbers x_0, x_1, x_2 gives integers such that $[x, y, z]$ is on the conic.

But what we have seen is that *any* point on the conic is either $[-1, 0, 1]$ or of the form

$$[x, y, z] = [1 - t^2, 2t, 1 + t^2]$$

for some rational number $t = p/q$, so we get all integer solutions by putting

$$x = q^2 - p^2, \quad y = 2pq, \quad z = q^2 + p^2.$$

For example, $p = 1, q = 2$ gives $3^2 + 4^2 = 5^2$ and $p = 2, q = 3$ gives $5^2 + 12^2 = 13^2$.

One other consequence of Theorem 11 is that we can express a point (x, y) on the general conic

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

in the form

$$x = \frac{p(t)}{r(t)}, \qquad y = \frac{q(t)}{r(t)}$$

where p, q and r are quadratic polynomials in t. Writing x, y as rational functions of t is why the process we have described is sometimes called the *rational parametrization of the conic.* It has its uses in integration. We can see, for example, that

$$\int \frac{dx}{x + \sqrt{ax^2 + bx + c}}$$

can be solved by elementary functions because if $y = x + \sqrt{ax^2 + bx + c}$ then

$$(y - x)^2 - ax^2 - bx - c = 0$$

and this is the equation of a conic. We can solve it by $x = p(t)/r(t), y = q(t)/r(t)$ and with this substitution, the integral becomes

$$\int \frac{r'(t)p(t) - p'(t)r(t)}{q(t)r(t)} dt$$

and expanding the rational integrand into partial fractions we get rational and logarithmic terms after integration.

3.4 Polars

We used a nondegenerate symmetric bilinear form B on a vector space V to define a quadric in $P(V)$ by the equation $B(v, v) = 0$. Such forms also define the notion of orthogonality $B(v, w) = 0$ and we shall see next what geometrically this corresponds to. First the linear algebra: given a subspace $U \subseteq V$ we can define its *orthogonal subspace* $U^\perp$ by

$$U^\perp = \{v \in V : B(u, v) = 0 \text{ for all } u \in U\}.$$

Note that unlike the Euclidean inner product, U and $U^\perp$ can intersect non-trivially – indeed a point with representative vector v lies on the quadric if it is orthogonal to itself. Note also that $U^\perp$ is the same if we change B to λB.

Orthogonal subspaces have a number of properties:

- $U = (U^\perp)^\perp$
- if $U_1 \subseteq U_2$, then $U_2^\perp \subseteq U_1^\perp$
- $\dim U^\perp + \dim U = \dim V$

These can be read off from the properties of the annihilator $U^0 \subseteq V'$, once we realize that a nondegenerate bilinear form on V defines an isomorphism between V and its dual V'. This is the map $\beta(v) = f_v$ where

$$f_v(w) = B(v, w).$$

The map $\beta : V \to V'$ defined this way is obviously linear in v and has zero kernel since $\beta(v) = 0$ implies $B(v, w) = 0$ for all w which means that $v = 0$ by nondegeneracy. Since $\dim V = \dim V'$, β is therefore an isomorphism, and one easily checks that

$$\beta(U^{\perp}) = U^{o}.$$

Definition 10 *If $X \in P(V)$ is represented by the one-dimensional subspace $U \subset V$, then the polar of X is the hyperplane $P(U^{\perp}) \subset P(V)$.*

At this stage, life becomes much easier if we work with the field of complex numbers $F = \mathbf{C}$. We should retain our intuition of conics, for example, as circles but realize that these really are just pictures for guidance. It was Jean-Victor Poncelet (1788-1867) who first systematically started to do geometry over $\mathbf{C}$ (he was also the one to introduce duality) and the simplifications it affords are really worthwhile. Poncelet's work on projective geometry began in Russia. As an officer in Napoleon's army, he was left for dead after the battle of Krasnoe, but was then found and spent several years as a prisoner of war, during which time he developed his mathematical ideas.

We consider then a complex projective plane $P(V)$ with a conic $C \subset P(V)$ defined by a non-degenerate symmetric bilinear form B. A *tangent* to C is a line which meets C at one point.

Proposition 12 *Let C be a nonsingular conic in a complex projective plane, then*

- *each line in the plane meets the conic in one or two points*
- *if $P \in C$, its polar line is the unique tangent to C passing through P*
- *if $P \notin C$, the polar line of P meets C in two points, and the tangents to C at these points intersect at P.*

Proof: Let $U \subset V$ be a 2-dimensional subspace defining the projective line $P(U)$ and let u, v be a basis for U. Then the point $[\lambda u + \mu v]$ lies on the conic if

$$0 = B(\lambda u + \mu v, \lambda u + \mu v) = \lambda^2 B(u, u) + 2\lambda\mu B(u, v) + \mu^2 B(v, v) \qquad (5)$$

Over the complex numbers this can be factorized as

$$0 = (a\lambda - b\mu)(a'\lambda - b'\mu)$$

giving the two (possibly coincident) points of intersection of the line and the conic

$$[bu + av], \quad [b'u + a'v].$$

Suppose the point P lies on C, and let u be a representative vector for P, so that $B(u, u) = 0$. Then any line through P is $P(U)$ where U is spanned by u and v. Then from (5) the points of intersection are given by

$$2\lambda\mu B(u, v) + \mu^2 B(v, v) = 0.$$

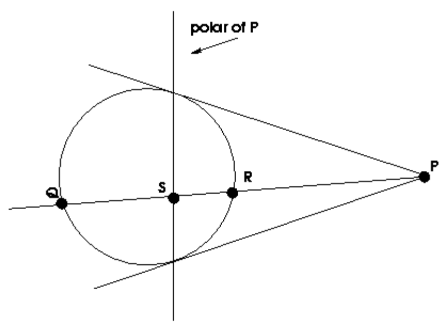
If the only point of intersection is $[u]$ then $\mu = 0$ is the only solution to this equation which means that $B(u, v) = 0$. Since any vector $w \in U$ is a linear combination of u and v and $B(u, u) = B(u, v) = 0$ this means $B(u, w) = 0$ and $P(U)$ is the polar line of X.

From the above, if P does not lie on C, its polar must meet C in two distinct points with representative vectors v_1, v_2. We then have

$$B(u, v_1) = 0 = B(v_1, v_1) \qquad (6)$$

Since $B(u, u) \neq 0$ and $B(v_1, v_1) = 0$, u and v_1 are linearly independent and span a 2-dimensional space U_1. From (6) $P(U_1)$ is the polar of $[v_1] \in C$ and hence is the tangent to C at $[v_1]$. Similarly for $[v_2]$. $\square$

The picture to bear in mind is the following real one, but even that does not tell the full story, since if P is inside the circle, its polar line intersects it in two complex conjugate points, so although we can draw the point and its polar, we can't see the two tangents.



Quadrics are nonlinear subsets of $P(V)$ but they nevertheless contain many linear subspaces. For example if $Q \subset P(V)$ is a nonsingular quadric, then $P(U) \subset Q$ if and only if $B(u, u) = 0$ for all $u \in U$. This implies

$$2B(u_1, u_2) = B(u_1 + u_2, u_1 + u_2) - B(u_1, u_1) - B(u_2, u_2) = 0$$

and hence

$$U \subset U^{\perp}.$$

Since

$$\dim U + \dim U^{\perp} = \dim V$$

this means that

$$2 \dim U \leq \dim U + \dim U^{\perp} = \dim V.$$

In fact, over $\mathbf{C}$ the maximum value always occurs. If $\dim V = 2m$, then from Theorem 10 there is a basis in which $B(v, v)$ can be written as

$$x_1^2 + x_2^2 + \ldots + x_{2m}^2 = (x_1 + ix_2)(x_1 - ix_2) + \ldots + (x_{2m-1} + ix_{2m})(x_{2m-1} - ix_{2m})$$

and so if U is defined by

$$x_1 - ix_2 = x_3 - ix_4 = \ldots = x_{2m-1} - ix_{2m} = 0,$$

then $\dim U = \dim V/2$ and $U \subseteq U^{\perp}$ so $U = U^{\perp}$. Over $\mathbf{R}$ this occurs when $p = q = m$ and the form can be reduced to

$$x_1^2 + \ldots + x_m^2 - x_{m+1}^2 \ldots - x_{2m}^2.$$

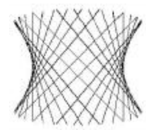
We can see this in more detail for the quadric surface

$$x_1^2 - x_2^2 = x_3^2 - x_4^2$$

in $P^3(\mathbf{R})$. This intersects the copy of $\mathbf{R}^3 \subset P^3(\mathbf{R})$ defined by $x_4 \neq 0$ in the hyperboloid of revolution

$$x^2 + y^2 - z^2 = 1$$

where $x = x_2/x_4, y = x_3/x_4, z = x_1/x_4$. This is the usual "cooling tower" shape



There are two one-parameter families of lines in the quadric given by:

$$\begin{aligned} \lambda(x_1 - x_2) &= \mu(x_3 - x_4) \\ \mu(x_1 + x_2) &= \lambda(x_3 + x_4) \end{aligned}$$

and

$$\begin{aligned} \lambda(x_1 - x_2) &= \mu(x_3 + x_4) \\ \mu(x_1 + x_2) &= \lambda(x_3 - x_4). \end{aligned}$$

In fact these two families of lines provide "coordinates" for the projective quadric: the map

$$F : P^1(\mathbf{R}) \times P^1(\mathbf{R}) \to Q$$

defined by

$$F([u_0, u_1], [v_0, v_1]) = [u_0v_0 + u_1v_1, u_1v_1 - u_0v_0, u_0v_1 + u_1v_0, u_1v_0 - u_0v_1]$$

is a bijection.

3.5 Pencils of quadrics

The previous sections have dealt with the geometrical interpretation of a symmetric bilinear form. Now we look at the theory behind a *pair* of bilinear forms and we shall see how the geometry helps us to determine algebraic properties of these.

We saw in Theorem 10 that over $\mathbf{C}$, any quadratic form B can be expressed in some basis as

$$B(v, v) = x_1^2 + \ldots + x_n^2.$$

In particular the matrix of B is diagonal (actually the identity matrix). If we have a pair A, B of symmetric bilinear forms we ask whether we can simultaneously diagonalize them both. The answer is:

Proposition 13 *Let α, β be symmetric $n \times n$ matrices with complex entries, and suppose α is non-singular. Then if the equation $\det(\lambda\alpha - \beta) = 0$ has n distinct solutions $\lambda_1, \ldots, \lambda_n$ there is an invertible matrix P such that*

$$P^T \alpha P = I, \qquad P^T \beta P = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Proof: From Theorem 10 we can find an invertible matrix Q such that $Q^T \alpha Q = I$. Write $\beta' = Q^T \beta Q$, so that β' is also symmetric. Then

$$\det Q^T \det(\lambda\alpha - \beta) \det Q = \det(Q^T(\lambda\alpha - \beta)Q) = \det(\lambda I - \beta')$$

and so the roots of $\det(\lambda\alpha - \beta) = 0$ are the eigenvalues of β'. By assumption these are distinct, so we have a basis of eigenvectors $v_1, \ldots, v_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$.

If $v_\ell = (x_1, \ldots, x_n)$ and $v_k = (y_1, \ldots, y_n)$ then

$$\lambda_\ell \sum_i x_i y_i = \sum_{i,j} \beta'_{ij} x_j y_i = \sum_{i,j} \beta'_{ji} x_j y_i = \lambda_k \sum_i x_i y_i$$

and since $\lambda_\ell \neq \lambda_k$, we have

$$\sum_i x_i y_i = 0.$$

Thus v_k and v_ℓ are orthogonal if $k \neq \ell$. We also must have $(v_i, v_i) \neq 0$ since otherwise v_i is orthogonal to each element of the basis $v_1, \ldots, v_n$, so we can write

$$w_i = \frac{1}{\sqrt{(v_i, v_i)}} v_i$$

and obtain an orthonormal basis. With respect to this basis, β' is diagonal so if R is the invertible matrix defining the change of basis, $R^T \beta' R = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $R^T R = I$. Putting $P = QR$ we get the result. $\square$

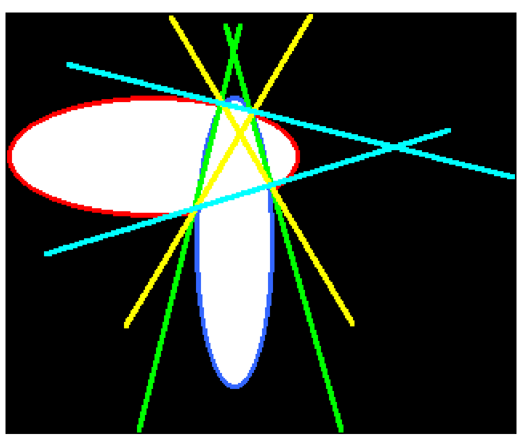
Now let us try and set this in a geometric context. Let A, B be symmetric bilinear forms on a complex vector space V which define different quadrics Q and Q'. Then

Definition 11 *The pencil of quadrics in $P(V)$ generated by Q and Q' is the set of quadrics defined by $\lambda A + \mu B$ where $(\lambda, \mu) \neq (0, 0)$.*

Another way of saying this is to let SV denote the vector space of symmetric bilinear forms on V, in which case a pencil of quadrics is a line in $P(SV)$ – a family of quadrics parametrized by a projective line. The singular quadrics in this pencil are given by the equation $\det(\lambda\alpha+\mu\beta) = 0$. If α is nonsingular then for a solution to this equation $\mu \neq 0$, and so under the hypotheses of the proposition, we can think of the points $[\lambda_i, -1] \in P^1(\mathbf{C})$ as defining n singular quadrics in the pencil. The geometry of this becomes directly visible in the case that $P(V)$ is a plane. In this case a singular quadric has normal form x_1^2 or $x_1^2+x_2^2 = (x_1-ix_2)(x_1+ix_2)$ and so is either a double line or a pair of lines.

Theorem 14 *Let C and C' be nonsingular conics in a complex projective plane and assume that the pencil generated by C and C' contains three singular conics. Then*

- *the pencil consists of all conics passing through four points in general position*
- *the singular conics of the pencil consist of the three pairs of lines obtained by joining disjoint pairs of these four points*
- *each such pair of lines meets in a point with representative vector v_i where $\{v_1, v_2, v_3\}$ is a basis for V relative to which the matrices of C and C' are simultaneously diagonalizable.*



Proof: The proof consists of reducing to normal form and calculating. Since by hypothesis there are three singular conics in the pencil, Proposition 13 tells us that there is a basis in which the two conics are defined by bilinear forms A, B with equations:

$$x_1^2 + x_2^2 + x_3^2 = 0, \qquad \lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2 = 0$$

and we directly find the points of intersection are the four points

$$[\sqrt{\lambda_2 - \lambda_3}, \pm\sqrt{\lambda_3 - \lambda_1}, \pm\sqrt{\lambda_1 - \lambda_2}]$$

where we fix one choice of square root of $\lambda_2 - \lambda_3$. To show that these are in general position we need to show that any three are linearly independent, but, for example

$$\det \begin{pmatrix} \sqrt{\lambda_2 - \lambda_3} & \sqrt{\lambda_3 - \lambda_1} & \sqrt{\lambda_1 - \lambda_2} \\ \sqrt{\lambda_2 - \lambda_3} & -\sqrt{\lambda_3 - \lambda_1} & \sqrt{\lambda_1 - \lambda_2} \\ \sqrt{\lambda_2 - \lambda_3} & -\sqrt{\lambda_3 - \lambda_1} & -\sqrt{\lambda_1 - \lambda_2} \end{pmatrix} = 4\sqrt{\lambda_1 - \lambda_2}\sqrt{\lambda_2 - \lambda_3}\sqrt{\lambda_3 - \lambda_1}$$

and since the λ_i are distinct this is non-zero.

Now clearly if $[u]$ is one of these four points $A(u, u) = 0, B(u, u) = 0$ and so $(\lambda A + \mu B)(u, u) = 0$, and every conic in the pencil passes through them. Conversely by Theorem 3 we can take the points to be

$$[1, 0, 0], [0, 1, 0], [0, 0, 1], [1, 1, 1].$$

The conics which pass through these points have matrices β_{ij} where

$$\beta_{11} = \beta_{22} = \beta_{33} = \beta_{12} + \beta_{23} + \beta_{31} = 0 \tag{7}$$

The vector space of symmetric 3×3 matrices is of dimension 6 spanned by

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

and the four equations (7) are linearly independent, so define a 2-dimensional space of bilinear forms spanned by A and B – this is the pencil.

We need to understand the singular quadrics in the pencil, for example

$$\lambda_1(x_1^2 + x_2^2 + x_3^2) - (\lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2) = (\lambda_1 - \lambda_2)x_2^2 - (\lambda_3 - \lambda_1)x_3^2 = 0$$

But this factorizes as

$$(\sqrt{\lambda_1 - \lambda_2}x_2 - \sqrt{\lambda_3 - \lambda_1}x_3)(\sqrt{\lambda_1 - \lambda_2}x_2 + \sqrt{\lambda_3 - \lambda_1}x_3) = 0.$$

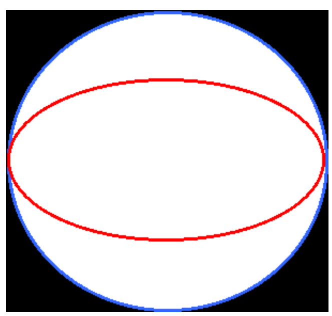
These two lines, by belonging to the pencil, pass through the points of intersection, but since those points are in general position, each line passes through only two of them, so these are the required pairs of lines.

The intersection of $\sqrt{\lambda_1 - \lambda_2}x_2 - \sqrt{\lambda_3 - \lambda_1}x_3 = 0$ and $\sqrt{\lambda_1 - \lambda_2}x_2 + \sqrt{\lambda_3 - \lambda_1}x_3 = 0$ is $[1, 0, 0]$, which is the first basis vector in which the two conics are diagonalized. Similarly the other two pairs of lines give the remaining basis vectors. □

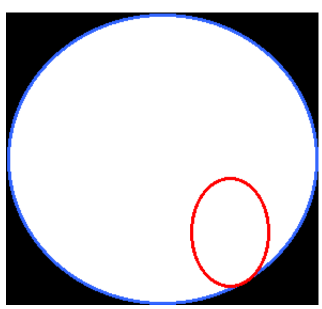
This geometrical approach tells us when two symmetric bilinear forms in three variables can and can not be simultaneously diagonalized. If the values λ_i are not all distinct, and the two forms *can* be simultaneously diagonalized, then they are either multiples of each other or are of the form

$$x_1^2 + x_2^2 + x_3^2 = 0, \qquad \lambda x_1^2 + x_2^2 + x_3^2 = 0.$$

In this case the two conics intersect where $x_1 = 0, x_2 = \pm i x_3$, i.e. at two points:



The non-diagonalizable case is where the intersection is at an odd number of points:



3.6 Exercises

1. Which of the following quadratic forms defines a non-singular conic?

- $x_0^2 - 2x_0x_1 + 4x_0x_2 - 8x_1^2 + 2x_1x_2 + 3x_2^2$
- $x_0^2 - 2x_0x_1 + x_1^2 - 2x_0x_2$.

2. Take five points in a projective plane such that no three are collinear. Show that there exists a unique non-singular conic passing through all five points.

3. Let C be a non-singular conic in a projective plane $P(V)$. Show that if $X \in P(V)$ moves on a fixed line ℓ, then its polar passes though a fixed point Y. What is the relationship between the point Y and the line ℓ?

4. Let $\tau : P^1(\mathbf{R}) \to P^1(\mathbf{R})$ be a projective transformation and consider its graph

$$\Gamma_\tau \subset P^1(\mathbf{R}) \times P^1(\mathbf{R})$$

i.e. $\Gamma_\tau = \{(X, Y) : Y = \tau(X)\}$. Using the one-to-one correspondence between $P^1(\mathbf{R}) \times P^1(\mathbf{R})$ and a quadric surface in $P^3(\mathbf{R})$, show that Γ_τ is the intersection of the quadric surface with a plane.

5. Prove that if $L \subseteq V$ and $M \subseteq V$ are vector subspaces of the same dimension then

$$\dim(L \cap M^\perp) = \dim(L^\perp \cap M).$$

6. Show that the two quadratic forms

$$x^2 + y^2 - z^2, \qquad x^2 + y^2 - yz$$

cannot be simultaneously diagonalized.

7. Let $P^5(\mathbf{R}) = P(\mathbf{R}^6)$ be the space of all conics in $P^2(\mathbf{R})$. Show that the conics which pass through three non-collinear points form a projective plane $P(V) \subset P^5(\mathbf{R})$. Show further that the conics parametrized by this plane and which are tangent to a given line form a conic in $P(V)$.

8. Prove that the set of tangent lines to a nonsingular conic in $P(V)$ is a conic in the dual space $P(V')$.

4 Exterior algebra

4.1 Lines and 2-vectors

The time has come now to develop some new linear algebra in order to handle the space of lines in a projective space $P(V)$. In the projective plane we have seen that duality can deal with this but lines in higher dimensional spaces behave differently. From the point of view of linear algebra we are looking at 2-dimensional vector subspaces $U \subset V$.

To motivate what we shall do, consider how in Euclidean geometry we describe a 2-dimensional subspace of $\mathbf{R}^3$. We could describe it through its unit normal $\mathbf{n}$, which is also parallel to $\mathbf{u}\times\mathbf{v}$ where $\mathbf{u}$ and $\mathbf{v}$ are linearly independent vectors in the space and $\mathbf{u}\times\mathbf{v}$ is the vector cross product. The vector product has the following properties:

- $\mathbf{u}\times\mathbf{v} = -\mathbf{v}\times\mathbf{u}$
- $(\lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2)\times\mathbf{v} = \lambda_1\mathbf{u}_1\times\mathbf{v} + \lambda_2\mathbf{u}_2\times\mathbf{v}$

We shall generalize these properties to vectors in any vector space V – the difference is that the product will not be a vector in V, but will lie in another associated vector space.

Definition 12 *An alternating bilinear form on a vector space V is a map $B : V \times V \to F$ such that*

- $B(v, w) = -B(w, v)$
- $B(\lambda_1 v_1 + \lambda_2 v_2, w) = \lambda_1 B(v_1, w) + \lambda_2 B(v_2, w)$

This is the skew-symmetric version of the symmetric bilinear forms we used to define quadrics. Given a basis $\{v_1, \ldots, v_n\}$, B is uniquely determined by the skew symmetric matrix $B(v_i, v_j)$. We can add alternating forms and multiply by scalars so they form a vector space, isomorphic to the space of skew-symmetric $n \times n$ matrices. This has dimension $n(n-1)/2$, spanned by the basis elements E^{ab} for $a < b$ where $E^{ab}_{ij} = 0$ if $\{a, b\} \neq \{i, j\}$ and $E^{ab}_{ab} = -E^{ab}_{ba} = 1$.

Definition 13 *The second exterior power $\Lambda^2 V$ of a finite-dimensional vector space is the dual space of the vector space of alternating bilinear forms on V. Elements of $\Lambda^2 V$ are called 2-vectors.*

This definition is a convenience – there are other ways of defining $\Lambda^2 V$, and for most purposes it is only its characteristic properties which one needs rather than what its objects are. A lot of mathematics is like that – just think of the real numbers.

Given this space we can now define our generalization of the cross-product, called the *exterior product* or *wedge product* of two vectors.

Definition 14 *Given $u, v \in V$ the exterior product $u \wedge v \in \Lambda^2 V$ is the linear map to F which, on an alternating bilinear form B, takes the value*

$$(u \wedge v)(B) = B(u, v).$$

From this definition follows some basic properties:

- $(u \wedge v)(B) = B(u, v) = -B(v, u) = -(v \wedge u)(B)$ so that

$$v \wedge u = -u \wedge v$$

  and in particular $u \wedge u = 0$.

- $((\lambda_1 u_1 + \lambda_2 u_2) \wedge v)(B) = B(\lambda_1 u_1 + \lambda_2 u_2, v) = \lambda_1 B(u_1, v) + \lambda_2 B(u_2, v)$ which implies

$$(\lambda_1 u_1 + \lambda_2 u_2) \wedge v = \lambda_1 u_1 \wedge v + \lambda_2 u_2 \wedge v.$$

- if $\{v_1, \ldots, v_n\}$ is a basis for V then $v_i \wedge v_j$ for $i < j$ is a basis for $\Lambda^2 V$.

This last property holds because $v_i \wedge v_j(E^{ab}) = E^{ab}_{ij}$ and in facts shows that $\{v_i \wedge v_j\}$ is the dual basis to the basis $\{E^{ab}\}$.

Another important property is:

Proposition 15 *Let $u \in V$ be a non-zero vector. Then $u \wedge v = 0$ if and only if $v = \lambda u$ for some scalar λ.*

Proof: If $v = \lambda u$, then

$$u \wedge v = u \wedge (\lambda u) = \lambda(u \wedge u) = 0.$$

Conversely, if $v \neq \lambda u$, u and v are linearly independent and can be extended to a basis, but then $u \wedge v$ is a basis vector and so is non-zero. $\square$

It is the elements of $\Lambda^2 V$ of the form $u \wedge v$ which will concern us, for suppose $U \subset V$ is a 2-dimensional vector subspace, and $\{u, v\}$ is a basis of U. Then any other basis is of the form $\{au + bv, cu + dv\}$, so, using $u \wedge u = v \wedge v = 0$, we get

$$(au + bv) \wedge (cu + dv) = (ad - bc) u \wedge v$$

and since the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is invertible $ad - bc \neq 0$. It follows that the 1-dimensional subspace of $\Lambda^2 V$ spanned by $u \wedge v$ for a basis of U is well-defined by U itself and is independent of the choice of basis. To each line in $P(V)$ we can therefore associate a *point* in $P(\Lambda^2 V)$.

The problem is, not every vector in $\Lambda^2 V$ can be written as $u \wedge v$ for vectors $u, v \in V$. In general it is a linear combination of such expressions. The task, in order to describe the space of lines, is to characterize such *decomposable* 2-vectors.

Example: Consider $v_1 \wedge v_2 + v_3 \wedge v_4$ in a 4-dimensional vector space V. Suppose we can write this as

$$v_1 \wedge v_2 + v_3 \wedge v_4 = (a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4) \wedge (b_1 v_1 + b_2 v_2 + b_3 v_3 + b_4 v_4).$$

Equating the coefficient of $v_1 \wedge v_2$ gives

$$a_1 b_2 - a_2 b_1 = 1$$

and so (a_1, b_1) is non-zero. On the other hand the coefficients of $v_1 \wedge v_3$ and $v_1 \wedge v_4$ give

$$\begin{aligned} a_1 b_3 - a_3 b_1 &= 0 \\ a_1 b_4 - a_4 b_1 &= 0 \end{aligned}$$

and since $(a_1, b_1) \neq 0$, $b_3 a_4 - a_3 b_4 = 0$. But the coefficient of $v_3 \wedge v_4$ gives $a_4 b_3 - a_3 b_4 = 1$ which is a contradiction. This 2-vector is not therefore decomposable. We shall find an easier method of seeing this by working with p-vectors and exterior products.

4.2 Higher exterior powers

Definition 15 *An alternating multilinear form of degree p on a vector space V is a map $M : V \times \ldots \times V \to F$ such that*

- $M(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_p) = -M(u_1, \ldots, u_j, \ldots, u_i, \ldots, u_p)$
- $M(\lambda_1 v_1 + \lambda_2 v_2, u_2, \ldots, u_p) = \lambda_1 M(v_1, u_2, \ldots, u_p) + \lambda_2 M(v_2, u_2, \ldots, u_p)$

Example: Let $u_1, \ldots, u_n$ be column vectors in $\mathbf{R}^n$. Then

$$M(u_1, \ldots, u_n) = \det(u_1 u_2 \ldots u_n)$$

is an alternating multilinear form of degree n.

The set of all alternating multilinear forms on V is a vector space, and M is uniquely determined by the values

$$M(v_{i_1}, v_{i_2}, \ldots, v_{i_p})$$

for a basis $\{v_1, \ldots, v_n\}$. But the alternating property allows us to change the order so long as we multiply by -1 for each transposition of variables. This means that M is uniquely determined by the values of indices for

$$i_1 < i_2 < \ldots < i_p.$$

The number of these is the number of p-element subsets of n, i.e. $\binom{n}{p}$, so this is the dimension of the space of such forms. In particular if $p > n$ this space is zero. We define analogous constructions to those above for a pair of vectors:

Definition 16 *The p-th exterior power $\Lambda^p V$ of a finite-dimensional vector space is the dual space of the vector space of alternating multilinear forms of degree p on V. Elements of $\Lambda^p V$ are called p-vectors.*

and

Definition 17 *Given $u_1, \ldots, u_p \in V$ the exterior product $u_1 \wedge u_2 \wedge \ldots \wedge u_p \in \Lambda^p V$ is the linear map to F which, on an alternating multilinear form M takes the value*

$$(u_1 \wedge u_2 \wedge \ldots \wedge u_p)(M) = M(u_1, u_2, \ldots, u_p).$$

The exterior product $u_1 \wedge u_2 \wedge \ldots \wedge u_p$ has two defining properties

- it is linear in each variable u_i separately
- interchanging two variables changes the sign of the product

- if two variables are the same the exterior product vanishes.

We have a useful generalization of Proposition 15:

Proposition 16 *The exterior product $u_1 \wedge u_2 \wedge \ldots \wedge u_p$ of p vectors $u_i \in V$ vanishes if and only if the vectors are linearly dependent.*

Proof: If there exists a linear relation

$$\lambda_1 u_1 + \ldots \lambda_p u_p = 0$$

with $\lambda_i \neq 0$, then u_i is a linear combination of the other vectors

$$u_i = \sum_{j \neq i} \mu_j u_j$$

but then

$$u_1 \wedge u_2 \wedge \ldots \wedge u_p = u_1 \wedge \ldots \wedge (\sum_{j \neq i} \mu_j u_j) \wedge u_{i+1} \wedge \ldots \wedge u_p$$

and expand this out by linearity, each term has a repeated variable u_j and so vanishes.

Conversely, if $u_1, \ldots, u_p$ are linearly independent they can be extended to a basis and $u_1 \wedge u_2 \wedge \ldots \wedge u_p$ is a basis vector for $\Lambda^p V$ and is thus non-zero. □

The exterior powers $\Lambda^p V$ have natural properties with respect to linear transformations: given a linear transformation $T : V \to W$, and an alternating multilinear form M on W we can define an induced one $T^* M$ on V by

$$T^* M(v_1, \ldots, v_p) = M(T v_1, \ldots, T v_p)$$

and this defines a dual linear map

$$\Lambda^p T : \Lambda^p V \to \Lambda^p W$$

with the property that

$$\Lambda^p T(v_1 \wedge v_2 \wedge \ldots \wedge v_p) = T v_1 \wedge T v_2 \wedge \ldots \wedge T v_p.$$

One such map is very familiar: take $p = n$, so that $\Lambda^n V$ is one-dimensional and spanned by $v_1 \wedge v_2 \wedge \ldots \wedge v_n$ for a basis $\{v_1, \ldots, v_n\}$. A linear transformation from a

1-dimensional vector space to itself is just multiplication by a scalar, so $\Lambda^n T$ is some scalar in the field. In fact it is the *determinant* of T. To see this, observe that

$$\Lambda^n T(v_1 \wedge \ldots \wedge v_n) = Tv_1 \wedge \ldots \wedge Tv_n$$

and the right hand side can be written using the matrix T_{ij} of T as

$$\sum_{i_1,\ldots,i_n} T_{i_1 1} v_{i_1} \wedge \ldots \wedge T_{i_n n} v_{i_n} = \sum_{i_1,\ldots,i_n} T_{i_1 1} \ldots T_{i_n n} v_{i_1} \wedge \ldots \wedge v_{i_n}.$$

Each of the terms vanishes if any two of $i_1, \ldots, i_n$ are equal by the property of the exterior product, so we need only consider the case where $(i_1, \ldots, i_n)$ is a permutation of $(1, \ldots, n)$. Any permutation is a product of transpositions, and any transposition changes the sign of the exterior product, so

$$\Lambda^n T(v_1 \wedge \ldots \wedge v_n) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) T_{\sigma(1)1} T_{\sigma(2)2} \ldots T_{\sigma(n)n} v_1 \wedge \ldots \wedge v_n$$

which is the definition of the determinant of T_{ij}. From our point of view the determinant is naturally defined for a linear transformation $T : V \to V$, and what we just did was to see how to calculate it from the matrix of T.

We now have vector spaces $\Lambda^p V$ of dimension $\binom{n}{p}$ naturally associated to V. The space $\Lambda^1 V$ is by definition the dual space of the space of linear functions on V, so $\Lambda^1 V = V'' \cong V$ and by convention we set $\Lambda^0 V = F$. Given p vectors $v_1, \ldots, v_p \in V$ we also have a corresponding vector $v_1 \wedge v_2 \wedge \ldots \wedge v_p \in \Lambda^p V$ and the notation suggests that there should be a product so that we can remove the brackets:

$$(u_1 \wedge \ldots \wedge u_p) \wedge (v_1 \wedge \ldots v_q) = u_1 \wedge \ldots \wedge u_p \wedge v_1 \wedge \ldots v_q$$

and indeed there is. So suppose $a \in \Lambda^p V, b \in \Lambda^q V$, we want to define $a \wedge b \in \Lambda^{p+q} V$. Now for fixed vectors $u_1, \ldots, u_p \in V$,

$$M(u_1, u_2, \ldots, u_p, v_1, v_2, \ldots, v_q)$$

is an alternating multilinear function of $v_1, \ldots, v_q$, so if

$$b = \sum_{j_1 < \ldots < j_q} \lambda_{j_1 \ldots j_q} v_{j_1} \wedge \ldots \wedge v_{j_q}$$

then

$$\sum_{j_1 < \ldots < j_q} \lambda_{j_1 \ldots j_q} M(u_1, \ldots, u_p, v_{j_1}, \ldots, v_{j_q})$$

only depends on b and not on the particular way it is written in terms of a basis $\{v_1, \ldots, v_n\}$. Similarly if

$$a = \sum_{i_1 < \ldots < i_p} \mu_{i_1 \ldots i_p} u_{i_1} \wedge \ldots \wedge u_{i_p}$$

then

$$\sum_{i_1 < \ldots < i_p} \mu_{i_1 \ldots i_p} M(u_{i_1}, \ldots, u_{i_p}, v_1, \ldots, v_q)$$

only depends on a. We can therefore unambiguously define $a \wedge b$ by its value on an alternating $p+q$-form M as

$$(a \wedge b)(M) = \sum_{i_1 < .. < i_p ; j_i, .. < j_q} \mu_{i_1 \ldots i_p} \lambda_{j_1 \ldots j_q} M(u_{i_1}, \ldots, u_{i_p}, v_{j_1}, \ldots, v_{j_q}).$$

The product just involves linearity and removing the brackets.

Example: Suppose $a = v_1 + v_2$, $b = v_1 \wedge v_3 - v_3 \wedge v_2$, with $v_1, v_2, v_3 \in V$ then

$$\begin{aligned} a \wedge b &= (v_1 + v_2) \wedge (v_1 \wedge v_3 - v_3 \wedge v_2) \\ &= v_1 \wedge v_1 \wedge v_3 - v_1 \wedge v_3 \wedge v_2 + v_2 \wedge v_1 \wedge v_3 - v_2 \wedge v_3 \wedge v_2 \\ &= -v_1 \wedge v_3 \wedge v_2 + v_2 \wedge v_1 \wedge v_3 \\ &= v_1 \wedge v_2 \wedge v_3 - v_1 \wedge v_2 \wedge v_3 = 0 \end{aligned}$$

where we have used the basic rules that a repeated vector from V in an exterior product gives zero, and the interchange of two vectors changes the sign.

Note that

$$u_1 \wedge u_2 \wedge \ldots \wedge u_p \wedge v_1 \wedge \ldots \wedge v_q = (-1)^p v_1 \wedge u_1 \wedge u_2 \wedge \ldots \wedge u_p \wedge v_2 \wedge \ldots \wedge v_q$$

because we have to interchange v_1 with each of the p u_i's to bring it to the front, and then repeating

$$u_1 \wedge u_2 \wedge \ldots \wedge u_p \wedge v_1 \wedge \ldots \wedge v_q = (-1)^{pq} v_1 \wedge \ldots \wedge v_q \wedge u_1 \wedge u_2 \wedge \ldots \wedge u_p.$$

This extends by linearity to all $a \in \Lambda^p V, b \in \Lambda^q V$. We then have the basic properties of the exterior product;

- $a \wedge (b + c) = a \wedge b + a \wedge c$

- $(a \wedge b) \wedge c = a \wedge (b \wedge c)$
- $a \wedge b = (-1)^{pq} b \wedge a$ if $a \in \Lambda^p V, b \in \Lambda^q V$

What we have done may seem rather formal, but it has many concrete applications. For example if $a = x \wedge y$ then $a \wedge a = x \wedge y \wedge x \wedge y = 0$ because $x \in V$ is repeated. So it is much easier to determine that $a = v_1 \wedge v_2 + v_3 \wedge v_4$ from the Exercise above is not decomposable:

$$(v_1 \wedge v_2 + v_3 \wedge v_4) \wedge (v_1 \wedge v_2 + v_3 \wedge v_4) = 2 v_1 \wedge v_2 \wedge v_3 \wedge v_4 \neq 0.$$

4.3 Decomposable 2-vectors

A line in $P(V)$ defines a point in $P(\Lambda^2 V)$ defined by a *decomposable* 2-vector

$$a = x \wedge y.$$

We need to characterize algebraically this decomposability, and the following theorem does just that:

Theorem 17 *Let $a \in \Lambda^2 V$ be a non-zero element. Then a is decomposable if and only if $a \wedge a = 0 \in \Lambda^4 V$.*

Proof: If $a = x \wedge y$ for two vectors x and y then

$$a \wedge a = x \wedge y \wedge x \wedge y = 0$$

because of the repeated factor x (or y).

We prove the converse by induction on the dimension of V. If $\dim V = 0, 1$ then $\Lambda^2 V = 0$, so the first case is $\dim V = 2$. In this case $\dim \Lambda^2 V = 1$ and $v_1 \wedge v_2$ is a non-zero element if v_1, v_2 is a basis for V, so any a is decomposable.

We consider the case $\dim V = 3$ separately now. Given a non-zero $a \in \Lambda^2 V$, define $A : V \to \Lambda^3 V$ by

$$A(v) = a \wedge v.$$

Since $\dim \Lambda^3 V = 1$, $\dim \ker A \geq 2$, so let u_1, u_2 be linearly independent vectors in the kernel and extend to a basis u_1, u_2, u_3 of V. We can then write

$$a = \lambda_1 u_2 \wedge u_3 + \lambda_2 u_3 \wedge u_1 + \lambda_3 u_1 \wedge u_2.$$

Now by definition $0 = a \wedge u_1 = \lambda_1 u_2 \wedge u_3 \wedge u_1$ so $\lambda_1 = 0$ and similarly $0 = a \wedge u_2$ implies $\lambda_2 = 0$. It follows that $a = \lambda_3 u_1 \wedge u_2$, which is decomposable.

Now assume inductively that the theorem is true for $\dim V \leq n-1$ and consider the case $\dim V = n$. Using a basis $v_1, \ldots, v_n$, write

$$\begin{aligned} a &= \sum_{1 \leq i < j}^{n} a_{ij} v_i \wedge v_j \\ &= (\sum_{i=1}^{n-1} a_{in} v_i) \wedge v_n + \sum_{1 \leq i < j}^{n-1} a_{ij} v_i \wedge v_j \\ &= u \wedge v_n + a' \end{aligned}$$

where $u \in U$ and $a' \in \Lambda^2 U$ and U is the $(n-1)$-dimensional space spanned by $v_1, \ldots, v_{n-1}$.

Now

$$0 = a \wedge a = (u \wedge v_n + a') \wedge (u \wedge v_n + a') = 2u \wedge a' \wedge v_n + a' \wedge a'.$$

But v_n doesn't appear in the expansion of $u \wedge a'$ or $a' \wedge a'$ so we separately obtain

$$u \wedge a' = 0, \qquad a' \wedge a' = 0.$$

By induction $a' \wedge a' = 0$ implies $a' = u_1 \wedge u_2$ and so the first equation reads

$$u \wedge u_1 \wedge u_2 = 0$$

which from Proposition 16 says that there is a linear relation

$$\lambda u + \mu_1 u_1 + \mu_2 u_2 = 0.$$

If $\lambda = 0$, then u_1 and u_2 are linearly dependent so $a' = u_1 \wedge u_2 = 0$. This means that $u = u \wedge v_n$ and is therefore decomposable. If $\lambda \neq 0$, $u = \lambda_1 u_1 + \lambda_2 u_2$, so

$$a = \lambda_1 u_1 \wedge v_n + \lambda_2 u_2 \wedge v_n + u_1 \wedge u_2$$

and this is the 3-dimensional case which is always decomposable as we showed above.

We conclude that a, in each case, is decomposable. □

4.4 The Klein quadric

The first case where we can apply Theorem 17 is when $\dim V = 4$, to describe the projective lines in the 3-dimensional space $P(V)$. In this case $\dim \Lambda^4 V = 1$ with a basis vector $v_0 \wedge v_1 \wedge v_2 \wedge v_3$ if V is given the basis $v_0, \ldots, v_3$.

For $a \in \Lambda^2 V$ we write

$$a = \lambda_1 v_0 \wedge v_1 + \lambda_2 v_0 \wedge v_2 + \lambda_3 v_0 \wedge v_3 + \mu_1 v_2 \wedge v_3 + \mu_2 v_3 \wedge v_1 + \mu_3 v_1 \wedge v_2$$

and then $a \wedge a = B(a,a) v_0 \wedge v_1 \wedge v_2 \wedge v_3$ where

$$B(a,a) = 2(\lambda_1\mu_1 + \lambda_2\mu_2 + \lambda_3\mu_3) \qquad (8)$$

This is a non-degenerate quadratic form, and so $B(a,a) = 0$ defines a nonsingular quadric $Q \subset P(\Lambda^2 V)$. Moreover, any other choice of basis rescales B by a non-zero constant and so Q is well defined in projective space.

We see then that a line $\ell \subset P(V)$ defines a decomposable 2-vector $a = x \wedge y$, unique up to a scalar and since $a \wedge a = 0$, it defines a point $L \in Q \subset P(\Lambda^2 V)$. Conversely, Theorem 17 tells us that every point in Q is represented by a decomposable 2-vector. Hence

Proposition 18 *There is a one-to-one correspondence $\ell \leftrightarrow L$ between lines ℓ in a 3-dimensional projective space $P(V)$ and points L in the 4-dimensional quadric $Q \subset P(\Lambda^2 V)$.*

It was Felix Klein (1849–1925), building on the work of his supervisor Julius Plücker, who first described this in detail and Q is usually called the *Klein quadric*. The equation of the quadric in the form (8) shows that there are linear subspaces inside it of maximal dimension 2 whatever the field. The linear subspaces all relate to intersection properties of lines in $P(V)$. For example:

Proposition 19 *Two lines $\ell_1, \ell_2 \subset P(V)$ intersect if and only if the line joining the two corresponding points $L_1, L_2 \in Q$ lies entirely in Q.*

Proof: Let $U_1, U_2 \subset V$ be the two-dimensional subspaces of V defined by ℓ_1, ℓ_2. Suppose the lines intersect in X, with representative vector $u \in V$. Then extend to bases $\{u, u_1\}$ for U_1 and $\{u, u_2\}$ for U_2. The line in $P(\Lambda^2 V)$ joining L_1 and L_2 is then $P(W)$ where W is spanned by $u \wedge u_1$ and $u \wedge u_2$.

Any 2-vector in W is thus of the form

$$\lambda_1 u \wedge u_1 + \lambda_2 u \wedge u_2 = u \wedge (\lambda_1 u_1 + \lambda_2 u_2)$$

which is decomposable and so represents a point in Q.

Conversely, if the lines do not intersect, $U_1 \cap U_2 = \{0\}$ so $V = U_1 \oplus U_2$. In this case choose bases $\{u_1, v_1\}$ of U_1 and $\{u_2, v_2\}$ of U_2. Then $\{u_1, v_1, u_2, v_2\}$ is a basis of V and in particular $u_1 \wedge v_1 \wedge u_2 \wedge v_2 \neq 0$. A point on the line joining L_1, L_2 is now represented by $a = \lambda_1 u_1 \wedge v_1 + \lambda_2 u_2 \wedge v_2$ so that

$$a \wedge a = 2\lambda_1 \lambda_2 u_1 \wedge v_1 \wedge u_2 \wedge v_2$$

which vanishes only if λ_1 or λ_2 are zero. Thus the line only meets Q in the points L_1 and L_2. □

Now fix a point $X \in P(V)$ and look at the set of lines passing through this point:

Proposition 20 *The set of lines $\ell \subset P(V)$ passing through a fixed point $X \in P(V)$ corresponds to the set of points $L \in Q$ which lie in a fixed plane contained in Q.*

Proof: Let x be a representative vector for X. The line $P(U)$ passes through X if and only if $x \in U$, so $P(U)$ is represented in the Klein quadric by a 2-vector of the form

$$x \wedge u.$$

Extend x to a basis $\{x, v_1, v_2, v_3\}$ of V, then any decomposable 2-vector of the form $x \wedge y$ can be written as

$$x \wedge (\mu x + \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3) = \lambda_1 x \wedge v_1 + \lambda_2 x \wedge v_2 + \lambda_3 x \wedge v_3.$$

Thus any line passing through X is represented by a 2-vector in the 3-dimensional space of decomposables spanned by $x \wedge v_1, x \wedge v_2, x \wedge v_3$, which is a projective plane in Q. Conversely any point in this plane defines a line in $P(V)$ through X. □

A plane in Q defined by a point $X \in P(V)$ like this is called an *α-plane*. There are other planes in Q:

Proposition 21 *Let $P(W) \subset P(V)$ be a plane. The set of lines $\ell \subset P(W)$ corresponds to the set of points $L \in Q$ which lie in a fixed plane contained in Q.*

A plane of this type contained in Q is called a *β-plane*.

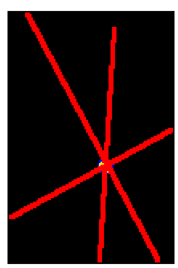
Proof: We just use duality here: if $U \subset V$ is 2-dimensional, then its annihilator $U^0 \subset V'$ is $4-2=2$-dimensional, so there is a one-to-one correspondence between lines in $P(V)$ and lines in $P(V')$. A point in Q therefore defines a line in either the projective space or its dual. Now the dual of the set of lines passing through a point is the set of lines lying in a (hyper)-plane. So applying Proposition 20 to $P(V')$ gives the result. □

In fact there are no more planes:

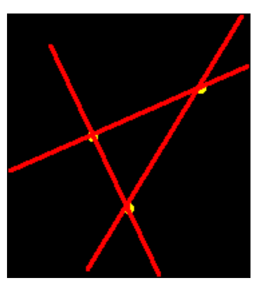
Proposition 22 *Any plane in the Klein quadric Q is either an α-plane or a β-plane.*

Proof: Take a plane in Q and three non-collinear points L_1, L_2, L_3 on it. We get three lines ℓ_1, ℓ_2, ℓ_3 in $P(V)$. Since the line joining L_1 to L_2 lies in the plane and hence in Q, it follows from Proposition 19 that each pair of ℓ_1, ℓ_2, ℓ_3 intersect. There are two possibilities:

- the three lines are concurrent:



- the three lines meet in three distinct points:



In the first case the three lines pass through a single point and so L_1, L_2, L_3 lie in an α-plane. But this must be the original plane since the three representative vectors for L_1, L_2, L_3 are linearly independent as the points are not collinear.

In the second case, if u_1, u_2, u_3 are representative vectors for the three points of intersection of ℓ_1, ℓ_2, ℓ_3, then L_1, L_2, L_3 are represented by $u_2 \wedge u_3, u_3 \wedge u_1, u_1 \wedge u_2$. A general point on the plane is then given by

$$\lambda_1 u_2 \wedge u_3 + \lambda_2 u_3 \wedge u_1 + \lambda_1 u_3 \wedge u_1$$

which is a general element of $\Lambda^2 U$ where U is spanned by u_1, u_2, u_3. Thus ℓ_1, ℓ_2, ℓ_3 all lie in the plane $P(U) \subset P(V)$. □

The existence of these two families of linear subspaces of maximal dimension is characteristic of even-dimensional quadrics – it is the generalization of the two families of lines we saw on the "cooling tower" quadric surface. In the case of the Klein quadric, two different α-planes intersect in a point, since there is a unique line joining two points. Similarly (and by duality) two β planes meet in a point. An α-plane and a β plane in general have empty intersection – if X is a point and π a plane with $X \notin \pi$, there is no line in π which passes through X. If $X \in \pi$, then the intersection is a line.

4.5 Exercises

1. If $a \in \Lambda^p V$ and p is odd, show that $a \wedge a = 0$.

2. Calculate $a \wedge b$ in the following cases:

- $a = b = v_1 \wedge v_2 + v_2 \wedge v_3 + v_3 \wedge v_1$
- $a = v_1 \wedge v_2 + v_3 \wedge v_1, \quad b = v_2 \wedge v_3 \wedge v_4$
- $a = v_1 + v_2 + v_3, \quad b = v_1 \wedge v_2 + v_2 \wedge v_3 + v_3 \wedge v_1.$

[v_1, v_2, v_3, v_4 *are linearly independent*]

3. Which of the following 2-vectors is decomposable?

- $v_1 \wedge v_2 + v_2 \wedge v_3$
- $v_1 \wedge v_2 + v_2 \wedge v_3 + v_3 \wedge v_4$
- $v_1 \wedge v_2 + v_2 \wedge v_3 + v_3 \wedge v_4 + v_4 \wedge v_1.$

[v_1, v_2, v_3, v_4 *are linearly independent*]

4. If $\dim V = n$ shown that every $a \in \Lambda^{n-1} V$ is decomposable.

5. Let $\ell \subset P(V)$ be a line and m another such that the corresponding point $M \in Q$ lies on the polar hyperplane to $L \in Q$. Show that ℓ and m intersect.

5 What is geometry?

5.1 The Erlanger Programm

Felix Klein moved on from his work on the Klein quadric and moved on physically from Bonn where he had been a student. Dodging the Franco-Prussian war in 1870 he passed through Paris, went briefly to Göttingen and then to Erlangen, near Nuremberg, in the south of Germany.

He prepared for his inaugural address in 1872 in Erlangen a paper which gave a very general view on what geometry should be regarded as. It was somewhat controversial at the time and in fact he spoke on something different for his lecture, but the point of view is still called the *Erlanger Programm.* Klein saw geometry as:

the study of invariants under a group of transformations.

This throws the emphasis on the group rather than the space, and was highly influential in a number of ways. Recall

Definition 18 *An action of a group G on a set Ω is a homomorphism $f : G \to \mathrm{Sym}(\Omega)$ to the group of all bijections of Ω to itself.*

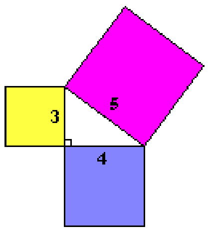
So for Klein, we have a set, say $\mathbf{R}^2$, and a group acting on it. For Euclidean geometry the group is the group of all transformations of the form

$$x \mapsto Ax + b$$

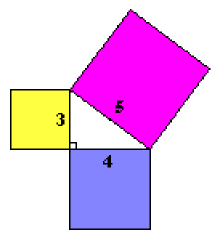
where A is an orthogonal transformation and $b \in \mathbf{R}^2$. This does form a group since

$$A(A'x + b') + b = AA'x + Ab' + b.$$

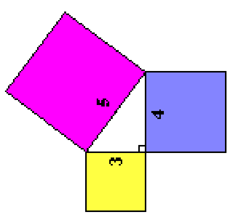
Every element in the group is the composition of a rotation or reflection $x \mapsto Ax$ and a translation $x \mapsto x + b$. So Pythagoras' theorem is a theorem of Euclidean geometry because if this is true:



so is this:



and this:



There are other groups that act on $\mathbf{R}^2$, though, and each of these gives rise to a different geometry. For example if we look at transformations

$$x \mapsto Ax + b$$

where A is just invertible and not orthogonal (this is called an *affine transformation*) then even the statement of Pythagoras's theorem is not invariant – a right angled triangle can be taken to any other triangle by an affine transformation. Even more interesting is the case of the hyperbolic plane, which we shall look at later. In fact it was the study of this, and Klein's realization that both Euclidean and hyperbolic geometry are special cases of projective geometry, which led him to formulate his proposal. We have to realize that in the early 19th century, as for most preceding centuries, Euclidean geometry had a special status, bound up with the logical structure of deductions from a set of axioms. Seeing it as just one of many geometries was as radical as Copernicus saying that the earth goes round the sun and not vice-versa.

The group that plays Klein's role for projective geometry is the group of projective transformations $\tau : P^n(F) \to P^n(F)$. Such a τ is defined by an invertible linear transformation $T : F^{n+1} \to F^{n+1}$, where we identify T and λT. Some names now:

Definition 19 *The group of all invertible linear transformations $T : F^n \to F^n$ is called the general linear group $GL(n, F)$.*

Definition 20 *The group of all invertible linear transformations $T : F^n \to F^n$ whose determinant is 1 is called the special linear group $SL(n, F)$.*

Definition 21 *The projective linear group $PGL(n, F)$ is the quotient of $GL(n, F)$ by the normal subgroup of non-zero multiples of the identity, and is the group of all projective transformations of $P^{n-1}(F)$ to itself.*

All the theorems we have proved so far are invariant under the projective linear group, but they have all been about lines, planes intersections etc. – just set-theoretical properties about linear subspaces. When Klein spoke about invariants he also meant numerical quantities, such as in Euclidean geometry the distance between two points $x, y \in \mathbf{R}^2$:

$$|(Ax + b) - (Ay + b)| = |A(x - y)| = |x - y|.$$

In projective geometry there is no invariant distance. Indeed Theorem 3 tells us that there is a projective transformation that takes any two distinct points to any other two. There are nevertheless more complicated invariants as we shall see next.

5.2 The cross-ratio

We shall deal with the geometry of the projective line $P^1(F)$ and the action of the group $PGL(2, F)$. We know from Theorem 3 that any three points go into any other three, but this is not the case for four points. There is a numerical invariant called the cross-ratio:

Definition 22 *Let P_1, P_2, P_3, P_4 be four distinct points on the projective line $P^1(F)$, with homogeneous coordinates $P_i = [x_i, y_i]$. The cross-ratio is defined by*

$$(P_1P_2; P_3P_4) = \frac{(x_1y_3 - x_3y_1)(x_2y_4 - x_4y_2)}{(x_1y_4 - x_4y_1)(x_2y_3 - x_3y_2)}.$$

First note that multiplying (x_i, y_i) by $\lambda_i \neq 0$ introduces the same factor in the numerator and denominator and leaves the formula for the cross-ratio unchanged. This shows it is well-defined. Similarly note that the linear transformation

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

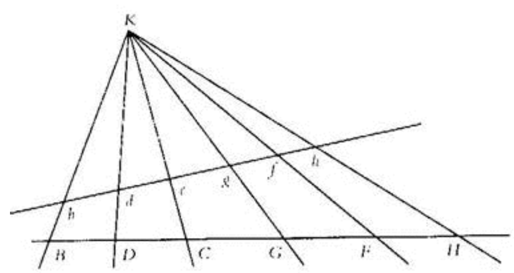
transforms $x_1y_2 - x_2y_1$ to

$$(ax_1 + by_1)(cx_2 + dy_2) - (ax_2 + by_2)(cx_1 + dy_1) = (ad - bc)(x_1y_2 - x_2y_1).$$

Again the factors in numerator and denominator cancel. We deduce that

- the definition of cross-ratio is independent of the choice of basis in which to write the homogeneous coordinates
- if $\tau : P(U) \to P(V)$ is a projective transformation between two projective lines then
$$(\tau(P_1)\tau(P_2); \tau(P_3)\tau(P_4)) = (P_1P_2; P_3P_4).$$

A particular case is projection from a point in the plane:



In the diagram, $(b\,d; c\,g) = (B\,D; C\,G)$.

If none of the y_i in the definition of cross-ratio is zero, then we can write it in terms of the scalars $z_i = x_i/y_i$ and obtain

$$(z_1\,z_2; z_3\,z_4) = \frac{(z_1 - z_3)(z_2 - z_4)}{(z_1 - z_4)(z_2 - z_3)} \tag{9}$$

Even when one of y_i is zero, an intelligent use of ∞ gives the correct results, for example if $y_1 = 0$, then the definition (22) provides

$$(\infty\,z_2; z_3\,z_4) = \frac{y_3(x_2y_4 - x_4y_2)}{y_4(x_2y_3 - x_3y_2)} = \frac{z_2 - z_4}{z_2 - z_3}$$

whereas the formula (9) gives

$$(\infty\, z_2; z_3\, z_4) = \frac{(\infty - z_3)(z_2 - z_4)}{(\infty - z_4)(z_2 - z_3)}.$$

On feature of the cross-ratio is the fact that you need to get the order right! Here (check for yourself the formulae) is what happens when you permute the variables:

- $(P_1P_2; P_3P_4) = (P_2P_1; P_4P_3) = (P_3P_4; P_1P_2)$
- $(P_2P_1; P_3P_4) = (P_1P_2; P_3P_4)^{-1}$
- $(P_1P_3; P_2P_4) = 1 - (P_1P_2; P_3P_4)$

Now, in the terminology of (9),

$$(\infty\, 0; 1\, x) = x$$

and so, since we assumed the four points were distinct, we see that $x \neq 0, 1$. But we have seen that there is a unique projective transformation τ that takes P_1, P_2, P_3 to $\infty, 0, 1$ so

$$(P_1P_2; P_3P_4) = (\tau(P_1)\tau(P_2); \tau(P_3)\tau(P_4)) = (\infty\, 0; 1\, \tau(P_4))$$

and thus any cross-ratio avoids the values $0, 1$. If we go through all the permutations of the four variables then in general we find the following six possible values of the cross-ratio:

$$x, \quad \frac{1}{x}, \quad 1-x, \quad \frac{1}{1-x}, \quad 1-\frac{1}{x}, \quad \frac{x}{x-1}.$$

There is a special case when these coincide in pairs: if $x = -1, 2$ or $1/2$ and this particular configuration of points has a name:

Definition 23 *The points P_1, P_2, P_3, P_4 are harmonically separated if the cross-ratio $(P_1P_2; P_3P_4) = -1$.*

Remark: There is some value in seeing the cross-ratio as a point in the projective line $P^1(F)\backslash\{0, 1, \infty\}$. We have seen that the symmetric group S_3 of order 6 permuting $\{0, 1, \infty\}$ acts as projective transformations of $P^1(F)$ to itself. The six values of the cross-ratio then constitute an orbit of this action.

5.3 Affine geometry

One of the features which motivated Klein was the fact that different types of geometry were special cases of projective geometry, and simply involved restricting the group. Affine geometry is a case in point. Consider a hyperplane $P(U)$ in a projective space $P(V)$ and the subgroup of the group of projective transformations of $P(V)$ to itself which takes $P(U)$ to itself. The complement $P(V)\backslash P(U)$ is acted on by this group and is called an *affine space*. Basically it is just a vector space without a distinguished origin, but it is the *group* which determines the way we look at it.

To get a hold on this, consider choosing a basis such that in homogeneous coordinates $[x_0, \ldots, x_n]$ the hyperplane in $P^n(F)$ is defined by $x_0 = 0$. Then a projective transformation τ which maps this to itself comes from a linear transformation T with matrix of the form

$$\begin{pmatrix} * & 0 & 0 & \ldots & 0 \\ * & * & * & \ldots & * \\ * & * & * & \ldots & * \\ * & * & * & \ldots & * \\ * & * & * & \ldots & * \end{pmatrix}$$

and we can unambiguously choose a representative for T by taking $T_{00} = 1$. The action of the subgroup of $GL(n+1, F)$ defined by these matrices on $P^n(F)\backslash P^{n-1}(F)$ which we identify with F^n as usual by

$$(x_1, x_2, \ldots, x_n) \mapsto [1, x_1, \ldots, x_n]$$

is of the form

$$x \mapsto Ax + b$$

where $A \in GL(n, F)$ and $b \in F^n$. The group of such transformations is called the *affine group* $A(n, F)$.

Example: The simplest affine situation is the real affine line, which is just $\mathbf{R}$ with the group of transformations

$$x \mapsto ax + b$$

with $a \neq 0$. The cross-ratio, which is a projective invariant of four points in $P^1(\mathbf{R})$ give an affine invariant of three points, since $A(1, \mathbf{R})$ is the subgroup of $PGL(2, \mathbf{R})$ which preserves the single point $\infty \in P^1(\mathbf{R})$. We have

$$(\infty\, x_1; x_2\, x_3) = \frac{x_1 - x_3}{x_1 - x_2}$$

so if $x_1 < x_2 < x_3$ the proportion in which x_2 divides the segment $[x_1, x_3]$ is the affine invariant $(x_1 \infty; x_2 x_3)$. The special value 1/2 (the concept of "harmonically separated" in projective geometry corresponds to the midpoint of the segment. So although there is no invariant notion of distance, we retain the notion of proportionality.

In higher dimensions, the centre of mass of m points $x_1, \ldots, x_m \in \mathbf{R}^n$

$$\bar{x} = \frac{1}{m}(x_1 + \ldots + x_m)$$

is a well-defined affine concept since

$$\frac{1}{m}(Ax_1 + b + \ldots + Ax_m + b) = A\bar{x} + b.$$

5.4 Gaussian optics

One application of this group-theoretical way of looking at geometry is in the simplest model of optical systems – those which have an axis of symmetry. If we track the effect on light rays which lie in a plane through that axis, then we have a 2-dimensional problem but with an extra symmetry – if the system was originally axially symmetric, and in $\mathbf{R}^2$ the x-axis is the axis of symmetry, then the effect on light rays must be symmetric with respect to the reflection in the x-axis

$$(x, y) \mapsto (x, -y).$$

To model optics this way we map $\mathbf{R}^2$ into $P^2(\mathbf{R})$ by

$$(x, y) \mapsto [1, x, y]$$

and the we consider first the subgroup of projective transformations which commutes with the reflection in the x-axis:

$$(x_0, x_1, x_2) \mapsto (x_0, x_1, -x_2)$$

These are given by linear transformations T whose matrix is of the form

$$\begin{pmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & * \end{pmatrix}.$$

We can choose $T_{22} = 1$, but now we need also the input of the physics. This tells us that the 2×2 matrix T_{ij}, $0 \leq i, j \leq 1$ must have determinant 1. This is part of what is called symplectic geometry which underlies many physical theories. So Gaussian optics can be put into geometrical form by considering an action of the group $SL(2, \mathbf{R})$.

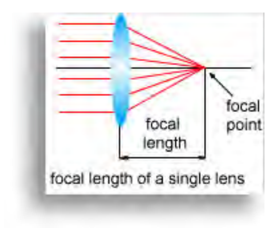
Note how the matrix in $SL(2, \mathbf{R})$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

acts:

$$(x, y) \mapsto (\frac{ax + b}{cx + d}, \frac{1}{cx + d})$$

so the axis of the system is treated as a projective line and the group acts as projective transformations. This is necessary because the image of a point on the axis through the action of a lens may be "at infinity": that point is called the focal point. From the point of view of projective geometry all those parallel light rays from infinity are the lines passing through the intersection of the x-axis with the line at infinity, the point $[0, 1, 0]$, and the projective transformation takes them to the lines passing through the focal point.



So let us consider the effect of a thin symmetric lens at the origin on objects on the optical axis. It is given by an element $T \in SL(2, \mathbf{R})$ which acts on the x axis as the projective transformation τ. If the focal point is $(x, y) = (-f, 0)$ then

$$\tau(\infty) = -f.$$

Because the lens is symmetric

$$\tau^{-1}(\infty) = f.$$

Light rays at the origin go straight through so

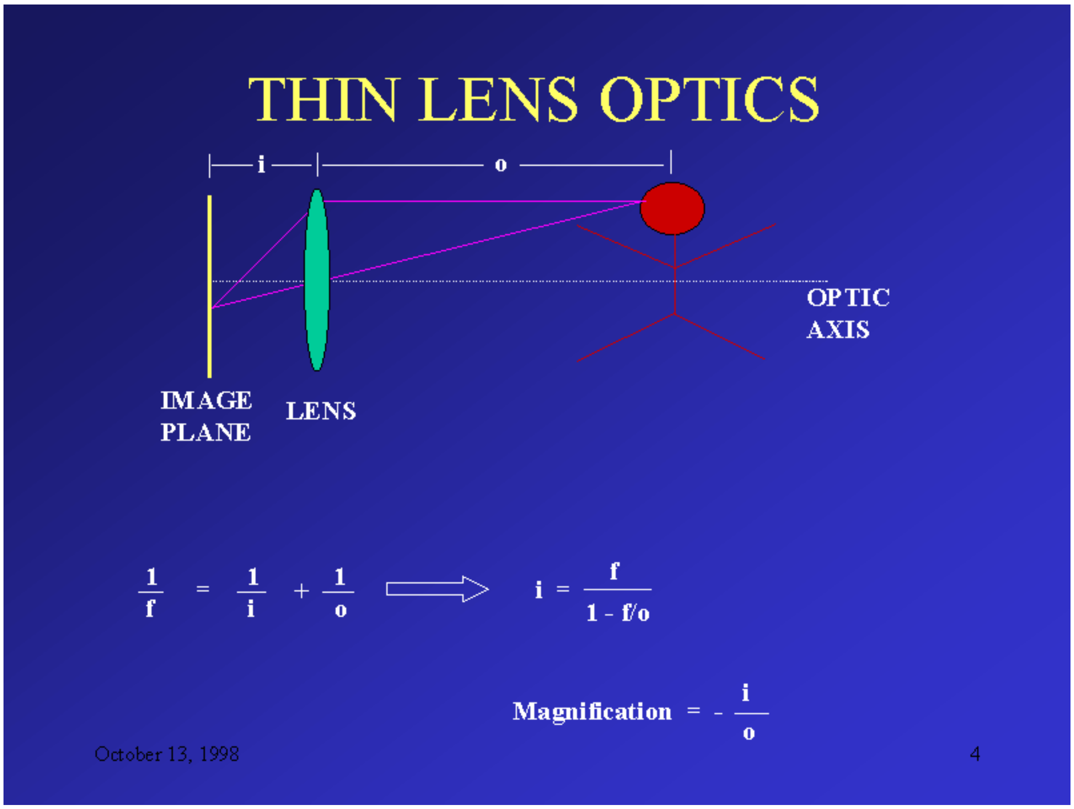
$$\tau(0) = 0.$$

The projective transformation of the line is is uniquely determined by this action on three points, and we easily find

$$\tau(x) = \frac{-fx}{x - f}.$$

An object at $x = \mathbf{o}$ then has image at $x = -\mathbf{i}$ where

$$\frac{1}{f} = \frac{1}{\mathbf{i}} + \frac{1}{\mathbf{o}}.$$



The group of projective transformations of the line $PGL(2, \mathbf{R})$ is obtained by the quotient of $GL(2, \mathbf{R})$ by the scalar matrices. In $SL(2, \mathbf{R})$ those scalars are just $\{\pm 1\}$, so we just need to choose a sign to get T:

$$T = \pm \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix}.$$

Find out which it is!

5.5 Non-Euclidean geometry

The group $SL(2, \mathbf{R})$ acts on other spaces than $P^1(\mathbf{R})$ or the optical space above, and if we study its invariants then we return to something close to the Euclidean geometry of the plane but also something essentially different, which is nowadays called *hyperbolic geometry*, or when it was first discovered *Non-Euclidean geometry*. A real matrix is a special type of complex matrix, so the group $PGL(2,\mathbf{R})$ is a subgroup of $PGL(2,\mathbf{C})$ and so acts on the complex projective line $P^1(\mathbf{C})$. It preserves the copy of $P^1(\mathbf{R}) \subset P^1(\mathbf{C})$ given by points with real homogeneous coordinates $[x_0, x_1]$, but we want to consider instead the complement of this subset. Since $\infty = [1, 0]$ is real, then

$$P^1(\mathbf{C})\backslash P^1(\mathbf{R}) \subset P^1(\mathbf{C})\backslash\{\infty\}$$

and is $\mathbf{C}\backslash\mathbf{R}$. This has two components, the upper and lower half-planes. An element of $PGL(2,\mathbf{R})$ may interchange these two (e.g. $z \mapsto -z$) but the subgroup $PSL(2,\mathbf{R}) = SL(2,\mathbf{R})/\{\pm 1\}$ preserves each one. We choose the upper one H. In fact the group acts transitively on H for the Möbius transformation

$$z \mapsto az + b$$

takes i to $ai + b$, and if $a > 0$, the transformation is defined by the 2×2 matrix

$$\frac{1}{\sqrt{a}}\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$$

which has determinant 1.

We can define an invariant function of a pair of distinct points in H using the cross-ratio:

$$(z\,w; \bar{w}\,\bar{z}).$$

Since z and w are in the upper half-plane $\bar{z}$ and $\bar{w}$ are in the lower half-plane so if $z \neq w$ the four points $z, w, \bar{w}, \bar{z}$ are distinct and the cross-ratio is well defined. Spelling it out we get

$$(z\,w; \bar{w}\,\bar{z}) = \frac{(z - \bar{w})(w - \bar{z})}{(z - \bar{z})(w - \bar{w})} = \frac{|z - \bar{w}|^2}{4\Im z \Im w}$$

which is positive, since $\Im z$ and $\Im w$ are positive. It is also symmetric in w and z. We shall use this to define a notion of the distance between z and w, but first note that we can allow $w = z$ in the cross-ratio $(z\,w; \bar{w}\,\bar{z})$ because

$$(z\,z; \bar{z}\,\bar{z}) = \frac{|z - \bar{z}|^2}{4\Im z \Im z} = 1.$$

Definition 24 *Let z, w be two points in the upper half plane. The hyperbolic distance between z and w is*

$$d(z, w) = 2\sinh^{-1}\sqrt{(z\,w; \bar{w}\,\bar{z})}.$$

This is clearly invariant under the action of $PSL(2,\mathbf{R})$. The particular analytical formula shows that if $z = ix, w = iy$ are imaginary points in H with $x > y$ then

$$e^{d/2} - e^{d/2} = 2\sqrt{\frac{(x-y)^2}{4xy}} = \left(\frac{x}{y}\right)^{1/2} - \left(\frac{y}{x}\right)^{1/2}$$

so that

$$d(ix, iy) = \log x - \log y \tag{10}$$

This notion of distance has the following property:

Proposition 23 *Let z_1, z_2 and w_1, w_2 be pairs of distinct points in the upper half plane and suppose that $d(z_1, z_2) = d(w_1, w_2)$. The there is a unique element $\tau \in PSL(2,\mathbf{R})$ such that $\tau(z_i) = w_i$ for $i = 1, 2$.*

Proof: Consider $z_1, z_2, \bar{z}_2$. These are distinct because $z_1 \neq z_2$ and $\bar{z}_2$ is in the lower half plane. By Theorem 3 there is a unique *complex* projective transformation τ such that

$$\tau(z_1) = w_1, \quad \tau(z_2) = w_2, \quad \tau(\bar{z}_2) = \bar{w}_2.$$

By projective invariance

$$(z_1\, z_2; \bar{z}_2\, \bar{z}_1) = (\tau(z_1)\, \tau(z_2); \tau(\bar{z}_2)\, \tau(\bar{z}_1)) = (w_1\, w_2; \bar{w}_2\, \tau(\bar{z}_1))$$

However, $d(z_1, z_2) = d(w_1, w_2)$ so

$$(w_1\, w_2; \bar{w}_2\, \tau(\bar{z}_1)) = (z_1\, z_2; \bar{z}_2\, \bar{z}_1) = (w_1\, w_2; \bar{w}_2\, \bar{w}_1)$$

and thus $\tau(\bar{z}_1) = \bar{w}_1$. We then have

$$\tau(\bar{z}_1) = \overline{\tau(z_1)}, \quad \tau(\bar{z}_2) = \overline{\tau(z_2)}$$

i.e. two solutions in the lower half-plane to the equation

$$\frac{az+b}{cz+d} = \frac{\bar{a}z+\bar{b}}{\bar{c}z+\bar{d}} \tag{11}$$

This is quadratic with imaginary coefficients:

$$(a\bar{c} - \bar{a}c)z^2 + (b\bar{c} - \bar{b}c + a\bar{d} - \bar{a}d)z + (b\bar{d} - \bar{b}d) = 0$$

which either vanishes identically or has conjugate roots. But since the imaginary part of both roots is negative, it must vanish. Thus (11) is identically true. This means that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \lambda \begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix}$$

and $|\lambda| = 1$, so $\lambda = e^{i\theta}$. It follows that

$$e^{-i\theta} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is real. Its determinant must be positive since it maps z_1 in the upper half plane to w_1 in the upper half plane. Thus rescaling we get an element in $SL(2, \mathbf{R})$. □

We noted above that $PSL(2, \mathbf{R})$ acts transitively on the upper half plane H. The theory of group actions tells us that there is a bijection

$$G/K \leftrightarrow H$$

from the space of cosets G/K of the subgroup K which fixes a point $x_0 \in H$. The map is just $gK \mapsto g \cdot x_0$. In our case, if we take $x_0 = i$, then K is the group of transformations

$$z \mapsto \frac{\cos\theta z + \sin\theta}{-\sin\theta z + \cos\theta}$$

which is isomorphic by $\theta \mapsto e^{2i\theta}$ to the group of unit complex numbers.

Klein's point of view then says that the geometry of the upper half plane is basically the study of the invariants of this space of cosets. It offers the opportunity to view this geometry in a different way – to find another space on which $PSL(2, \mathbf{R})$ acts transitively with the same stabilizer K. We look at such a space next.

We consider the space of quadrics in the projective line $P^1(\mathbf{R})$. By this, we mean the real projective plane defined by the 3-dimensional real vector space of quadratic forms

$$B(v, v) = ax_0^2 + bx_0x_1 + cx_1^2$$

on $\mathbf{R}^2$. The singular ones are given by the equation

$$b^2 - 4ac = 0$$

which, when we write it as

$$b^2 - (a + c)^2 + (a - c)^2$$

is clearly a non-empty non-singular conic in $P^2(\mathbf{R})$. We consider the subset $D \subset P^2(\mathbf{R})$ defined by

$$D = \{[a, b, c] \in P^2(\mathbf{R}) : b^2 - 4ac < 0\}.$$

If $a + c = 0$, $b^2 - 4ac = b^2 + (a - c)^2 > 0$ so D lies in the complement of the line $a + c = 0$ in $P^2(\mathbf{R})$ and we can regard D as the interior of the circle $(b/(a+c))^2 + ((a-c)/(a+c))^2 = 1$ in $\mathbf{R}^2$. The group $PGL(2, \mathbf{R})$, and its subgroup $PSL(2, \mathbf{R})$, acts naturally on this space of quadrics.

Now when $b^2 - 4ac < 0$, $a \neq 0$ and the quadratic equation $ax^2 + bx + c = 0$ has complex conjugate roots $z, \bar{z}$, one of which, say z, must be in the upper half plane. The quadratic form factorizes as $a(x_0 - zx_1)(x_0 - \bar{z}x_1)$. Conversely any $z \in H$ defines a point in D by $(x_0 - zx_1)(x_0 - \bar{z}x_1)$. We therefore have a bijection

$$F : H \to D$$

commuting with the action of $PSL(2, \mathbf{R})$. This means that the geometry of D, the interior of a conic in $P^2(\mathbf{R})$, is identical to that of H, the upper half plane. Historically, H is called the *Poincaré model* and D the *(Klein)-Beltrami model*. Each one has its distinct advantages.

How do we define distance in the Beltrami model? Two points $X, Y \in D$ define two quadrics and we take the pencil of quadrics generated by them – the unique projective line in $P^2(\mathbf{R})$ which joins them. This meets the conic in two points A and B – the two singular conics in the pencil. So we have four points on a line and we have an invariant, the cross-ratio.

To see how this relates to the hyperbolic distance, we need only consider z, w as imaginary because of Proposition 23 and (10) which shows that any distance can be realized by imaginary points in H. So if $z = ix, w = iy$, the two points $X = F(ix), Y = F(iy)$ in D are the quadrics

$$x_0^2 + x^2x_1^2, \qquad x_0^2 + y^2x_1^2$$

and the pencil generated by them is

$$\lambda(x_0^2 + x^2x_1^2) + \mu(x_0^2 + y^2x_1^2) = (\lambda + \mu)x_0^2 + (\lambda x^2 + \mu y^2)x_1^2.$$

With $[\lambda, \mu]$ as homogeneous coordinates on this line in $P^2(\mathbf{R})$ we have $X = [1, 0], Y = [0, 1]$ and the two singular quadrics occur when $\lambda + \mu = 0$ and $\lambda x^2 + \mu y^2 = 0$. So $A = [1, -1], B = [y^2, -x^2]$. Thus the cross-ratio

$$(YX; AB) = x^2/y^2.$$

From (10) we see that in this model

$$d(X,Y) = \frac{1}{2}\log|(XY;AB)| \tag{12}$$

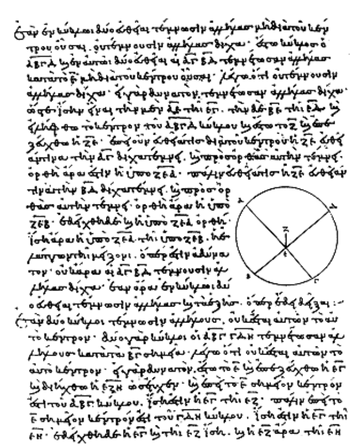
5.6 The parallel postulate

For centuries, Euclid's deduction of geometrical theorems from self-evident common notions and postulates was thought not only to represent a model of the physical space in which we live, but also some absolute logical structure. One postulate caused some problems though – was it really self-evident? Did it follow from the other axioms? This is how Euclid phrased it:

"That if a straight line falling on two straight lines makes the interior angle on the same side less than two right angles, the two straight lines if produced indefinitely, meet on that side on which the angles are less than two right angles".

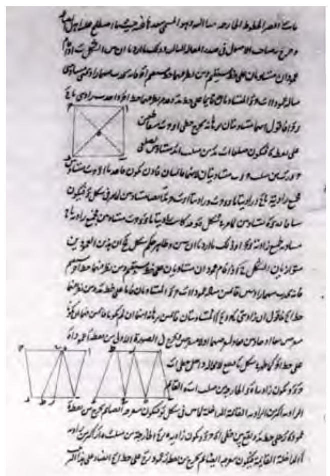
Some early commentators of Euclid's *Elements*, like Posidonius (1st Century BC), Geminus (1st Century BC), Ptolemy (2nd Century AD), Proclus (410 - 485) all felt that the parallel postulate was not sufficiently evident to accept without proof.

Here is a page from a medieval edition of Euclid dating from the year 888. It is handwritten in Greek. The manuscript, contained in the Bodleian Library, is one of the earliest surviving editions of Euclid.



The controversy went on and on; here is the Islamic mathematician Nasir al-Din al-Tusi (1201-1274)

struggling with the parallel postulate:



Finally Janos Bolyai (1802-1860) and Nikolai Lobachevsky (1793-1856)

discovered non-Euclidean geometry simultaneously. It satisfies all of Euclid's axioms except the parallel postulate, and we shall see that it is the geometry of H or D that we have started to look at.

Bolyai became interested in the theory of parallel lines under the influence of his father Farkas, who devoted considerable energy towards finding a proof of the parallel postulate without success. He even wrote to his son:

"I entreat you, leave the doctrine of parallel lines alone; you should fear it like a sensual passion; it will deprive you of health, leisure and peace – it will destroy all joy in your life."

Another relevant figure in the discovery was Carl Friedrich Gauss (1777-1855)

who was the first to consider the possibility of a geometry denying the parallel postulate. However, for fear of being ridiculed he kept his work unpublished. When he read Janos Bolyai's work he wrote to Janos's father:

"If I commenced by saying that I must not praise this work you would certainly be surprised for a moment. But I cannot say otherwise. To praise it, would be to praise myself. Indeed the whole contents of the work, the path taken by your son, the results to which he is led, coincide almost entirely with my meditations, which have occupied my mind partly for the last thirty or thirty-five years."

Yes, well..... he would say that wouldn't he?

Euclid's axioms were made rigorous by Hilbert. They begin with undefined concepts of

- "point"
- "line"
- "lie on" (a point lies on a line)
- "betweenness"
- "congruence of pairs of points"
- "congruence of pairs of angles".

Euclidean geometry is then determined by logical deduction from the following axioms:

EUCLID'S AXIOMS

I. AXIOMS OF INCIDENCE

1. Two points have one and only one straight line in common.

2. Every straight line contains a least two points.

3. There are at least three points not lying on the same straight line.

II. AXIOMS OF ORDER

1. Of any three points on a straight line, one and only one lies between the other two.

2. If A and B are two points there is at least one point C such that B lies between A and C.

3. Any straight line intersecting a side of a triangle either passes through the opposite vertex or intersects a second side.

III. AXIOMS OF CONGRUENCE

1. On a straightline a given segment can be laid off on either side of a given point (the segment thus constructed is congruent to the give segment).

2. If two segments are congruent to a third segment, then they are congruent to each other.

3. If AB and $A'B'$ are two congruent segments and if the points C and C' lying on AB and $A'B'$ respectively are such that one of the segments into which AB is divided by C is congruent to one of the segments into which $A'B'$ is divided by C', then the other segment of AB is also congruent to the other segment of $A'B'$.

4. A given angle can be laid off in one and only one way on either side of a given half-line; (the angle thus drawn is congruent to the given angle).

5. If two sides of a triangle are equal respectively to two sides of another triangle, and if the included angles are equal, the triangles are congruent.

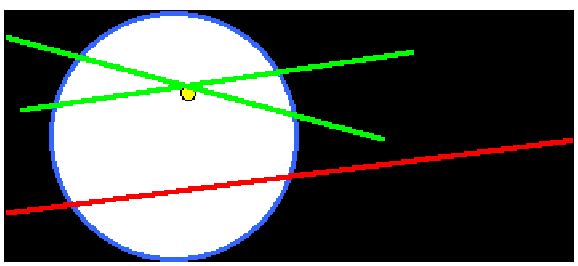
IV. AXIOM OF PARALLELS

Through any point not lying on a straight line there passes one and only one straight line that does not intersect the given line.

V. AXIOM OF CONTINUITY

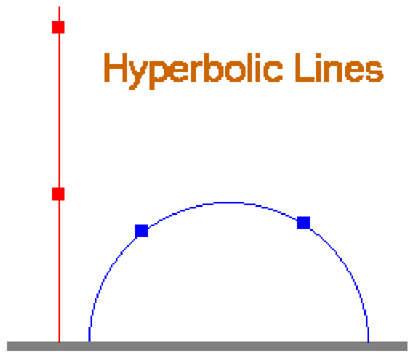
1. If AB and CD are any two segments, then there exists on the line AB a number of points $A_1, \ldots, A_n$ such that the segments $AA_1, A_1A_2, \ldots, A_{n-1}A_n$ are congruent to CD and such that B lies between A and A_n

The main property of hyperbolic geometry we have not introduced is the notion of a line. This is easiest in the Beltrami model – it is just the intersection of a projective line in $P^2(\mathbf{R})$ with the interior D of the conic. This obviously does *not* satisfy the parallel postulate:



In the Poincaré model we find:

Proposition 24 *In the upper half plane, a line is either a half-line parallel to the imaginary axis, or a semi-circle which intersects the real axis orthogonally.*



Proof: The map $F : H \to D$ is defined in homogeneous coordinates by

$$z \mapsto [1, -(z + \bar{z}), z\bar{z}]$$

so since a line in the Beltrami model is given by a projective line $\alpha_0 x_0 + \alpha_1 x_1 + \alpha_2 x_2 = 0$, in the upper half plane this is

$$\alpha_0 - \alpha_1(z + \bar{z}) + \alpha_2 z\bar{z} = 0.$$

If $\alpha_2 = 0$ this is a line of the form $ax + b = 0$, i.e. $x = \mathit{constant}$, which for $y > 0$ is a half-line parallel to the y-axis. Otherwise it is the intersection of the circle

$$\alpha_2(x^2 + y^2) - 2\alpha_1 x + \alpha_0 = 0$$

with D, and this is of the form

$$(x - a)^2 + y^2 = c$$

which is a circle centred at the point $(a, 0)$ on the x-axis, as required. $\square$

Remark: Note that from (10) that ix, iy, iz lie on a hyperbolic line and if $x > y > z$ then

$$d(x, z) = \log x - \log z = \log x - \log y + \log y - \log z = d(x, y) + d(y, z) \qquad (13)$$

But there is a unique line through any two points, from the Beltrami model. It follows from Proposition 23 that there is an element of $PSL(2, \mathbf{R})$ which takes any line to any other and so (13) holds for any line.

The fact that hyperbolic geometry satisfies all the axioms except the parallel postulate is now only of historic significance and the reader is invited to do all the checking. Often one model is easier than another. Congruence should be defined through the action of the group $PGL(2, \mathbf{R})$. We mentioned that this has a natural action on D and not just its index 2 subgroup $PSL(2, \mathbf{R})$. On H we have to add in transformations of the form

$$z \mapsto \frac{a\bar{z} + b}{c\bar{z} + d}$$

with $ad - bc = -1$ to get the action of $PGL(2, \mathbf{R})$.

Finally we should point out the link with the *Geometry of Surfaces* course. The upper half plane has a Riemannian metric, a first fundamental form, given by

$$\frac{dx^2 + dy^2}{y^2} \tag{14}$$

This is just a rescaling of the flat space fundamental form $dx^2 + dy^2$ by the positive function y^{-2}. This affects the lengths of curves but not the angles between them. In the coordinates x, y of H, angles between curves as measured by the metric (14) are just Euclidean angles. We say then that the metric is *conformally flat*. In the Beltrami model, we have affine coordinates in which the lines of the hyperbolic geometry become straight lines. In fact, the *geodesics* of the metric (14) are either half-lines parallel to the imaginary axis, or semi-circles which intersect the real axis orthogonally. In this case we say that the metric is *projectively flat*. The hyperbolic distance between two points is the metric distance, and then it follows that the basic axioms of a metric space hold for $d(z, w)$.