

# 1

## Hyperbolic space and its isometries

In this chapter we gather together basic information about the geometry of two- and three-dimensional hyperbolic spaces and their isometries. This will set the stage for our study of quotient manifolds and orbifolds which begins in the next chapter.

### 1.1 Möbius transformations

A *Möbius transformation* in the unit sphere  $\mathbb{S}^n$  of dimension  $n$  is, by definition, the result of a composition of reflections in  $(n-1)$ -dimensional spheres in  $\mathbb{S}^n$ . It will be orientation preserving if it is the composition of an even number of reflections. A defining property is that Möbius transformations send  $(n-1)$ -dimensional spheres onto  $(n-1)$ -dimensional spheres. Automatically, a symmetric pair of points (with respect to reflection) about one sphere gets sent to a symmetric pair about the other.

*From now on, the unqualified term **Möbius transformation** will be reserved for those that preserve orientation.* The orientation reversing kind will be called anti-Möbius transformations. For a discussion of the latter, see Exercise 1-31 at the end of the chapter.

The study of hyperbolic 3-manifolds is intimately connected with the study of Möbius and anti-Möbius transformations on the two-dimensional sphere  $\mathbb{S}^2$ . Via

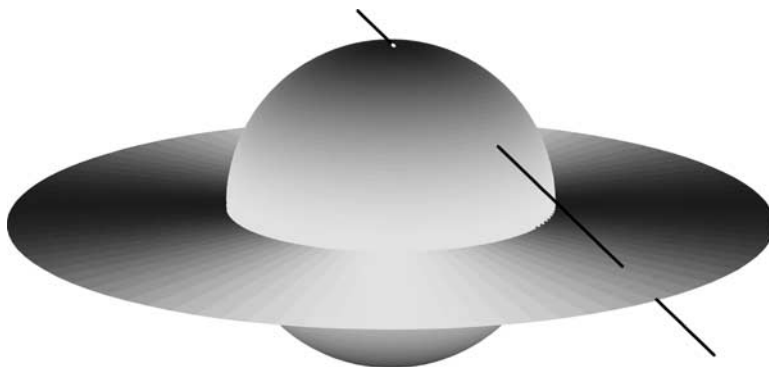


Fig. 1.1. Stereographic projection

stereographic projection (Figure 1.1),  $\mathbb{S}^2$  is homeomorphic to the extended plane  $\mathbb{C} \cup \infty$ , and we will freely use this fact to change points of view between the extended plane and the 2-sphere. Under stereographic projection, the collection of circles and straight lines in  $\mathbb{C}$  corresponds to the collection of circles on  $\mathbb{S}^2$ ; a straight line in  $\mathbb{C}$  corresponds to a circle on  $\mathbb{S}^2$  through the north pole. With this correspondence in mind, we can refer to the collection of circles and lines in  $\mathbb{C}$  simply as “circles”. Moreover stereographic projection is a conformal map, that is, it preserves angles between intersecting arcs — in particular, angles of intersection between circles.

Möbius transformations in two dimensions are *fractional linear transformations* of the extended plane. That is, a Möbius transformation acting on  $\mathbb{C} \cup \infty$  has the form

$$z \mapsto A(z) = \frac{az + b}{cz + d}, \quad \text{with } a, b, c, d \in \mathbb{C} \text{ such that } ad - bc \neq 0. \quad (1.1)$$

(When  $ad - bc = 0$  the expression on the right is a constant, so the map is not a Möbius transformation.) As we will see shortly, a map of this form can indeed be expressed as the composition of an even number of reflections in circles (in fact, two or four circles: see Exercise 1-7). The symmetry properties of such maps are established in Exercise 1-2.

Möbius transformations are conformal maps. In fact, the only conformal homeomorphisms of  $\mathbb{C} \cup \infty$  are Möbius transformations.

We will generally assume that the representation in (1.1) is *normalized*, meaning that  $ad - bc = 1$ . Then we can identify the group of Möbius transformations with the quotient  $\text{PSL}(2, \mathbb{C}) := \text{SL}(2, \mathbb{C}) / \pm I$ , where  $\text{SL}(2, \mathbb{C})$  is the group of  $2 \times 2$  matrices of determinant one and  $I$  is the identity matrix:

$$A(z) = \frac{az + b}{cz + d} \longleftrightarrow \pm \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A^{-1}(z) \longleftrightarrow \pm \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

The  $\pm$  ambiguity cannot be avoided. We will not keep inserting it, unless it plays an essential role. In any case the value of changing from transformations to matrices lies mainly in the algebra of composition. If  $A, B$  are Möbius transformations, the Möbius transformation resulting from the application of  $A$  followed by  $B$  is written  $BA$ ; the corresponding matrix is just the usual product  $BA$  of the component matrices, in the order written. The  $\pm$  ambiguity follows along. We will hop from one to the other, the representation as a transformation to the representation as a matrix, depending on which best suits the situation, without changing the labeling.

Two Möbius transformations  $A, B$  are *conjugate* if there is a Möbius transformation  $U$  such that  $B = UAU^{-1}$ . Conjugate transformations have the same geometry:  $U$  effects transfer of the geometry of  $A$  to that of  $B$ .

The expression  $ABA^{-1}B^{-1}$  is called the *commutator* of  $A$  and  $B$  and written as  $[A, B]$ . Two elements commute if and only if their commutator is the identity. \*

\* The alternative conventions  $[A, B] = B^{-1}A^{-1}BA$  or  $A^{-1}B^{-1}AB$  are preferred by some authors; they do the same job, but the formulas come out differently.

The *trace* of a Möbius transformation  $A$  is, by definition, the trace of the normalized matrix of  $A$ :

$$\tau_A = \text{tr } A = \pm(a + d).$$

It is invariant under conjugation. The  $\pm$  ambiguity can be avoided either by using  $\tau_A^2$  or by specifying  $0 \leq \arg \tau_A < \pi$ .

By solving the equation  $A(z) = z$ , we find that a nontrivial Möbius transformation has one or two fixed points in  $\mathbb{S}^2$ , namely  $(a - d \pm \sqrt{\tau_A^2 - 4})/2c$ , when  $c \neq 0$ , or otherwise the points  $\infty$  and  $b/(d - a) = ab/(1 - a^2)$ . Here  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ,  $ad - bc = 1$ . Only the identity can have three fixed points.

Given three distinct points  $(p_2, p_3, p_4) \in \mathbb{S}^2$ , there exists a necessarily unique Möbius transformation sending  $p_2$  to 1,  $p_3$  to 0,  $p_4$  to  $\infty$ . It is given by

$$z \mapsto \frac{(z - p_3)(p_2 - p_4)}{(z - p_4)(p_2 - p_3)} = (z, p_2, p_3, p_4),$$

when none of the points  $p_i$  is  $\infty$ . By taking the limit as some  $p_i \rightarrow \infty$ , we obtain the correct expression for  $p_i = \infty$ . The expression  $(z, p_2, p_3, p_4)$  is called the *cross ratio* of the four points. \* Cross ratios are invariant under Möbius transformations:

$$(Az, Ap_2, Ap_3, Ap_4) = (z, p_2, p_3, p_4) \quad \text{for any } A.$$

This is a consequence of the fact that  $T(z) = (z, p_1, p_2, p_3)$  satisfies  $T \circ A^{-1}(z) = (z, Ap_1, Ap_2, Ap_3)$ .

Apart from the identity, Möbius transformations fall into one of three types:

$A$  is *parabolic* if the following equivalent properties hold.

- $A$  is conjugate to  $z \mapsto z + 1$ .
- $A$  has exactly one fixed point in  $\mathbb{S}^2$ .
- $\tau_A = \pm 2$  and  $A \neq \text{id}$ .

$A$  is *elliptic* if the following equivalent properties hold.

- $A$  is conjugate to  $z \mapsto e^{2i\theta} z$ , with  $2\theta \not\equiv 2\pi$ .
- $\tau_A \in (-2, +2)$ .
- $A$  has exactly two fixed points, and the derivative of  $A$  has absolute value 1 at each of them.

$A$  is *loxodromic* if the following equivalent properties hold.

- $A$  is conjugate to  $z \mapsto \lambda^2 z$ , with  $|\lambda| > 1$ .
- $\tau_A \in \mathbb{C} \setminus [-2, +2]$ .
- $A$  has exactly two fixed points, one attracting and one repelling.

We will use the term *standard forms* for the conjugates for the conjugates just listed. The geometry of a general normalized Möbius transformation  $A$  is most easily read off from the conjugate standard form. Note that the elliptic  $z \mapsto 1/z$  is conjugate to  $z \mapsto -z$ .

\* The definition given has the property  $(z, 1, 0, \infty) = z$ . A common alternate definition results in  $(z, 0, 1, \infty) = z$ .

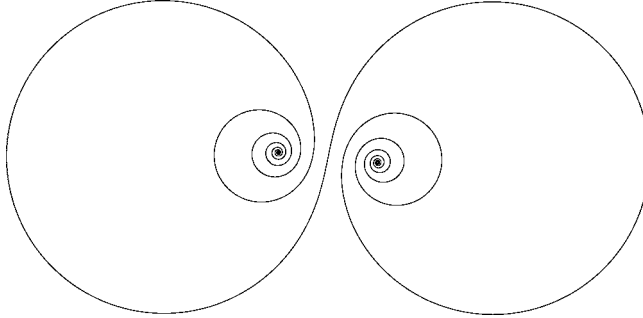


Fig. 1.2. Invariant spiral of a loxodromic with trace  $\lambda + \lambda^{-1} = 1.976 + 0.005i$ .

A loxodromic Möbius transformation  $A$  has a collection of *loxodromic curves* or *invariant spirals* in  $\mathbb{S}^2$ . (In navigation, a *loxodromic curve* or *rhumb line* is a path of constant bearing: it makes equal oblique angles with all meridians, and so coils around the poles without ever reaching them.) For the standard form  $z \mapsto \lambda^2 z$ , one such spiral is given by

$$z(t) = \lambda^{2t}, \quad -\infty < t < \infty.$$

If  $\sigma$  denotes the segment  $0 \leq t < 1$  of the spiral, the various images  $\{A^n(\sigma)\}$  cover the spiral without overlap. See Figure 1.2.

For additional structure in special cases see [Wright 2006].

The term *hyperbolic transformation* has historically been used to designate a loxodromic transformation whose trace is real. Such a transformation is conjugate to  $z \mapsto \lambda^2 z$  with  $\lambda > 1$ . Nowadays the term “hyperbolic” is also used for a loxodromic element acting in hyperbolic 3-space.

The classification is proved by first conjugating  $A$  so that one fixed point lies at  $\infty$  and the other, if there is one, at 0. The further conjugation  $z \mapsto 1/z$  that interchanges 0 and  $\infty$  may be needed to put the attracting fixed point at  $\infty$ .

If  $p \in \mathbb{C}$  is a fixed point of  $A \neq \text{id}$ ,  $p$  is attracting if and only if  $|A'(p)| < 1$  and repelling if and only if  $|A'(p)| > 1$ . The transformation  $A$  is parabolic if and only if  $A'(p) = 1$ ;  $A$  is elliptic if and only if  $|A'(p)| = 1$  but  $A'(p) \neq 1$ .

Upon referring to the normalized matrix  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , we find that the *eigenvalues* are  $\lambda, \lambda^{-1} = \frac{1}{2}(\text{tr } A \pm \sqrt{\text{tr}^2 A - 4})$ . The corresponding eigenvectors  $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$  satisfy

$$\frac{\alpha}{\beta} = \frac{\lambda - d}{c} = p \quad \text{and} \quad \frac{\alpha}{\beta} = \frac{\lambda^{-1} - d}{c} = q,$$

where  $p, q$  are the fixed points. Like the trace, the eigenvalues are invariant under conjugation. The eigenvalues of an elliptic transformation have the form  $e^{\pm i\theta}$  and the trace is  $2 \cos \theta$ . A loxodromic transformation has eigenvalues  $\lambda^{\pm 1}$  and trace  $\lambda + \lambda^{-1}$ . We can choose  $\lambda$  so that  $|\lambda| > 1$ , that is, so that  $\lambda$  is the expanding eigenvalue.

The expanding eigenvalue of a loxodromic element  $A$  can be expressed as a cross ratio by the formula

$$\lambda^2 = (z, A(z), p_+, p_-),$$

where  $p_+, p_-$  are the attracting and repelling fixed points. (It is enough to confirm this when  $p_+ = \infty$  and  $p_- = 0$ .)

We can write  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  as

$$Az = \frac{1}{-c^2(z + d/c)} + \frac{a}{c} \quad \text{if } c \neq 0, \quad Az = \frac{a}{d} \left( z + \frac{b}{a} \right) \quad \text{if } c = 0. \quad (1.2)$$

This expresses  $A$  in terms of simple building blocks: maps in standard form, plus the map  $z \mapsto 1/z$ . Each of these has the property of preserving (generalized) circles. Therefore any Möbius transformation preserves circles, as mentioned earlier. Likewise each building block is easily seen to be a composition of two reflections, so a Möbius transformation is the composition of an even number of reflections.

Three distinct points  $p_2, p_3, p_4$  uniquely determine a circle  $C$ , with an orientation determined by their order. When  $C$  is a proper circle, we say that the orientation thus defined is *positive* if the interior of the circle lies to the left as  $p_2, p_3, p_4$  are encountered in that order. Let  $q_2, q_3, q_4$  be another set of distinct points, and  $C'$  the circle through them. The Möbius transformation  $T$  that sends  $p_i \rightarrow q_i$  automatically sends  $C$  onto  $C'$ . If both are proper circles,  $T$  sends the interior of  $C$  to the interior of  $C'$  if and only if the triples give both circles positive (or negative) orientations. The transformation  $T : z \rightarrow w$  can be expressed in terms of cross ratios as

$$(w, q_2, q_3, q_4) = (z, p_2, p_3, p_4).$$

But if we focus simply on sending  $C$  to  $C'$ , and a designated side of  $C$  to a designated side of  $C'$ , it is more efficient to find  $T$  by cross ratio using the symmetry property: A Möbius transformation sends points symmetric with respect to reflection in one circle, to a pair of points symmetric in the image (Exercise 1-2). For a proper circle, the most conspicuous symmetric points are its center and  $\infty$ .

A cross ratio  $(p, p_2, p_3, p_4)$  is real if and only if the four points lie on a circle in  $\mathbb{S}^2$ . The cross ratio is positive if and only if  $(p, p_3, p_4)$  gives the circle the same orientation as  $(p_2, p_3, p_4)$ .

---

We are now ready to show that Möbius transformations in  $\mathbb{C} \cup \infty$  can be extended to Möbius transformations acting in upper half-space  $\{\vec{x} = (z, t) : z \in \mathbb{C}, t > 0\}$ . The simplest way to see this is by applying the following observation. Each Möbius transformation is the composition of an even number of reflections in circles or lines in  $\mathbb{C}$ . A reflection in a circle extends naturally to the reflection in the upper hemisphere bounded by that circle. Likewise the reflection in a straight line extends to the reflection in the vertical half-plane bounded by that line. (The same argument shows that Möbius transformations on  $\mathbb{S}^n = \mathbb{R}^n \cup \{\infty\}$  extend to upper half  $(n+1)$ -space.)

A Möbius transformation acting on  $\mathbb{C} \cup \infty$  sends a given circle to another circle or line. Its extension to upper half-space will therefore map the hemisphere bounded by the circle to the hemisphere or half-plane bounded by the image of the circle. We conclude that the extension to upper half-space maps the totality of hemispheres and vertical half-planes onto itself.

If two hemispheres intersect, or a hemisphere and a vertical half-plane intersect, the intersection is a semicircle which is orthogonal to  $\mathbb{C}$ . If two vertical half-planes intersect, they intersect in a vertical half-line orthogonal to  $\mathbb{C}$ . The extension of a Möbius transformation thus maps the totality of half-lines and semicircles orthogonal to  $\mathbb{C}$  onto itself. The dihedral angles between intersecting hemispheres is the same as the angle of intersection between their bounding circles in  $\mathbb{C}$ .

It is useful to explicitly work out the formula for extension to upper half-space  $\{\vec{x} = (z, t) : z \in \mathbb{C}, t > 0\}$ . We first extend the building blocks. First,

$$\begin{aligned} z \mapsto az & \text{ becomes } (z, t) \mapsto (az, |a|t); \\ z \mapsto z + b & \text{ becomes } (z, t) \mapsto (z + b, t). \end{aligned}$$

The inversion  $z \mapsto z^{-1}$  is most easily dealt with as the composition of two anti-Möbius transformations:  $z \mapsto \bar{z}$  (reflection in a line) and  $z \mapsto z/|z|^2 = \bar{z}^{-1}$  (reflection in the unit circle). Extending to reflections in a vertical plane and the unit hemisphere, we get respectively  $(z, t) \mapsto (\bar{z}, t)$  and

$$\vec{x} \mapsto \frac{\vec{x}}{|\vec{x}|^2} \quad \text{or} \quad (z, t) \mapsto \left( \frac{z}{|z|^2 + t^2}, \frac{t}{|z|^2 + t^2} \right).$$

Therefore,

$$z \mapsto \frac{1}{z} \quad \text{becomes} \quad (z, t) \mapsto \left( \frac{\bar{z}}{|z|^2 + t^2}, \frac{t}{|z|^2 + t^2} \right).$$

Composing the building blocks we find that the extension of  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  is

$$(z, t) \mapsto \left( -\frac{\overline{z + d/c}}{c^2(|z + d/c|^2 + t^2)} + \frac{a}{c}, \frac{t}{|c|^2(|z + d/c|^2 + t^2)} \right) \quad \text{when } c \neq 0.$$

$$(z, t) \mapsto \left( \frac{a}{d}(z + b/a), \left| \frac{a}{d} \right| t \right) \quad \text{when } c = 0.$$

### 1.2 Hyperbolic geometry

In the euclidean plane, there is exactly one line through a given point and not meeting a given line disjoint from the point; this is the famous fifth postulate of Euclid. It gradually became clear in the nineteenth century that one can have a self-consistent and interesting geometry where this postulate is not valid — where “parallel” lines are not unique and indeed exist in uncountable abundance. This became known as *hyperbolic geometry*. Though the name was bestowed in connection with conics and projective geometry [Klein 1871, p. 72], it is a doubly felicitous choice, because the Greeks had named the hyperbola after the word for excess (compare “hyperbole”,

from the same Greek word). Hyperbolic geometry certainly has an excess of lines — and of “room” — compared to euclidean geometry!

Here are some of the salient features that distinguish hyperbolic geometry from the familiar euclidean and spherical geometry.

- (i) The angle sum  $\Sigma$  of a hyperbolic triangle  $\Delta$  satisfies  $0 < \Sigma < \pi$ ; in fact,  $\Sigma$  equals  $\pi - \text{area } \Delta$ . The limiting case  $\Sigma = 0$  is achieved by *ideal triangles* whose vertices are “at infinity”: we will have more to say about such *ideal vertices* soon (page 14). At the other extreme, the case  $\Sigma = \pi$  is the limiting case of hyperbolic triangles of very small area. Indeed, on the infinitesimal scale, hyperbolic geometry is euclidean.
- (ii) There are no similarities in hyperbolic space — one cannot scale a figure up or down without changing its angles and shape. It follows, for instance, that all hyperbolic triangles with the same angles are isometric (hyperbolic triangles are “rigid”), and also that the choice of a unit of length is not arbitrary, as in euclidean space; one can privilege a unit having some special property, say the side length of an equilateral triangle whose vertex angles are  $\pi/4$ .
- (iii) For any  $0 \leq \theta < \pi/(n-2)$  there is a regular  $n$ -sided hyperbolic polygon with vertex angles  $\theta$ . More generally, a necessary and sufficient condition for the existence of an  $n$ -sided convex polygon with vertex angles  $\theta_i$  (with  $0 \leq \theta_i < \pi$ ) in clockwise order is that  $\sum \theta_i < (n-2)\pi$ . The polygon is uniquely determined up to isometry and its area is  $(n-2)\pi - \sum \theta_i$ .
- (iv) Two convex hyperbolic polyhedra that are combinatorially the same with the same dihedral angles and valence 3 at all vertices are isometric [Rivin 1996; Bobenko and Springborn 2004].
- (v) The hyperbolic volume  $V$  of a ball and the surface area  $S$  of its bounding sphere grow exponentially with the hyperbolic radius  $\rho$ . The ratio of the surface area to the volume approaches 2 as  $\rho \rightarrow \infty$ .

In short, in the hyperbolic plane and space there are more geometric shapes, they have a tendency toward rigidity, and there is a lot more space in which to build them — in the estimate of Dick Canary, a baseball game played in the hyperbolic plane would require more than  $10^{100}$  ballplayers to provide the same level of outfield coverage as in euclidean space!

Most 2-dimensional abstract surfaces and 3-dimensional manifolds can be modeled using hyperbolic geometry, but not euclidean or spherical geometry. Hyperbolic space is a good place to embed exponentially growing graphs, like a graph representing interconnected web sites. In fact PARC has patented an algorithm for laying out such graphs in  $\mathbb{H}^2$  [Lamping et al. 1995]. A different, unpatented, algorithm for laying out graphs in  $\mathbb{H}^3$  is presented in [Munzner 1997]. The change of focus from one site to another is effected by a hyperbolic isometry.

By studying the ancient microwave radiation that pervades the universe, astrophysicists hope to get clues about the topology and large-scale curvature of our cosmic

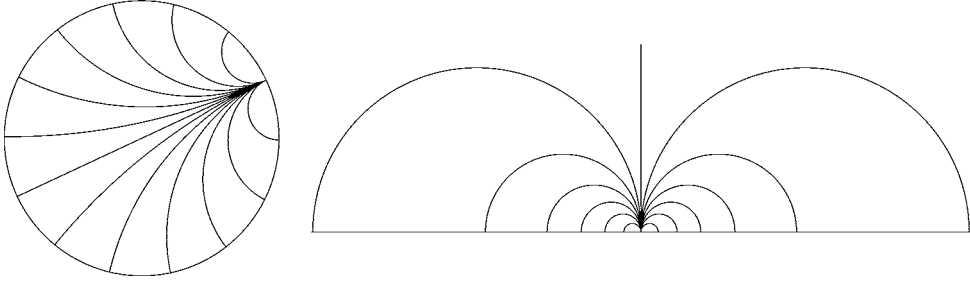


Fig. 1.3. Disk and upper half-plane models of  $\mathbb{H}^2$  showing the same geodesics.

home. An earlier proposal that we live in a hyperbolic universe appears to be incompatible with recent data from the Wilkinson Microwave Anisotropy Probe (WMAP), which found the total density (matter plus vacuum energy) to have essentially the value expected for flat space. To the extent that there may be deviation, it is toward a spherical universe (positive curvature); see the discussion in [Weeks 2004]. If the universe is a closed manifold with positive curvature, it can have one of only a few topological types. \* To establish that the universe is not simply connected would be astounding!

We now discuss the most commonly used *models* of the hyperbolic plane and of hyperbolic space. These are subsets of  $\mathbb{R}^n$  with appropriate riemannian metrics.

### *The hyperbolic plane*

The *upper half-plane model* is  $\{z \in \mathbb{C} : \text{Im } z > 0\}$  with the metric

$$ds = \frac{|dz|}{\text{Im } z}.$$

Here  $\text{Im } z$  is the notation for the imaginary part. The *unit disk model* is  $\{z \in \mathbb{C} : |z| < 1\}$  with the metric

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

The two models are equivalent under any Möbius transformation that maps the upper half-plane onto the unit disk. We will denote either one of these models by  $\mathbb{H}^2$ , the notation for the *hyperbolic plane*. These models have the following properties.

- (i) The metrics are *infinitesimally euclidean*; at each point they equal a rescaled euclidean metric. Thus the angle between two curves in the disk or upper half-plane is the same whether measured in the hyperbolic or the euclidean geometry;

\* For example, it might conceivably be Poincaré dodecahedral space, the famous first example found by Henri Poincaré of a closed manifold with zero homology which is not homeomorphic to  $\mathbb{S}^3$ . He had initially believed that such a manifold must be  $\mathbb{S}^3$ ; the example led him to the Poincaré Conjecture. A good explanation of this space and of the classification of spherical three-manifolds can be found in [Thurston 1997].



as a result these models are often called *conformal*. (For other models see Exercise 1-25 and following.)

- (ii)  $\mathbb{H}^2$  is *complete* in its metric. Every arc tending to the boundary has infinite length.
- (iii) The metrics are invariant under any Möbius transformation that maps the model onto itself. In fact these transformations comprise the full group of orientation preserving isometries of the model.
- (iv) The hyperbolic lines (geodesics) in the upper half-plane model are semicircles orthogonal to  $\mathbb{R}$  and vertical half lines. In the disk model they are diameters and circular arcs orthogonal to  $\{|z| = 1\}$ .

### Hyperbolic space

The *upper half-space model* is  $\{(z, t) : z \in \mathbb{C}, t > 0\}$  with the metric

$$ds = \frac{|d\vec{x}|}{t}, \quad |d\vec{x}|^2 = |dz|^2 + dt^2.$$

The *ball model* is  $\{\vec{x} \in \mathbb{R}^3 : |\vec{x}| < 1\}$  with the metric

$$ds = \frac{2|d\vec{x}|}{1 - |\vec{x}|^2}.$$

The two models are equivalent by a Möbius transformation that maps one to the other. Stereographic projection extends to such a Möbius transformation (Exercise 1-11). We will refer to either of these models with its metric as *hyperbolic space* and denote it by  $\mathbb{H}^3$ .

We repeat our list of properties:

- (i) The metrics are infinitesimally euclidean and correctly represent the angles in  $\mathbb{H}^3$ .
- (ii)  $\mathbb{H}^3$  is complete in its metric.
- (iii) The metrics are invariant under any Möbius transformation that maps the model onto itself. These transformations form the full group of orientation preserving isometries of the models.

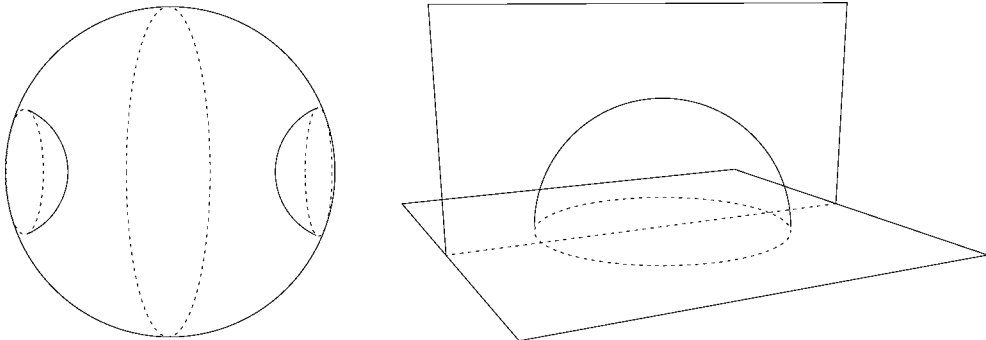


Fig. 1.4. Ball and upper half-space model of  $\mathbb{H}^3$  showing geodesic planes.

- (iv) The hyperbolic planes in the upper half-space model are hemispheres orthogonal to  $\mathbb{C}$  and vertical euclidean half-planes. The lines (geodesics) are semicircles orthogonal to  $\mathbb{C}$  and vertical euclidean half-lines. In the ball model the hyperbolic planes are spherical caps orthogonal to the unit sphere, and equatorial planes. The lines are circular arcs orthogonal to the unit sphere, and euclidean diameters.

Restricting the hyperbolic metric to a hyperbolic plane in the model yields the 2-dimensional hyperbolic metric on that plane. Particular cases are the vertical half-plane rising from  $\mathbb{R}$  in the upper half-space model and the equatorial plane in the ball model, where the restriction of the metrics give rise to our models of  $\mathbb{H}^2$ .

*Proof of property (iii).* For the proof that the Möbius transformations are orientation preserving isometries of the models, see Exercises 1-9 and 1-12. Here we show that there are no other such isometries, concentrating on the hyperbolic plane.

Given three positive distances  $d_1, d_2, d_3$  satisfying the triangle inequality, and a point  $z$  on an oriented line  $\ell \in \mathbb{H}^2$ , there are exactly two triangles with a vertex at  $z$ , a side of length  $d_1$  lying on the positive side of  $\ell$ , a side of length  $d_2$  sharing the vertex  $z$ , and a third side of length  $d_3$ . They are reflections of each other in  $\ell$  and one of the two is uniquely determined if an ordering of the vertices is given and required to give the positive orientation of the triangle they bound.

Given an orientation preserving isometry  $T$ , the  $T$ -images of three points not on a line are not on a hyperbolic line either. There is a Möbius transformation  $A$  such that  $A \circ T$  fixes the three points. It then pointwise fixes the sides of the triangle they determine, and then fixes the whole triangle  $\Delta$ . That is,  $T(z) = A^{-1}(z)$ , for  $z \in \Delta$ . If  $\Delta'$  is a triangle sharing an edge with  $\Delta$ , there is Möbius transformation  $A_1$  such that  $T(z) = A_1^{-1}(z)$  on  $\Delta'$ . Necessarily  $A_1 = A$ . Continuing on, building up the whole plane  $\mathbb{H}^2$  by adding in succession adjacent triangles, we conclude that  $T \equiv A$ .  $\square$

*Proof of property (iv).* In view of (iii) we need only prove that the vertical axis  $\ell$  is itself a geodesic. We will work in the upper half-space model. Let  $\ell$  denote the vertical axis rising from  $z = 0$ . Given  $\vec{x} = (z, t) \in \mathbb{H}^3$ , define the map  $r : \mathbb{H}^3 \rightarrow \ell$  as  $r(\vec{x}) = (0, t)$ . This map is called a *retraction* since in the hyperbolic distance  $d(r(\vec{x}), r(\vec{y})) \leq d(\vec{x}, \vec{y})$ . There is equality if and only if both  $\vec{x}, \vec{y}$  lie on a vertical line. This is an immediate consequence of the differential inequality

$$ds^2 = \frac{dx^2 + dy^2 + dt^2}{t^2} \geq \frac{dt^2}{t^2}.$$

Now suppose  $\gamma(u)$ , with  $0 \leq u \leq 1$ , is a differentiable path both of whose endpoints lie on  $\ell$ . Its length strictly exceeds the length of  $r(\gamma)$ , unless the path is the segment on  $\ell$  between its endpoints. That is,  $\ell$  is a geodesic: the unique shortest path between two points lying on  $\ell$  is the segment of  $\ell$  between the two points. Therefore all images of  $\ell$  by the isometries are also geodesics. In particular, through any two points there passes a unique geodesic.

Likewise the vertical half-plane resting on  $\mathbb{R}$  is a hyperbolic plane: the geodesic through any two points of the plane also lies in the plane. Therefore the totality of