Chapter 10

Conics in projective geometry

Recall that in projective geometry the meaningful questions deal with notions which are invariant under projections. For instance the circle is not a meaningful figure - after projection from one plane to another it can be transformed to an ellipse, a parabola or a hyperbola. Indeed, consider the cone over this circle with vertex at the center of projections. The image of the circle under projection to a plane π' is the intersection of this cone with the plane π'. The figures that we can get by intersecting a cone with a plane are called conic sections. One can prove that they are in fact just ellipses, parabolas and hyperbolas.

Thus instead of many different notions (circles, ellipses, parabolas, hyperbolas) in Euclidean geometry, there is only one figure, a conic section, in the projective world. This observation can be often used when we are solving problems about conics. If we happen to notice that the problem can be formulated using notions that are invariant under projective transformations, we can choose a transformation that makes our problem simpler! (e.g. sends the conic to a circle).

Our first goal is to learn one of the notions that are preserved by projective transformations: cross ratio of points on a conic.

10.1   Cross-ratio of four points on a conic

In this section we define the notion of cross ratio of four points on a conic and prove its projective invariance.

Consider a conic E and four points A, B, C, D on it. Choose any other

point O on the conic E and any line l in the plane not passing through O. The lines OA,OB,OC and OD intersect the line l' in points A',B',C' and D'. We define the cross ratio of points A, B, C, D on the conic E to be equal to the cross-ratio of points A', B', C', D'.

This definition looks rather dubious: what if we choose a different point X on E and a different line l'? Will we get the same answer or a different one?

We shall prove now that in fact the answer in fact doesn't depend on the choices of X and l and the cross-ratio of points on a conic is well-defined.

First note that the definition doesn't depend on a choice of line l': if l' is a different line and A'', B'', C'', D'' are the points of its intersection with the lines XA, XB, XC, XD, then the cross-ratios (A', B', C', D') and (A'', B'', C'', D'') are equal, as we proved in the last lecture.
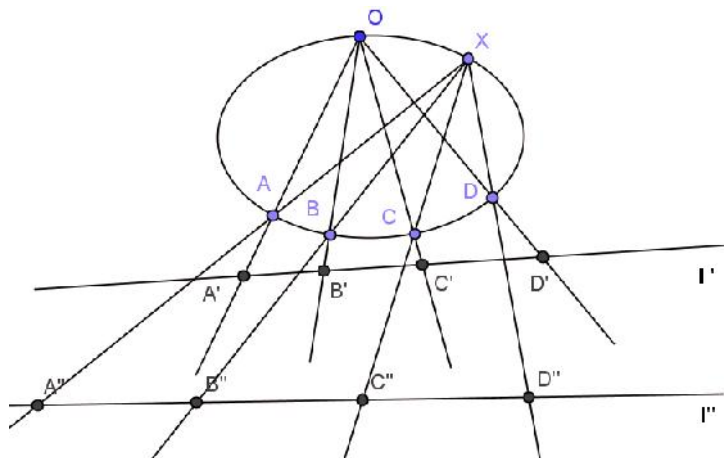
Now we want to prove that if X_1 and X_2 are different points on the ellipse and A'_1, B'_1, C'_1, D'_1 and A'_2, B'_2, C'_2, D'_2 are the points of intersection of line l' with the lines X_1A, X_1B, X_1C, X_1D and X_2A, X_2B, X_2C, X_2D respectively then the cross ratios (A'_1, B'_1, C'_1, D'_1) and (A'_2, B'_2, C'_2, D'_2) are equal. Note that the question whether the cross ratios thus constructed are equal is invariant under projective transformations. Thus we can apply any projective transformation to our picture and prove the claim for the picture we get after the application of this transformation. A natural transformation to apply in this case would be such that the conic gets mapped to a circle.
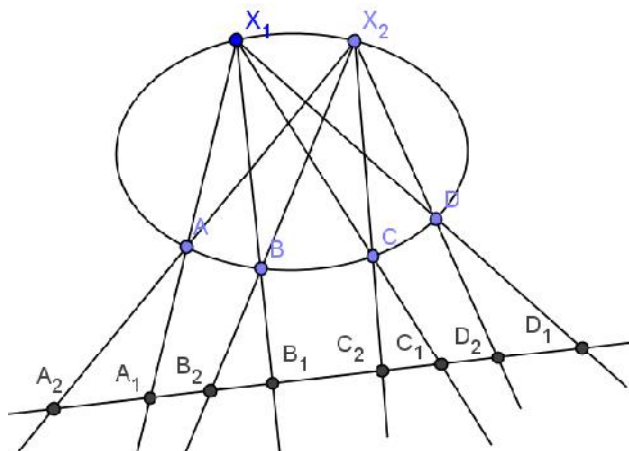
So now we are left with the same question as before, but instead of an arbitrary conic we have a circle. Now for the circle the question is rather simple: as we have seen the cross ratio (A_1', B_1', C_1', D_1') is equal to $\frac{\sin \angle A_1' X_1 C_1'}{\sin \angle B_1' X_1 C_1'} \div \frac{\sin \angle A_1' X_1 D_1'}{\sin \angle B_1' X_1 D_1'}$, which is the same as $\frac{\sin \angle AX_1 C}{\sin \angle BX_1 C} \div \frac{\sin \angle AX_1 D}{\sin \angle BX_1 D}$. Similarly the cross-ratio (A_1', B_1', C_1', D_1') is equal to $\frac{\sin \angle AX_2 C}{\sin \angle BX_2 C} \div \frac{\sin \angle AX_2 D}{\sin \angle BX_2 D}$. Now the equality of the two cross ratios follos from the equalities of angles $\angle AX_1 C = \angle AX_2 C$, $\angle BX_1 C = \angle BX_2 C$, $\angle AX_1 D = \angle AX_2 D$ and $\angle BX_1 D = \angle BX_2 D$.

10.1.1 A different definition of cross-ratio

We can define the cross ratio of four points on a quadric in a quite different way: choose a line l_X tangent to the quadric E at point X on it. Let l_A, l_B, l_C, l_D denote the lines tangent to the quadric at points A,B,C,D respectively. Let A', B', C', D' denote the points of intersection of lines l_A, l_B, l_C, l_D with the line l_X. Then we can define the cross-ratio of points A, B, C, D as the cross-ration of points A', B', C', D' on the line l_X. We claim that this definition is in fact the same as the definition we had before - if we compute the cross ratio of points A, B, C, D in the previous way and in this way we get the same answer.

The proof of this claim is similar to what we did before: first we note that the question is invariant under projective transformations. Then we apply a transformation to make the conic a circle. And finally for the circle we use some angle chasing. Namely we will use the following obvious observation:

Let E be a circle with center O and A, X be two points on it. Let l_X, l_A be the tangent lines to the circle E at the points X and A. Let A' be the

intersection point of these tangent lines. Then the angle $\angle A'OX$ is equal to one half of the angle $\angle AOX$.

Thus we find that in our situation angles $\angle A'OC', \angle A'OD', \angle B'OC', \angle B'OD'$ are equal to $\frac{1}{2}\angle AOC, \frac{1}{2}\angle AOD, \frac{1}{2}\angle BOC$ and $\frac{1}{2}\angle BOD$ respectively (e.g. $\angle A'OC' = \angle A'OX - \angle C'OX = \frac{1}{2}\angle AOX - \frac{1}{2}\angle COX = \frac{1}{2}\angle AOC$). Thus the double ratio (A', B', C', D') is equal to $\frac{\sin \angle A'OC'}{\sin \angle B'OC'} \div \frac{\sin \angle A'OD'}{\sin \angle B'OD'}$ which is the same as $\frac{\sin \frac{1}{2}\angle AOC}{\sin \frac{1}{2}\angle BOC} \div \frac{\sin \frac{1}{2}\angle AOD}{\sin \frac{1}{2}\angle BOD}$. Now this quantity is equal to the cross-ratio we defined previously, since for any choice of point X on the circle E we have $\angle AXC = \frac{1}{2}\angle AOC$ and similarly for the other relevant angles.

10.2   Projective maps from a line to itself

As we have seen a projective mapping of a space to itself preserves the cross ratio of any four collinear points. This prompts the following definition: a mapping $f : l \rightarrow l'$ from a projective line l to a projective line l' is projective if it preserves cross-ratios: for any four points A, B, C, D on l we have $(A, B, C, D) = (f(A), f(B), f(C), f(D))$.

Notice that this defines what a projective transformation of a line to another line is for any two lines - they don't have to be in the same plane, not even in the same projective space!

Let us give two examples of projective transformations in the case the two lines l and l' do belong to the same projective plane.

In the first example let E be a point in the plane not lying on either of the two lines. Consider the transformation f sending any point $A \in l$ to the point $f(A) \in l'$, which is the intersection of line AE and l'. This is a central projection from line l to line l' with center E. To show it is in fact a projective transformation, we should only show that if A, B, C, D are four points on the line l and the lines EA, EB, EC, ED intersect the line l' at points $f(A), f(B), f(C), f(D)$, then $(A, B, C, D) = (f(A), f(B), f(C), f(D))$. But this is just the cross ratio of lines EA, EB, EC, ED computed in two different ways - by intersecting them with line l or intersecting with line l'.

Note that in this example the point of intersection of lines l and l' gets mapped to itself.

Another example of a projective transformation is as follows: suppose that a conic E is tangent to both lines l and l'.Then we define the image of point $A \in l$ as the intersection point of line l' with the tangent line to the

conic E from point A which is different from l.

The fact that such transformation is projective follows from the fact that the definition of cross ratio of four points on a quadric using tangent lines is well-defined.

Indeed, if lines $Af(A), Bf(B), Cf(C), Df(C)$ are tangent to the conic E at points A', B', C', D', then the cross ratios (A, B, C, D) and $(f(A), f(B), f(C), f(D))$ are both equal to the cross ratio of the points A', B', C', D' on the conic E.

In fact for two lines l and l' in the same plane every projective transformation is of one of the two kinds. We will prove this statement in full generality later, after we study projective duality. For now let's prove only the following statement:

Theorem 22. *Let l and l' be two lines in projective plane and let $f : l \to l'$ be a projective transformation from l to l' which maps the point of intersection of l with l' to itself. Then there exists a point E in the plane so that the transformation f is just the central projection of l onto l' with center E.*

Proof. Let A, B be any two points on line l which are distinct and not equal to O. Let E be the point of intersection of lines $Af(A)$ and $Bf(B)$. The central projection with center E from l to l' sends O to itself and A and B to $f(A)$ and $f(B)$. Suppose now $X \in l$ is any point on l and X' is the image of X under central projection with center E from l to l'. Then we know that (O, A, B, C) is equal to $(O, f(A), f(B), f(X))$ because f is projective and we also know it is equal to $(O, f(A), f(B), X')$, because the central projection with center E is projective. Since $(O, f(A), f(B), f(X)) = (O, f(A), f(B), X')$, we see that $X' = f(X)$, i.e. f coincides with the central projection with center at E. $\qquad\square$

!take the lemma that projective transformation is defined by where it sends three points out of the proof!

10.3   Pascal's theorem

!!!Change notations to be consistent with Pappus theorem - "odd vertices O, even vertices P"!!!

The fact that we have a well-defined notion of cross-ratio of four points on a conic shows the following statement:

Let l_1 and l_2 be any two lines in the plane and E any conic. Suppose that O_1 and O_2 are two distinct points on the conic E. Define the mapping

f from l_1 to l_2 as following: first do a projection from line l_1 to the conic E with center at O_1 and then project from the conic to line l_2 with center O_2.

The mapping f is projective.

Indeed, if A, B, C, D are four points on the line l_1 and A', B', C', D' are the points of intersection of lines O_1A, O_1B, O_1C, O_1D with the conic $\mathfrak{E}$, then the cross ratio (A', B', C', D') is equal both to (A, B, C, D) and to $(f(A), f(B), f(C), f(D))$.

Now suppose that the lines l_1 and l_2 in fact intersect at a point O on the conic $\mathfrak{E}$. Then we see that f sends the point O to itself (check it to be sure you understand well how f is defined!). But then the theorem from he previous section tells us that there is some point E in the plane such that f is just the central projection from l_1 to l_2 with center E. Let's see whether we can find this point E.

Let P_1 denote the point of intersection of l_1 with $\mathfrak{E}$ and P_2 be the point of intersection of l_2 with $\mathfrak{E}$. The point P_1 gets mapped by the first projection to itself, and then by the second projection to the point of intersection of l_2 with O_2P. Thus the point $f(O_1)$ lies on the line O_2P_1.

Similarly the point of intersection of line l_1 with O_1P_2 gets mapped by the first projection to point P_2 and then point P_2 gets mapped to itself by the second projection. Thus the point of intersection of the line O_1P_2 with the line l_1 gets mapped to point P_2.

Thus the center of projection E must be the point of intersection of O_2P_1 and O_1P_2.

Now let P be any point on conic $\mathfrak{E}$, and let E_1 and E_2 be the intersection points of lines PO_1 and PO_2 with l_1 and l_2 respectively. Then $E_2 = f(E_1)$ and hence E_1, E_2 and E are in fact collinear!

Let's summarize what we proved. Let $\mathfrak{E}$ be any ellipse and let $OP_1O_2PO_1P_2O$ be any hexagon inscribed in it. Let E, E_1 and E_2 be the points of intersection of pairs of opposite sides of this hexagon, namely $E_1 = OP_1 \cap PO_1$, $E_2 = OP_2 \cap PO_2$ and $E = O_1P_2 \cap P_1O_2$. Then the points E, E_1, E_2 are collinear!

Points E_1, E, E_2 are collinear
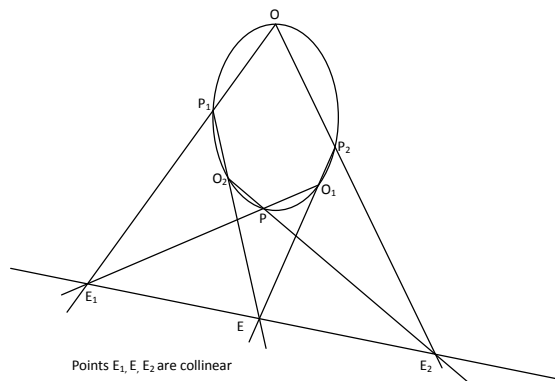
Thanks to Chang Yu and Siming Zhao for illustration.

This theorem was discovered by B. Pascal when he was only sixteen years old! He called this theorem "The Mystic Hexagram" and wrote a treatise on conic sections where this theorem and its numerous corollaries were the leit-motif.

We can prove a much older theorem, due to Pappus, along the same lines:

Theorem 23. *Let $\mathfrak{L}_O$ and $\mathfrak{L}_P$ be two lines. Let O_1, O_2, O_3 be three points on line $\mathfrak{L}_O$ and let P_1, P_2, P_3 be three points on line $\mathfrak{L}_P$. Let E_1 be the point of intersection of lines O_2P_3 and O_3P_2. Similarly let $E_2 = O_1P_3 \cap O_3P_1$ and $E_3 = O_1P_2 \cap O_2P_1$. Then the three points E_1, E_2, E_3 are collinear.*
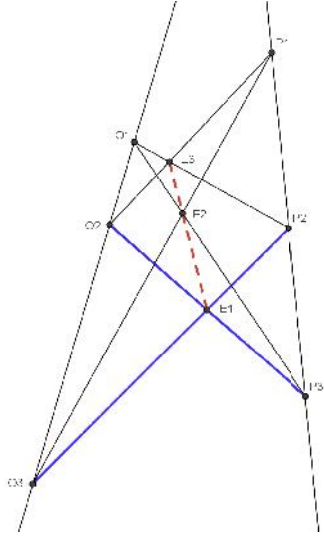
Thanks to Chenchen Li for the illustration

Proof. Consider the mapping $f : O_1P_2 \to O_1P_3$ which is the composition of the central projection from O_1P_2 to $\mathfrak{L}_P$ with center at O_2 and the central projection from $\mathfrak{L}_P$ to O_1P_3 with center O_3. This mapping is clearly projective transformation. It maps the point O_1 to itself (the first central projection maps the point O_1 to the intersection point of $\mathfrak{L}_O$ and $\mathfrak{L}_P$, while the second central projection maps this intersection point back to O_1). Thus it must itself be a central projection with some center. Let us locate this center.

Consider point P_2. The first central projection maps it to itself. Thus the point $f(P_2)$ lies on the line O_3P_2.

Consider now the point P_3. Its preimage under the second central projection is itself, hence its preimage under f lies on the line O_2P_3.

Thus the center of the central projection f is the point E_1 - the intersection of lines O_3P_2 and O_2P_3.

Now let's look at point E_3. Its image under the first central projection is the point P_1 and the image of P_1 under the second central projection is E_2. Thus $f(E_3) = E_2$. Since f is the central projection with center E_1, the points E_1, E_2, E_3 must be collinear.

$\square$

10.4 Conics as rational curves

In this section we will show that conic sections have a very useful property: they admit a parametrization by rational functions of one parameter. This property can be used in quite diverse applications: from explanation why cross-ratio can be defined for points on a conic section, to finding all Pythagorean triples, to computations of integrals of functions that involve square roots. It's amazing to realize how the same idea can be useful in such different contexts!

10.4.1 Rational parametrization

The parametrization we are looking for is just a rational map from a line to a conic, which is one-to-one and onto and such that the inverse map is also a rational map. A curve that admits such parametrization is usually called a rational curve. Since rational curves in general are not the subject of this lecture, we won't make our definition completely precise, but instead just study the rational parametrization of a conic.

Let O be a point on a conic E. Let l be any line not passing through point O. For any point A on the line l the line OA intersects the conic E at two points - O and some other point. To see this, we can restrict the equation of a conic, which is a quadratic equation, to the line OA. What we get is a quadratic equation that has one root corresponding to point O. Hence it has exactly one other root! (in case the line OA is tangent to the conic the other root should be thought of coinciding with point O - the quadratic equation will have a double root corresponding to point O).

Our parametrization $\phi : l \to E$ takes the point $A \in l$ to the other point of intersection of line OA with the conic E.

The inverse map $\psi : E \to l$ is obvious: it is the central projection from the conic E to line l with center O.

Now we claimed that both these maps are in fact rational maps. The reason it is so is that to find the point of intersection of a line with a quadric, which is different from the point which is already given to us, we don't in fact have to solve any quadratic equations. Instead we can just use Vieta's formulas.

Let us do an example to make this idea both clear and precise.

Let E be the circle $x^2 + y^2 = 1$, point O be the point $(0, 1)$ and the line l to be the line $y = 0$.

We can describe a point on a line l withe just one coordinate, say the x-coordinate of the corresponding point. So the point $A = (0, t) \in l$ will have coordinate t.

Now the line OA has the equation $t(1 - y) = x$ (check that it passes through both points $(0, 1)$ and $(0, t)$). To intersect it with the circle E we have to solve the system of equations

$$x^2 + y^2 = 1$$
$$t(1 - y) = x$$

By plugging the second equation to the first we do the operation we described before: restrict the equation of a quadric E to the line OA. What we get is the equation $t^2(1 - y)^2 + y^2 = 1$ or $(t^2 + 1)y^2 - 2t^2y + t^2 - 1 = 0$. It seems that it is a quadratic equation in y and in order to solve it we need to extract square roots. The beautiful idea here is that we already know one of the solutions of this equation: point $(0, 1)$ belongs to the intersection of OA and E, and thus $y = 1$ is a solution. To find the other solution all we have to do is to notice that Vieta's formaula tells us that the product of the two solutions is equal to $\frac{t^2-1}{t^2+1}$. Since one of them is 1, the other one is $\frac{t^2-1}{t^2+1}$. From $x = t(1 - y)$ we find that corresponding to $y = \frac{t^2-1}{t^2+1}$ we have $x = \frac{2t}{t^2+1}$.

Thus we have that the parametrization of the circle E we were after is $\phi(t) = (\frac{2t}{t^2+1}, \frac{t^2-1}{t^2+1})$.

The inverse map ψ can be found by looking at triangles with vertices at O,$(0, 0)$ and $(0, t)$ and with vertices O,$(0, y)$,and (x, y) (where $(x, y) \in E$ is the point on E that gets mapped to the point with coordinate t by ψ). Since these triangles are similar, we have $\frac{x}{t} = \frac{1-y}{1}$, or $t = \frac{x}{1-y}$. Thus $\psi(x, y) = \frac{x}{1-y}$ for any point $(x, y) \in E$. This map is clearly rational and is the inverse of ϕ.

Let's now use this parametrization to solve the question of finding all triples of integer numbers (a, b, c) such that $a^2 + b^2 = c^2$ (such triples are called Pythagorean triples, since they correspond naturally to right-angles triangles with integer side lengths).

Instead of solving this question we will solve the question of finding all rational solutions of $x^2 + y^2 = 1$ (clearly if a, b, c is a Pythagorean triple, then $\frac{a}{c}, \frac{b}{c}$ is a rational solution of $x^2 + y^2 = 1$; to go in the opposite direction one should be a little careful and deal with only primitive Pythagorean triples - those without common factors. But this can also be done quite easily). Now (x, y) is a pair of rational numbers satisfying $x^2 + y^2 = 1$ if and only if $t = \psi(x, y)$ is a rational number! Now we have a complete answer: every rational

point (x, y) on teh unit circle can be obtained by applying ϕ to a rational number t. If we take $t = \frac{p}{q}$ with integer numbers p, q without common factors, then we get the corresponding point $(x, y) = \phi(t) = (\frac{2pq}{p^2+q^2}, \frac{p^2-q^2}{p^2+q^2})$. From this parametrization it's not hard to gt that all Pythagorean triples are of the form $(2pqk, k(p^2 - q^2), k(p^2 + q^2))$ for some integer numbers p, q, k.

Another problem we can solve using rational parametrization of conics is the problem of finding integrals of functions of the form $\int R(x, \sqrt{ax^2 + bx + c}) \, dx$, where R is some rational function of two variables and the expression under the square root is not a complete square. We can look at the conic E given by the equation $y^2 = ax^2 + bx + c$. Then we have to integrate $R(x, y) \, dx$, where (x, y) is restricted to the conic E. To do so we use the rational parametrization of the conic. Let $\phi(t) = (\phi_1(t), \phi_2(t))$ be the rational parametrization by parameter t. Make the substitution $t = \psi(x, y) = \psi(x, \sqrt{ax^2 + bx + c})$ with the inverse being $x = \phi_1(y)$. Then $dx = \frac{d\phi_1}{dt} dt$ and thus we have to integrate the rational form $R(\phi_1(t), \phi_2(t)) \frac{d\phi_1}{dt} dt$ - no square roots are involved now! Once we find this integral we plug in $t = \psi(x, y) = \psi(x, \sqrt{ax^2 + bx + c})$ to find the answer to the original question!

Finally in our context we used the central projection from a point on a conic to define cross-ratio of points on a quadric. Essentially the rational parametrization of a quadric makes it indistinguishable (as an abstract, non-embedded curve) from the projective line. Thus one should be able to define every notion that can be defined abstractly for the rational line also for conic sections. We, however, defined the cross-ratio in terms of the embedding, and thus we can't see this point clearly yet.

In courses of algebraic geometry you will study how the ideas we described here evolved to curves of higher degrees, to surfaces and so on.

10.5 Homogeneous coordinates

In this section we will study a way to describe projective geometry and projective transformations which lends itself readily to introducing coordinates and relating geometry to algebra. This introduction of algebra gives us the opportunity to study some geometric questions over different fields, e.g. the field of complex numbers. It also allows us to formulate the general notion of a projective transformation from one projective space to another.

We will start with a rather unusual definition: a projective space on

dimension n is the space of lines passing through the origin in a vector space of dimension $n + 1$.

Let's think about this definition for a little bit. How does it reflect what we imagine when we talk about projective space?

If we have a hyperplane π in the vector space V which doesn't pass through the origin, then each line that does pass through the origin intersects it in a point. Well, almost: the lines that lie in the plane π_∞ parallel to π and passing through the origin do not intersect π. Instead of defining a point on π, such lines define a family of parallel lines inside the plane π: line l in π_∞ defines the family of lines inside π which are parallel to it. Thus the space of lines through origin in V is the same as space of points in π, to which we added a "hyperplane at infinity" - the hyperplane, whose points are families of parallel lines in π. We see that our new definition is indeed consistent with the way we thought previously about projective space.

Moreover, if we choose a different plane π' instead of π, then to go from one of them to the other all we have to do is to project from point $0 \in V$.

We see that our new definition incorporates all our previous intuitions and also is more "self-contained" - it doesn't require the choice of some esoteric "line at infinity". The price we have to pay is the slightly higher level of abstraction - the points in our new definition of a projective space are lines through the origin in a projective space. We will use the following notation below: for every $0 \neq x \in V$ the line through the origin containing x will be denoted by $[x]$. The lines $[x]$ and $[y]$ are equal to each other if and only if $x = \lambda y$ for some nonzero λ.

We can define now what a projective transformation is: if V, W are two vector spaces and $f : V \rightarrow W$ is an invertible linear transformation, then we can define the map $\mathbb{P}f : \mathbb{P}V \rightarrow \mathbb{P}W$ by saying that $\mathbb{P}f([x]) = [f(x)]$ for any $x \neq 0$. We call transformations that can be obtained in this way projective transformations.

Let us give an example of these notions in the simplest case of a projective line.

By its definition projective line is the space of lines through the origin in a two-dimensional vector space. Choose a basis in this vector space, so that any point in it can be given by a pair of coordinates (x, y). We will denote by $[x : y]$ the line passing through the origin and the point (x, y), provided that $(x, y) \neq (0, 0)$. For any nonzero λ we have then that $[x : y] = [\lambda x, \lambda y]$. In particular if $y \neq 0$, then $[x : y] = [\frac{x}{y} : 1]$. And if y is equal to zero, then

$[x : y] = [x : 0] = [1 : 0]$. Thus inside our projective line we have a copy of a usual line of points of the form $[z, 1]$ and one extra point at infinity $[1 : 0]$, which we call ∞.

Any invertible linear transformation from this space to itself can be written as $f(x, y) = (ax + by, cx + dy)$ for some numbers a, b, c, d such that $ad - bc \neq 0$. Let's see how the projective transformation $\mathbb{P}f$ induced from f looks like. The image of a "finite point" $[z : 1]$ under $\mathbb{P}f$ is $[az + b : cz + d]$, or $[\frac{az+b}{cz+d} : 1]$, if $cz + d \neq 0$. The point at infinity gets mapped to $[a : c] = [\frac{a}{c} : 1]$.

In the case our base field is complex numbers we have just recovered the Mobius transformations! The Riemann sphere is now nothing else but the complex projective line $\mathbb{CP}^1$ and the Mobius transformations are the projective transformations from this line to itself!

10.6   Projective classification of real quadrics

If we think of the projective space as the space of lines through the origin in some vector space, then a quadric in this projective space is the space of lines contained in a quadratic cone, i.e. the set of solutions of a homogeneous equation of degree 2.

A degree 2 homogeneous equation can be written as $x \in V : x^T A x = 0$ for some symmetric matrix A.

Let's see how it is done in $\mathbb{R}^3$: a homogeneous equation of degree 2 is of the form $ax^2 + bxy + cy^2 + dxz + eyz + fz^2 = 0$. We can write the left-hand side of this equation as $\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$.

Now a projective transformation is a transformation coming from a linear transformation of the form $x \to Bx$ for some invertible matrix B. The solutions of the equation $x^T A x = 0$ get mapped by this transformation to the solutions of the equation $y^T (B^{-1})^T A B^{-1} y = 0$ (y and x are related by $y = Bx$). Thus the study of the classification of quadrics up to projective transformation is the same as the study of symmetric matrices up to transformations of the form $A \to C^T A C$.

But this is a classical question of linear algebra, with a very simple answer to it: by transformation of the form $A \to C^T A C$ a symmetric matrix can be brought in a unique way to a diagonal matrix with k ones along the diagonal, m minus ones and all the rest zeros. The integer m is called the signature of

the matrix and the integer $n + m$ is called its rank.

Let's see what this result gives us geometrically for the projective plane. According to the result by a projective transformation every quadric (solution of a quadratic equation) can be brought to one of the following forms (note that we can multiply the equation of a quadric by -1 without changing the solution set, so we can always assume that $n \geq m$ above):

Signature 0, rank 3: $x^2 + y^2 + z^2 = 0$, the empty quadric (this equation doesn't have any non-zero solutions)

Signature 1, rank 3: $x^2 + y^2 - z^2 = 0$, the non-degenerate quadric (in different models of the projective plane can look like a circle, an ellipse, a parabola or a hyperbola)

Signature 0, rank 2: $x^2 + y^2 = 0$, one point (only the point $[0 : 0 : 1]$ is the solution of this equation in the projective space)

Signature 1, rank 2: $x^2 - y^2 = 0$, pair of lines $x = \pm y$ (in different models of projective plane they can be parallel or intersecting)

Signature 0, rank 1: $x^2 = 0$, double line (we see only the line $[0 : y : z]$, but it's natural to think of it as a pair of coinciding lines, since it is given by equation of degree two $x^2 = 0$, rather than the linear equation $x = 0$)

Signature 0, rank 0: $0 = 0$, the whole plane.

In particular this linear algebra shows that in projective geometry there is only one non-degenerate planar quadric, which in different coordinate charts could appear as an ellipse, hyperbola or parabola.

10.7   The only conic through five points

We know that through any two points in (Euclidean, affine or projective) plane lie on a line. Moreover, if these points are distinct, the line is unique.

In a similar vein, any three non-collinear points lie on a unique circle (in Euclidean plane - the notion of circle in not well-defined in affine or projective geometry).

In this section we will try to prove the following theorem:

Theorem 24. *Any five points lie on a quadric. If these points are distinct and no four of them lie on one line, the quadric passing through them is unique. Moreover, the quadric containing the five distinct points is non-degenerate if and only if no three points among these five are collinear.*

Proof. Let's start with the simpler part of the statement, which deals with degenerate cases. We first prove that if three distinct points are collinear, then every quadric passing through them is degenerate. Indeed, take the equation of a quadric passing through three collinear points and restrict it to the line containing these points. This equation is quadratic, but it has three roots! (the three points on the line that are assumed to lie on the quadric). Hence this quadratic equation must vanish identically on this line, i.e. the line must be contained in the quadric. Hence this quadric is degenerate (either union of two lines, a double line or the whole plane).

Moreover, in the case three of the five points are collinear, but no four among them are, it's easy to see that the degenerate quadric passing through the five points is unique. Indeed, what we have to do is take the union of the line that contains three collinear points with the line (distinct from the first one) that contains the other two.

Finally if we do have four collinear points among the five, then there are many degenerate quadrics passing through these five points: we can take the union of the line passing through the four collinear points and any line passing through the other one.

Now we show existence of a quadric passing through five given points by a very general argument.

Let Q be any n-dimensional space of functions on any space X. Let $a_1, \ldots, a_{n-1}$ be any $n-1$ points of this space X. Then there exists a nonzero function $f \in Q$ which vanishes on $a_1, \ldots, a_{n-1}$.

For our application we will take X to be $\mathbb{R}^2$, the space Q will be the space of quadratic functions, i.e. functions of the form $ax^2 + bxy + cy^2 + dx + ey + f$. Since this space is six-dimensional (the parameters a, b, c, d, e, f can be arbitrary, so that the functions $x^2, xy, y^2, x, y, 1$ are a basis for this space of functions), the claim above tells us that for every five points $a_1, \ldots, a_5 \in \mathbb{R}^2$ there exists a non-zero quadratic function, which vanishes at these five points. The zero set of this function is the quadric passing through the five points $a_1, \ldots, a_5$.

To prove the general case all we have to do is to notice that the condition $f(a_i) = 0$ defines a hyperplane in the space of functions Q. Indeed, the space of functions vanishing at a point a_i is linear subspace (if f and g both vanish at a_i, then all their linear combinations do). Moreover, it is of codimension at most 1 because $f(a_i) = 0$ is just one equation on the function f (more precisely, the quotient space of Q by the space $f | f(a_i) = 0$ is one-dimensional, since the value of a function at point a_i defines its coset uniquely). If this

is somewhat vague, consider the example we are interested in: the condition that a quadratic function $f(x, y) = ax^2 + bxy + cy^2 + dx + ey + f$ vanishes at point $a_i = (x_i, y_i)$ defines a linear equation on the coefficients a, b, c, d, e, f: $x_i^2 a + x_i y_i b + y_i^2 c + x_i d + y_i e + f = 0$. In this case it's obvious that it is a hyperplane in the space of coefficients a, b, c, d, e, f.

Now if we have $n - 1$ linear subspaces of Q of codimension at most 1, their intersection must be of codimension at most $n - 1$, i.e. is of dimension at least 1. Hence it must contain at least one non-zero function.

Exercise: prove that there exists a cubic passing through any given collection of 9 points.

What we proved by now is that there exists a quadric passing through any five points in the plane and we investigated the case when at least three of the points are collinear to the end. Now we want to prove that for five points in general position the quadric passing through them is unique.

Suppose to the contrary that there are two quadrics passing through five given points a_1, a_2, a_3, a_4, a_5 in general position. Since these points are in general position, both these quadrics must be non-degenerate (if five points lie on a union of two lines, then at least three of them must lie on the same line). Suppose their equations are $Q_1(x) = 0$ and $Q_2(x) = 0$. First we notice that we can find a whole pencil (i.e. a one-parameter family) of quadrics passing through the five points, namely the family of quadrics of the form $(Q_1 + \lambda Q_2)(x) = 0$. The crucial claim is that this family must contain a degenerate quadric! Indeed, take the point a of intersection of lines through points a_1, a_2 and through points a_3, a_4 (if these happen to be parallel then choose the point of intersection of lines through a_1, a_3 and $a - 2, a_4$). Since Q_2 is non-degenerate, it can't pass through any triple of collinear points. In particular $Q_2(a) \neq 0$. So we can take $\lambda = -\frac{Q_1(a)}{Q_2(a)}$ and for this choice of λ we have $(Q_1 + \lambda Q_2)(a) = 0$. The quadric $(Q_1 + \lambda Q_2)(a) = 0$ is degenerate, since it passes through a triple of collinear points! But this is a contradiction: the degenerate quadric $(Q_1 + \lambda Q_2)(x) = 0$ passes through five points $a_1, \ldots, a_5$, no three of which are collinear!

$\square$

We note that the trick we used in the end of the proof to show that two non-degenerate quadrics can intersect at most at four points can be used to help solving quadratic equations!

Namely, let $ax^4 + bx^3 + cx^2 + dx + e = 0$ be a quartic equation we want to solve. Set $y = x^2$. Then the solutions of the quartic equation are the same

as the x-coordinates of the system of two quadratic equations $Q_1(x,y) = ay^2 + bxy + cx^2 + dx + e = 0$ and $Q_2(x,y) = y - x^2$. Now every solution of these two equations is also a solution of the equation $(Q_1 + \lambda Q_2)(x,y) = 0$ for any choice of λ. We want to choose λ in such a way that the quadric defined by equation $(Q_1 + \lambda Q_2)(x,y) = 0$ is degenerate. How can we find such λ without finding the intersection points?

It turns out that the quadric $ax^2 + bxy + cy^2 + dx + ey + f = 0$ is degenerate if and only if the determinant of the matrix $\begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix}$ is equal to zero[1]. If quadrics Q_1 and Q_2 are defined by matrices $\tilde{Q}_1$ and $\tilde{Q}_2$, then to find the λ's for which the quadric $(Q_1 + \lambda Q_2)(x,y) = 0$, we have to solve the cubic in the unknown λ equation $\det(\tilde{Q}_1 + \lambda \tilde{Q}_2) = 0$.

So to solve a quartic equation we have to intersect two quadrics $Q_1(x,y) = 0$ and $Q_2(x,y) = 0$. Instead we can solve a cubic equation $\det(\tilde{Q}_1 + \lambda \tilde{Q}_2) = 0$ to find λ for which the quadric $(Q_1 + \lambda Q_2)(x,y) = 0$ becomes a pair of two lines and then intersect the quadric $Q_1(x,y) = 0$ with this pair of two lines. To do so we have to solve only a pair of quadratic equations!

---

[1]If matrix A has zero determinant, then we can find point $x \neq 0$ with $Ax = 0$, and then for every y on the quadric defined by the matrix A (i.e. quadric $y^T A y = 0$) the whole line of points $\lambda x + \mu y$ is on the quadric - $(\lambda x + \mu y)^T A (\lambda x + \mu y) = \mu^2 y^T A y = 0$, so the quadric is degenerate and x is the vertex of the corresponding cone. Conversely, if the quadric defined by A is degenerate, then it looks like a cone with vertex x for some x and then for every y on the quadric we have $x^T A y = \frac{1}{2}((x+y)^T A(x+y) - x^T A x - y^T A y) = 0$ and hence $A^T x$ is orthogonal to every y on the quadric. Thus $Ax = 0$.

Chapter 11

Projective duality

11.1 Definition of projective duality

We are going to see now a very funny transformation, called projective duality. What it allows us to do is to take a projective theorem (like Pascal's), apply this transformation, and get a new theorem, which is guaranteed to be true! We saw something similar with inversion: we could prove a theorem, apply inversion and get a new one. Projective duality, however, is not a construction that requires any choices, however (like a choice of circle for inversion), but rather is an internal feature of projective geometry.

Let us proceed straight to the definition. Let V be a vector space and let V^* denote the space dual to V, i.e. the space of all linear functions on V. First we define duality as a transformation from the set of linear subspaces of V to the set of linear subspaces of V^*: the space $W \subset V$ gets sent to the space $W^* = \{f \in V^* | \langle f, w \rangle = 0\}$, i.e. to the space of all functionals evaluating to zero on W. We used the notation $\langle f, w \rangle$ for the value of functional f on vector w to emphasize the symmetry of roles of f and w: we can think of a vector w as a functional on the space V^* evaluating on the element $f \in V^*$ to $\langle f, w \rangle$. This way of thinking shows how to identify $(V^*)^*$ with V itself (for finite dimensional V the embedding of V into $(V^*)^*$ we described is an isomorphism, because of equality of dimensions of the space and its dual). If we identify V with $(V^*)^*$ in this way, applying duality twice becomes identity: $(W^*)^* = W$ for any $W \subset V$.

The duality we described is easily seen to reverse inclusions: if the linear subspace W_1 contains W_2, then every function vanishing on it, must vanish

on W_2 as well, hence W_1^* is contained in W_2^*.

Now we can projectivise duality we defined above to get the notion of projective duality: projective duality is the transformation from projective subspaces of $\mathbb{P}V$ to projective subspaces of $\mathbb{P}V^*$ sending the subspace $\mathbb{P}W$ (for $W \subset V$ a linear subspace) to $\mathbb{P}W^*$.

Since in the usual duality lines through origin get sent to hyperplanes through origin, in projective duality points are dual to hyperplanes. More generally spaces of dimension k are dual to spaces of codimension $k+1$.

By now we learned the following important properties of projective duality: it sends points to hyperplanes, hyperplanes to points and reverses inclusion relations.

For instance if before the duality we had a picture of n points lying on the same hyperplane, after applying duality we will get the picture of n hyperplanes passing through the same point.

To get some really powerful applications of projective duality we will have to extend its definition to more than just linear subspaces.

11.2   Dual polygon and dual curve

We now restrict our attention to the projective plane.

Let C be a polygon in $mathbbRP^2$ whose vertices are $P_1, \ldots, P_n$. Let $L_1, \ldots, L_n$ denote the lines on which the sides of polygon C lie, so that L_i is the line through P_i and P_{i+1} and P_i is the point of intersection of lines L_{i-1} and L_i.

We define the dual polygon C^* to be the polygon whose vertices are points L_i^* - the points in $\mathbb{RP}_2^*$, which are dual to the sides of the polygon C. What are its sides? The side containing the points L_{i-1}^* and L_i^* should be dual to the point of intersection of L_{i-1} and L_i, i.e. to P_i.

Thus we can think of the dual polygon in two ways: as the polygon whose sides are the lines dual to the vertices of the original polygon, or as the polygon, whose vertices are the points dual to the sides of the original one.

By combining these two descriptions it's easy to see that the dual polygon of the dual polygon is the original one.

Indeed, the vertices of the dual polygon of C^* are the points dual to P_i^*, hence are equal to P_i.

We introduced the notion of a dual polygon to help us understand better the following definition of a dual curve.

Suppose C is a curve in the projective plane $\mathbb{RP}^2$. We define the dual curve to C to be the curve C^*, whose points are the points dual to tangent lines to C.

We can think of the space $\mathbb{RP}^2$ as the space parameterizing lines in $\mathbb{RP}^2$. Then the curve C^* should be thought of as the curve whose points are tangent lines to C.

We can imagine the curve C as a limit of polygons, whose sides get shorter and shorter and whose vertices lie on the curve C. Then the sides of this polygon tend to the lines tangent to the curve C. We see then that the dual curve C^* is approximated well by the dual polygons to the polygons approximating C.

The observations we made about dual polygons have their direct analogues for dual curves. Namely, dual curve to a curve C can be thought of in two ways:

The curve whose points are dual to the tangent lines of C

or

The curve whose tangent lines are dual to the points of C

This observation lead immediately to the claim that the dual curve of the dual curve of C is C itself.

Let's see a couple of examples: if the curve C has a double tangent, i.e. a line tangent to C at two points, then the dual curve will have a self-intersection point. Indeed, imagine a point P going along C. Then the tangent line at P is changing with P. As P goes from one tangency point of the double tangency line to C to the other one, the tangent line makes a loop and comes back to itself. Thus the dual curve intersects itself before closing up.

Dually, if a curve has a double point, then its dual has a double tangent.

Space for animation! (worth 5 points if done with asymptote animation module: the animation should depict a curve with point P moving along it and another curve C^*, which is dual to the first, on which point dual to the tangent line at P is moving)

In a similar vein, if a curve has an inflection point (i.e. point where its concavity changes), then the dual curve will have a cusp. (Define cusp!) As point P moves along C and passes the inflection point, the tangent line changes the direction of its movement.

Dually, if a curve has a cusp then the dual curve has an inflection point.

These two last statements make it necessary to define a little bit more precisely what kind of curves we are talking about. Talking about smooth curves is not enough - even if the original curve is smooth, its dual might have cusps. On the other hand we do need the notion of the tangent line at every point to define the dual of the curve at all. (*The class of curves, whose only singularities are cusps and double points seems to be the most natural class to talk about.*) In particular all algebraic curves fall into this class, and hence projective duality becomes a valuable tool in the study of projective algebraic geometry.

We could have defined the notion of dual hypersurface to a given hypersurface C in $\mathbb{RP}^n$ in essentially the same way we did it for curves. The points of the dual hypersurface C^* are the dual points of tangent hyperplanes to C and then we could prove that the tangent hyperplanes to C^* are the hyperplanes dual to the points of C. We won't use this theorem below, however, so we don't explain why it is true.

11.3   Dual to a quadric

In this section we want to understand what is the dual hypersurface to a quadric.

Let A be a symmetric matrix defining the quadric $\{x | \langle Ax, x \rangle = 0\}$. We want to understand what is the equation of the hyperplane tangent to this quadric at point x_0 on it. To do so, let's try to move the point x_0 infinitesimally in direction ξ and see whether we go off the quadric or stay on it. To do so we evaluate the function $\langle Ax, x \rangle$ at point $x_0 + \xi$:

$$\langle A(x_0 + \xi), x_0 + \xi \rangle = \langle Ax_0, x_0 \rangle + \langle A\xi, x_0 \rangle + \langle Ax_0, \xi \rangle + \langle A\xi, \xi \rangle$$

The first term $\langle Ax_0, x_0 \rangle$ in this formula vanishes because the point x_0 lies on the quadric. The last term is an infinitesimal of second order (it is of order $|\xi|^2$) and hence can be ignored in linear approximation. Thus if we want to stay on the quadric to the first order of magnitude, we should have $\langle A\xi, x_0 \rangle + \langle Ax_0, \xi \rangle = 0$. Since the matrix A is symmetric, $\langle A\xi, x_0 \rangle = \langle \xi, Ax_0 \rangle$ and thus the equation of the hyperplane tangent to the quadric at point x_0 is $\langle Ax_0, \xi \rangle = 0$. The point dual to this hyperplane clearly is Ax_0.

As the point x varies along the the quadric $\langle Ax, x \rangle = 0$, the point $y = Ax$ varies along the quadric $\langle y, A^{-1}y \rangle = 0$. Thus we have proved the following theorem.

Theorem 25. *The dual hypersurface to a non-degenerate quadric is a non-degenerate quadric. Moreover, the dual of the quadric given by equation $x \in \mathbb{RP}^n | \langle Ax, x \rangle = 0$ is the quadric given by equation $y \in \mathbb{RP}^{n*} | \langle A^{-1}y, y \rangle = 0$.*

11.4 Brianchon's theorem

Projective duality gives us a very simple way to generate new theorems from old ones: all we have to do is apply duality.

Let's do it for Pascal's theorem.

Recall that Pascal's theorem tells us that if $O_1, O_2, O_3, P_1, P_2, P_3$ are points on a quadric $\mathfrak{E}$, then the points E_1, E_2, E_3 defined as the points of intersection of O_2P_3 with O_3P_2, of O_1P_3 with O_3P_1 and of O_2P_1 with O_1, P_2 respectively, are collinear.

Let's try to apply duality to this claim.

The dual to a quadric $\mathfrak{E}$ is a quadric $\mathfrak{E}^*$. The dual lines to the points $O_1, O_2, O_3, P_1, P_2, P_3$ are lines $O_1^*, O_2^*, O_3^*, P_1^*, P_2^*, P_3^*$, which are tangent to the quadric $\mathfrak{E}^*$.

The point dual to the line O_2P_3 is the point of intersection of lines O_2^* and P_3^*. Similarly the point dual to the line O_3P_2 is the point of intersection of lines O_3^* and P_2^*. Thus the line dual to point E_1 is the line E_1^* containing the points of intersection of O_2^* with P_3^* and of O_3^* with P_2^*. The definition of E_2^* and E_3^* is similar.

Finally the statement dual to "E_1, E_2, E_3 are collinear" is "E_1^*, E_2^*, E_3^* are concurrent.

What we get is the wonderful theorem of Brianchon.

Theorem 26. *The diagonals connecting opposite vertices of a hexagon circumscribed around a conic are concurrent.*

Notice that we know this theorem is valid because we know the theorem of Pascal and projective duality. We could however prove this theorem without using projective duality, if we wanted to. To do so we would need to dualize all the arguments we used for proving Pascal's theorem.

The key argument for the proof of Pascal's theorem was the following:

Let $f : l_1 \to l_2$ be a mapping from line l_1 to line l_2 in a projective plane. Suppose f is projective, (i.e. preserves cross ratios). Let O be the point of intersection of l_1 and l_2. Suppose that $f(O) = O$. Then there exists a point

C in the plane so that the mapping f coincides with the central projection from line l_1 to line l_2 with center C.

In the dual statement instead of a mapping taking in points on line l_1 and spitting out points of line l_2, we should talk about a mapping that takes in lines passing through one point and spitting out lines passing through another line.

So our dual statement will start with "Let $f : \mathfrak{P}_1 \to \mathfrak{P}_2$ be a mapping from the pencil $\mathfrak{P}_1$ of lines passing through point P_1 to the pencil $\mathfrak{P}_2$ of lines passing through point P_2. Suppose that f is projective (i.e. preserves cross ratios)."

Instead of point O - the point lying on both lines l_1, l_2, we will have the line P_1P_2 that belongs to both pencils.

Finally we should understand what is dual to a central projection with center C. Recall that it is sends the point X on line l_1 to the point of l_2 lying on the line CX.

We now define a "central projection from pencil of lines $\mathfrak{P}_1$ to the pencil of lines $\mathfrak{P}_2$ with center at line l". The image of line $X \in \mathfrak{P}_1$ under this projection will be the line in $\mathfrak{P}_2$ passing through point of intersection of X with l.

Finally we can formulate the dual statement to the end:

Let $f : \mathfrak{P}_1 \to \mathfrak{P}_2$ be a mapping from the pencil $\mathfrak{P}_1$ of lines passing through point P_1 to the pencil $\mathfrak{P}_2$ of lines passing through point P_2. Suppose that f is projective (i.e. preserves cross ratios). Suppose also that $f(P_1P_2) = P_1P_2$. Then there exists a line l in the plane so that the mapping f coincides with a mapping sending line $X \in \mathfrak{P}_1$ to the line in $\mathfrak{P}_2$ passing through point of intersection of X with l.

We hope the reader now has acquired some sense for what does it mean to dualize statements/definitions of projective geometry.

11.5   Projective mappings from a line to a line, revisited

Now we have enough tools to describe all projective mappings from a line l_1 in a plane to another line l_2 in the same plane.

Recall that we proved already that through any five points one can draw a unique quadric provided no three points among the five are collinear. More-

over in this case the quadric is non-degenerate.

Duality gives us the fact that there exists exactly one quadric tangent to given five lines in general position (i.e. no three meet at a point), and, moreover, this quadric is non-degenerate.

Now we will prove the following theorem.

Theorem 27. *Let $f : l_1 \to l_2$ be a projective mapping from line l_1 to line l_2 in the plane. Let O be the point of intersection of l_1 and l_2.*

If $f(O) = O$, then there exists a point C in the plane so that the mapping f coincides with the central projection with center C.

If $f(O) \neq O$, then there exists a non-degenerate conic C tangent to l_1 and l_2 such that the mapping f sends every point X on l_1 to the point on l_2 lying on the line passing through X and tangent to C, which is different from l_1.

Proof. We have dealt with the proof of the first statement in chapter ...

Now suppose $f(O) \neq O$. Take any three points on l_1 which are distinct and do not coincide with O or its preimage under f. Call these points A, B, C. Now the lines $l_1, l_2, Af(A), Bf(B), Cf(C)$ are five lines in general position. Hence there exists a unique non-degenerate quadric C tangent to all five of them. Now define the map $\tilde{f}$ to be the map that takes a point X on l_1 and outputs the point of l_2 that lies on the tangent line to conic C through point X and which is different from l_1. The map $\tilde{f}$ is projective (see the proof that cross ratio of four points on a quadric is well-defined). Moreover, the maps f and $\tilde{f}$ agree on three points: A,B,C! Hence f must be the same as $\tilde{f}$. $\square$