Alhazen's Hyperbolic Billiard Problem

Nathan Poirier and Michael McDaniel
Aquinas College
May 2011

Alhazen's billiard problem, first posed in 150 BC, starts with a given circle and two points A and B inside. We then try to construct an inscribed, isosceles triangle with A on one leg and B on the other using compass and straightedge. The billiard aspect comes from imagining a round billiard table with two points marked on the felt. Setting up a three-cushion shot which passes through the two points and stops where it started turns out to be the same as solving the Euclidean problem. Until Riede in 1989, various mathematicians attempted the general construction but only obtained constructions for specific locations of the given points. Riede proved that the general construction could not be obtained. His proof puts Alhazen's billiard problem in the non-constructible category with cube roots, angle trisection and other mathematical quantities unreachable with compass and unmarked straightedge.

In this paper, we consider the hyperbolic version of Alhazen's billiard problem and show that the hyperbolic situation matches the Euclidean. For the hyperbolic version, we have the same given: two points A and B inside a given circle. We will use the Poincaré disk model of hyperbolic geometry, where hyperbolic lines are diameters of the disk and arcs of circles orthogonal to the disk's boundary. A satisfactory isosceles triangle would have its sides on such lines, with its vertices on the given circle and with A and B on the legs. We will see that the hyperbolic and Euclidean solutions correspond. Since hyperbolic triangles can have angle sums between 0 and π, this correspondence was unexpected.

1 Euclidean example

An obvious construction exists for A and B on a diameter of the given circle and equidistant from the center, our first figure. We can construct the perpendicular bisector of $\overline{AB}$, which intersects the given circle at two points, each a potential vertex of the desired isosceles triangle. We can also prove that these two triangles are the only two non-degenerate solutions. First, we place a point V on the given circle and construct the legs of a potentially isosceles triangle through A and B. Now consider the difference of the squares of the lengths of these potential legs as a function of V. We can make this function explicit using the Law of Cosines.

$$VU^2 - VW^2 = r^2 + r^2 - 2r^2 \cos VOU - (r^2 + r^2 - 2r^2 \cos VOW)$$
$$VU^2 - VW^2 = 2r^2(\cos VOW - \cos VOU)$$

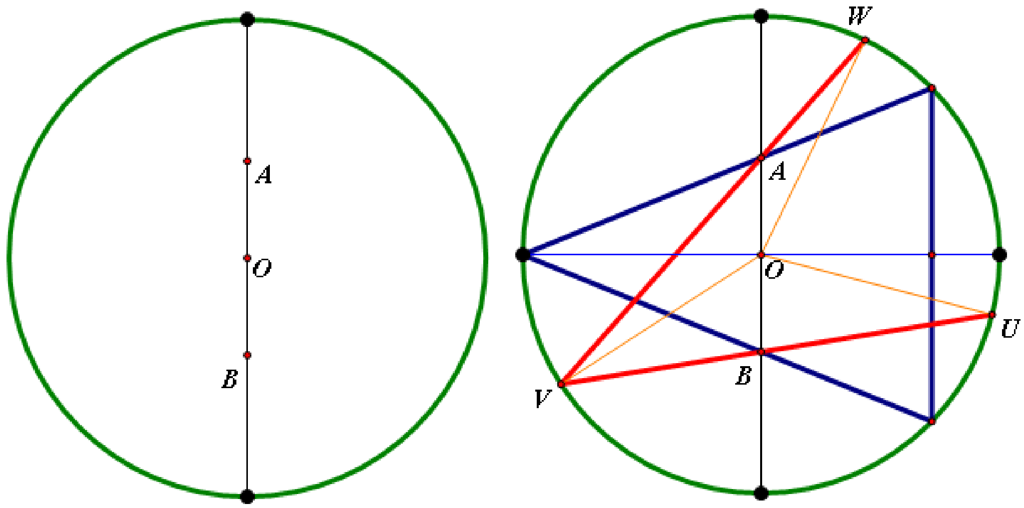

Figure 1. Two solutions from this A and B.

Since cosine is a strictly decreasing function on the interval of our angle sizes, the difference $\cos VOW - \cos VOU = 0$ only at the points emphasized with large dots in Figure 1. Only the two points on the perpendicular bisector of AB give non-degenerate solutions. We include such a simple case because its corresponding hyperbolic case will illustrate our bijection between Euclidean and hyperbolic cases later.

1.0.1 Alhazen's Billiard Problem in Hyperbolic Geometry

Hyperbolic geometry has several models; we will use the Poincaré disk and the Klein model. The Poincaré disk is a hyperbolic space made of the Euclidean points inside, not including, a fixed boundary circle. Euclidean diameters and arcs of circles orthogonal to the boundary circle are hyperbolic lines. Hyperbolic distance cannot be discerned from appearances in the model because the hyperbolic distance formula gives different sizes for segments which look the same in Euclidean size, depending on their proximity to the boundary. One consequence of the hyperbolic distance formula is that a circle in the disk has a Euclidean center and a hyperbolic center which are the same only when the circle's Euclidean center coincides with the center of the boundary circle, called O. Another property which surprises people in their first visit to the Poincaré disk is that the angle sums of triangles lie between 0 and π. The Klein model also uses a disk for the boundary. The hyperbolic lines are Euclidean chords. In this model, Euclidean appearances are even more deceiving because hyperbolic angle sizes are not the Euclidean sizes visible in the model. The Poincaré disk has its angle size visible, using tangents at the vertex of the angle. Both models allow infinitely many intersecting hyperbolic lines parallel to a given line. There exists an isomorphism between these two models, called stereographic projection, which, by coincidence, fits our Alhazen problem.

The Alhazen billiard problem in the hyperbolic disk model has more given than the Euclidean version because we have the disk, with the given circle inside the disk. The given circle has a Euclidean center E, the one we would use to draw the circle with a compass. The given circle also has a hyperbolic center H, the point which is hyperbolic equidistant from the points on the circle. When E is not O, then H ends up closer to the boundary than E, on the Euclidean ray $\overrightarrow{OE}$. (We can construct H by constructing a hyperbolic line perpendicular to both the boundary and the given circle.) Figure 2 summarizes the properties of the hyperbolic situation. For points A and B conveniently placed, the construction is possible. (Our bijection will define what we mean by "convenient.") As a quick look at the hyperbolic version, the given points have been placed hyperbolic equidistant from $\overrightarrow{OH}$ and chosen so that the base of the constructed triangle contains H. The triangle JKL illustrates one easy way to get congruent legs: the hyperbolic altitude, $\overline{HL}$, to the base is also a median.

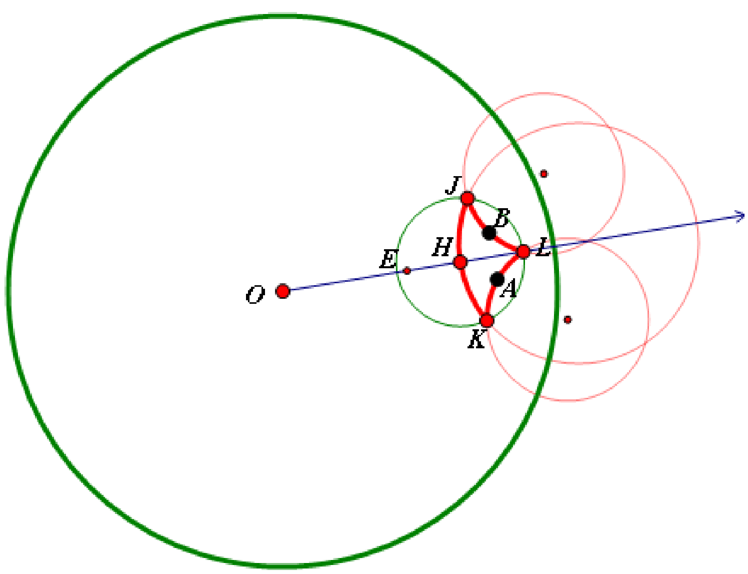

Figure 2. A hyperbolic Alhazen triangle.

2 Non-constructibility

Theorem *There exists a bijection between constructible Euclidean Alhazen triangles and constructible hyperbolic Alhazen triangles, as well as a bijection between the non-constructible cases.*

Proof. Suppose we can construct any of the hyperbolic isosceles triangles whose legs contain A and B. Construction of the translation of the given circle to the center O is always possible [3]. One way to see this is, the distances AH, BH, length of the radius of the given circle and the angle AHB may all be copied with compass and straightedge with O taking the place of H. When the given circle and its points have been translated to the center of the disk, the hyperbolic and Euclidean centers become the point O and any arcs through H become Euclidean straight. Without loss of generalization, we can consider our given circle centered at O from the start.

Besides being easier to see and construct, our given is now easily transformed through stereographic projection to the Klein model of hyperbolic geometry.

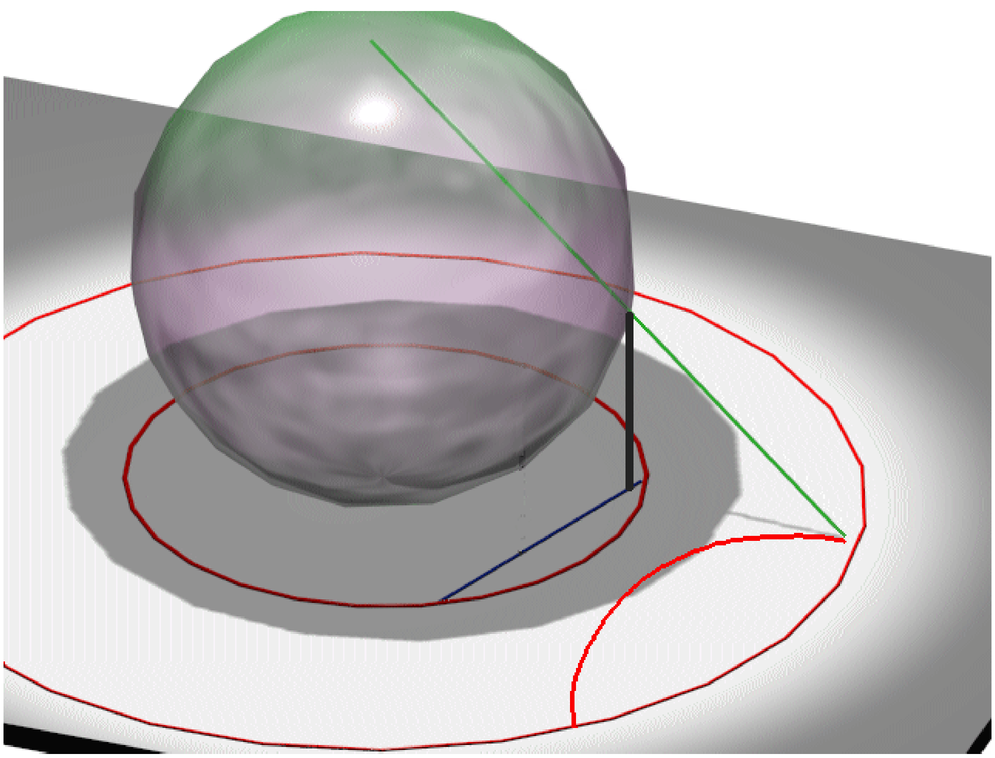

Figure 3. Stereographic projection of a hyperbolic line in both models.

The projection starts with a sphere placed on the hyperbolic disk, tangent at O, the center of the disk. Euclidean segments connecting the top point of this sphere to the points of a hyperbolic line in the Poincaré model pass through the sphere itself. When we project the intersection points straight down, we get Euclidean segments, the chords of the Klein model. Stereographic projection sends circles to circles and preserves congruence and incidence [4]. The legs of the constructed hyperbolic isosceles triangle of the Poincaré disk become Euclidean segments in this model. The given circle transforms to a circle. The projection can be accomplished with compass and straightedge because the vertices of the Poincaré disk triangle remain on the given circle, so we can join them with our ruler to get the Klein model triangle. Then the Euclidean rays $\overrightarrow{OA}$ and $\overrightarrow{OB}$ intersect the Euclidean segments at A' and B', the Euclidean points on the legs.

Now, when we look at this twice-transformed picture without the hyperbolic interpretation, we see an isosceles triangle with the transformed images of points A and B, one on each leg. In other words, our hyperbolic construction yields a Euclidean construction for an Alhazen solution. The process is reversible because translation and stereographic projection are bijections. If we could find any hyperbolic construction, we could then find any Euclidean construction. But the Euclidean solution is, in general, unobtainable. Therefore our initial supposition cannot be true. $\square$

Two examples will illustrate our theorem. First, we recall Figure 1. Our bijection says exactly two non-degenerate triangles must be constructible in hyperbolic geometry. We will start with the given circle centered at O. As before, we place a point V on the given circle and construct the segments $\overline{VW}$ and $\overline{VU}$. Figure 4 illustrates our situation (we have hidden the diameter through A and B for clarity.) The hyperbolic Law of Cosines applies, where r is a radius of the given circle. We get the same conclusion as the Euclidean version: two triangles constructible for the given A and B, and no others possible.

$$\cosh VU - \cosh VW = \cosh^2 r - \sinh^2 r \cos VOU - (\cosh^2 r - \sinh^2 r \cos VOW)$$
$$\cosh VU - \cosh VW = \sinh^2 r(\cos VOW - \cos VOU)$$

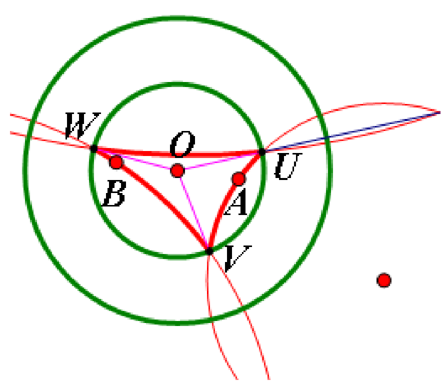

Figure 4. Hyperbolic version of Figure 1.

Figures 1 and 4 illustrate a detail from Alhazen's history. Dörrie's concise presentation [2] gives an analytic proof for there being as many as four Alhazen triangles possible, depending on the placement of the given points A and B. Our bijection said there would be exactly two constructible hyperbolic triangles from this special position of A and B because the Euclidean case had two. We constructed the hyperbolic triangle (the second triangle uses the vertex on the other side of the given circle, just like the Euclidean case.) The hyperbolic Law of Cosines verified that our two triangles were the only two.

In Figure 2, where the base of the hyperbolic triangle passed through the hyperbolic center of the given circle, we can show how such triangles turn out after the bijection. The base remains a diameter, but also looks Euclidean straight when translated to the center. We finish the bijection as pictured in Figure 5. Our disk boundary and given circle remain circles. Our Euclidean diameter, which is also the base of our triangle, remains visibly unchanged as well. However, the legs of our triangle go from hyperbolic lines to Euclidean segments as if the end points of the hyperbolic lines were connected with a straightedge. Lastly, we need to transform our image from the Poincaré model to the Klein model by simply removing the disk boundary from the picture. Since stereographic projection preserves incidence and congruence, our image triangle is an isosceles Euclidean triangle with its two legs passing through points A' and B'.

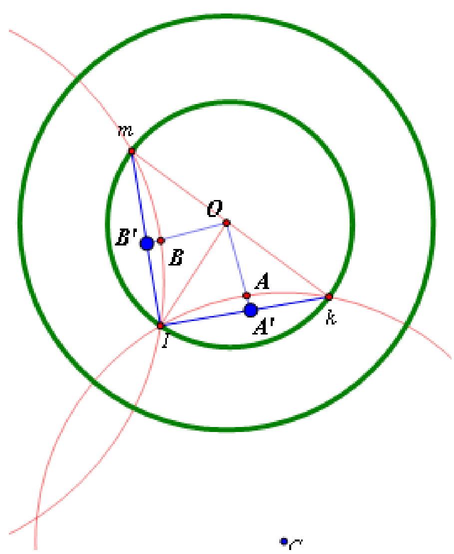

Figure 5. Last step of bijection.

In summary, regarding the hyperbolic Alhazen problem, we are in the same situation as the Euclidean plane: unless we have special information about the locations of A and B, we do not have a construction for an inscribed isosceles triangle. If the Alhazen given information allows a construction in one geometry, there is a corresponding constructible case in the other geometry. Likewise for the non-constructible cases.

3 Acknowledgement

We must thank the Mohler-Thompson Fund for sponsoring our research. The referee's suggestions transformed our mish-mash of ideas into an article: a valuable contribution.

3.1 References

1. H. Riede, Reflexion am Kugelspiegel. Oder: das Problem des Alhazen, *Praxis der Mathematik*, **31**, 1989.
2. H. Dörrie, *100 Great Problems of Elementary Mathematics: Their History and Solution,* (Tr. D. Antin), Dover, 1965.
3. C. Goodman-Strauss, Compass and straightedge in the Poincaré disk, *American Mathematical Monthly*, 108 **1**, January 2001.
4. H. S. M. Coxeter, *Introduction to Geometry*, John Wiley and Sons, Inc., 1961.
5. P. M. Neumann, Reflections on reflection in a spherical mirror, *American Mathematical Monthly,* 105 **1**, 1998.