The Arithmetics of the Hyperbolic Plane
Talk 1: Hyperbolic Geometry

Valentino Delle Rose

Introduction

These notes are a short introduction to the geometry of the hyperbolic plane.

We will start by building the **upper half-plane model** of the hyperbolic geometry. Here and in the continuation, a *model* of a certain geometry is simply a space including the notions of point and straight line in which the axioms of that geometry hold.

Then we will describe the **hyperbolic isometries**, i.e. the class of trasformations preserving the hyperbolic distance, and the **geodesics**, that are the shortest paths connecting two point in the hyperbolic plane.

After a brief introduction to the **Poincaré disk model**, we will talk about **geodesic triangle** and we will give a **classification of the hyperbolic isometries**.

In the end, we will explain **why the hyperbolic geometry is an example of a non-Euclidean geometry**.

For more details (and for the missing proofs) see

> R.E. Schartz, *Mostly surfaces*, A.M.S. Student library series, Volume 60.
>
> **https://www.mathbrown.edu/ res/Papers/surfacebook.pdf**

1 The Upper Half-Plane Model

We start defining the hyperbolic plane as a set of points with a metric.

Definition 1.1. We call *upper half-plane* the set $U = \{z \in \mathbb{C} \mid Im(z) > 0\}$.

Definition 1.2. Let V be a real vector space. An *inner product* on V is a map

$$\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{R}$$

such that:

1. $\langle av + w, x\rangle = a\langle v, x\rangle + \langle w, x\rangle \quad \forall v, w, x \in V, \ \forall a \in \mathbb{R}$;
2. $\langle x, y\rangle = \langle y, x\rangle \quad \forall x, y \in V$
3. $\langle x, x\rangle \geq 0 \quad \forall x \in V \quad and \quad \langle x, x\rangle = 0 \Leftrightarrow x = 0$

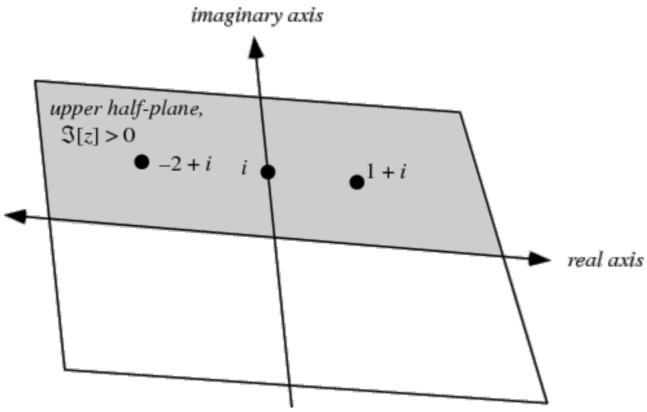

Figure 1: the upper half-plane

Let $z = x + iy \in \mathbb{C}$. At the point z, we introduce the inner product

$$\langle v,\, w\rangle_z = \frac{1}{y^2}\,(v \cdot w)\,,$$

where $\cdot$ is the usual dot product. We mean to apply this inner product to vectors v and w "based" at z.

Now we can define the *hyperbolic norm* as the norm induced by the inner product $\langle\,\cdot\,,\,\cdot\,\rangle_z$, i.e.

$$\|v\|_z = \sqrt{\langle v,\, v\rangle_z}.$$

Definition 1.3. Let $\gamma\,:\,[a,b] \longrightarrow U$ be a differentiable curve. The *length* of γ is

$$L\,(\gamma) = \int_a^b \|\gamma'\,(t)\|_{\gamma(t)}\,dt = \int_a^b \frac{|\gamma'\,(t)|}{Im\,(\gamma\,(t))}\,dt.$$

Example 1.4. Consider the curve $\gamma\,:\,\mathbb{R} \longrightarrow U$ defined by

$$\gamma\,(t) = ie^t.$$

Let $a < b$. Then the length of the portion of γ connecting $\gamma\,(a)$ and $\gamma\,(b)$ is given by

$$\int_a^b \frac{|ie^|}{e^t}\,dt = \int_a^b dt = b - a.$$

Note that the image of γ is an open vertical ray, but the formula tells us that this ray, measured hyperbolically, is infinite in both directions.

Definition 1.5. 1. Let $p, q \in U$. The *hyperbolical distance* between p and q is

$$d\,(p,\,q) = \inf_{\gamma \in \Gamma_{p,q}} L\,(\gamma)\,,$$

where $\Gamma_{p,q}$ is the set of the differentiable curves connecting p and q.

2. The pair $(U,\,d) = \mathbb{H}^2$ is called the *hyperbolic plane*.

3. The *angle* between two differentiable and regular curves in $\mathbb{H}^2$ is defined to be the ordinary Euclidean angle between them, i.e. the Euclidean angle between the two tangent vectors at the point of intersections. (That means: in the upper half-plane model of the hyperbolic geometry, the distances are distorted from the Euclidean model, but the angles are not).

4. The *hyperbolic area* of a region $D \subset \mathbb{H}^2$ is defined by

$$\int_D \frac{dxdy}{y^2}.$$

2 Hyperbolic Isometries

Definition 2.1. Let $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. A *complex linear fractional transformation* (or *Möbius transformation*) is a map $T_A : \hat{\mathbb{C}} \longrightarrow \hat{\mathbb{C}}$ such that

$$T_A(z) = \begin{cases} \frac{az+b}{cz+d} & z \in \mathbb{C} \setminus \left\{-\frac{d}{c}\right\} \\ \infty & z = -\frac{d}{c} \\ \frac{a}{c} & z = \infty \end{cases}$$

where $\begin{bmatrix} a & b \\ c & d \end{bmatrix} = A \in SL_2(\mathbb{C}) = \{A \in \mathrm{M}_2 \mid det(A) = 1\}$.

By direct calculation, it is easy to prove that

$$T_{AB} = T_A \circ T_B$$

where $A, B \in SL_2(\mathbb{C})$. In particular (since $detA = 1$ and then A^{-1} exists) we have the existence of $T_A^{-1} = T_{A^{-1}}$. We now focus on a special kind of Möbius transformation.

Definition 2.2. We call *real linear fractional transformation* a Möbius transformation T_A such that $A \in SL_2(\mathbb{C})$.

Theorem 2.3. *Let T be a real linear fractional transformation. Then T preserves $\mathbb{H}^2$, i.e. $T(\mathbb{H}^2) = \mathbb{H}^2$.*

(To prove this statement, just check that $Im(z) > 0$ implies $Im(T(z)) > 0$.)

Definition 2.4. We say taht a real linear fractional transformation is *basic* if it has one of the following forms:

1. $T_b(z) = z + b$ (*translations*);

2. $R_r(z) = rz$ (*homothetis*);

3. $I(z) = -\frac{1}{z}$ (*circular inversion*).

It is easy to prove the following

Theorem 2.5. *Let T be a real linear fractional transformation. Then T is a composition of basic ones.*

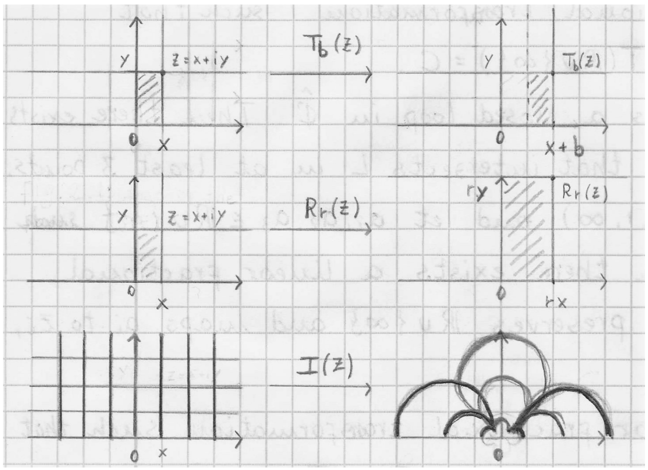

Figure 2: basic real linear fractional transformation

Now we want to prove that every real linear fractional transformation is a *hyperbolic isometry*, i.e. preserves the hyperbolic metric. Obviusly, translations $T_b(z) = z+b$ are isometries. Just by using the definition of length of a curve (1), one can easily check that $R_r(z)$ and $I(z)$ are also hyperbolic isometries. Then, by 2.5 follows immediately

Theorem 2.6. *Any real linear fractional transformation is a hyperbolic isometry.*

3 Geodesics

Our next goal is to describe the shortest curves connecting two points in $\mathbb{H}^2$, i.e. the *geodesics* in $\mathbb{H}^2$.

To do this, we need first to prove the *circle preserving property* of the Möbius transformations.

Definition 3.1. A *generalized circle* in $\hat{\mathbb{C}}$ in either a circle in $\mathbb{C}$ or a set $L \cup \{\infty\}$, where L is a straight line in $\mathbb{C}$.

We will use the following properties (given without proofs):

Lemma 3.2.

1. *Let be $C \subset \hat{\mathbb{C}}$ any generalized circle. Then there exists a Möbius transformation T such that*

$$T(\mathbb{R} \cup \{\infty\}) = C.$$

2. *Suppose that L is a closed loop in $\hat{\mathbb{C}}$. Then there exists a generalized circle C that intersects L in at least 3 points.*

3. *Let $(z_1, z_2, z_3) = (0, 1, \infty)$ and let $a_1, a_2, a_3 \in \mathbb{R} \cup \{\infty\}$ such that $a_1 \neq a_2 \neq a_3$. Then there exists a Möbius transformation that preserves $\mathbb{R} \cup \{\infty\}$ and maps a_i to z_i, $i = 1, 2, 3$.*

4. *Let T be a Möbius transformation such that $T(0) = 0$, $T(1) = 1$ and $T(\infty) = \infty$. Then T is the identity map.*

Theorem 3.3 (circle preserving property)**.** *Let C be a generalized circle and let T be a Möbius transformation. Then $T(C)$ is also a generalized circle.*

Proof. Suppose that there exist a Möbius transformation T and a generalized circle C such that $T(C)$ is not a generalized circle. By lemma 3.2 1., we can assume $C = \mathbb{R} \cup \{\infty\}$. By lemma 3.2 2., there is a generalized circle D such that $\exists p_1, p_2, p_3 \in D \cap T(\mathbb{R} \cup \{\infty\})$ and $p_1 \neq p_2 \neq p_3$. Again by lemma 3.2 1., there exists a Möbius transformation S such that $S(\mathbb{R} \cup \{\infty\}) = D$. There are points $a_1, a_2, a_3 \in \mathbb{R} \cup \{\infty\}$ such that $S(a_i) = p_i$, for $i = 1, 2, 3$. Also, there are points $b_1, b_2, b_3 \in \mathbb{R} \cup \{\infty\}$ such that $T(a_i) = p_i$, for $i = 1, 2, 3$. By lemma 3.2 3., there are two Möbius transformations A and B such that both preserve $\mathbb{R} \cup \{\infty\}$ and $A(a_i) = B(b_i) = z_i$, where $i = 1, 2, 3$ and $(z_1, z_2, z_3) = (0, 1, \infty)$. Now we have $T \circ B^{-1}(z_i) = S \circ A^{-1}(z_i) = p_i$, $i = 1, 2, 3$. Hence, by lemma 3.2 4., $T \circ B^{-1} \equiv S \circ A^{-1}$. But now we have $T \circ B^{-1}(\mathbb{R} \cup \{\infty\}) = T(\mathbb{R} \cup \{\infty\})$ that, for assumption, is not a generalized circle, while $S \circ A^{-1}(\mathbb{R} \cup \{\infty\}) = S(\mathbb{R} \cup \{\infty\}) = D$ is a generalized circle. This is a contradiction. □

Back to the geodesics. We first consider the case of points that lie on the imaginary axis.

Lemma 3.4. *Let $p, q \in \mathbb{H}^2$ being points that lie on the imaginary axis. Then the portion of the imaginary axis connecting p and q is the unique shortest curve in $\mathbb{H}^2$ that connects p and q.*

(To prove this lemma, it is sufficient to remember example 1.4 and consider the map $F(x + iy) = iy$).

Remark 3.5. It follows from symmetry that the vertical rays in $\mathbb{H}^2$ are all geodesics: a vertical ray is the unique shortest path in $\mathbb{H}^2$ connecting any pair of points on that ray.

Moreover one can prove, with a simple (but long and boring...) calculation, the following

Lemma 3.6. *Let $p, q \in \mathbb{H}^2$. Then there is a real linear fractional transformation that carries p and q to points that lie in the same vertical ray.*

Now, we are ready to prove the general case.

Theorem 3.7. *Any two distinct points in $\mathbb{H}^2$ can be joined by a unique shortest path. This path is either a vertical line segment or else an arc of a circle that is centered on the real axis.*

Proof. We have already proved (lemma 3.4) this result for points that lie on the same vertical ray. So, in light of lemma 3.6, it suffices to prove that the image of a vertical ray ρ under a real linear fractional transformation T is one of the two kinds of curves described in the theorem.

From theorem 3.3, we know that $T(\rho)$ is an arc of a (generalized) circle. Since T preserves $\mathbb{R} \cup \{\infty\}$, both endpoints of this arc of a circle lie on $\mathbb{R} \cup \{\infty\}$. Finally, since T preserves the angles, $T(\rho)$ meets $\mathbb{R}$ at the right angles, at any point where $T(\rho)$ intersects $\mathbb{R}$. If $T(\rho)$ limits on ∞, then $T(\rho)$ is another vertical ray. Otherwise, $T(\rho)$ is a semicircle, contained in a circle that is centered on the real axis. □

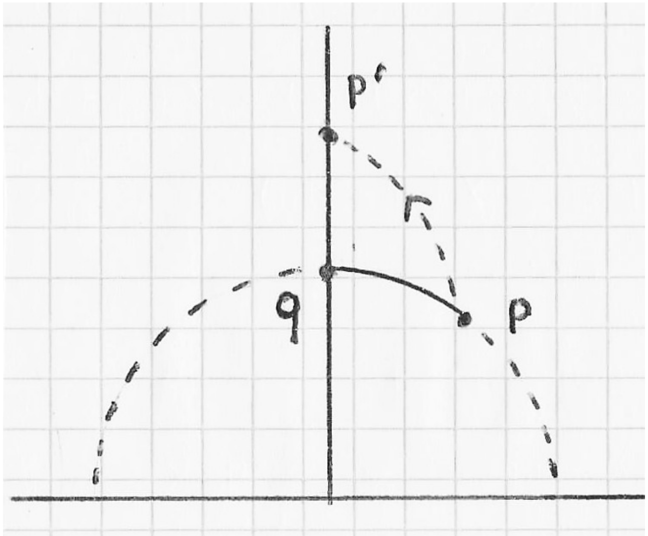

Figure 3: geodesics in the upper half-plane

4 The Disk Model

Now we want to introduce another model for the hyperbolic geometry, that is more convenient to draw pictures. Let Δ be the open unit disk. Then we have the map $M : \mathbb{H}^2 \longrightarrow \Delta$ defined by

$$M(z) = \frac{z-i}{z+i}.$$

Note that if $z \in \mathbb{H}^2$, then $|z+i| > |z-i|$, hence $M(z) < 1$. Since M maps circle to circle and preserves angles, M maps geodesics in $\mathbb{H}^2$ to circular arcs in Δ that meet the unit circle at right angles.

To make Δ a proper model of $\mathbb{H}^2$, we need to give to Δ a metric that makes M an isometry. Let $z \in \Delta$. Then we define the inner product

$$\langle v, w \rangle_z = \frac{4(v \cdot w)}{(1-|z|^2)^2}.$$

Once we have this inner product, we can define lengths of curves and the hyperbolic distance in Δ as in the upper half-plane model. Now, a simple calculation (it suffices to prove that $\langle v, w \rangle_z = \langle dM(v), dM(w) \rangle_{M(z)}$) proves the following

Theorem 4.1. *$M : \mathbb{H}^2 \longrightarrow \Delta$ given by $M(z) = \frac{z-i}{z+i}$ is an isometry.*

Definition 4.2. The pair (Δ, d), where d is the hyperbolic distance on Δ, si called the *Poincaré disk model* of the hyperbolic plane.

Remark 4.3.

1. Let T be a real linear fractional transformation. Then the map $M \circ T \circ M^{-1}$ is an isometry of Δ.

2. Since M preserves angles, the hyperbolic angle between two curves in Δ is the same as the Euclidean angle between them. So, in both our models, Euclidean and hyperbolic angles coincide.

5 Geodesic Triangles

Definition 5.1. The *ideal boundary* of $\mathbb{H}^2$ is

$$\partial\,\mathbb{H}^2 = \begin{cases} \mathbb{R}\cup\{\infty\} & \text{in the upper half-plane model} \\ S^1 & \text{in the disk model} \end{cases}.$$

Points $p \in \partial\,\mathbb{H}^2$ are called *ideal points.*

Definition 5.2. A *geodesic triangle* in $\mathbb{H}^2$ is a simple closed path made from 3 geodesic segments.

Some of the "vertexes" of the triangle are allowed to be ideal points: such vertexes are called *ideal vertexes.* The interior angle at an ideal vertex is 0: the two geodesics both meet the ideal point perpendicular to the ideal boundary.

We call an *ideal triangle* a geodesic triangle having 3 infinite geodesic sides and 3 ideal vertexes.

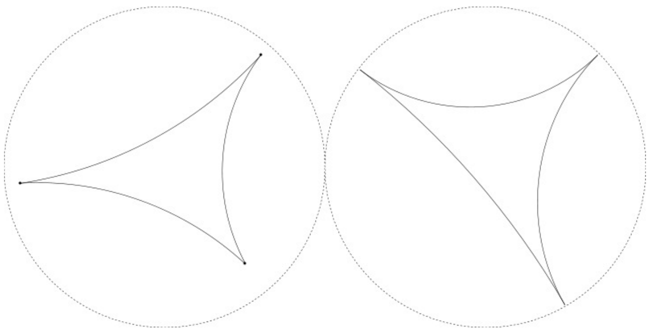

Figure 4: triangle and ideal triangle on Δ

We want to prove the following

Theorem 5.3. *Let T be a geodesic triangle in the hyperbolic plane. Then the area of T is $\pi - X$, where X is the sum of the interior angles of T. In particular, $X < \pi$.*

We need first some preliminary results:

Lemma 5.4. *Theorem 5.3 holds for ideal triangles.*

(To prove this lemma, it suffices to take the triangle T in the upper half-plane model with vertexes in 1, -1, ∞ and check that $Area\,(T) = \int_T \frac{dxdy}{y^2}$.)

Lemma 5.5. *Let $T\,(\theta)$ be a geodesic triangle having two ideal vertexes and one interior vertex with an interior angle θ. Then theorem 5.3 holds for $T\,(\theta)$.*

(The idea is to prove by induction that $f\,(r\pi) = \pi - Area\,(T\,(r\pi)) = r\pi \quad \forall r \in \mathbb{Q}\cap(0,1)$. Then, since f is continuous, $f\,(\theta) = \theta \quad \forall\theta \in [0,\pi]$).

Then theorem 5.3 follows immediately: we can extend the sides of any geodesic triangle T so that they hit $\partial\,\mathbb{H}^2$, and we get an ideal triangle $\bar{T}$ and three triangles $T\,(\alpha)$, $T\,(\beta)$ and $T\,(\gamma)$. By lemmas 5.4 and 5.5 we have

$$Area\,(T) = \pi - \alpha - \beta - \gamma$$

as desired.

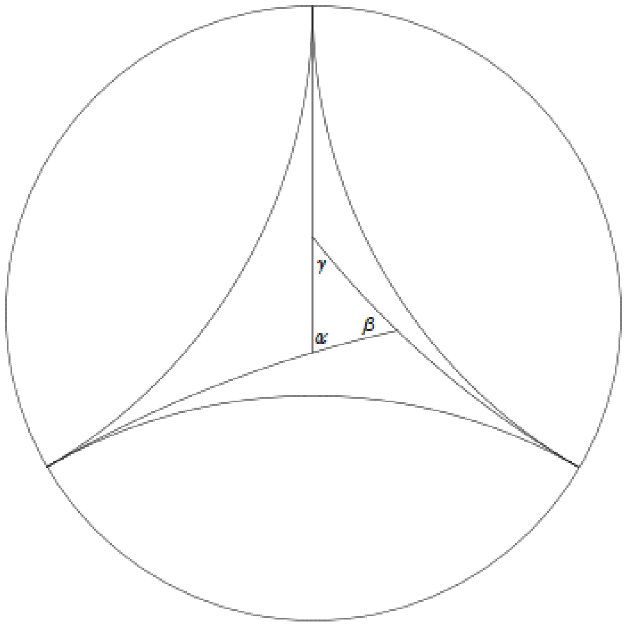

Figure 5: geometric construction to prove theorem 5.3

6 Classification of Isometries

Let T be a real linear fractional transformation:

$$T(z) = \frac{az + b}{cz + d}.$$

If $T(\infty) = \infty$, then $T(Z) = az + b$. If $T(\infty) \neq \infty$, then the equation $T(z) = z$ leads to a quadratic equation. The solutions of this equation are the fixed points of T. So, if T is not the identity map, there are 3 possibilities:

1. T fixes one point in $\mathbb{H}^2$, and no other points. In this case, T is called *elliptic*.
2. T fixes no points in $\mathbb{H}^2$ and one point in $\mathbb{R} \cup \{\infty\}$. In this case, T is called *parabolic*.
3. T fixes no points in $\mathbb{H}^2$ and two point in $\mathbb{R} \cup \{\infty\}$. In this case, T is called *hyperbolic*.

Let g and T be hyperbolic isometries and define

$$S = g \circ T \circ g^{-1}$$

a *conjugate* of T. Note that g maps the fixed point of T onto the fixed points of S.

Suppose T is elliptic. Working in the disk model, we can conjugate T so that the result fixes the origin. It also maps geodesic through the origin to geodesics through the origin and preserves distances along these geodesics. So, in the disk model, all the elliptic isometries are conjugate to ordinary rotations.

Suppose T is parabolic. Working in the upper half-plane model, we can conjugate T so that the result fixes ∞. Since it cannot fixes any other point, the conjugate of T must be a translation $T_b(z) = z + b$.

Suppose T is hyperbolic. Working in the upper half-plane model, we can conjugate T so that the result fixes 0 and ∞. Then it must be a homothety $R_r(z) = rz$.

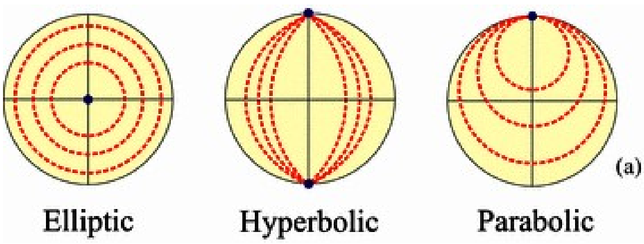

Figure 6: isometries of Δ

7 Hyperbolic Geometry as an Example of Non-Euclidean Geometry

We call *non-Euclidean geometry* a geometry in which the *parallel postulate* does not hold. Recall that the parallel postulate say:

> If a line segment intersects two straight lines forming two interior angles on the same side that sum less than two right angles, then the two straight lines meet on that side on which the angles sum to less then two right angles.

The parallel postulate is clearly equivalent to the statement:

> the sum of the angles in every triangle is π.

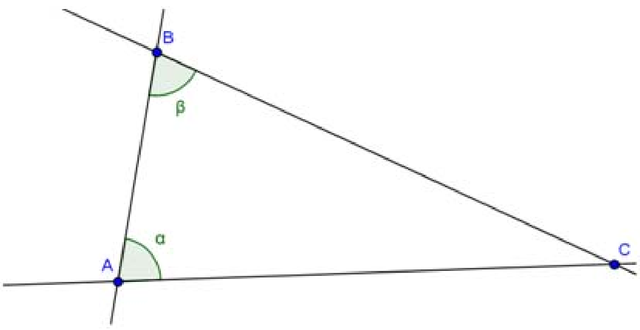

Figure 7: the parallel postulate

We have seen (theorem 5.3) that this statement does not hold in hyperbolic geometry. It means that hyperbolic geometry is a non-Euclidean geometry.

Note that the parallel postulate is also equivalent to the *Playfair´s axiom*, that is the well known statement:

> For any given line R and point p not on R, there is one and only one line R' through p that does not intersects R.

The hyperbolic geometry is built replacing this statement with:

> For any given line R and point p not on R, in the plane containing both line R and point p there are at least two distinct lines through p that does not intersect R.

This implies that there are through p an infinity of coplanar lines that does not intersect R.

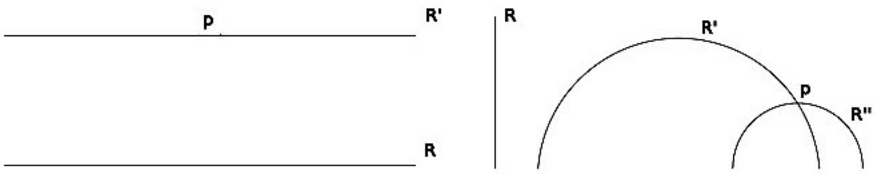

Figure 8: Playfair's axiom vs hyperbolic axiom