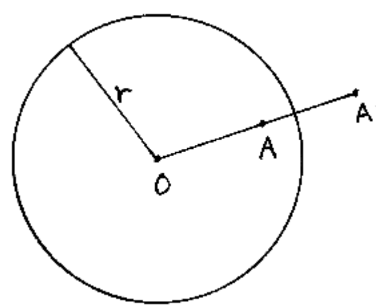
37 Circular Inversion

In this section we will study circular inversion, which is a kind of transformation of the plane that leaves the points of a given circle fixed, and sends points inside the circle to points outside and vice versa. While this study belongs in the context of Euclidean geometry, it is a technique not used by Euclid. Perhaps this is because the idea of a *transformation* of the plane, moving points to other points, was foreign to the way of thinking of the Greeks. Euclid does use the "method of superposition" to compare triangles in his proof of (SAS), but there is no evidence that he thought of a rigid motion moving the whole plane onto itself. Given that perspective it seems even less likely that the Greeks would have seen any value in a transformation of the plane that does not even preserve distances, in fact that does not even preserve proportion in figures. It seems that the notion of transformation in geometry is a relatively recent notion, which has come to serve very important roles, as we have seen in the usefulness of the existence of rigid motion (ERM).

The theory of circular inversion can be developed purely geometrically using the results of Euclid's Book III, but it will be more efficient to use the theory of proportion (similar triangles) in our proofs. So in this section we work for convenience in the Cartesian plane over a Euclidean ordered field F. Thus we have Hilbert's axioms, including (P) and (E), and we can use the theory of similar triangles (Section 20). The Euclidean hypothesis on F can be slightly relaxed (Exercises 37.16, 37.17).

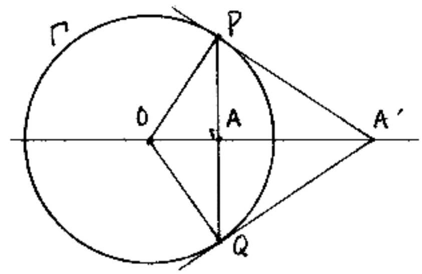
Definition

Let Γ be a fixed circle in the plane (over the field F as above...), with center O and radius r. For any point $A \neq O$, draw the ray OA, and let A' be the unique point on the ray OA such that $OA \cdot OA' = r^2$. (The dot in this equation means products of lengths in the field F.) Then we say that A' is obtained from A by *circular inversion* with respect to the circle Γ.



Briefly, we will say that A' is the *inverse* of A (with respect to the circle Γ). Since the condition $OA \cdot OA' = r^2$ is symmetric in A and A', this notion is reciprocal: A' is the inverse of A if and only if A is the inverse of A'. We think of circular inversion in Γ as a transformation $\rho = \rho_\Gamma$ defined for all points $A \neq O$, which thus transforms the plane $\Pi - \{O\}$ into itself by sending each point to its inverse. From the definition it is clear that any point *on* Γ is sent to itself. Points inside Γ are sent to points outside Γ, and vice versa. As a point A approaches the center of the circle O, its inverse gets farther and farther away, so in the limit, the point O would have to go to infinity. Since we do not have infinity in our geometry, we simply say that ρ is *undefined* at O. (Or if you like, you can imagine completing our plane by a single point called infinity, and then O and ∞ are interchanged—cf. Exercise 37.1 on stereographic projection for another interpretation of this idea.)

Proposition 37.1
Let A be a point inside the circle Γ. Draw the ray OA. Let PQ be the chord of the circle perpendicular to OA at A. Then the tangents to Γ at P and Q will meet the ray OA at the point A' that is the inverse of A with respect to Γ.



Proof First note that the two tangents will both meet OA in the same point A', by symmetry. Now, the right triangles ΔOAP and $\Delta OPA'$ have the angle at O in common, so they are similar. Hence corresponding sides are proportional. In particular,

$$\frac{OA}{OP} = \frac{OP}{OA'}.$$

Cross multiplying, we obtain

$$OA \cdot OA' = OP^2 = r^2.$$

Hence A and A' are circular inverses in Γ.

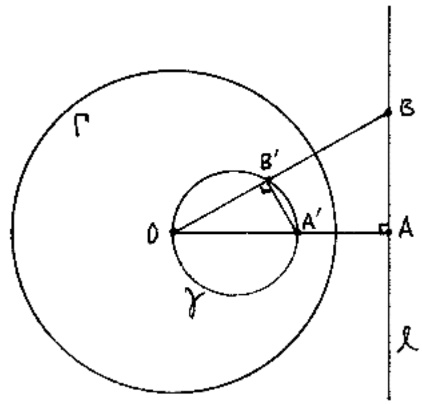
Remark 37.1.1
This proposition gives us a method of constructing circular inverses by ruler and compass: If A is given inside Γ, draw OA, construct the perpendicular to OA at A, let it meet Γ at P, draw the radius OP, draw the perpendicular to OP at P, which will be the tangent line, and let this line meet OA at A' (9 steps). Conversely, if the point A' is given outside the circle, draw the two tangent lines from A' to Γ, join their points of tangency P, Q, and let the line PQ meet OA' at A (6 steps).

Next we will investigate what circular inversion does to lines and circles in the plane.

Proposition 37.2
A line through O is transformed into itself by circular inversion. A line not passing through O will be transformed into a circle passing through O, and conversely.

Proof A line through O is transformed into itself by definition of circular inversion.

Now let l be a line not through O. Let OA be the perpendicular from O to l. Let A' be the inverse of A, and let γ be the circle with diameter OA'. I claim that the inverses of points on l all lie on γ and vice versa. So let B be any point on l. Draw OB and let it meet γ at B'. Then $OB'A'$ is a right triangle (III.31). It has the angle at O in common with the right triangle OAB, so we have similar triangles. Therefore, the sides are proportional:



$$\frac{OB'}{OA'} = \frac{OA}{OB}.$$

Cross multiplying, we obtain

$$OB \cdot OB' = OA \cdot OA'.$$

But A' was chosen to be the inverse of A, so $OA \cdot OA' = r^2$. Hence also $OB \cdot OB' = r^2$, so B and B' are inverse to each other. This shows that the circular inversion transforms the points of the line l to the points of the circle γ (except O) and vice versa.

Definition

When two circles meet (or when a circle meets a line) by the *angle* between them we mean the angle between their tangent lines at that point (resp. the angle between the tangent line and the other line).

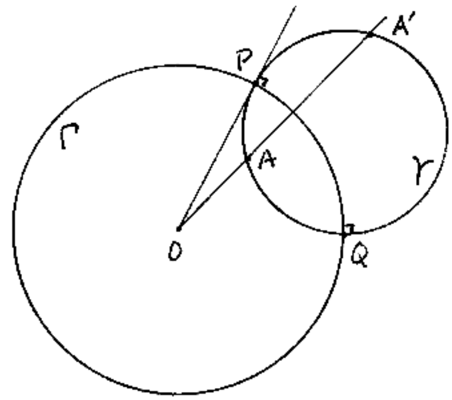
Note that when two circles meet in two points, the angle between them is the same at both points, because the two circles are symmetrical about the line joining their two centers.

Proposition 37.3

If a circle γ is perpendicular to Γ (at its intersection points), then γ is transformed into itself by circular inversion in Γ. Conversely, if a circle γ contains a single pair A, A' of inverse points, then γ is perpendicular to Γ and is sent into itself.

Proof First suppose that γ is perpendicular to Γ, and let γ meet Γ at P and Q. Then the radius OP is tangent to γ, because radius and tangent of any circle are perpendicular (III.18). Let A be another point of γ and let OA meet γ again at A'. Applying (III.36) to γ we obtain $OP^2 = OA \cdot OA'$. (Actually, Euclid meant that the square on OP has content equal to the rectangle formed by OA and OA', but since we are working over the field F, we interpret this statement as lengths and products (20.9).) Since $OP = r$, this shows that A and A' are inverses. This holds for any A on γ, so γ is sent into itself.

Now (using the same picture) suppose, conversely, that γ is any circle passing through some pair of inverse points A, A'. Let γ meet Γ at P, and draw OP. Since OP is a radius and A, A' are inverse points, we have $OA \cdot OA' = OP^2$. But now by (III.37), the converse of (III.36), it follows that OP is tangent to γ, which means that γ and Γ are perpendicular at P (and hence also at their other point of intersection Q).

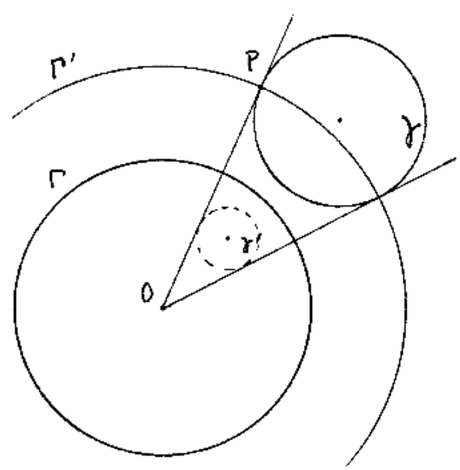


Proposition 37.4

If γ is a circle not passing through the center O of Γ, then the transform of γ by circular inversion is another circle γ'.

Proof This result is not so easy to prove directly (you can try if you like), so we will resort to a trick. Suppose we are given a circle γ not passing through O, and assume that O is outside γ, for the moment.

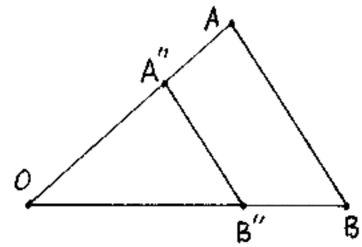
Draw OP tangent to γ, and let Γ' be a new circle with center O, passing through P. Then by construction γ is sent into itself by $\rho_{\Gamma'}$. Thus $\rho_\Gamma(\gamma) = \rho_\Gamma \cdot \rho_{\Gamma'}(\gamma)$, and we are led to consider the new transformation of the plane $\theta = \rho_\Gamma \cdot \rho_{\Gamma'}$. Let $r =$ radius of Γ, and $r' =$ radius of Γ'. For any point A, let $A' = \rho_{\Gamma'}(A), A'' = \rho_\Gamma(A')$. Then $\theta(A) = A''$. By definition of inversion, $OA \cdot OA' = r'^2$ and $OA' \cdot OA'' = r^2$. Dividing, we find that



$$\frac{OA''}{OA} = \frac{r^2}{r'^2},$$

so

$$OA'' = k \cdot OA, \qquad \text{where} \quad k = \frac{r^2}{r'^2}.$$



This is a transformation that leaves O fixed, and stretches points toward (or away from) O in a fixed ratio k. It is called a *dilation* with center O and ratio k. In rectangular coordinates with center at the origin it would be expressed by

$$\begin{cases} x' = kx, \\ y' = ky. \end{cases}$$

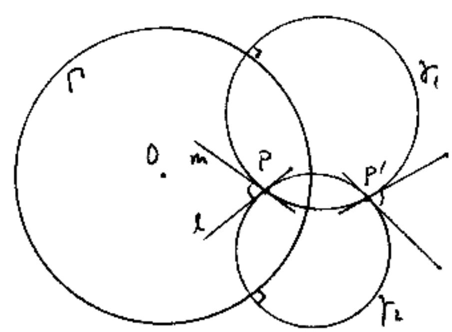
It follows, either from thinking of the distance formula in terms of coordinates, or by using the (SAS) criterion for similar triangles (20.4), that all distances are changed by the same ratio k.

In particular, a dilation sends any circle (and its center) into another circle and its center. It follows that $\rho_\Gamma(\gamma) = \rho_{\Gamma} \cdot \rho_{\Gamma'}(\gamma) = \theta(\gamma)$ is a circle. (Warning: Even though $\rho_\Gamma(\gamma)$ is a circle, in general ρ_Γ does not send the center of γ to the center of γ'.)

In this proof we were assuming O outside γ. If O is inside γ, we leave you to construct an analogous proof in Exercise 37.4.

Now that we have seen that circular inversion preserves lines and circles (every line or circle is transformed into another line or circle), sometimes turning a line into a circle, or a circle into a line, the next step is to show that circular inversion is *conformal*, i.e., preserves angles.

Proposition 37.5
Circular inversion is conformal: *Whenever two curves meet (here "curve" means line or circle), their transforms under circular inversion meet again at the same angles.*



Proof First suppose that $P \notin \Gamma$, and that two curves (not shown) meet at P with tangent lines l, m. Let P' be the inverse of P. Then we can find a circle γ, through P, P' and with tangent m at P; and we can find a circle γ_2 through P, P' with tangent line l at P. Now, by (37.3) γ_1 and γ_2 are transformed into themselves. Therefore, the original curves are transformed into curves at P' tangent to γ_1 and γ_2, so they make the same angle as at P, because when two circles γ_1, γ_2 intersect they have the same angle at both intersections. For this proof we need to observe that a line and a circle, or two circles, are tangent if and only if they have just one point in common. Hence the property of tangency is preserved by inversion.

If $P \in \Gamma$, we leave the special case to you (Exercise 37.5).

For our last general result about the properties of circular inversion, we look at what happens to distances. Of course, distances are not preserved, because a very small distance near O will be transformed into a very large distance far away. Even ratios of distances are not preserved, as you can see by simple examples. However, a remarkable fact is that if we take four points, a certain ratio of ratios of distances, called their cross-ratio, is preserved.

Definition

Let A, B, P, Q be four distinct points in the Cartesian plane. Their *cross-ratio* (an element of the field F) is defined to be the ratio of ratios

$$(AB, PQ) = \frac{AP}{AQ} \div \frac{BP}{BQ},$$

which can also be written

$$\frac{AP}{AQ} \cdot \frac{BQ}{BP}.$$

Proposition 37.6

Let A, B, P, Q be four distinct points in the plane, different from O. Then circular inversion in Γ preserves the cross-ratio: If their inverses are A', B', P', Q', then

$$(AB, PQ) = (A'B', P'Q').$$

Proof Given two points A, P and their inverses A', P', we know by definition that

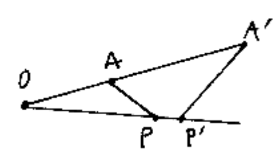
$$OA \cdot OA' = r^2 = OP \cdot OP'.$$

Thus

$$\frac{OA}{OP} = \frac{OP'}{OA'}.$$

Case 1 Suppose O, A, P are not collinear. Since the triangles ΔOAP and $\Delta OP'A'$ have the angle at O in common, they are similar (20.4), and we conclude that

$$\frac{AP}{A'P'} = \frac{OA}{OP'}. \tag{1}$$



Case 2 If O, A, P are collinear, then $AP = OP - OA$ and $A'P' = OA' - OP'$. Using the fact that in a field F,

$$\frac{a}{b} = \frac{c}{d} \Rightarrow \frac{a}{b} = \frac{c-a}{d-b},$$

we conclude the same result (1).

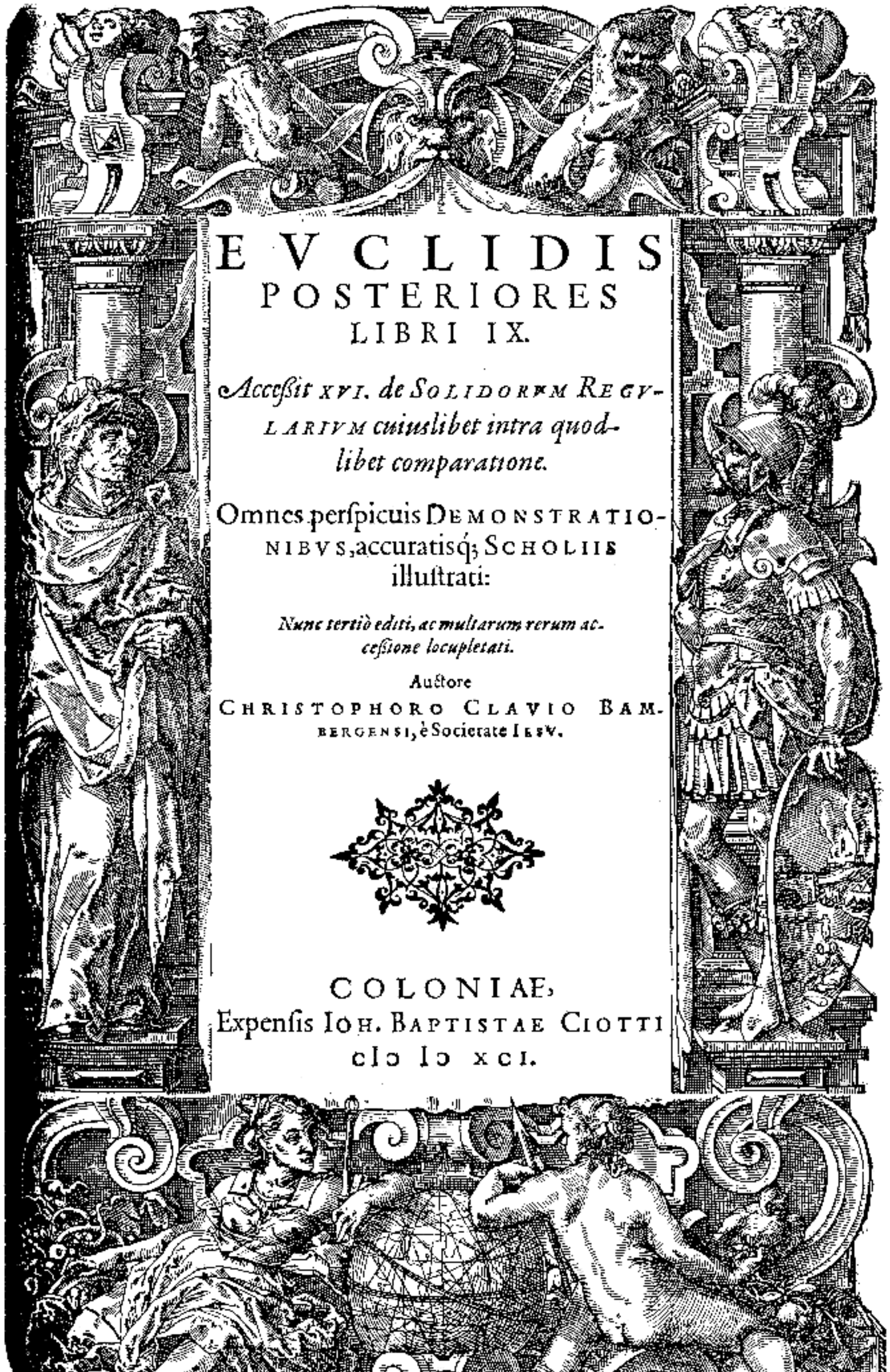


Plate XIII. Title page of volume II of the important edition of the *Elements* by the Jesuit mathematician Christopher Clavius (1591).

Now if Q is another point, we find similarly that

$$\frac{AQ}{A'Q'} = \frac{OA}{OQ'}. \tag{2}$$

Dividing, we get

$$\frac{AP}{A'P'} \div \frac{AQ}{A'Q'} = \frac{OQ'}{OP'}. \tag{3}$$

Now let B be another point. Working with P and Q as before, we obtain similarly

$$\frac{BP}{B'P'} \div \frac{BQ}{B'Q'} = \frac{OQ'}{OP'}. \tag{4}$$

So the expressions (3) and (4) are equal. Moving the primed letters to one side and the unprimed letters to the other side shows that the cross-ratios (AB, PQ) and $(A'B', P'Q')$ are equal.

Remark 37.6.1

At this point I can just hear someone asking, "What is the geometrical significance of the cross-ratio?" Although I first encountered cross-ratios as a senior in high school, and have dealt with them many times since then, I must say frankly that I cannot visualize a cross-ratio geometrically. If you like, it is magic. Here is this algebraic quantity whose significance it is impossible to understand, and yet it turns out to do something very useful. It works. You might say it was a triumph of algebra to invent this quantity that turns out to be so valuable and could not be imagined geometrically. Or if you are a geometer at heart, you may say that it is an invention of the devil and hate it all your life.

Let me say a few words in defense of the poor cross-ratio.

In the present context of transformations of the Euclidean plane, there are rigid motions, which preserve distance. Then there are dilations, which do not preserve distance, but do preserve ratios of distances. Then there is circular inversion, which does not preserve distances or even ratios of distances. Since it does preserve the cross-ratio, that particular ratio of ratios, it is the best we can do. It is something to hang on to, a pillar of support, when the distances and their ratios are changing all around us. In Section 39 we will use the cross-ratio to define the notion of distance in the Poincaré model of non-Euclidean geometry: It plays an essential role there.

In projective geometry the cross-ratio is also important. A *projectivity* from one line to another is defined as a composition of a finite number of projections from one line to another in the plane. A projectivity preserves neither distances nor ratios of distances, but it does preserve the cross-ratio (Exercise 37.14). In fact, a fundamental theorem of projective geometry is that a transformation of one line to another in the projective plane is a projectivity *if and only if* it preserves the cross-ratio of every set of four distinct points on the line.

If you have studied complex variables, the projectivities of the projective line over the complex numbers correspond to the fractional linear transformations of the Riemann sphere $\mathbb{C} \cup \{\infty\}$, given by

$$z' = \frac{az+b}{cz+d}, \quad ad - bc \neq 0.$$

If you are given four points A, B, C, D, there is always a fractional linear transformation sending A, B, C to $0, 1, \infty$. In that case the image of D, say λ, is the cross-ratio of the original four points, in a suitable order.

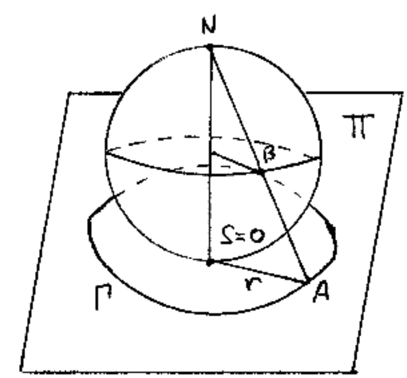
Finally, if you have four points on a line, and you take signed distances (+ or − depending on a chosen preferred direction), then the cross-ratio is equal to -1 if and only if the four points form a set of *four harmonic points* (Exercise 37.15). This notion of harmonic points is also important in projective geometry.

The notion of cross-ratio already occurs in the work of Pappus (300 A.D.). It came into prominence again in the early nineteenth century with the projective geometry of Poncelet and Monge.

Exercises

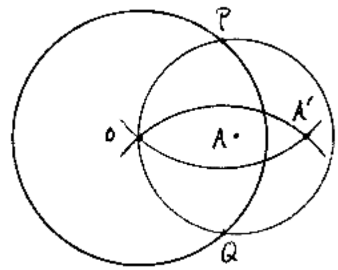
Unless otherwise noted, the following exercises take place in the Cartesian plane over a Euclidean ordered field F.

37.1 *Stereographic projection.* In three-dimensional space, imagine our plane Π and a circle Γ of radius r and center O. Now take a sphere of radius $\frac{1}{2}r$, and set it on the plane Π so that its south pole is at O. Then *stereographic projection* associates to each point B of the sphere, $B \neq N =$ north pole, that point of Π obtained by drawing the line NB and intersecting with Π. (In the limit, N would go to infinity, so you can think of the sphere as a completion of the plane by adding the point N.) Under this projection, the equator of the sphere is mapped to the circle Γ.

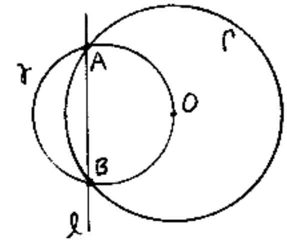


Show that circular inversion in the circle Γ corresponds to the operation of *reflection in the equator* of the sphere, which interchanges the northern and southern hemispheres. In other words, if B is a point on the sphere, and B' its reflection in the equator (same longitude, but latitude has changed from north to south or from south to north), then the projected points A, A' are inverses under inversion in Γ.

37.2 Prove that the following construction with compass alone gives the inverse of A in Γ (provided that $OA > \frac{1}{2}r$): Draw a circle with center A through O to meet Γ at P and Q. Then draw circles with centers P and Q through O to meet at A' (3 steps). (The diagram shows A inside Γ, but the construction works equally well if A is outside Γ.)



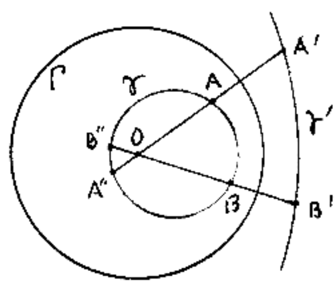
37.3 Let l be a line that meets the circle Γ in two points A, B. Let γ be the (unique) circle through O, A, B. Prove that γ is the transform of l under inversion in Γ.



37.4 Prove the other case of (37.4), namely, if γ is a circle containing O, then $\rho_\Gamma(\gamma)$ is a circle. For any points A, B on γ, let A', B' be their inverses in Γ, and let A'', B'' be the points where the lines OA, OB meet γ again. By (III.35) (cf. (20.8)),

$$OA \cdot OA'' = OB \cdot OB'' = c$$

is a constant independent of the points A, B, depending only on O, γ.



Let us use signed lengths from O, so that $c < 0$. Since $OA \cdot OA' = r^2$, then show that $OA' = k \cdot OA''$ for a certain constant $k < 0$. Thus γ' is obtained from the circle γ by dilation with a negative constant k. Conclude that $\rho_\Gamma(\gamma) = \gamma'$ is a circle.

37.5 If two lines or circles meet at a point $P \in \Gamma$, show that their two transforms by circular inversion in Γ meet at the same angle at P.

37.6 If we identify the real Euclidean plane $\mathbb{R}^2$ with the complex numbers $\mathbb{C}$, show that the transformation $z' = 1/\bar{z}$ (where $z = a + bi$, $\bar{z} = a - bi$) is just inversion in the unit circle $|z| = 1$.

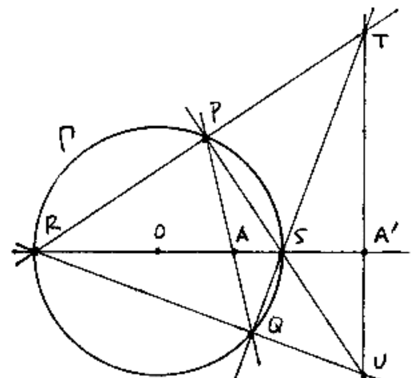
37.7 If A is a point inside the circle Γ, improve the ruler and compass construction of the inverse of A given in (37.1.1) by constructing the circle through O, P, Q instead of constructing the tangent line at P (par = 7 steps).

37.8 Given the circle Γ and its center O, and given a line l, give a ruler and compass construction of the circle $\rho_\Gamma(l)$ (par = 7 steps).

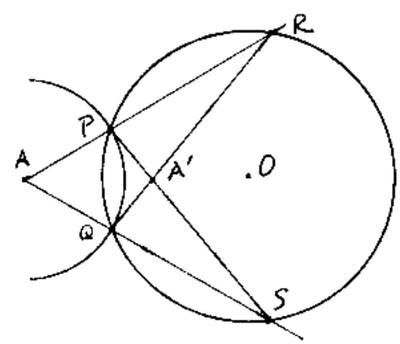
37.9 Given Γ, given its center O, and given a circle γ passing through O (but not given the center of γ), construct the line $\rho_\Gamma(\gamma)$ (par = 15 steps).

37.10 Given Γ and its center O, and given a circle γ not through O, construct the circle $\rho_\Gamma(\gamma)$ (par = 15 steps).

37.11 Verify the following ruler-only construction of the inverse of a point A (7 steps): Draw OA, get R, S. Draw any line l through A meeting Γ in P, Q. Draw RP and SQ to meet at T. Draw RQ and PS to meet at U. Draw TU to meet OA at A'. Show also that TU is perpendicular to OA.



37.12 Verify the following 5-step construction for the inverse of a point A with respect to the circle Γ. Take a circle of any radius with center A, to meet Γ at P and Q. Let AP and AQ meet Γ in further points R, S. Join PS and RQ. Their intersection is A'. (This works equally well if A is inside Γ.)

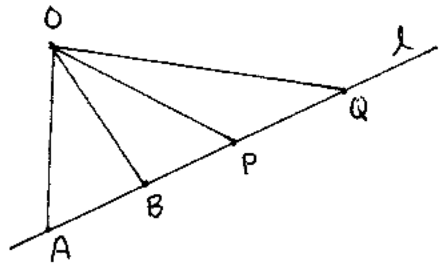


37.13 (a) Given four points A, B, P, Q, if you permute A and B, or if you permute P and Q, the cross-ratio is replaced by its inverse: $(BA, PQ) = (AB, QP) = (AB, PQ)^{-1}$.

(b) More generally, if A, B, P, Q are four points on a line, and if $(AB, PQ) = \lambda$, then the 24 possible permutations of the points give rise to 6 possible values of the cross-ratio, namely

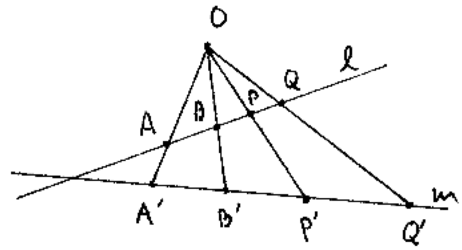
$$\lambda, \quad \frac{1}{\lambda}, \quad 1-\lambda, \quad \frac{1}{1-\lambda}, \quad \frac{\lambda-1}{\lambda}, \quad \frac{\lambda}{\lambda-1}.$$

37.14 (a) Given four points on a line l, and given a point O not on l, let the angles at O subtended by AP, AQ, BP, BQ be $\alpha_P, \alpha_Q, \beta_P, \beta_Q$. Use the law of sines to show that



$$(AB, PQ) = \frac{\sin \alpha_P}{\sin \alpha_Q} \div \frac{\sin \beta_P}{\sin \beta_Q}.$$

(b) If the four points A, B, P, Q on l are projected from O to four points A', B', P', Q' on another line m, then the cross-ratio is preserved:



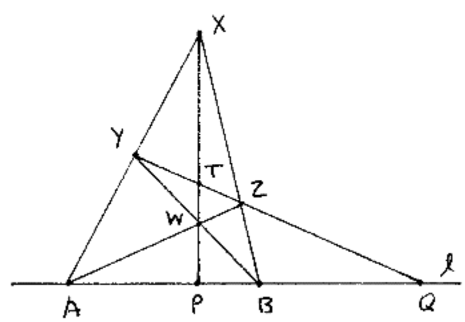
$$(AB, PQ) = (A'B', P'Q').$$

Conclude that cross-ratio is preserved by any projectivity, that is, a finite succession of projections from one line to another.

37.15 We say that four points A, B, P, Q on a line form a set of *four harmonic points* if their cross-ratio (AB, PQ) is equal to -1.

(a) Given A, B, P, show that the fourth harmonic point Q is uniquely determined.

(b) Verify the following ruler-only construction of the fourth harmonic point: Given A, B, P on a line l, take a point X not in the line. Draw XA, XB, XP. Take any point Y on AX. Draw BY, get W. Draw AW, get Z. Draw YZ, get Q. *Hint*: Project the four points A, B, P, Q from X to the line YQ, and then from W back to the original line l, and use Exercise 37.14.



(c) If A, B, P, Q are four harmonic points, show that Q is the inverse of P in the circle with diameter AB.

37.16 If F is a Pythagorean ordered field, we can still define inversion in a circle Γ by the same method as at the beginning of this section. If γ is a circle with $\rho_\Gamma(\gamma) = \gamma$, show that γ still meets Γ in two points, even though we do not have the axiom (E).

37.17 Let F be a Pythagorean ordered field, and let d be a positive element of F that has no square root in F. We consider the *virtual circle* Γ defined by the equation $x^2 + y^2 = d$. Since F is Pythagorean and $\sqrt{d} \notin F$, this equation has no solutions. So Γ has no points. Nevertheless, it is useful to refer to Γ as a virtual circle, because we can still define circular inversion ρ in Γ by the formula $OA \cdot OA' = d$.

(a) Show that the results (Propositions 37.2, 37.4, 37.5, 37.6) still hold for ρ.

(b) Show that a part of Proposition 37.3 holds: A circle γ is sent to itself by ρ if and only if it contains a pair of inverse points. In this case we say by abuse of language that γ is orthogonal to Γ.

(c) Given O and given one pair A, A' of inverse points by ρ, give a construction with Hilbert's tools (Section 10) for the inverse B' of a point B. (Par $= 4$ if O, A, B are not collinear.)

37.18 Give a new proof of Exercise 1.15 by doing a circular inversion with center P and suitable radius, and solving the transformed problem.

Poor John has lost his ruler. Can you help him do his construction problems (following) using his compass alone?

37.19 Given two points A, B construct a third collinear point C with $AB = BC$ (par $= 4$ steps).

37.20 Given two points A, B, construct the midpoint C of the segment AB (par $= 7$ steps). *Hint*: Use Exercise 37.2.

37.21 Given two points A, B, construct C, D such that $ABCD$ will be a square (par $= 8$ steps).

37.22 Given points A, B, C, O, with O, A, B not collinear, construct the intersection points of the line AB with the circle OC (assuming that they meet) (par $= 4$ steps).

37.23 Given noncollinear points A, B, C, construct the foot of the perpendicular from C to the line AB (par $= 9$ steps).

37.24 Given noncollinear points O, A, B, show that it is possible to construct the intersection of the circle OA with the line OB using compass alone. *Hint*: Perform a circular inversion that leaves the circle OA fixed and transforms the line OB into a circle. (Par $= 13$ to get one intersection point.)

37.25 Given points A, B, C, D show that it is possible to construct the intersection point of the lines AB and CD using compass alone. *Hint*: Use a circular inversion to transform the two lines into circles. (Par $= 13$ steps if the points are in favorable position; otherwise 18 steps.)

37.26 Using the experience gained in the previous exercises, prove the following theorem of Mascheroni: Any point that can be constructed from given data by ruler and compass construction can also be constructed using compass alone.

38 Digression: Circles Determined by Three Conditions

In order to specify a circle in the plane, you give its center, which is a point, and its radius, which is a line segment, or distance. A point moves in a 2-dimensional