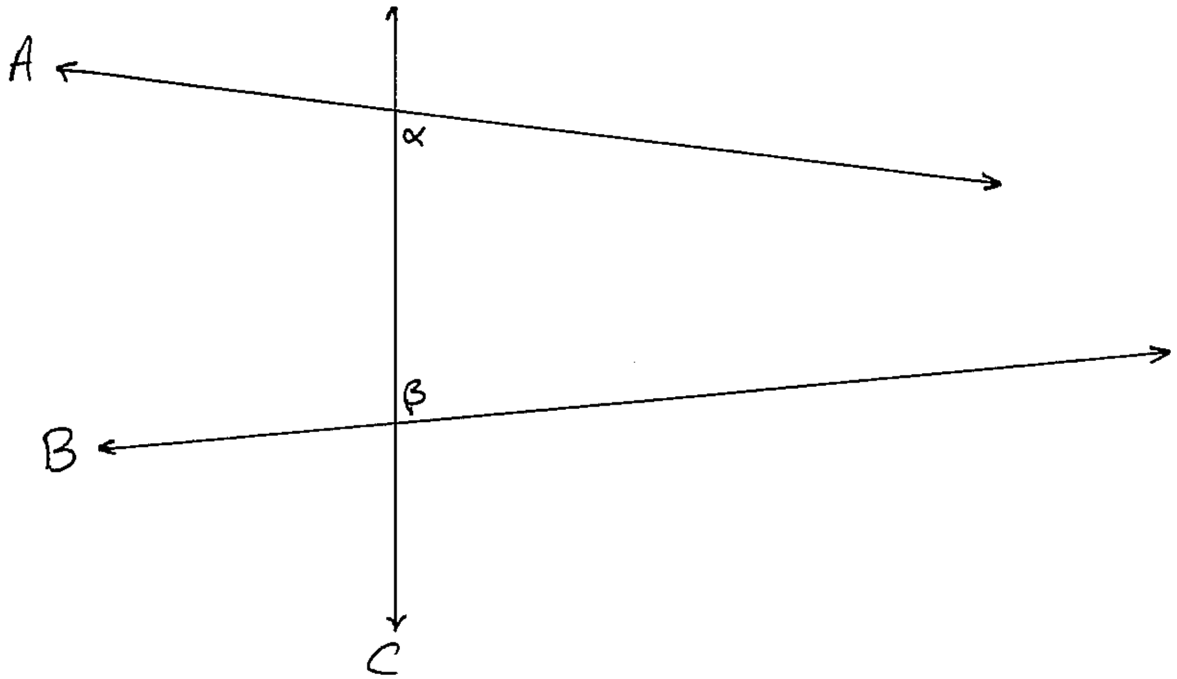
Euclid's fifth postulate is independent

Undoubtedly the most famous mathematics book ever written is Euclid's "Elements", written around 300 BC. The "Elements" is a virtual encyclopedia of geometry, algebra, and number theory as understood by the Greeks, but its real importance lies in its method of presentation. Euclid begins with five "postulates" or axioms. These are five simple self-evident truths of geometry, and from these five postulates, in a tour de force of logic and ingenuity, Euclid deduced all other known mathematics. The five postulates are

1. Two points determine a segment
2. A segment may be extended to become a line
3. A segment may be used as a radius to create a circle
4. Any two right angles are equal
5. If two lines cross a third line and the sum of facing angles so made is less than two right angles, then the two lines meet on the side of those facing angles.

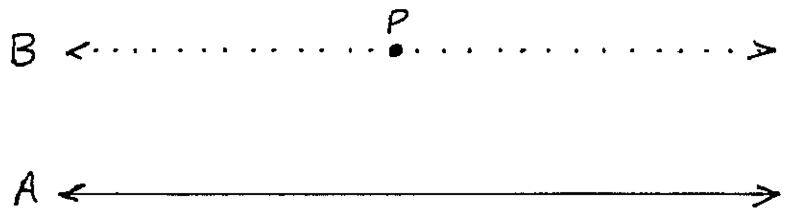
The fifth postulate needs a picture



Here the two lines A and B cross the third line C, while α and β are facing angles. Euclid says that if $\alpha+\beta < 180^\circ$ (as drawn above), then lines A and B meet, that is A and B are not parallel. This is known as the "parallel postulate".

Euclid's successors were not happy with the fifth postulate. It is much less transparent than the other four, much less intuitively obvious. Euclid's postulate is now usually replaced with "Playfair's axiom":

5'. Given a line and a point not on the line, there is one line through the point parallel to the given line.



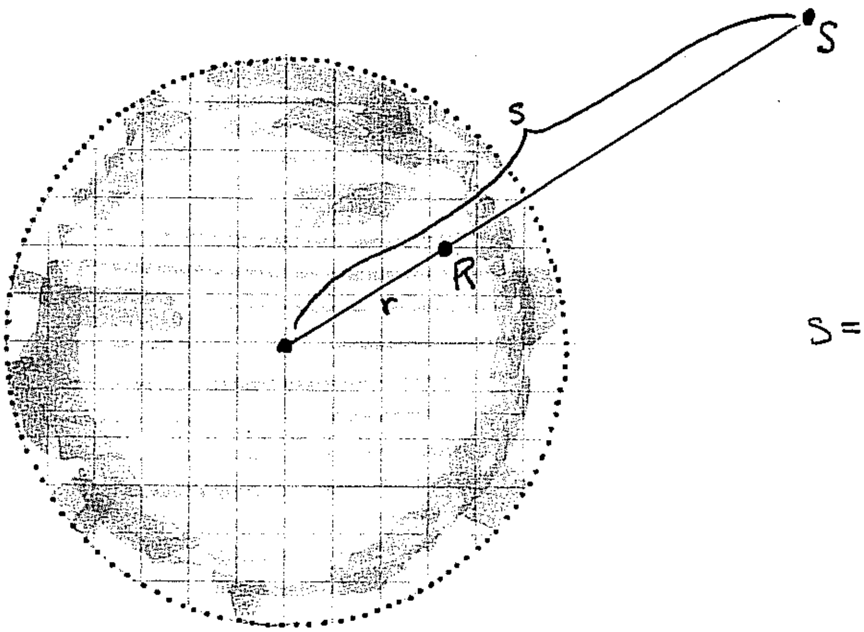
Given the line A and the point P we can draw one line B parallel to A. Though this is easier to understand than Euclid's postulate, it is logically equivalent; from postulate 5 we can deduce 5' and vice versa. Many people tried to show that the fifth postulate could be deduced from the first four, or that geometry could be developed without using the fifth postulate, but all attempts failed. Finally, in the early nineteenth century, Gauss, Lobachevsky and others showed that a perfectly consistent geometry could be developed assuming the parallel postulate is false! The parallel postulate is independent of the other postulates, and its truth or falsity is a matter of choice, not deduction. There is a universe in which it is true and an equally valid universe in which it is false.

There are basically two types of these non-Euclidean universes – ones with too many parallels and ones with too few parallels. We will look at "projective" geometry, where seemingly parallel lines meet "at infinity".

The difficulty is understanding "at infinity". "Infinity" is just too far away for us to see, but Renaissance artists found a method to project the entire Euclidean plane into a disk, which will assist our visualization. (They used these methods to paint the infinite heavens onto the ceiling of a church, or the Earth onto a canvass).

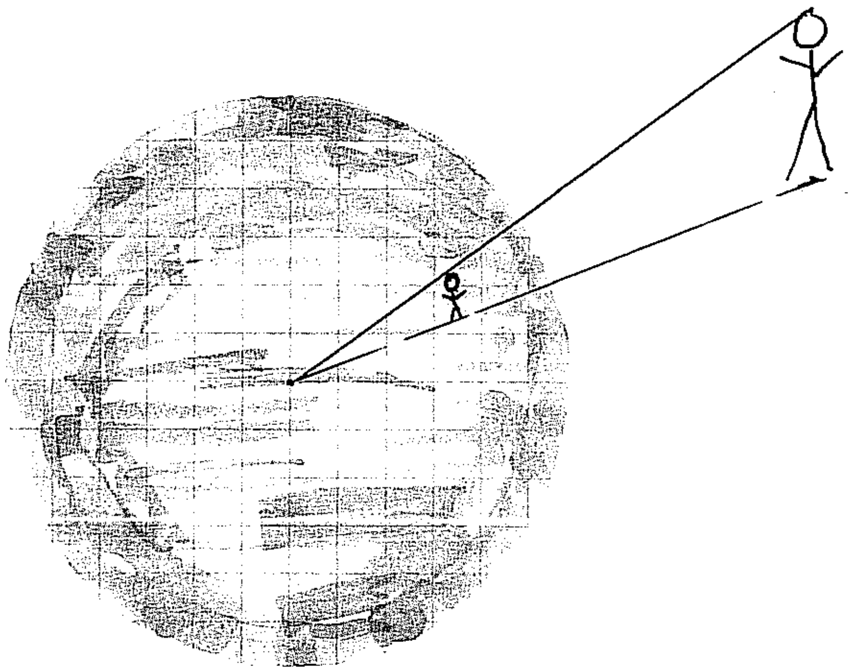
Consider an open disk of radius one, i.e. the interior of a unit circle, not including the circle itself. To each point R in the disk we will associate a point S as follows: If the distance from R to the center of the disk is r, S is the point in the same direction as R but with distance from the origin s where

$s = \frac{2r}{1-r^2}$ or $r = \frac{\sqrt{s^2+1}-1}{s}$ (which is the same thing)

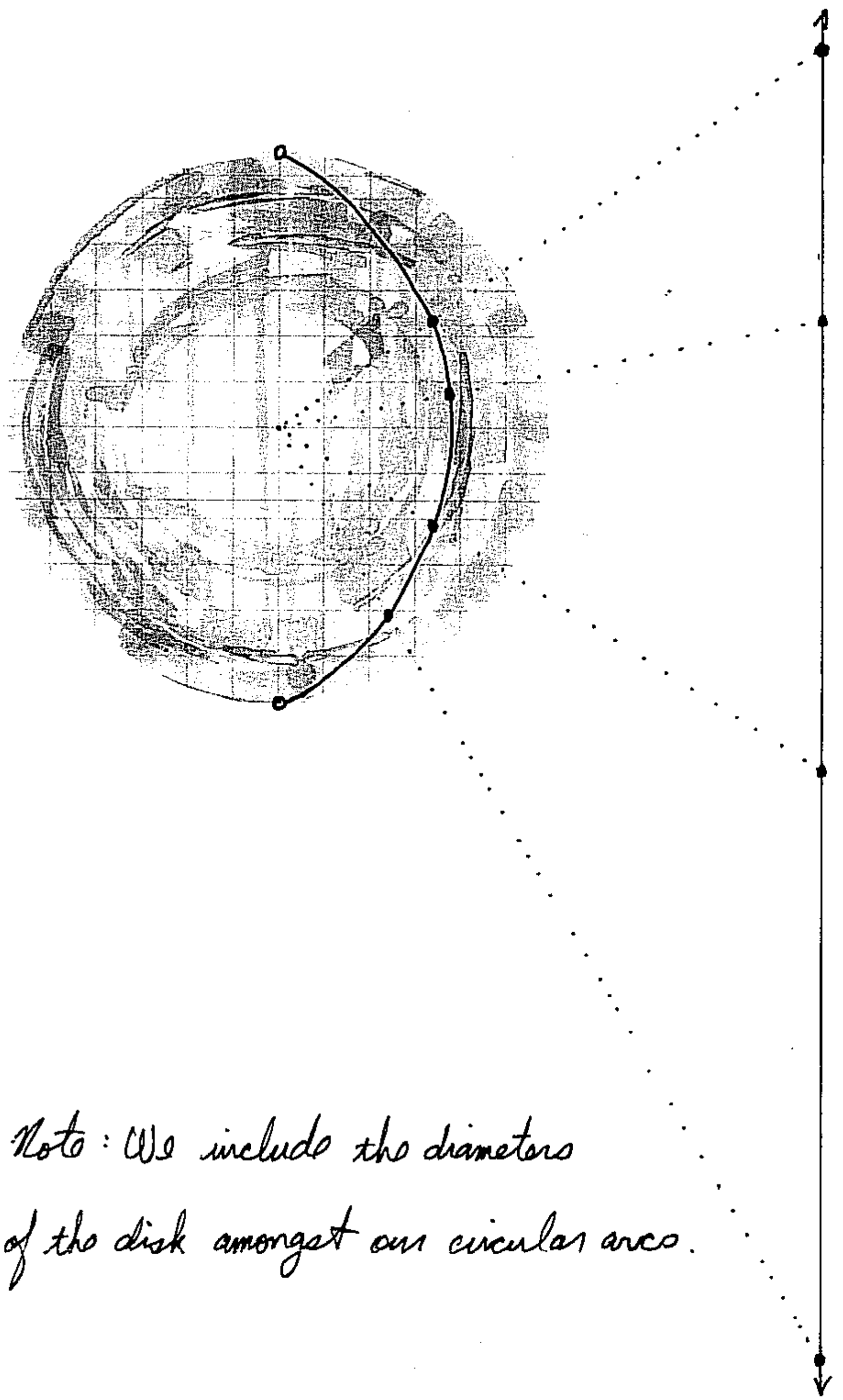


$$S = \frac{2r}{1-r^2}$$

This associates each point S in the whole plane to a point R within the disk. In fact any diagram in the plane is projected onto a miniaturized diagram in the disk.



In particular, any line in the Euclidean plane is projected onto an antipodal circular arc within the disk, i.e. a circular arc that connects the ends of a diameter of the disk.

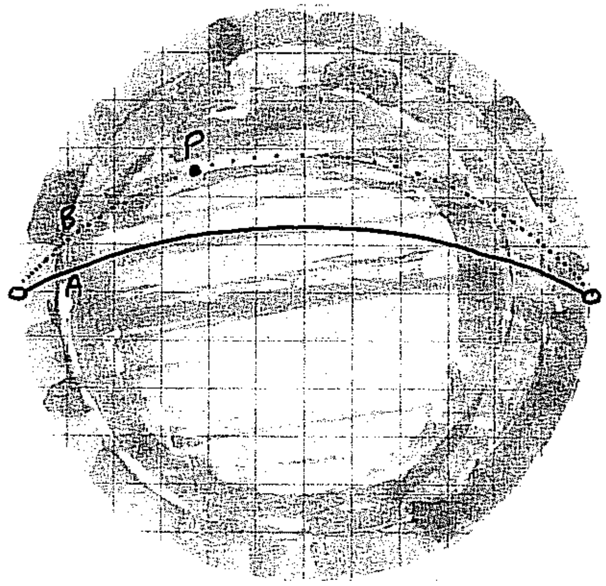


Note: We include the diameters of the disk amongst our circular arcs.

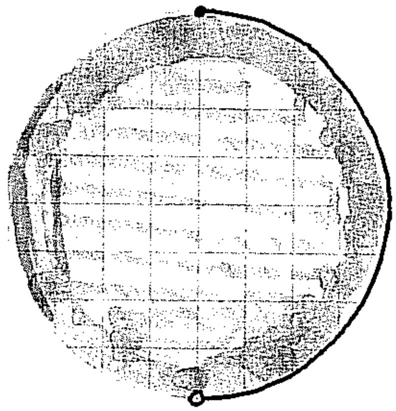
Note that as the points on the line move further from the center of the disk the corresponding points on the circular arc move closer to the edge of the disk, though they never reach the edge of the disk (as indicated by the little holes). We will show in a supplement below that if the line is a distance A from the center of the disk, then the center of the circular arc is a distance $\frac{1}{A}$ on the opposite side.

The open unit disk is thus a model for Euclidean geometry. If we interpret the word "point" to mean "point in the open disk", and the word "line" to mean "antipodal circular arc", and the word "angle" to mean "angle of the corresponding diameters", we find all the axioms of Euclid hold in our miniature world. The disadvantage to doing geometry in the model is that figures, like circles, will be badly distorted, and we can't measure distance in the usual way, but the great advantage is that we can see the horizon line, the "line at infinity", the boundary of the open disk.

In this setting the diagram which illustrates Playfair's axiom looks like this:

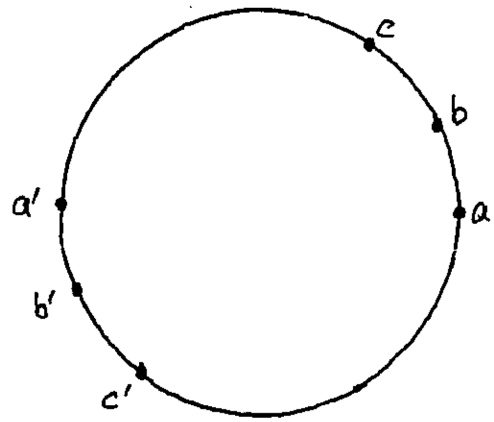


The line B is the unique parallel to line A through the point P. These lines do not meet because there are no points on the edge of the disk. But what would happen if there were points on the boundary of the disk, i.e. if there were a line at infinity? It is easy enough to complete our model by adding the boundary points to get a closed disk, but then Euclid's first axiom (two points determine a segment) would fail because there would be many segments connecting two antipodes (lines A and B above, for example). There are two ways to fix this. We could add only half the boundary points, so our model would look like this:



Now two points will determine a unique antipodal circular arc as before, but we've made one side of the universe look different from the other side. This is not only cosmologically unsatisfying, but it will wreak havoc with our other axioms (for example, think about a circle whose center is the north pole).

Here, then, is the right way to build our model for geometry: Add all the points on the boundary of the disk, and then sew opposite sides of the boundary together matching antipodes, so those antipodes become one point, i.e. paste a to a′, b to b′, c to c′ below:



We don't have to actually sew the sides together; it is enough to understand that a and a' are the same point in our new model, which is known as the "real projective plane", as opposed to the "real Euclidean plane" that we started with.

The important feature of the projective plane is that the first four axioms of Euclid still hold. We can draw lines and circles and angles just like before, but the fifth axiom is not true in the projective plane. There are no parallel lines. The lines that used to be parallel meet. This shows that the fifth axiom cannot be a logical consequence of the first four.

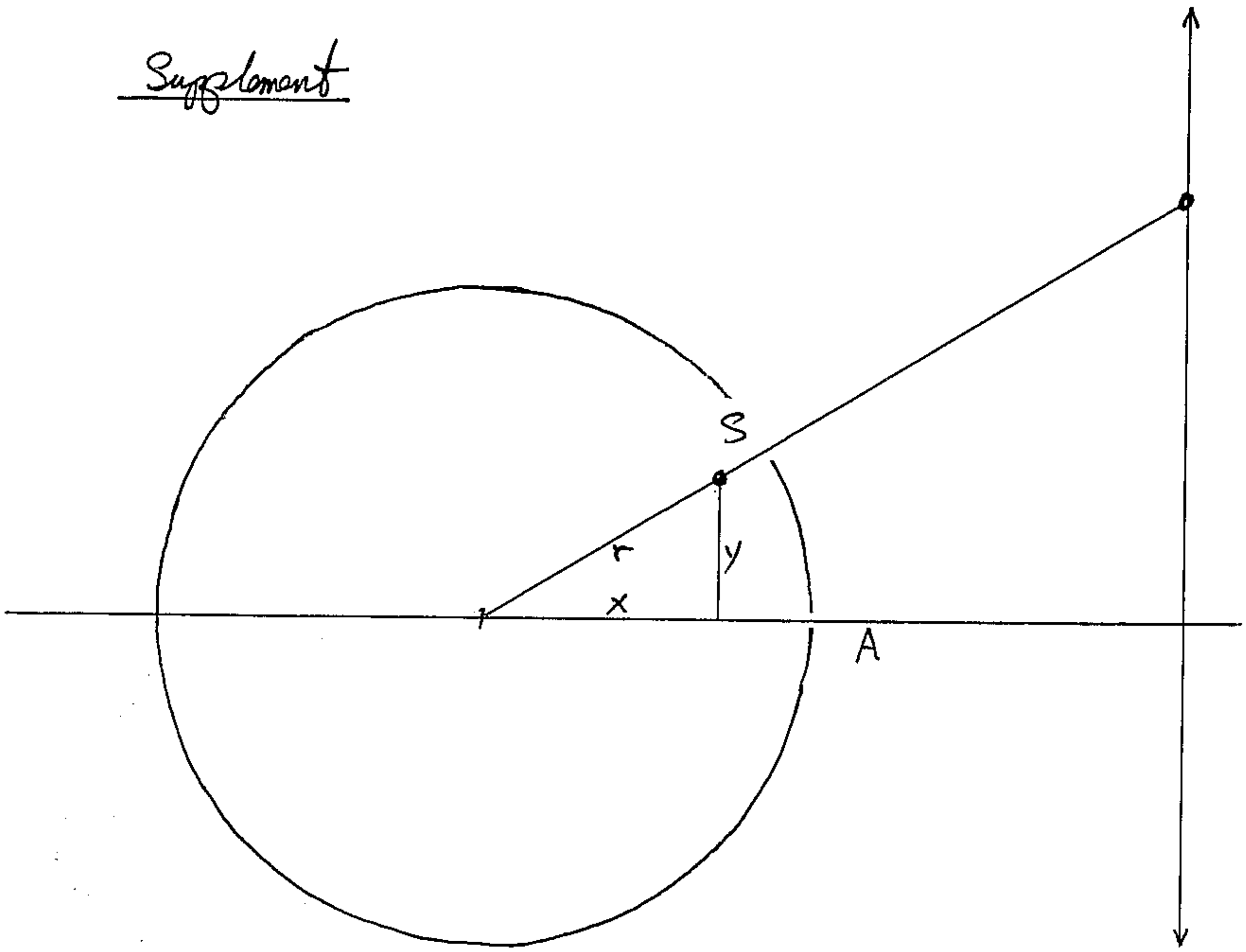
You may object that this has no relevance to the real world, in which we can "clearly" draw parallel lines. But how do you know what happens "at infinity"? The physicists tell us that the universe is in fact finite, and so the geometry of the universe cannot be Euclidean. Perhaps a three dimensional projective space is a better representation of the real universe, or perhaps "hyperbolic" space is better yet.

Mathematicians are not concerned with our corporeal universe; they are only concerned with the idealized mathematical universe, and projective geometry is just as valid as Euclidean geometry from a mathematical perspective. If you want to do geometry, you must choose which geometric universe you want to work in, which set of axioms you want to start with. This idea confronts all our ideas of mathematical truth and reality. What is true in one universe is not necessarily true in an alternate universe, even though all our "facts" are the result of logical deduction.

By the way, modern mathematicians studying "algebraic geometry" have chosen to work in complex projective space, an amalgamation of the last chapter and this. It is there that the clearest harmonies between geometry and algebra are revealed.

"Finite man cannot claim to be able to regard the infinite as something to be grasped by means of ordinary methods of observation"
— C.F. Gauss

Supplement



Coordinatize the disk and line as above. Note that

$r^2 = x^2 + y^2$ (Pythagoras) and $\frac{r}{S} = \frac{x}{A}$ (similar triangles).

Then from $S = \frac{2r}{1-r^2}$ we have

$$1 - r^2 = 2\frac{r}{S}$$

$$r^2 + 2\frac{r}{S} = 1$$

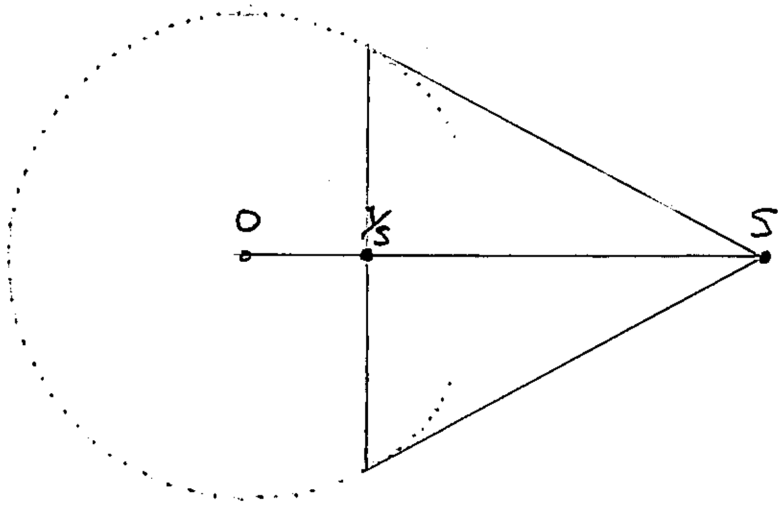
$$x^2 + y^2 + \frac{2}{A}x = 1$$

$$\left(x + \frac{1}{A}\right)^2 + y^2 = 1 + \frac{1}{A^2}$$

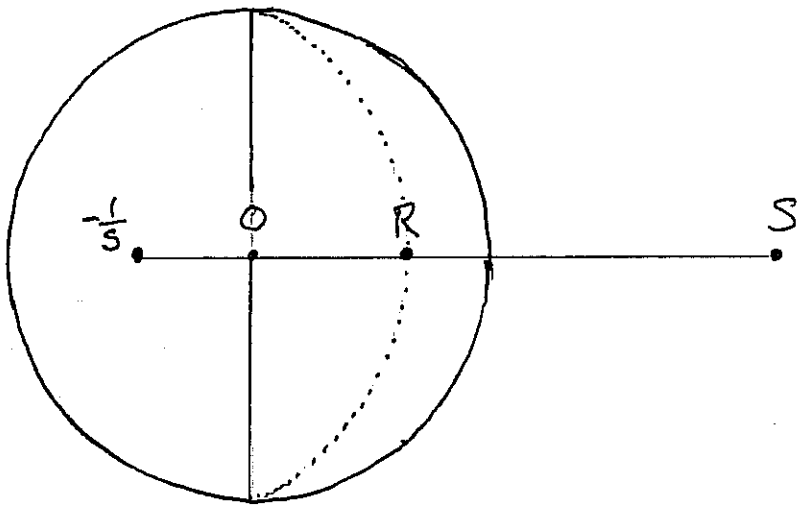
Thus the point (x, y) is on the circle with center $\left(-\frac{1}{A}, 0\right)$

and radius $\sqrt{1 + \frac{1}{A^2}}$.

This gives the key to constructing the projection using only ruler and compass geometry. Given a point S outside the disk, draw the two tangents to the circle. The cord connecting them cuts the line from S to the center O at a distance $\frac{1}{S}$ from the center as below (this is the well known "inversion" of S):



Now find the point $-\frac{1}{S}$ on the other side of the center and draw a diameter perpendicular to OS. Finally, draw an antipodal arc centered at $-\frac{1}{S}$ as below. Its intersection with the line OS is the point R. We leave it to the reader to work out what to do if S is inside the disk.



The picture above should also make it clear how to find S given R: Find the center of the arc passing through R. That's $-\frac{1}{s}$, and S is just the inverse of $\frac{1}{s}$.