HYPERBOLIC GEOMETRY

KRISHNENDU GONGOPADHYAY

1. Introduction

Hyperbolic geometry was created in the first half of the nineteenth century in order to prove the dependence of Euclid's fifth postulate on the first four. Euclid wrote his famous Elements around 300 B. C. In this thirteen volume work he brilliantly organized and presented the fundamental propositions of Greek geometry and number theory. In the first book of elements Euclid develops plane geometry starting with basic geometric terms, five "common notions" concerning with magnitudes and five postulates. In modern language those can be stated as follows.

Let X be a set. $\mathcal{L}$ and $\mathcal{C}$ are sets of certain subsets of X. We call the elements of X as *points*, the elements of $\mathcal{L}$ as *lines* and the elements of $\mathcal{C}$ as *circles*. Then $\mathbb{E} = (X, \mathcal{L}, \mathcal{C})$ satisfies the following postulates.

(1) For all $A, B \in X, A \neq B$ there exists a unique $l \in \mathcal{L}$ such that $A \in l, B \in l$.
(2) Given a $l \in \mathcal{L}$ there exists at least three points which do not belong to l.
(3) For all $l, m \in \mathcal{L}$, we have either $l \cap m =$ a single point or, $l \cap m = \phi$. If $l \cap m = \phi$, then l and m are called *parallel* lines.
(4) For a ordered pair (A, B) of points, there exists a unique $C \in \mathcal{C}$ with *center* A and passing through B.
(5) There is an intuitive notion of angle between two lines. Euclid's classical fourth postulate says that *all right angles are equal*.
(6) (*Modern version of Euclid's classical Fifth postulate*) For all $l \in \mathcal{L}$ and $P \in X$ such that P does not belong to l, there exists a unique $m \in \mathcal{L}$ such that $P \in m$ and l is parallel to m.

In other words Euclid's classical fifth postulate can be stated as follows:

Through a point outside a given line there is one and only one line parallel to the given line.

For two thousand years mathematicians attempted to establish Euclid's fifth postulate from the other simple postulates. In each case one reduced the proof of the fifth postulate to the conjunction of the other postulates with an additional natural postulate which proved to be equivalent to the fifth. As an incidence how much people tried it we note the reference

Date: October 12, 2012.

of Göttingen Mathematician Kästner (1719–1800) who directed a thesis of student Klügel (1739–1812) which considered approximately thirty proof attempts for the parallel postulate!

Decisive progress came in the 19th century, when mathematicians abandoned the effort to find a contradiction in the denial of the fifth postulate and instead workout carefully and completely the consequence of such a denial. It was found that a coherent theory arises if instead one assumes that:

Given a line and a point not on it, there is more than one line going through the given point that is parallel to the given line.

Unusual consequences of this change came to be recognized as fundamental and surprising properties of non-Euclidean geometry: geodesics were not straight lines, but curved; the sum of the angles of a triangle were not equal to π and so forth.

History has associated five names with this enterprise, those of three professionals and two amateurs. The amateurs were jurist Schweikart and his nephew Taurins. The professionals were Carl Freidrich Gauss (1777–1855), Nikolai Lobachevskii (1793–1856) and Johann Bolyai (1802-1860).

Gauss began his meditation on the theory of parallels about 1792. After trying to prove the fifth postulate over twenty years, Gauss discovered that the denial of the fifth postulate leads to a new strange geometry which he called 'non-Euclidean geometry'. After investing its properties for over ten years and discovering no inconsistencies, Gauss was fully convinced of its consistency. In a letter to F. A. Taurinus in 1824, he wrote:

"*The assumption that the sum of three angles of a triangle is smaller that 180 degrees leads to a geometry which is quiet different from our (Euclidean) geometry, but which is in itself completely consistence.* "

Gauss's assumption that the sum of the angles of a triangle is less than 180 degrees is equivalent to the denial of Euclid's fifth postulate. Unfortunately, Gauss never published his results on non-Euclidean geometry.

Only a few years passed before non-Euclidean geometry was rediscovered independently by Nikolai Lobachevsky and J. Bolyai. Lobachevsky published the first account of non-Euclidean geometry in 1829 in a paper entitled "*on the principles of geometry*". A few years later, in 1932, Bolyai published an independent account of non-Euclidean geometry in a paper entitled "*the absolute science in space*".

Gauss, Bolyai and Lobachevskii developed non-Euclidean geometry axiomatically on a synthetic basis. They didn't prove the consistency of their geometries. The basis necessary for an analytic study of hyperbolic non-Euclidean geometry was laid by Leonhard Euler, Gaspard Monge, and Gauss in their investigation of curved surfaces. Later on many people tried to find an analytic model of hyperbolic geometry where these five postulates can be proved and mathematicians found several of them. We shall consider in this exposition two of the most famous analytic models of the hyperbolic geometry which are known as Poincaré models in the name of its inventor, the one and only H. Poincaré.

Models serve primarily a logical purpose. They are useful when exploring the geometric properties of the hyperbolic plane. The same object may 'look' differently in each of the models, but its geometric properties (such as lengths, angles, area) will be the same. In the Poincaré disk model, a *hyperbolic line* is an arc of a circle that is orthogonal to the unit (boundary) circle. In the upper-half plane model, a *hyperbolic line* is a semicircle with center on the x-axis. We can choose any model and can define 'points', 'lines', 'distance' and 'angles' as anything we want. If we can prove that the relations that exist among these 'points', 'lines', 'distance' and 'angles' satisfy all the axioms of the hyperbolic geometry, then we have a model of the hyperbolic plane. Using the respective *hyperbolic distance* in each model, we identify these models up to *isometry*. This gives us a coherent approach to hyperbolic geometry independent of the choice of a model.

2. Inversion in a circle

2.1. **Reflections in $\mathbb{C}$.** Identify $\mathbb{C}$ with the Euclidean space. Let $\Im$ denote the imaginary axis. A complex number $z = x + iy$ represents a vector (x, y) in $\mathbb{R}^2$. For $z = x_1 + iy_1,\ w = x_2 + iy_2$, the Euclidean inner product on $\mathbb{R}^2$ is given by

$$\langle z, w\rangle_e = x_1x_2 + y_1y_2.$$

The Euclidean distance is given by $d(z, w) = |z-w|$. Two vectors $z = (x_1, y_1)$ and $w = (x_2, y_2)$ are *orthogonal* if $\langle z, w\rangle_e = 0$. A *line* on $\mathbb{R}^2$ is an one-dimensional vector subspace spanned by a vector. The notion of orthogonality may be extended naturally to define *orthogonal lines.*

Recall that a line in $\mathbb{R}^2$ is given by the equation:

$$l : ax + by + c = 0,\ (a, b) \neq (0, 0).$$

Writing $x = \frac{1}{2}(z + \bar{z})$, $y = \frac{1}{2i}(z - \bar{z})$, the line l is represented by the complex numbers as

$$l : \bar{\alpha}z + \alpha\bar{z} + c = 0,$$

where $\alpha = \frac{1}{2}(a + ib)$, i.e. $|\alpha| \neq 0$. Thus the complex number α corresponds to the normal vector $v_l = (a, b)$ to the line l. This normal vector is determined up to a real multiple.

Definition 2.1. Let l be a line in $\mathbb{C}$. A map $\sigma_l : \mathbb{C} \to \mathbb{C}$ is said to be *the reflection in* l if

(i) $\sigma_l(z) = z$ whenever $z \in l$.
(ii) If z does not belong to l, then the segment joining z to $\sigma_l(z)$ is orthogonal to l and is bisected by l.

Proposition 2.2. Given a line $l : \bar{\alpha}z + \alpha\bar{z} + c = 0,\ |\alpha| \neq 0$, the reflection in l, $\sigma_l : \mathbb{C} \to \mathbb{C}$ is given by

$$\sigma_l(z) = \frac{-\alpha}{\bar{\alpha}}\bar{z} - \frac{c}{\bar{\alpha}}.$$

Proof. Let $z_0 \in l$. Then

$$\sigma_l(z_0) = \frac{-\alpha}{\bar{\alpha}}\bar{z}_0 - \frac{c}{\bar{\alpha}} = \frac{-\alpha\bar{z}_0 - c}{\bar{\alpha}} = z_0,$$

since $\bar{\alpha}z_0 = -\alpha\bar{z}_0 - c$. If z_0 does not belong to l, then $\sigma_l(z_0) \neq z_0$ and we have

$$z_0 - \sigma_l(z_0) = z_0 + \frac{\alpha}{\bar{\alpha}}z_0 + \frac{c}{\bar{\alpha}} = \frac{\bar{\alpha}z_0 + \alpha\bar{z}_0 + c}{\bar{\alpha}} = \alpha\left\{\frac{1}{\alpha\bar{\alpha}}(\bar{\alpha}z_0 + \alpha\bar{z}_0 + c)\right\}.$$

Thus $z_0 - \sigma_l(z_0)$ is a non-zero real multiple of α. Since α corresponds to the normal vector to the line l, $z_0 - \sigma_l(z_0)$ is orthogonal to l. Hence the line joining z_0 to $\sigma_l(z_0)$ is orthogonal to l.

The mid-point of the segment joining z_0 to $\sigma_l(z_0)$ is

$$\xi = \frac{1}{2}(z_0 - \frac{\alpha}{\bar{\alpha}}\bar{z}_0 - \frac{c}{\bar{\alpha}}) = \frac{1}{2\bar{\alpha}}(\bar{\alpha}z_0 - \alpha\bar{z}_0 - c).$$

The following observation implies that ξ is a point on l:

$$\alpha\bar{\xi} + \bar{\alpha}\xi + c = \frac{1}{2}(\bar{\alpha}z_0 - \alpha\bar{z}_0 - c) + \frac{1}{2}(\alpha\bar{z}_0 - \bar{\alpha}z_0 - c) + c = 0.$$

Hence σ_l is indeed the reflection in the line l. □

Remark 2.3. The above proposition gives us a mnemonic rule for determining the formula of a reflection σ_l in a line l in terms of complex variables. First write down the complex equation of the line $l : \bar{\alpha}z + \alpha\bar{z} + c = 0$. Solve this equation for z. We get after solving :

$$z = \frac{-\alpha}{\bar{\alpha}}\bar{z} - \frac{c}{\bar{\alpha}}.$$

Now replacing the equality by an arrow we get the required formula for σ_l as in the proposition.

2.2. **Inversion in a circle.** Let $S(a, r)$ be the circle in $\mathbb{C}$ given by:

$$S(a, r) = \{z \in \mathbb{C} : |z - a| = r\}$$

Definition 2.4. An Inversion σ of $\mathbb{C}$ with respect to the circle $S(a, r)$ is defined by

$$\sigma(z) = a + \left(\frac{r}{|z - a|}\right)^2 (z - a).$$

It is clear from the definition that σ satisfies the following two properties:
(i) $\sigma(z) = z$ if and only if z is in $S(a, r)$
(ii) $\sigma^2(z) = z$ for all $z \neq a$ in $\mathbb{C}$.

There is a nice geometric construction of the point $\sigma(z)$. Assume first that z is inside $S(a, r)$. Erect a chord of $S(a, r)$ passing through z perpendicular to the line joining a to z. Let u and v be the endpoints of the chord. Then $\sigma(z)$ is the point z' of intersection of the lines tangent to $S(a, r)$ at the points u and v, see Figure 1. Observe that the right triangles $T(a, z, v)$ and $T(a, v, z')$ are similar. Consequently, we have

$$\frac{|z' - a|}{r} = \frac{r}{|z - a|}.$$

Therefore $z' = \sigma(z)$ as claimed.

Now assume that z is outside $S(a, r)$. Let y be the midpoint of the line segment $[a, z]$ and let C be the circle centered at y of radius $|zy|$. Then C intersects $S(a, r)$ in two points u, v, and $\sigma(z)$ is the point z' of intersection of the line segments $[a, z]$ and $[u, v]$, as in Figure 2.

FIGURE 1. when z is inside S(a, r)

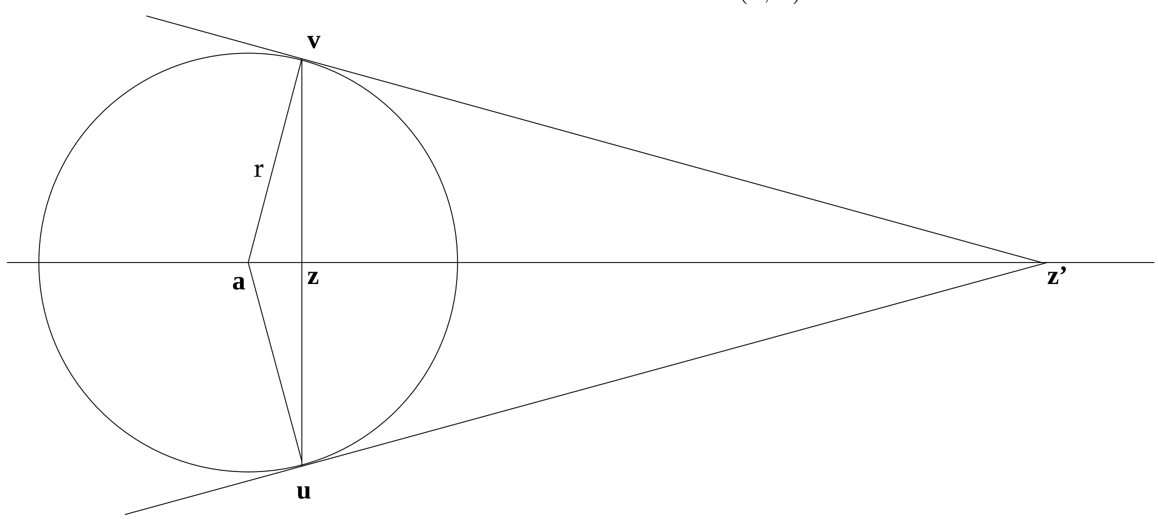
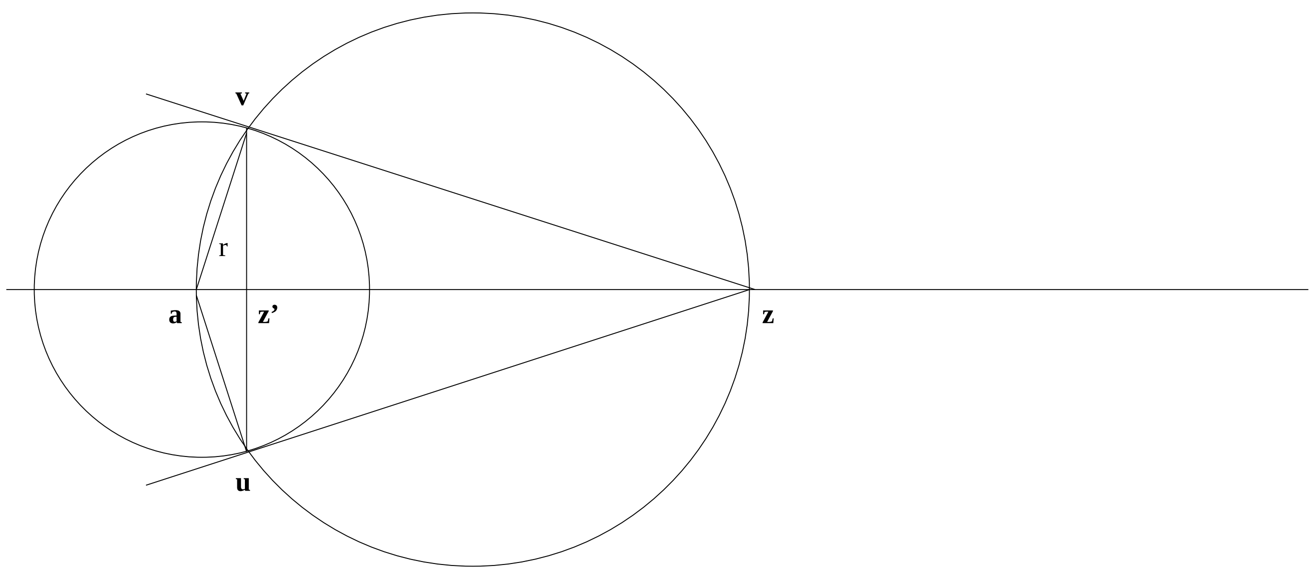


FIGURE 2. when z is outside S(a, r)



The inversion of a point z outside the circle S(a, r)

Let $\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ be the extended complex plane, called the *Riemann sphere*. We can extend σ continuously to $\widehat{\mathbb{C}}$ by defining $\sigma(a) = \infty$ and $\sigma(\infty) = a$. The extension of σ to $\widehat{\mathbb{C}}$ is denoted by σ again. The map σ is a homeomorphism of $\widehat{\mathbb{C}}$. An inversion with respect to a circle $S(a, r)$ sometimes also refer to as the reflection of $\widehat{\mathbb{C}}$ in the circle $S(a, r)$.

Lemma 2.5. *If σ is a reflection of $\widehat{\mathbb{C}}$ in the circle $S(a,r)$ and σ_1 is the reflection in $S(0,1)$, and $\phi : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ is defined by $\phi(z) = a + rz$, then $\sigma(z) = \phi\sigma_1\phi^{-1}$.*

Proof. Observe that for $\sigma_1(z) = \frac{z}{|z|^2}$ and $\phi(z) = a + rz$ we have,

$$\sigma(z) = \phi(\frac{r(z-a)}{|z-a|^2}) = \phi\sigma_1(\frac{z-a}{r}) = \phi\sigma_1\phi^{-1}(z).$$

□

Lemma 2.6. *Any inversion of $\mathbb{C}$ in a circle maps circles and lines onto circles or lines. Hence inversions in circles in $\hat{\mathbb{C}}$ always maps circles onto circles.*

Proof. The lines and circles in $\mathbb{C}$ are represented by the following general equation:

$$E(z) = a|z|^2 + bz + \bar{b}\bar{z} + c = 0, \ a, c \in \mathbb{R}, \ b \in \mathbb{C}. \tag{2.1}$$

If $a = 0$ then (2.1) represents a line in $\mathbb{C}$, otherwise, a circle. Let σ be a reflection in a circle in $\hat{\mathbb{C}}$. Then $\sigma = \phi\sigma_1\phi^{-1}$, where ϕ, σ_1 as in Lemma 2.5. Since $E(\phi(z))$ and $E(\sigma_1(z))$ again represent a circle in $\hat{\mathbb{C}}$, hence σ maps circles onto circles in $\hat{\mathbb{C}}$. □

For example, the circle with diameter $[a,b]$ on the real line can be obtained from the line $x = a$ by the inversion in the circle $S(b, |a-b|)$.

3. Elements of Arc-length

Let U be a path-connected subspace in the plane $\mathbb{R}^2$. A *path* in U is a differentiable function $f : [a,b] \to U$ such that $f'(t)$ is continuous on (a,b). In coordinates, we can write $f(t) = (x(t), y(t))$ where $x(t)$ and $y(t)$ are continuous on $[a,b]$ and differentiable on (a,b) with continuous derivative. The image of an interval (either open, closed, or semi-open) under a path is a *curve* in $\mathbb{R}^2$.

The *Euclidean length* of f is given by the integral

$$length(f) = \int_a^b \sqrt{(x'(t)^2 + y'(t)^2)}dt,$$

where $\sqrt{(x'(t)^2 + y'(t)^2)}$ is the *element of arc length* in $\mathbb{R}^2$. If we identify $\mathbb{R}^2$ with $\mathbb{C}$, write $f(t) = x(t) + iy(t)$, then $f'(t) = x'(t) + iy'(t)$ and $|f'(t)| = \sqrt{x'(t)^2 + y'(t)^2}$. In particular, the integral for the length of f becomes

$$length(f) = \int_a^b \sqrt{(x'(t)^2 + y'(t)^2)}dt = \int_a^b |f'(t)|dt.$$

At this point, introduce a new notation and abbreviate the integral as

$$\int_a^b |f'(t)|dt = \int_f |dz|,$$

where we write the standard Euclidean element of arc-length in $\mathbb{C}$ as

$$|dz| = |f'(t)|dt.$$

Using this notation, one may write any path integral. That is, let ρ be a continuous function $\rho : \mathbb{C} \to \mathbb{R}$. The *path integral* of ρ along a path $f : [a, b] \to \mathbb{C}$ is the integral

$$\int_f \rho(z)|dz| = \int_a^b \rho(f(t))|f'(t)|dt.$$

This path integral can be interpreted as a new element of arc-length, denoted $ds = \rho(z)|dz|$, obtained by scaling the Euclidean element of arc-length $|dz|$ at every point $z \in \mathbb{C}$, where the amount of scaling is described by the function ρ. This gives us the following definition.

Definition 3.1. For a path $f : [a, b] \to \mathbb{C}$ the *length of f with respect to the element of arc-length $ds = \rho(z)|dz|$* is defined to be the integral

$$length_\rho(f) = \int_f ds = \int_f \rho(z)|dz| = \int_a^b \rho(f(t))|f'(t)|dt.$$

Example 3.2. Let $\rho(z) = \frac{1}{1+|z|^2}$, and consider the element of arc length $\rho(z)|dz|$ on $\mathbb{C}$. For $r > 0$, consider the path $f : [0, 2\pi] \to \mathbb{C}$ given by $f(t) = re^{it}$, which parametrizes the Euclidean circle with center 0 and radius r. The length of f with respect to the element of arc-length $\rho(z)|dz|$ is

$$length_\rho(f) = \int_f \frac{1}{1+|z|^2}|dz| = \int_0^{2\pi} \frac{1}{1+|f(t)|^2}|f'(t)|dt = \frac{2\pi r}{1+r^2}.$$

Remark 3.3. The above notions also hold for piecewise C^1 path. A path $f : [a, b] \to \mathbb{C}$ is *piecewise* C^1 if f is continuous and if there is a partition of $[a, b]$ into subintervals $[a_0, a_1], [a_1, a_2], ..., [a_{n-1}, a_n]$, $a = a_0$, $b = a_n$, such that f is a path when restricted to each subinterval $[a_i, a_{i+1}]$. Any calculation or operation that we can perform on a path, we can also perform on a piecewise C^1 path, by expressing it as the concatenation of the appropriate number of C^1 paths. From now on, by a path we shall mean a piecewise C^1 path.

Let U be a path-connected subset of $\mathbb{C}$ which is equipped with an element of arc-length $\rho(z)|dz|$. Let $x, y \in U$, and let $P(x, y)$ be the set of paths joining x and y. Define the function $d : U \times U \to \mathbb{R}$ by

$$d(x, y) = inf\{length_\rho(f) \mid f \in P(x, y)\}.$$

It can be proved that d satisfies the conditions to be a metric on U, and hence (U, d) is a metric space. If x, y are two points on U, then the *geodesic segment* joining x and y is the shortest path between x and y. Thus a geodesic segment between x and y is a path $f \in P(x, y)$ such that $length_\rho(f) = d(x, y)$. It is not necessary that there always exists a geodesic segment joining two points in U. The length space (U, ρ) is called a *geodesic metric space* if any two points can be joined by a geodesic segment. The Hopf-Renow theorem states that (U, ρ) is a geodesic metric space if and only if (U, d) is a complete metric space, i.e. every Cauchy sequence in (U, d) is convergent.

4. The hyperbolic Space

4.1. The Upper-Half Space Model. The complex upper-half space is given by

$$H = \{z = x + iy \mid \Im z = y > 0\}.$$

The complex upper-half space H equipped with the element of arc-length $ds = \frac{|dz|}{\Im z}$ is the *upper-half space model of the hyperbolic space*, and is denoted by $\mathbf{H}^2$. Note that the boundary of $\mathbf{H}^2$ in $\widehat{\mathbb{C}}$ is the circle $\widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$. It is often called the *boundary at infinity* or *the circle at infinity* of the hyperbolic plane. Let $\gamma : [0,1] \to \mathbf{H}^2$ be a path

$$\gamma = \{z(t) = x(t) + iy(t) \mid t \in [0,1],\ y(t) > 0\}.$$

As in the above section, its *hyperbolic length* is given by

$$length(\gamma) = \int_\gamma \frac{|dz|}{y(t)} = \int_0^1 \frac{\frac{|dz|}{|dt|}}{y(t)} dt.$$

A *geodesic* or *a hyperbolic line* is a path which minimizes the length. The *hyperbolic distance* between two points z and w is given by

$$\rho(z, w) = inf\ length(\gamma),$$

where the infimum is taken over all γ joining z and w in $\mathbf{H}^2$. It is easy to check that ρ is non-negative, symmetric and satisfies the triangle inequality

$$\rho(z, w) \leq \rho(z, u) + \rho(u, w),$$

i.e. ρ is a metric on $\mathbf{H}^2$.

4.2. The geodesics in $\mathbf{H}^2$. To understand the geometry of $\mathbf{H}^2$, a first step is to know which curves in $\mathbf{H}^2$ are geodesics. The first step is to show that the vertical lines are geodesics in $\mathbf{H}^2$:

Let P_0 and P_1 are two points in $\mathbf{H}^2$ with same x-coordinate. Let their y-coordinate be y_0 and y_1 respectively. The length of the vertical segment γ joining P_0 to P_1 is

$$\int_\gamma ds = \left| \int_{y_0}^{y_1} \frac{dy}{y} \right| = \left| log \frac{y_1}{y_0} \right|.$$

If l is a different path from P_0 to P_1 then length of l is

$$\begin{aligned} &= \left| \int_{y_0}^{y_1} \frac{1}{y} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt \right| \\ &\geq \left| \int_{y_0}^{y_1} \frac{1}{y} \left[\left(\frac{dy}{dt}\right)^2\right]^{\frac{1}{2}} dt \right| \\ &= \left| \int_{y_0}^{y_1} \frac{1}{y} dy \right| \end{aligned}$$

Thus the length of l is greater than the length of γ. So, γ is the shortest path between P_0 and P_1. Hence it follows that *all vertical lines in* $\mathbf{H}^2$ *are geodesics.*

The second step is to find some isometries of $\mathbf{H}^2$. As isometric image of a geodesic is again a geodesic, this will give us some more geodesics in $\mathbf{H}^2$. One obvious kind of isometry is to reflect $\mathbf{H}^2$ in vertical lines. Recall that the reflection in the line $x = c$ is given by

$$r(z) = 2c - \bar{z}.$$

If $z = x + iy$ then $dz = dx + idy$. Also

$$\frac{|dz|}{\Im z} = \frac{dx^2 + dy^2}{y}.$$

For $w = 2c - (x - iy)$, $dw = -(dx - idy)$ and $|dw| = dx^2 + dy^2 = |dz|$. This implies,

$$\Im z = \Im w, \text{ hence, } \frac{|dw|}{\Im w} = \frac{|dz|}{\Im z}.$$

This shows that the reflections in vertical lines are isometries of $\mathbf{H}^2$.

Lemma 4.1. *An inversion σ of $\widehat{\mathbb{C}}$ in $S(a, r)$ restricts to an isometry of $\mathbf{H}^2$ provided a is in $\mathbb{R}$.*

Proof. First observe that σ_1 is an isometry of the hyperbolic plane. Let

$$w = \sigma_1(z) = \frac{z}{|z|^2} = \frac{1}{\bar{z}}.$$

Clearly, it preserves the upper-half space, and $dw = -\frac{\overline{dz}}{\bar{z}^2}$. Therefore,

$$ds^2 = \frac{|dw|^2}{(\Im w)^2} = \frac{4dw\overline{dw}}{(w - \bar{w})^2} = \frac{4dz\overline{dz}}{(z - \bar{z})^2} = \frac{|dz|^2}{(\Im z)^2}.$$

Thus σ_1 is an isometry of $\mathbf{H}^2$. If $y = a + rz$, then $dy = rdz$ and $\Im y = r\Im z$ if and only if $a \in \mathbb{R}$. This shows that the map $\phi(z) = a + rz$ is an isometry of $\mathbf{H}^2$ if and only if $a \in \mathbb{R}$. Since composition of two isometries is again an isometry, by Lemma 2.5 the lemma follows. $\square$

Since, the semi-circle with ends at a and b on the real line can be obtained from the line $x = a$ by inverting in the circle $S(b, |a - b|)$, hence it follows that every semi-circle in $\mathbf{H}^2$ centered at a point on the real line is a geodesic. We claim that there is no more geodesics in $\mathbf{H}^2$ other than the semi-circles centered at the real line, and the vertical lines.

Theorem 4.2. *A subset L of $\mathbf{H}^2$ is a geodesic if and only if L is the intersection of $\mathbf{H}^2$ with either a (straight) line, or a circle orthogonal to the real line.*

Proof. Suppose that there is a geodesic l of $\mathbf{H}^2$ which is not of the types mentioned above. Take any two points P and Q on l. Then through P, Q there will be either a vertical line or a semi-circle with center on the real-line. For, if P, Q are not joined by vertical lines, join them by a straight line. Take the perpendicular bisector of PQ. Suppose it cuts the real axis at the point C. Then the semi-circle with center C and radius CP or CQ will justify the assertion. But then the length of PQ along the above semi-circle or vertical length will be greater than the length of PQ along l, which cannot be possible as we have shown earlier. Hence l cannot be any other path than a vertical line, or a semi-circle of the above type. $\square$

Corollary 4.3. *Any two points z, w in $\mathbf{H}^2$ can be joined by a unique geodesic.*

4.3. **The Isometry Group.** It is easy to see that the isometries of $\mathbf{H}^2$ form a group under composition of maps. We denote the isometry group by $I(\mathbf{H}^2)$.

Lemma 4.4. $PSL(2,\mathbb{R}) \subset I(\mathbf{H}^2)$

Proof. We have already seen that the translations $z \mapsto z+b,\ b \in \mathbb{R}$ and the dialations $z \mapsto az,\ a \in \mathbb{R},\ a > 0$ are isometries of $\mathbf{H}^2$. Thus all transformations of the form $z \mapsto az+b,\ a,b \in \mathbb{R},\ a \neq 0$, are isometries. Also the maps $\iota : z \mapsto \frac{1}{\bar{z}}$ and $\eta : z \mapsto -\bar{z}$ are in $I(\mathbf{H}^2)$. Thus the maps of the form $z \mapsto \frac{1}{cz+d},\ c,d \in \mathbb{R},\ c \neq 0$ are in $I(\mathbf{H}^2)$. It follows that all the linear fractional transformations of the form

$$S : z \mapsto \frac{az+b}{cz+d},\ a,b,c,d \in \mathbb{R}$$

are in $I(\mathbf{H}^2)$ provided with some condition on a,b,c,d, which follow from the relation: $\Im(\frac{az+b}{cz+d}) > 0$. Now

$$\Im\left(\frac{az+b}{cz+d}\right) = \frac{\Im z}{|cz+d|^2} det \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

which is > 0 if and only if

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} > 0.$$

Thus $z \mapsto \frac{az+b}{cz+d}$ is an isometry of $\mathbf{H}^2$ if and only if $a.b,c,d \in \mathbb{R}$ and $ad-bc > 0$. Since the map ι is also in $I(\mathbf{H}^2)$,the map

$$\iota \circ S : z \mapsto \frac{c\bar{z}+d}{a\bar{z}+b}$$

is in $I(\mathbf{H}^2)$ if and only if

$$det \begin{pmatrix} c & d \\ a & b \end{pmatrix} = -det \begin{pmatrix} a & b \\ c & d \end{pmatrix} < 0.$$

The group

$$\mathcal{M}^+(1) = \{z \mapsto \frac{az+b}{cz+d} : ad-bc > 0\}$$

may be identified with the group $SL(2,\mathbb{R})$ in a natural way: just divide the numerator and denominator of the linear fractional by $\sqrt{ad-bc}$. After this identification we see that $SL(2,\mathbb{R})$ acts on $\mathbf{H}^2$ as a subgroup of isometries under the action:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{az+b}{cz+d}.$$

Note that for a given matrix A in $SL(2,\mathbb{R})$, A and $-A$ produce the same isometry under the above action. Hence $\mathcal{M}^+(1)$ is isomorphic to $PSL(2,\mathbb{R}) = SL(2,\mathbb{R})/\{\pm I\}$. □

Lemma 4.5. *$PSL(2,\mathbb{R})$ acts transitively on $\mathbf{H}^2$.*

Proof. Let $ai + b \in \mathbf{H}^2$, $a > 0$. Then $T(z) = \frac{\frac{a}{\sqrt{a}}z + \frac{b}{\sqrt{a}}}{\frac{1}{\sqrt{a}}}$ is the action of

$$A_T = \begin{pmatrix} \frac{a}{\sqrt{a}} & \frac{b}{\sqrt{a}} \\ 0 & \frac{1}{\sqrt{a}} \end{pmatrix}$$

on $\mathbf{H}^2$ and $T(i) = ai + b$. Thus for all z in $\mathbf{H}^2$, there exists A in $PSL(2, \mathbb{R})$ such that $A(z) = i$. Thus any point on $\mathbf{H}^2$ can be mapped onto i by the action of a suitable element of $PSL(2, \mathbb{R})$. This completes the proof. □

Lemma 4.6. *Let $[z, w]$ denote the closed segment on the geodesic joining distinct points z and w on $\mathbf{H}^2$. Then*

$$\rho(z, w) = \rho(z, \zeta) + \rho(\zeta, w)$$

if and only if $\zeta \in [z, w]$.

Proof. Then if $\zeta \in [z, w]$,

$$\rho(z, w) = \rho(z, \zeta) + \rho(\zeta, w).$$

If ζ does not belong to $[z, w]$, let γ be the path consisting of the segments $[z, \zeta]$ and $[\zeta, w]$. Then γ is a path from z to w other than the geodesic. Hence length of γ along this path will be greater than $\rho(z, w)$. Thus

$$\rho(z, w) < \rho(z, \zeta) + \rho(\zeta, w).$$

□

The above lemma implies that for any isometry ϕ of $\mathbf{H}^2$, the points $\phi(\zeta)$ is between $\phi(z)$ and $\phi(w)$ if and only if ζ is between z and w. So ϕ maps $[z, w]$ onto $[\phi(z), \phi(w)]$. Hence

Corollary 4.7. *A transformation in $PSL(2, \mathbb{R})$ maps geodesics onto geodesics.*

Lemma 4.8. *$PSL(2, \mathbb{R})$ acts transitively on the set of all geodesics in $\mathbf{H}^2$.*

Proof. It suffices to show that any geodesic l can be mapped onto the y-axis $x = 0$ by the action of an appropriate element of $PSL(2, \mathbb{R})$. If l is perpendicular to the real line, suppose it intersects the real line at the point c. Then the element of $PSL(2, \mathbb{R})$ corresponding to either $z \mapsto z + c$, or, $z \mapsto z - c$ does the job. If l is a semi-circle perpendicular to the real axis, let the point of intersections with the real line be $a,\ b,\ b > a$. Then the element of $PSL(2, \mathbb{R})$ corresponding to the isometry $z \mapsto \frac{z-b}{z-a}$ maps l onto the line $x = 0$. This completes the proof. □

Lemma 4.9. *$PSL(2, \mathbb{R})$ acts triply transitively on $\widehat{\mathbb{R}}$.*

Proof. Suppose a, b, c in $\mathbb{R}$ are mutually distinct. We assume $a > b > c$.

Case (i). Let $a \neq \infty$. Let

$$A = \begin{pmatrix} a - c & -b(a - c) \\ b - c & -c(b - c) \end{pmatrix}.$$

Consider the action of $\frac{A}{|A|}$ on $\mathbf{H}^2$:

$$z \mapsto \frac{(a-c)z - b(a-c)}{(b-c)z - c(b-c)}.$$

Then $Aa = 1,\ Ab = 0,\ Ac = \infty$.

Case (ii). Let $a = \infty$.

$$\text{The transformation} \quad z \mapsto \frac{z-c}{z-b} \quad \text{maps } (\infty, b, c) \text{ onto } (1, 0, \infty).$$

Thus any triplet $\{a, b, c\} \subset \widehat{\mathbb{R}}$ is mapped onto $\{1, 0, \infty\}$ by a suitable element of $PSL(2, \mathbb{R})$. This completes the proof. □

Recall that the *cross ratio* of distinct points $z_1,\ z_2,\ z_3,\ z_4$ on $\hat{\mathbb{C}}$ is given by

$$(z_1, z_2; z_3, z_4) = \frac{(z_1 - z_2)(z_4 - z_3)}{(z_3 - z_2)(z_4 - z_1)}.$$

It is well-known that the cross ratios are invariant under linear fractional transformations.

Theorem 4.10. *Let $z,\ w$ be two distinct points in $\mathbf{H}^2$. Let p_z and p_w be the end points of the geodesic on $\widehat{\mathbb{R}}$. Then*

$$\rho(z, w) = |ln(w, p_z; z, q_w)|,$$

where $(p_z, q_w; z, w)$ denotes the cross-ratio of $p_z,\ q_w,\ z,\ w$.

Proof. First consider the case when $z = ia,\ w = ib$, that is, both the points lies on the y-axis. Then $\{p_z, q_w\} = \{0, \infty\}$. Let $p_z = 0, q_w = \infty$. Let γ be the segment of the y-axis joining ia and ib. Then

$$\rho(z, w) = length(\gamma) = \left| \int_a^b \frac{dy}{y} \right| = |ln\frac{b}{a}|.$$

Note that $(w, 0; z, \infty) = \frac{b}{a}$. Suppose σ is some other arc from ia to ib, then

$$\begin{aligned} &= \left| \int_\sigma \frac{1}{y} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt \right| \\ &\geq \left| \int_a^b \frac{1}{y} \left[\left(\frac{dy}{dt}\right)^2 \right]^{\frac{1}{2}} dt \right| \\ &= |ln\frac{b}{a}| \end{aligned}$$

Therefore $\rho(z, w) = |ln(w, 0; z, \infty)|$,where $z = ia,\ w = ib$.

Suppose $z,\ w$ are points other than pure imaginary. Then there is a unique geodesic joining them. Let $A \in PSL(2, \mathbb{R})$ be such that, for $a, b \in \mathbb{R}$, $Az = ia,\ Aw = ib,\ Ap_z = 0, Aq_w = \infty$. Since linear fractional transformations preserve the cross ratio,we have

$$(0, \infty; ia, ib) = (Aw, Ap_z; Az, Aq_w) = (w, p_z; z, q_w).$$

Therefore $\rho(z, w) = |ln(w, p_z; z, q_w)|$. This proves the theorem. □

Corollary 4.11.

$$\cosh \rho(z,w) = 1 + \frac{|z-w|^2}{2\Im z \Im w}$$

.

Proof. Since $PSL(2,\mathbb{R})$ acts transitively on the geodesics of $\mathbf{H}^2$, it suffices to show the result for points on the imaginary axis. So suppose without loss of generality, $z = ia,\ w = ib,\ a < b$. For such points, the above relation is immediate. □

Theorem 4.12. *The group of isometries of $I(\mathbf{H}^2)$ is given by the linear fractional transformations over the reals, i.e.*

$$I(\mathbf{H}^2) = \mathcal{M}^+(1) \cup \mathcal{M}^-(1),$$

where

$$\mathcal{M}^+(1) = \left\{ z \mapsto \frac{az+b}{cz+d} \mid a,b,c,d \in \mathbb{R},\ det\begin{pmatrix} a & b \\ c & d \end{pmatrix} > 0 \right\},$$

$$\mathcal{M}^-(1) = \left\{ z \mapsto \frac{a\bar{z}+b}{c\bar{z}+d} \mid a,b,c,d \in \mathbb{R},\ det\begin{pmatrix} a & b \\ c & d \end{pmatrix} < 0 \right\}.$$

Proof. Given any isometry ϕ in $I(\mathbf{H}^2)$ there is an isometry

$$g(z) = \frac{az+b}{cz+d}, a,b,c,d \in \mathbb{R},\ ad-bc = 1$$

such that $g\phi$ leaves the positive imaginary axis $\Im^+$ invariant:

Simply choose g such that it maps $\phi(\Im^+)$ onto $\Im^+$. Such g exists by Proposition 4.8.

For if, Suppose $g\phi(i) \neq i$ then compose $g\phi$ with a reflection σ sending $g\phi(i)$ to i. If this composition doesn't leave the rays $(0,i)$ and (i,∞) invariant, compose it with a reflection through a geodesic l through i, such that the circle containing l is $S(0,1)$. Thus we can assume $g\phi(i) = i$ and $g\phi$ leaves the rays $(0,i)$, (i,∞) invariant. Thus, without loss of generality, we assume that, $g\phi(i) = i$, and $g\phi$ leaves the rays $(0,i)$ and (i,∞) invariant.

Let $z, w \in \Im^+$. A point on $\Im^+$ is determined by its distance from i, by the formula

$$\rho(z,w) = \rho(g\phi(z), g\phi(w)) = \rho(g\phi(z), i) + \rho(i, g\phi(w)).$$

Therefore we must have $g\phi(z) = z,\ g\phi(w) = w$. Thus $g\phi$ fixes $\Im^+$ point-wise.

Select $z \in \mathbf{H}^2$. Let $z = x+iy, y > 0$. Let $g\phi(z) = u + iv$. Then

$$\begin{aligned} & \rho(z,it) = \rho(g\phi(z), g\phi(it)) = \rho(u+iv, it) \\ \Rightarrow\quad & cosh\rho(z,it) = \cosh\rho(u+iv,it) \\ \Rightarrow\quad & v[x^2 + (y-t)^2] = y[u^2 + (v-t)^2]. \end{aligned}$$

This relation is hold for all $t > 0$. Hence we must have $y = v$ and $x^2 = u^2$. Thus $g\phi(z) = z$ or $-\bar{z}$.

On $\Im^+$, $g\phi(z) = z = -\bar{z}$. Since $g\phi$ is continuous on $\mathbf{H}^2$, it maps connected components of $\mathbf{H}^2 - \Im^+$ onto connected components. Hence one and only one of the possibilities $g\phi(z) = z$ or $g\phi(z) = -\bar{z}$ must hold for all $z \in \mathbf{H}^2$.

If $g\phi(z) = z$, then $a\phi(z) + b = zc\phi(z) + dz$. This implies,

$$\phi(z) = \frac{dz - b}{a - cz}, \quad det \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = ad - bc = 1$$

If $g\phi(z) = -\bar{z}$ then

$$\phi(z) = \frac{-d\bar{z} - b}{a - cz}, \quad det \begin{pmatrix} -d & -b \\ c & a \end{pmatrix} = -1.$$

Given any transformation of the form $z \mapsto \frac{a\bar{z}+b}{c\bar{z}+d}$, $ad-bc < 0$ we can consider it as the following transformation after dividing the numerator or the denominator by $\sqrt{-(ad-bc)}$:

$$z \mapsto \frac{a\bar{z} + b}{c\bar{z} + d}, \ ad - bc = -1.$$

This proves the theorem. □

Corollary 4.13. $I(\mathbf{H}^2)$ *is generated by the reflections in geodesics.*

Proof. We have seen in the proof of Lemma 4.4 that every isometry in $\mathcal{M}^+(1)$ is generated by reflections in geodesics. The elements in $\mathcal{M}^-(1)$ are obtained by composing the reflection $z \to \bar{z}$ with elements of $\mathcal{M}^-(1)$. □

Remark 4.14. Given three distinct points on the circle $\widehat{\mathbb{R}}$, we can order them by the usual cyclic ordering on a circle. There are exactly two such choices of order, either clockwise or anti-clockwise. A preferred choice of an order is called an *orientation* on the boundary circle of $\mathbf{H}^2$. If an isometry ϕ preserves a chosen orientation, i.e. for all triples, (z_1, z_2, z_3) and $(\phi(z_1), \phi(z_2), \phi(z_3))$ belong to the same ordered class, it is called *orientation-preserving*. It follows that the group of orientation-preserving isometries of $\mathbf{H}^2$ is $\mathcal{M}^+(1)$. This is because $det\ \phi'(z) > 0$. We identify $\mathcal{M}^+(1)$ with the group $PSL(2, \mathbb{R})$. This is an index 2 subgroup in $I(\mathbf{H}^2)$.

4.4. **The Poincaré Disk Model.** We shall now describe the Poincaré disk model of the hyperbolic plane. We shall deduce this model naturally from the upper-half space model by the map

$$f : z \to \frac{z - i}{z + i}.$$

Note that f maps $\mathbf{H}^2$ onto the unit disk $\mathbf{D}^2$ and $f^{-1} : \mathbf{D}^2 \to \mathbf{H}^2$ is given by

$$f^{-1}(z) = \frac{i(1 + z)}{1 - z}.$$

The metric on $\mathbf{D}^2$ is given by $d(z, w) = \rho(f^{-1}(z), f^{-1}(w))$. Thus f is an isometry between $\mathbf{H}^2$ and $\mathbf{D}^2$.

Let $w = f^{-1}(z) = \frac{i(1+z)}{1-z}$. Then $ds = \frac{|dw|}{\Im w}$, where $dw = \frac{1+i}{(1-z)^2}dz$. Since, $w = \frac{i(1+z)}{(1-z)} = \frac{i(1+z)(1-\bar{z})}{|1-z|^2}$, this implies $\Im w = \frac{1-|z|^2}{|1-z|^2}$. Therefore,

$$ds = \frac{|dw|}{\Im w} = \frac{2|dz|}{1 - |z|^2}.$$

Thus the element of arc-length on $\mathbf{D}^2$, induced from $\mathbf{H}^2$ by the isometry f, is

$$ds = \frac{2|dz|}{1-|z|^2}.$$

Proposition 4.15.

$$\cosh d(z,w) = 1 + \frac{2|z-w|^2}{(1-|z|^2)(1-|w|^2)}.$$

Proof. We have,

$$\begin{aligned} \cosh d(z,w) &= \cosh \rho(f^{-1}(z), f^{-1}(w)) \\ &= 1 + \frac{|f^{-1}(z) - f^{-1}(w)|^2}{2\Im f^{-1}(z)\Im f^{-1}(w)} \\ &= 1 + \frac{\left(\frac{2|z-w|}{|1-z||1-w|}\right)^2}{\frac{2(1-|z|^2)(1-|w|^2)}{|1-z|^2|1-w|^2}} \\ &= 1 + \frac{2|z-w|^2}{(1-|z|^2)(1-|w|^2)} \end{aligned}$$

□

The geodesics of $\mathbf{D}^2$ are clearly the images of the geodesics of $\mathbf{H}^2$ under f. It follows that the geodesics are either the open diameters of $\mathbf{D}^2$ or the intersection of $\mathbf{D}^2$ with circles orthogonal to the boundary circle. The isometry group of $\mathbf{D}^2$ is

$$I(\mathbf{D}^2) = \{f\phi f^{-1} : \phi \in I(\mathbf{H}^2)\}.$$

Let $\phi \in I(\mathbf{H}^2)$ be such that $\phi(z) = \frac{az+b}{cz+d},\ ad-bc=1,\ a,b,c,d \in \mathbb{R}$. By a simple computation, it follows that

$$f\phi f^{-1}(z) = \frac{\alpha z - \beta}{\bar{\beta} z - \bar{\alpha}},\ |\alpha|^2 - |\beta|^2 = 1.$$

Since ϕ is an orientation preserving isometry of $\mathbf{H}^2$, hence $f\phi f^{-1}$ is an orientation-preserving isometry of $\mathbf{D}^2$. If

$$\phi(z) = \frac{a\bar{z}+b}{c\bar{z}+d},$$

then

$$f\phi f^{-1} = \frac{\alpha \bar{z} - \beta}{\bar{\beta} z - \bar{\alpha}},\ |\alpha|^2 - |\beta|^2 = -1,$$

which is an orientation reversing isometry of $\mathbf{D}^2$. Thus we have:

Theorem 4.16. *The isometries of the Poincaré disk model $\mathbf{D}^2$ of the hyperbolic plane are given by*

$$\frac{\alpha z - \beta}{\bar{\beta} z - \bar{\alpha}},\ |\alpha|^2 - |\beta|^2 = 1,$$

$$\frac{\alpha \bar{z} - \beta}{\bar{\beta} z - \bar{\alpha}},\ |\alpha|^2 - |\beta|^2 = -1.$$

Thus the matrix group

$$SU(1,1) = \left\{ \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} \; : \; |\alpha|^2 - |\beta|^2 = 1 \right\}$$

acts on $\mathbf{D}^2$ as the group of orientation-preserving isometries.

4.5. **The Hyperbolic Triangles.** Let us define the angle between two hyperbolic lines in $\mathbf{H}^2$ as the interior angle between their tangents at the point of intersection. Let x, y, z be three hyperbolically non-collinear points in $\mathbf{H}^2$. Let $L(x, y)$ be the unique geodesic in $\mathbf{H}^2$ containing x and y. Let $H(x, y, z)$ be the closed half-space of $\mathbf{H}^2$ with $L(x, y)$ as its boundary and z in its interior. The *hyperbolic triangle* with vertices x, y, z is defined to be

$$T(x, y, z) = H(x, y, z) \cap H(y, z, x) \cap H(z, x, y).$$

Let $[x, y]$ be the segment of $L(x, y)$ joining x and y. The *sides* of $T(x, y, z)$ are defined to be $[x, y],\ [y, z],\ [z, x]$. Let a, b, c be the hyperbolic lengths of $[z, y],\ [z, x]$ and $[x, y]$ respectively. Suppose $f : [0, a] \to \mathbf{H}^2$, $g : [0, b] \to \mathbf{H}^2$ and $h : [0, c] \to \mathbf{H}^2$ are the geodesic arcs from y to z, z to x, and x to y respectively. The angle α between the sides $[z, x]$ and $[x, y]$ of $T(x, y, z)$ is the interior angle between $-g'(b)$ and $h'(0)$, which is the interior angle between the tangents at the point of intersection of the sides. Similarly, angles between the other pair of sides are obtained. Now we shall allow the vertices of a triangle to belong to the circle at infinity. The angle between two geodesics is defined to be zero if they intersect at the circle at infinity. If all the three vertices of a hyperbolic triangle lie on the circle at infinity, it is called an *ideal* triangle.

The area of a set X in $\mathbf{H}^2$ is defined by

$$\text{Area}(X) = \iint_X \frac{dxdy}{y^2}.$$

The area in the unit-disk model $\mathbf{D}^2$ is

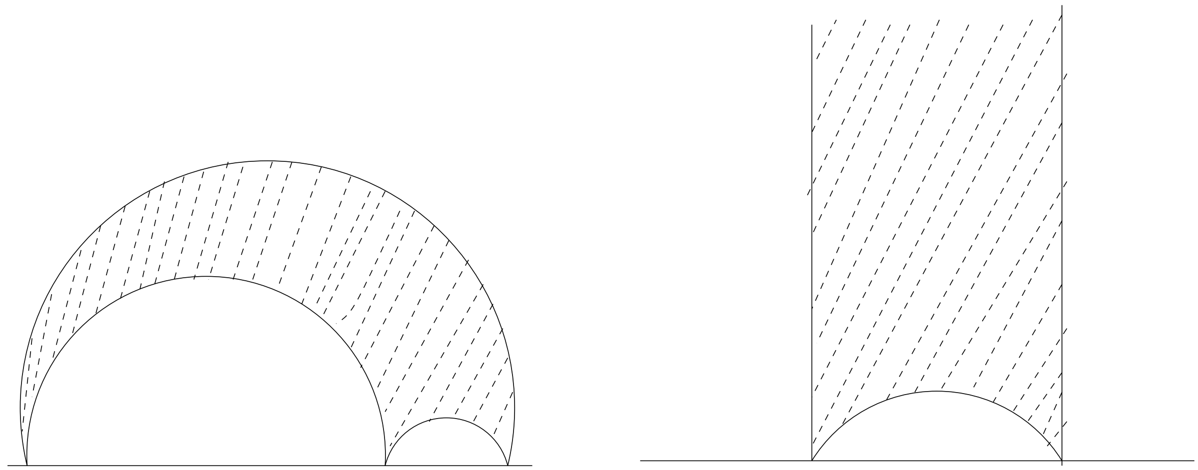
$$\iint_X \frac{2dxdy}{1 - x^2 - y^2}.$$

It can be proved that *The hyperbolic area is invariant under the isometries of* $\mathbf{H}^2$.

Theorem 4.17. *Any ideal triangle in the hyperbolic space has area* π.

Proof. Suppose $T(x, y, z)$ is any ideal triangle in $\mathbf{H}^2$. Since $PSL(2, \mathbb{R})$ acts triply transitively on $\widehat{\mathbb{R}}$, x, y, z can be mapped onto $\infty, 1, -1$ by a suitable isometry in $PSL(2, \mathbb{R})$. Thus the triangle we start with, will be mapped onto the triangle Δ with vertices $(-1, 0), (1, 0)$ and ∞. Therefore any ideal triangle $T(x, y, z)$ can be mapped onto Δ by a transformation in $PSL(2, \mathbb{R})$. This shows that all ideal triangles are congruent to each-other. Consequently, the

FIGURE 3. Ideal triangles in the hyperbolic plane



An ideal triangle with real vertices

An ideal triangle with one vertex at infinity

ideal triangles have equal area. Hence it is sufficient to compute the area of Δ. Observe that

$$\begin{aligned} Area(\Delta) &= \iint_T \frac{dxdy}{y^2} \\ &= \int_{-1}^{1} \int_{\sqrt{1-x^2}}^{\infty} \frac{dy}{y^2} dx \\ &= \int_{-1}^{1} \left[-\frac{1}{y}\right]_{\sqrt{1-x^2}}^{\infty} dx \\ &= \int_{-1}^{1} \frac{dx}{\sqrt{1-x^2}} \\ &= [\sin^{-1} x]_{-1}^{1} \\ &= \pi. \end{aligned}$$

This completes the proof. □

The following formula shows that the hyperbolic area of a hyperbolic triangle depends only on its angles.

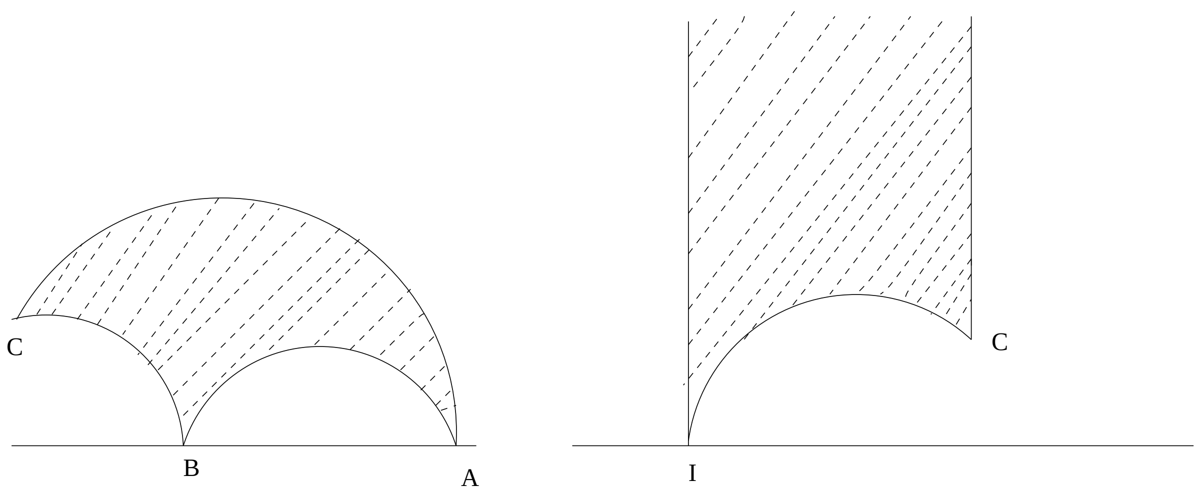
Theorem 4.18. (Gauss-Bonnet) *Let Δ be a hyperbolic triangle with angles α, β and γ. Then*

$$Area(\Delta) = \pi - (\alpha + \beta + \gamma).$$

Proof. Case (i). Suppose Δ is an ideal triangle. Then the theorem follows from Theorem 4.17.

Case (ii). Two vertices of Δ are on the circle at infinity. Let α be the angle of the triangle

FIGURE 4. Two vertices on the circle at infinity



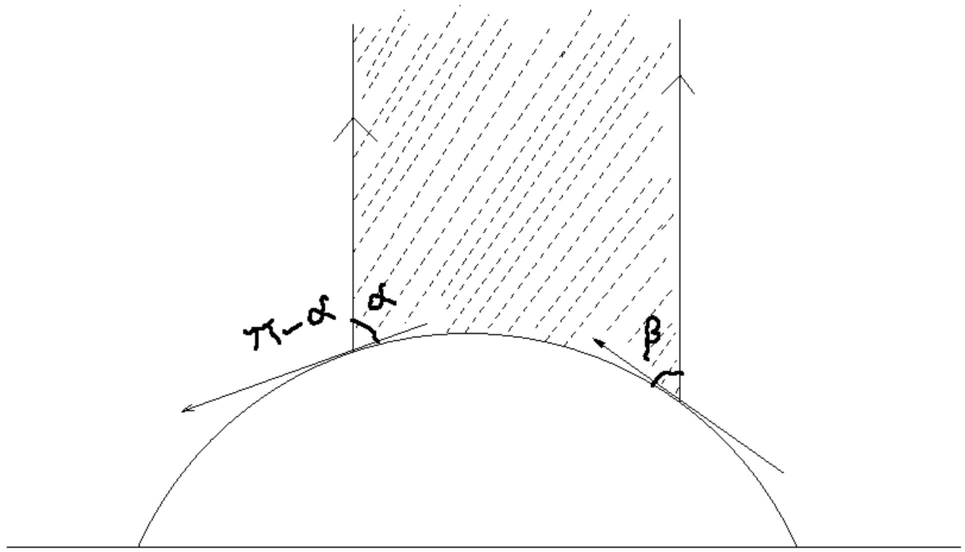
at the finite vertex. Now any such triangle can be mapped by a suitable transformation of $PSL(2,\mathbb{R})$ onto the triangle T with ideal vertices at $A(1,0)$, ∞ and the other vertex on the geodesic segment $S(0,1)\cap\mathbf{H}^2$. For example, in Figure 4 first use one element in $PSL(2,\mathbb{R})$ which maps A to ∞. Then use suitable transformations in $PSL(2,\mathbb{R})$ to map the resulting triangle onto the triangle with vertices at $(1,0)$, ∞ and $P\in S(0,1)$. Note that the finite vertex at the triangle has angle α. So the finite vertex will be $e^{i(\pi-\alpha)}$. Then the area of Δ is

$$\begin{aligned}
Area(\Delta) &= \iint_T \frac{dxdy}{y^2} \\
&= \int_{\cos(\pi-\alpha)}^{1}\int_{\sqrt{1-x^2}}^{\infty}\frac{dy}{y^2}dx \\
&= \int_{\cos(\pi-\alpha)}^{1}\big[-\frac{1}{y}\big]_{\sqrt{1-x^2}}^{\infty}dx \\
&= \int_{\cos(\pi-\alpha)}^{1}\frac{dx}{\sqrt{1-x^2}} \\
&= \pi-\alpha.
\end{aligned}$$

Case(iii). Two vertices are finite.

Let v_1, v_2 be the finite vertices and let l be the geodesic joining them, α, β are the angles at the vertices. By a suitable transformation in $PSL(2,\mathbb{R})$ we can map the infinite vertex on ∞

FIGURE 5. Two finite vertices



and v_1, v_2 on the unit circle. Thus we further assume that $v_1 = e^{i(\pi-\alpha)}, v_2 = e^{i\beta}$. Therefore,

$$\begin{aligned} Area(\Delta) &= \iint_T \frac{dxdy}{y^2} \\ &= \int_{\cos(\pi-\alpha)}^{\cos\beta} \int_{\sqrt{1-x^2}}^{\infty} \frac{dy}{y^2} dx \\ &= \int_{\cos(\pi-\alpha)}^{\cos\beta} \big[-\frac{1}{y}\big]_{\sqrt{1-x^2}}^{\infty} dx \\ &= \int_{\cos(\pi-\alpha)}^{\cos\beta} \frac{dx}{\sqrt{1-x^2}} \\ &= \pi - (\alpha+\beta) \end{aligned}$$

Case(iv): All vertices are finite, with angles α, β, γ.

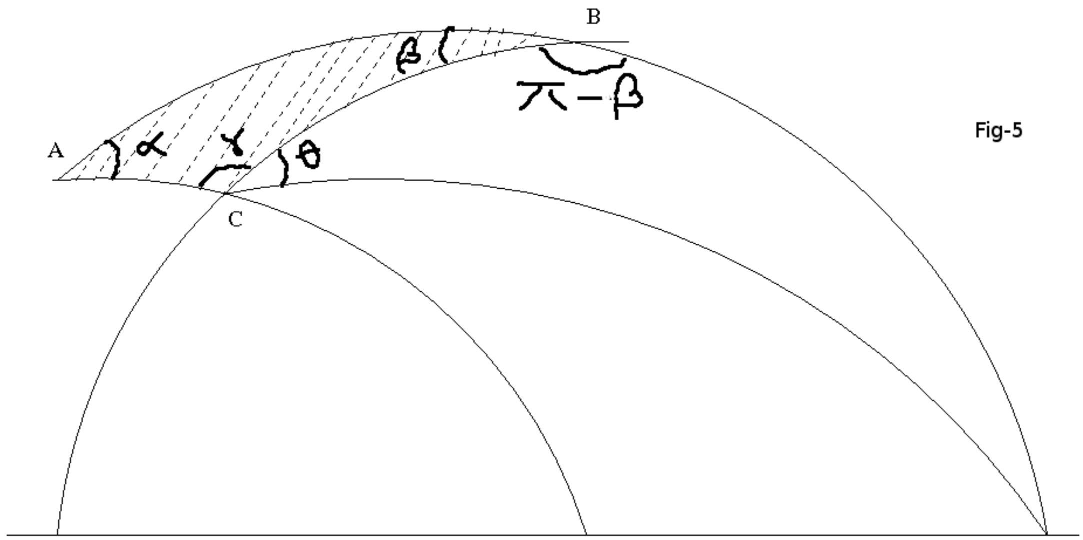
By the action of $PSL(2,\mathbb{R})$, we can assume that Δ has all the vertices on semi-circles orthogonal to the real line. Thus we can express it as a difference of two hyperbolic triangles each with one vertex at the circle at infinity as shown in the figure:

$$\Delta ABC = \Delta ACDA - \Delta BCDB.$$

Thus

$$\begin{aligned} Area(\Delta) &= Area(\Delta ACDA) - Area(\Delta BCDB) \\ &= \pi - (\alpha + (\gamma+\theta)) - (\pi - (\pi-\beta) - \theta) \\ &= \pi - (\alpha+\beta+\gamma). \end{aligned}$$

FIGURE 6. All vertices are finite



Thus the area of a hyperbolic triangle with angles α, β, γ is equal to be $\pi - (\alpha + \beta + \gamma)$, where the angle at the vertex at infinity, if any, is assumed to be zero. This formula is called the Gauss-Bonnet formula for hyperbolic triangles. □

The following are basic references in hyperbolic geometry. The reader may look at any of these texts for further reading.

References

[1] J. Anderson, *Hyperbolic Geometry*, Springer Undergraduate Mathematics Series, Springer-Verlag, 1999.
[2] Alan F. Beardon, *The Geometry Of Discrete Groups*, Graduate Texts in Mathematics 91, Springer-Verlag, 1983.
[3] H. S. M. Coxeter, *Non-Euclidean geometry*, Mathematical Expositions 2, University of Toronto Press, Toronto, 1978.
[4] W. Fenchel, *Elementary Geometry in Hyperbolic Space*, de Gruyter Studies in Mathematics 11, Walter de Gruyter, New York, 1989.
[5] B. Iversen, *Hyperbolic geometry*, London Mathematical Society Student Texts 25, Cambridge University Press, Cambridge, 1992.
[6] G. A. Jones and D. Singerman, *Complex functions : An algebraic and geometric viewpoint*, Cambridge University Press, 1987.
[7] P. J. Kelly and G. Matthews, The Non-Euclidean, Hyperbolic Plane, Universitext, Springer-Verlag, New York, 1981.
[8] John G. Ratcliffe, *Foundation Of Hyperbolic manifolds*, Graduate Texts in Mathematics 149, Springer-Verlag, 1994.
[9] S. Stahl, *The Poincaré Half-Plane*, Jones and Bartlett, Boston, 1993.

Indian Institute of Science Education and Research (IISER) Mohali, Knowledge City, S.A.S. Nagar, Phase 9, Sector 81, Mohali 140306, India

E-mail address: krishnendu@iisermohali.ac.in, krishnendug@gmail.com