A Hyperbolic View of the Seven Circles Theorem

Kostiantyn Drach and Richard Evan Schwartz *

October 4, 2019

1 Introduction

Cecil John Alvin Evelyn, or simply "Jack" to friends, was born in 1904 in the United Kingdom in the aristocratic Evelyn family.

Figure 1: Cecil John Alvin Evelyn

A true "gentleman of leisure", he had hobbies rather than jobs. Among those hobbies was a genuine passion for elementary geometry. Jack and some friends, also gentlemen of leisure, often spent time in a cafe hand-plotting various lines and circles on large sheets of paper in pursuit of new configuration theorems. These plots, which today might be routine manipulations

with modern geometry software, back then were acts of scientific inquiry. One result of these meetings was a self-published book "The Seven Circles Theorem and other new theorems", [**EMCT**]. (He also co-authored several papers in number theory; see the bibliography in [**Tyr**].) This book did indeed contain a new theorem about a configuration of touching circles, the Seven Circles Theorem. This theorem reads as follows:

Theorem 1.1 *For every chain $H_1, ..., H_6$ of consequently touching circles inscribed in and touching the unit circle the three main diagonals of the hexagon comprised of the points at which the chain touches the unit circle intersect at a common point.*

Figure 2 shows the Seven Circles Theorem in action.
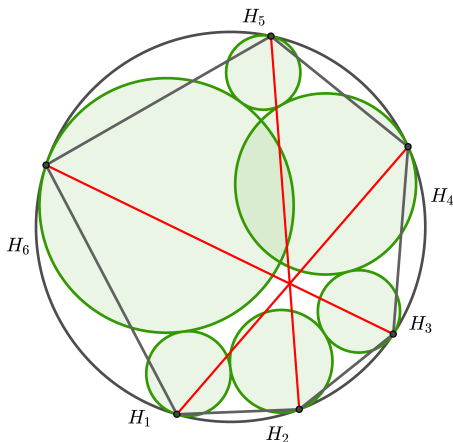
Figure 2: The Seven Circles Theorem

There are several proofs of this result. See, for instance [**Cu**], [**EMCT**], or [**Ra**]. We noticed that the Seven Circles Theorem fits naturally into the setting of hyperbolic geometry because everything in sight takes place inside the unit disk, and the open unit disk is a common model for the hyperbolic plane. What is interesting is that actually the open unit disk is a model for hyperbolic plane in *two* ways, as the Klein model and as the Poincaré model. The key to decoding the Seven Circles Theorem is understanding the conversion between these two models. In this note, we will explain the connection between the Seven Circles Theorem and hyperbolic geometry, then prove a stronger result about hyperbolic geometry hexagons which implies the Seven Circles Theorem as a special case.

2 Hyperbolic Geometry without Distances

On the simplest level, the hyperbolic plane is a system of points and lines which satifies the first 4 of Euclid's axioms and not the 5th axiom – the parallel postulate. About 180 years ago, Bolyai, Gauss, and Lobachevsky all discovered that one really can make a system like this. Here we discuss 3 models for the hyperbolic plane. The first two models are quite common and the third one is specially introduced in order to better examine the relationship between the first two models. Let Δ be the open unit disk $x^2 + y^2 < 1$ in the plane and let Σ denote the open northern hemisphere in the unit sphere $x^2 + y^2 + z^2 = 1$.

The Klein model: In this model, the *points* are the points of Δ and the *lines* are the intersections of straight lines with Δ.

The Poincaré model: In this model, the *points* are the points of Δ, and the *lines* are the intersections of circles with Δ, provided that the circle intersects the boundary of Δ at right angles.

The Hemisphere model: In this model, the *points* are the points of Σ and the *lines* are intersections of vertical planes with Σ.

Consider the maps $f : \Sigma \to \Delta$ and $f^{-1} : \Delta \to \Sigma$ given by the formulas

$$f(x, y, z) = (x, y), \qquad f^{-1}(x, y) = (x, y, \sqrt{1 - x^2 - y^2}).$$

Geometrically, the map f is just vertical projection. The map f carries lines in the Hemisphere model to lines in the Klein model, and the map f^{-1} does the reverse. As the notation suggests, f and f^{-1} are inverse maps. So, f and f^{-1} give the conversion between the Klein and Hemisphere models.

Now consider the maps $g : \Sigma \to \Delta$ and $g^{-1} : \Delta \to \Sigma$ given by the formulas

$$g(x, y, z) = \left(\frac{x}{1+z}, \frac{y}{1+z} \right), \quad g^{-1}(x, y) = \frac{1}{1 + x^2 + y^2}(2x, 2y, 1 - x^2 - y^2).$$

Geometrically, the points $(0, 0, -1)$ and (x, y) and $g(x, y)$ are collinear. That is, we get $g(x, y)$ by taking the line through the south pole of the unit sphere and (x, y) and intersecting it with Σ. This map is known as *Stereographic Projection*. So, g and g^{-1} give the conversion between the Poincaré and Hemisphere models.

3

The maps gf^{-1} and fg^{-1} give an equivalence between the Klein and Poincaré models. Rather than write formulas for this maps we explain what they do geometrically. The map from the one model to the other just replaces each line of one kind with the line of the other kind that has the same "endpoints" on the unit circle. What is miraculous about this description is that there are underlying maps, on the level of points, which do the right things to the lines. So, if we have 3 Klein-lines all containing the same *triple point*, as we do in the conclusion of the Seven Circles Theorem, the corresponding Poincaré-lines also intersect in a triple point. With this in mind, we show what Figure 2 looks like when we replace the three relevant Klein-lines with the three relevant Poincaré-lines.
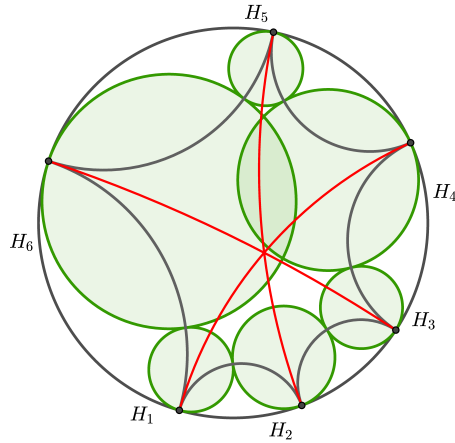
Figure 3: The Seven Circles Theorem translated to the Poincaré model

Let $\boldsymbol{H}^2$ denote the Poincaré model of the hyperbolic plane. We hereby call Poincaré-lines *geodesics*, as is commonly done. Second, we denote the unit circle as $\partial \boldsymbol{H}^2$ and call it the *ideal boundary* of $\boldsymbol{H}^2$, as is commonly done. The points of $\partial \boldsymbol{H}^2$ are known as *ideal points*, even though technically they are not points of $\boldsymbol{H}^2$. Finally, we say that a disk contained in $\boldsymbol{H}^2$, except for a single ideal point, is a *horodisk*. We will explain the geometric significance of such horodisks below. In Figure 3, the 6 points where the horodisks intersect $\partial \boldsymbol{H}^2$ are ideal points. The figure made by connecting the 6 ideal points by geodesics is known as an *ideal hexagon*. This ideal hexagon is shown in black. Here is our first (but not last) reformulation of the Seven Circles Theorem.

Theorem 2.1 *Let P be an ideal hexagon. Suppose that there are horodisks $H_1, ..., H_6$ placed at the ideal vertices of P in such a way that every two consecutive horodisks are tangent. Then the hyperbolic geodesics connecting opposite vertices of P meet at a triple point.*

Once we make the conversion to the Poincaré model, our first reformulation is mostly an exercise in using new terminology. However, after we go more deeply into hyperbolic geometry, we will prove a different reformulation that actually says something new.

3 Hyperbolic Geometry with Distances

Like the Euclidean plane, the hyperbolic plane is not only a system of points and lines but also a *metric space*.

The Hyperbolic Metric: The distance between two points $b, c \in \boldsymbol{H}^2$ is computed as follows. The geodesic containing b and c intersects $\partial \boldsymbol{H}^2$ in points a, d. These points are labeled so that a, b, c, d occur in order. The quantity

$$\text{dist}(b, c) = \log \frac{(a - c)(b - d)}{(a - b)(c - d)}. \tag{1}$$

is known as the *hyperbolic distance* between b and c. Here we are taking advantage of the fact that we can represent points in the plane as complex numbers. Thus $(a - c)(b - d)$ is the product of the complex numbers $a - c$ and $b - d$. Conveniently, the quantity inside the log function is always a real number greater than or equal to 1.

Hyperbolic Isometries: Just as the Euclidean distance function, typically defined in terms of sums of squares, exhibits a surprising rotational symmetry, so does the hyperbolic distance function. A *linear fractional transformation* of the complex plane $\boldsymbol{C}$ is a map of the form

$$T(z) = \frac{az + b}{cz + d}, \qquad ad - bc \neq 0 \tag{2}$$

Such a map is called a *hyperbolic isometry* when both T and T^{-1} map $\boldsymbol{H}^2$ to itself. The reason for the name, in this case, is that T preserves hyperbolic distances: $\text{dist}(T(b), T(c)) = \text{dist}(b, c)$ for all points $b, c \in \boldsymbol{H}^2$. One can verify

this claim with a straightforward algebraic manipulation. One can map any point of $\boldsymbol{H}^2$ to any other point by such a hyperbolic isometry. For instance, the map

$$T(z) = \frac{z - r}{rz - 1}$$

is a hyperbolic isometry which has the property that $T(\pm 1) = \pm 1$ and $T(r) = 0$. The rotations about the origin are also hyperbolic isometries. Using these two kinds of maps, you can probably convince yourself of our claim that one can map any point of $\boldsymbol{H}^2$ to any other point using a hyperbolic isometry.

Hyperbolic Disks: A *hyperbolic disk* in $\boldsymbol{H}^2$ any set of the form

$$B(z_0, r) = \{z | \ \mathrm{dist}(z, z_0) \leq r\}.$$

Here z_0 is a point of $\boldsymbol{H}^2$, called the *hyperbolic center* of the hyperbolic disk. The boundary of the hyperbolic disk is called a *hyperbolic circle*. The boundary of a horodisk is a *horocircle*.

The hyperbolic disks centered at the origin are round Euclidean disks, by symmetry. Also, the hyperbolic isometries map Euclidean disks to Euclidean disks. Since these hyperbolic isometries also preserve the hyperbolic distance, we see that as sets the hyperbolic disks and circles are exactly the same as the Euclidean disks and circles contained in $\boldsymbol{H}^2$. Note, however, that the hyperbolic and Euclidean center of a disk are typically distinct, as are the hyperbolic and Euclidean radii. If we keep the Euclidean size of a disk the same and start shifting it over until it becomes tangent to $\partial \boldsymbol{H}^2$, the hyperbolic center moves to the tangency point and the radius tends to ∞. So, one can view a horodisk as a disk of infinite radius centered at an ideal point. The first two pictures in Figure 4 show hyperbolically concentric circles, and the last picture shows hyperbolically concentric horocircles.
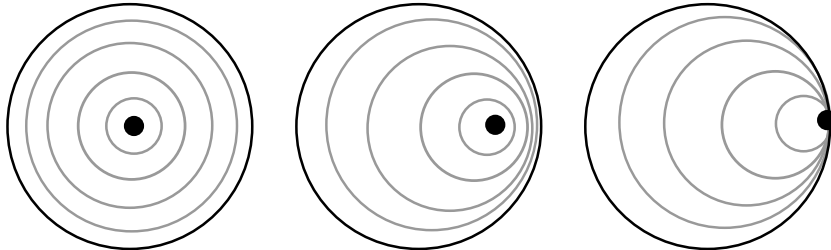
Figure 4: Hyperbolically concentric circles and horocircles.

Alternating Perimeter: A *hyperbolic polygon* has the same definition as a Euclidean polygon except that everything takes place in $\boldsymbol{H}^2$. Namely, it is a closed loop made by connecting together finitely many hyperbolic geodesic segments. The endpoints of these segments are the *vertices* and the segments themselves are the *edges*. A hyperbolic polygon always has a perimeter, namely the sum of the hyperbolic distances between the vertices. Equivalently, we could say that the perimeter is the sum of the lengths of the edges. When the polygon has an even number of sides, it also has an *alternating perimeter*. This is defined to be alternating sum of the lengths of the edges, namely $(S_1 + S_3 + S_5 + ...) - (S_2 + S_4 + S_6 + ...)$.

An *ideal polygon* has the same definition as above, except that the vertices are all ideal points. Surprisingly, the alternating perimeter of an ideal polygon makes sense, even though the individual terms in the sum are infinite. To see this for hexagons (which is the case we care about) imagine that we have a sequence $\{P_n\}$ of hyperbolic hexagons converging to some ideal hexagon P_∞. We place disks at the vertices of P_n, in such a way that consecutive ones do not overlap, and we compute the alternating perimeter instead as

$$(S_1' + S_3' + S_5') - (S_2' + S_4' + S_6') \tag{3}$$

where S_k' is the length of the portion of the kth side of P_n that lies outside the two disks centered at its endpoints. If we replace one of our disks by another one with the same center, we are adding some amount to some S_k' and subtracting the same amount to S_{k+1}'. So, the modified sum does not depend on which disks we choose. Taking a limit as $n \to \infty$, we see that the alternating perimeter of P_∞ can be defined as the same kind of sum as in Equation 3, except that S_k' denotes the (finite!) length of the portion of the kth side of P_∞ that lies outside the two horodisks centered at its endpoints.

3.1   The Main Result

An ideal hexagon P determines a small triangle T_P, the geodesic triangle bounded by the three geodesics connecting opposite sides of P. In Figure 5 below, the red triangle is T_P. Here is the main result.

Theorem 3.1 *For any ideal hexagon P, the alternating perimeter of P is, up to sign, twice the perimeter of T_P.*

7

Proof: Define a *semi-ideal triangle* to be a hyperbolic triangle with two ideal vertices and one vertex in $\boldsymbol{H}^2$. The yellow triangles Y_1, Y_2, Y_3 on the left side in Figure 5 are semi-ideal triangles. The green triangles G_1, G_2, G_3 on the right side in Figure 5 are also semi-ideal triangles. Note that the green triangles overlap. They all contain T_P.
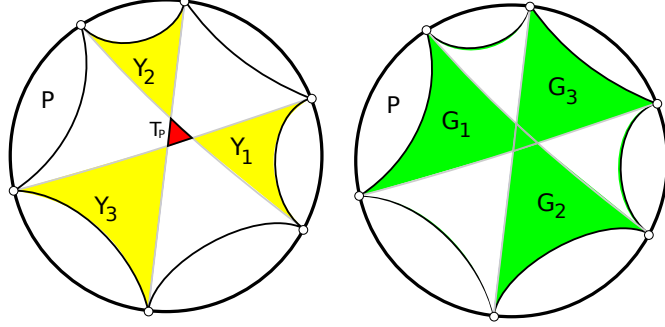
Figure 5: The ideal hexagon P and the small triangle T_P.

Let L_1, L_2, L_3 denote the sides of a semi-ideal triangle V, with the convention that L_3 connects the two ideal vertices of V. We delete disjoint horodisks from the two ideal vertices of V and define

$$A(V) = L_1' + L_2' - L_3', \tag{4}$$

Where L_j' is the length of the portion of L_j outside the two horodisks. This definition is very similar to the definition of the alternating perimeter above, and it does not depend on which horodisks we remove.

Note that G_k and Y_k share a vertex for $k = 1, 2, 3$. We have a hyperbolic isometry I_k such that $I_k(Y_k) = G_k$. This is most easily seen if we normalize the picture so that the vertex $Y_k \cap G_k$ is the Euclidean origin; in this case I_K is just reflection in the origin. Hence $A(Y_k) = A(G_k)$. Summing up, we have

$$A(Y_1) + A(Y_2) + A(Y_3) - (A(G_1) + A(G_2) + A(G_3)) = 0. \tag{5}$$

The terms in the sum on the left side of Equation 5 can be divided into two types: Those which come from geodesics connecting ideal points, and the rest. When we sum up the terms of the first kind, we get $A(P)$ (up to sign) When we sum up the rest of the terms, we get twice the perimeter of T_P because, so to speak, the lengths of the rest of the geodesics bounding T_P are counted in pairs with opposite signs. ♠

Now that we have Theorem 3.1, we can give our final reformulation of the Seven Circles Theorem. Say that a hexagon P has *point reflection symmetry* if there is a nontrivial hyperbolic isometry T, which has a single fixed point, such that $T(P) = P$ and T^2 is the identity. We call T a *point reflection*. The map T swaps the opposite vertices of P. Here is our final reformulation of the Seven Circles Theorem.

Theorem 3.2 *The following are equivalent for an ideal hexagon P.*

1. *The geodesics connecting opposite vertices of P meet at a triple point.*

2. *P has point reflection symmetry.*

3. *The alternating perimeter of P is 0.*

Proof: $(1 \to 2)$: Suppose that the first condition holds. The conditions above are unchanged if we move the picture by a hyperbolic isometry. We do this in such a way that the triple point lies at the origin of the unit disk. But then, in the Poincaré model, the three geodesics connecting the opposite vertices of P are all Euclidean diameters of the unit disk. This means that the Euclidean symmetry $z \to -z$ is a symmetry of P. But this particular symmetry is also a hyperbolic isometry. Hence P has point reflection symmetry.

$(2 \to 3)$: We choose the horidisks H_1, H_2, H_3 arbitrarily, then afterwards we set $H_{k+3} = T(H_k)$ for $k = 1, 2, 3$. This makes all of Figure 3 invariant under the point reflection T. But then, referring to Equation 3, we have $S'_{k+3} = S'_k$ for $k = 1, 2, 3$. Hence the sum in Equation 3 is 0.

$(3 \to 1)$: If P has alternating perimeter 0 then, by Theorem 3.1, the triangle T_P has perimeter 0. This means that T_P is actually a single point.
♠

If we have an ideal hexagon as in Theorem 2.1, then its alternating perimeter is 0 because in Equation 3 we gave $S'_1 = ... = S'_6 = 0$. So, Theorem 3.2 implies Theorem 2.1, and hence the Seven Circles theorem. Theorem 3.2 also reveals that all the instances of the Seven Circles Theorem involve ideal hexagons having point reflection symmetry, and this makes the Seven Circles Theorem obvious.