11
Hyperbolic geometry

One of the most remarkable discoveries of nineteenth century mathematics is that the pseudosphere discussed in Section 8.3 has a geometry that closely resembles Euclidean geometry, with geodesics playing the role of straight lines. In fact, the closest correspondence with Euclidean geometry is obtained by 'embedding' the pseudosphere in a larger geometry, which is called *hyperbolic* or *non-Euclidean* geometry. When this is done, we find that all the axioms of Euclidean geometry hold in hyperbolic geometry, except the so-called 'parallel axiom': this states that if $\mathbf{p}$ is a point that is not on a straight line l, there is a unique straight line passing through $\mathbf{p}$ that does not intersect l (i.e., which is 'parallel' to l in the usual sense).

Hyperbolic geometry was discovered independently and almost simultaneously by the Hungarian mathematician Janos Bolyai and the Russian Nicolai Lobachevsky, although the formulations of it that we shall describe in this chapter are due to Eugenio Beltrami, Felix Klein and Henri Poincaré. David Hilbert, one of the greatest mathematicians of the twentieth century, wrote that the discovery of non-Euclidean geometry was 'one of the two most suggestive and notable achievements of the last century'. It ended centuries of attempts by Greek, Arab and later Western mathematicians to deduce the parallel axiom from the other axioms of Euclidean geometry, and it profoundly changed our view of what geometry is.

Andrew Pressley, *Elementary Differential Geometry: Second Edition*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-1-84882-891-9_11,

11.1 Upper half-plane model

We saw in Example 9.3.3 that if the pseudosphere is parametrized as

$$\tilde{\boldsymbol{\sigma}}(v,w) = \left(\frac{1}{w}\cos v, \frac{1}{w}\sin v, \sqrt{1-\frac{1}{w^2}} - \cosh^{-1} w\right),$$

where we must have $w > 1$ for $\tilde{\boldsymbol{\sigma}}$ to be well defined and smooth, its geodesics correspond to arcs of circles and straight lines in the vw-plane that intersect the v-axis perpendicularly. The line $w = 1$ appears to be a rather artificial boundary in the vw-plane, since the geodesics are well defined in the entire region $w > 0$. On the other hand, the line $w = 0$ is a 'real' boundary since the first fundamental form

$$\frac{dv^2 + dw^2}{w^2} \tag{11.1}$$

of the pseudosphere is undefined when $w = 0$. It is therefore natural to ask if there is a surface that corresponds to the whole of the half-plane $w > 0$ with this first fundamental form. In fact, there is no such surface for a celebrated theorem of Hilbert shows that there is no surface with constant negative Gaussian curvature that is 'geodesically complete', i.e., a surface for which all geodesics can be extended indefinitely in both directions (see Exercise 10.1.3).

One possible response to Hilbert's theorem is essentially to ignore it: all those properties of surfaces that depend only on the first fundamental form (lengths, angles, areas, geodesics, local isometries, ...) can still be studied for the half-plane

$$\mathcal{H} = \{(v,w) \in \mathbb{R}^2 \,|\, w > 0\}$$

equipped with the first fundamental form (11.1). They will then be called hyperbolic lengths, hyperbolic angles, etc. (We shall see the reason for the adjective 'hyperbolic' later.)

It will often be convenient to identify $\mathbb{R}^2$ with the set of complex numbers $\mathbb{C}$ via $(v,w) \leftrightarrow v + iw$, so that

$$\mathcal{H} = \{z \in \mathbb{C} \,|\, \mathfrak{Im}(z) > 0\}$$

is the set of complex numbers with positive imaginary part.

The first thing to observe is that $\mathcal{H}$ is a 'conformal model':

Proposition 11.1.1

Hyperbolic angles in $\mathcal{H}$ are the same as Euclidean angles.

Proof

This is just because the first fundamental form (11.1) of $\mathcal{H}$ is a multiple of $dv^2 + dw^2$ (see Corollary 6.3.4). $\square$

The 'hyperbolic lines' are the geodesics in $\mathcal{H}$, which were determined in Example 9.3.3.

Proposition 11.1.2

The geodesics in $\mathcal{H}$ are the half-lines parallel to the imaginary axis and the semicircles with centres on the real axis.

Here are some simple properties of hyperbolic lines.

Proposition 11.1.3

(i) There is a unique hyperbolic line passing through any two distinct points of $\mathcal{H}$.

(ii) The parallel axiom does not hold in $\mathcal{H}$.

In the following proof, and later in this chapter, 'lines' and 'circles' will mean Euclidean lines and circles ('hyperbolic line' means 'geodesic'). On the other hand, 'lengths' and 'angles' will always mean hyperbolic lengths and angles, unless explicitly stated otherwise.

Proof

(i) Let $a, b \in \mathcal{H}$, $a \neq b$. If the line passing through a and b is parallel to the imaginary axis, the unique hyperbolic line passing through the points a and b is the half-line containing them. If the line through a and b is not parallel to the imaginary axis, its perpendicular bisector intersects the real axis at some point c, say, and the unique hyperbolic line passing through a and b is the semicircle with centre c and radius $|a - c| = |b - c|$.
(ii) Take l to be the imaginary axis and let $a \in \mathcal{H}$ be any point not on l. For definiteness, assume that the real part $\mathfrak{Re}(a) > 0$. Then, the perpendicular bisector of the line joining a to the origin intersects the real axis at some point $b > 0$. Let c be a real number greater than b; then the semicircle with centre c passing through a is a hyperbolic line in $\mathcal{H}$ that does not intersect l. Of course,

the half-line through a parallel to the imaginary axis is another hyperbolic line with the same property (it can be regarded as the limiting case when $c \to \infty$). $\square$

Since there is a unique hyperbolic line passing through any two points $a, b \in \mathcal{H}$, it makes sense to define the *hyperbolic distance* $d_{\mathcal{H}}(a, b)$ between a and b to be the length of the hyperbolic line segment joining them. It is shown in Exercise 11.2.1 that this is actually the hyperbolic length of the shortest curve joining a and b.

Proposition 11.1.4

The hyperbolic distance between two points $a, b \in \mathcal{H}$ is

$$d_{\mathcal{H}}(a, b) = 2 \tanh^{-1} \frac{|b - a|}{|b - \bar{a}|}.$$

In this formula, $\bar{a}$ denotes the complex conjugate of the complex number a. The appearance of the hyperbolic tangent gives an indication of the reason why the geometry of $\mathcal{H}$ is called 'hyperbolic geometry'.

Proof

There are two cases, depending on whether the hyperbolic line joining a and b is a semicircle or a half-line. We shall deal with the semicircle case, leaving the simpler case of the half-line to Exercise 11.1.2.

Suppose then that a and b lie on the semicircle with centre c on the real axis and radius r. The semicircle can be parametrized by

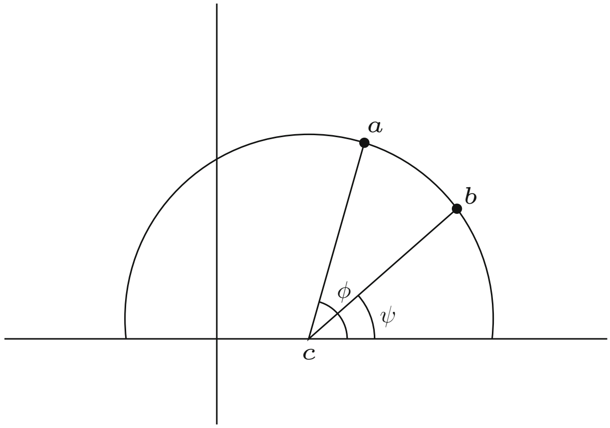
$$v = c + r \cos\theta, \quad w = r \sin\theta.$$

Writing d for $d_{\mathcal{H}}(a, b)$ and denoting $d/d\theta$ by a dot, we have

$$d = \int_{\varphi}^{\psi} \sqrt{\frac{\dot{v}^2 + \dot{w}^2}{w^2}}\, d\theta = \int_{\varphi}^{\psi} \sqrt{\frac{r^2 \sin^2\theta + r^2 \cos^2\theta}{r^2 \sin^2\theta}}\, d\theta = \int_{\varphi}^{\psi} \frac{d\theta}{\sin\theta},$$

where $\varphi = \arg(a - c)$, $\psi = \arg(b - c)$ (note that d is independent of the radius of the semicircle). Hence,

$$d = \ln \frac{\tan\frac{\psi}{2}}{\tan\frac{\varphi}{2}}.$$



Now,

$$\tanh\frac{d}{2} = \frac{e^d - 1}{e^d + 1} = \frac{\tan\frac{\psi}{2} - \tan\frac{\varphi}{2}}{\tan\frac{\psi}{2} + \tan\frac{\varphi}{2}} = \frac{\sin\frac{\psi}{2}\cos\frac{\varphi}{2} - \cos\frac{\psi}{2}\sin\frac{\varphi}{2}}{\sin\frac{\psi}{2}\cos\frac{\varphi}{2} + \cos\frac{\psi}{2}\sin\varphi 2} = \frac{\sin\frac{\psi-\varphi}{2}}{\sin\frac{\psi+\varphi}{2}}. \tag{11.2}$$

On the other hand,

$$\begin{aligned} |b-a|^2 &= r^2((\cos\psi - \cos\varphi)^2 + (\sin\psi - \sin\varphi)^2) \\ &= 2r^2(1 - \cos(\psi - \varphi)) = 4r^2 \sin^2\frac{\psi - \varphi}{2}, \end{aligned}$$

and similarly

$$|b - \bar{a}|^2 = 4r^2 \sin^2\frac{\psi + \varphi}{2}.$$

Combining the last two equations with Eq. 11.2 gives

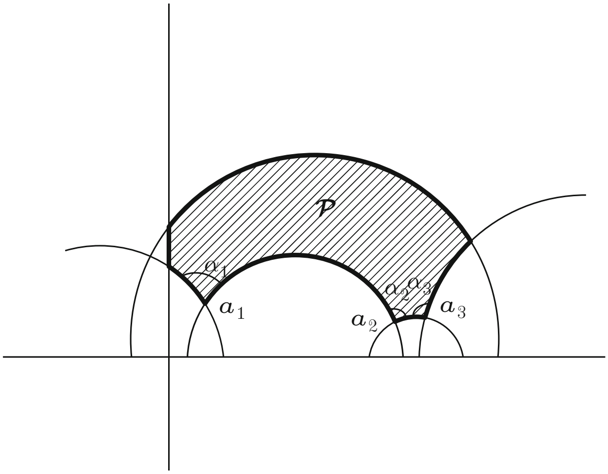
$$\tanh\frac{d}{2} = \frac{|b-a|}{|b-\bar{a}|}. \qquad \square$$

We conclude this section with another beautiful formula, this time for the area of a hyperbolic polygon, i.e., a polygon whose sides are hyperbolic lines.

Theorem 11.1.5

Let $\mathcal{P}$ be a n-sided hyperbolic polygon in $\mathcal{H}$ with internal angles $\alpha_1, \alpha_2, \ldots, \alpha_n$. Then, the hyperbolic area of the polygon is given by

$$\mathcal{A}(\mathcal{P}) = (n-2)\pi - \alpha_1 - \alpha_2 - \cdots - \alpha_n.$$



In particular, for a triangle with angles α, β, γ, the area is

$$\pi - \alpha - \beta - \gamma.$$

This should be compared with the well-known formula

$$\alpha + \beta + \gamma = \pi$$

for the sum of the angles of a Euclidean triangle with straight line sides, and the formula

$$\alpha + \beta + \gamma - \pi$$

for the area of a triangle on the unit sphere with geodesic (i.e., great circle) sides (Theorem 6.4.7).

Proof 11.1.5 Let $a_1, \ldots, a_n$ be the vertices of $\mathcal{P}$ and $\mathcal{C}$ its boundary, consisting of n hyperbolic line segments $a_1a_2, a_2a_3, \ldots, a_na_1$ (we assume that α_1 is the internal angle of $\mathcal{P}$ at the vertex a_1, etc.). Since the first fundamental form is $(dv^2 + dw^2)/w^2$, the area of $\mathcal{P}$ is

$$\int_{\mathcal{P}} \frac{dvdw}{w^2}.$$

We evaluate this integral by using Green's theorem (Section 3.2):

$$\int_{\mathcal{C}} pdv + qdw = \int_{\mathcal{P}} \left(\frac{\partial q}{\partial v} - \frac{\partial p}{\partial w} \right) dvdw,$$

where p and q are smooth functions of (v, w). Taking $p = 1/w$ and $q = 0$ gives

$$\int_{\mathcal{P}} \frac{dvdw}{w^2} = \int_{\mathcal{C}} \frac{dv}{w}. \tag{11.3}$$

To evaluate this integral we first prove the following lemma.

Lemma 11.1.6

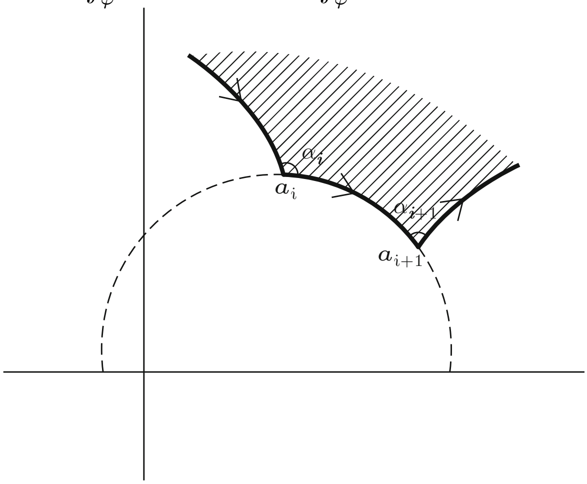
Let a and b be the endpoints of a segment l of a hyperbolic line in $\mathcal{H}$ that forms part of a semicircle with centre p on the real axis, and suppose that the radius vectors joining p to a and p to b make angles φ and ψ, respectively, with the positive real axis (see the diagram in the proof of Proposition 11.1.4). Then,

$$\int_l \frac{dv}{w} = \varphi - \psi.$$

Note that the integral is independent of the radius of the semicircle, and that the formula is correct even if the hyperbolic line is part of a half-line, for in that case the integral vanishes since v is constant along the hyperbolic line.

Proof 11.1.6 We parametrize the hyperbolic line by $v = r\cos\theta$, $w = r\sin\theta$, where r is the radius of the semicircle. Then, the integral is

$$\int_{\varphi}^{\psi} \frac{-r\sin\theta\, d\theta}{r\sin\theta} = -\int_{\varphi}^{\psi} d\theta = \varphi - \psi. \qquad \square$$



Returning to the proof of Theorem 11.1.5, let φ_i and ψ_i be the angles defined in the lemma corresponding to the side with endpoints a_i and a_{i+1}, for $i = 1, \dots, n$ (it is understood that a_{n+1} means a_1). By Eq. 11.3 and the lemma,

$$\int_{\mathcal{P}} \frac{dvdw}{w^2} = \sum_{i=1}^{n} (\varphi_i - \psi_i). \tag{11.4}$$

We can simplify this sum by considering the change in direction of the outward-pointing normal of $\mathcal{P}$ as we traverse its boundary in an anticlockwise direction. As we traverse the side with endpoints a_i and a_{i+1}, the outward normal rotates anticlockwise through an angle $\psi_i - \varphi_i$, while at the vertex a_i it rotates by $\pi - \alpha_i$. Hence, as we traverse the boundary of $\mathcal{P}$, the outward normal rotates through an angle

$$n\pi + \sum_{i=1}^{n} (\psi_i - \varphi_i - \alpha_i).$$

But this angle of rotation is 2π (cf. Theorem 3.1.4), so we have the equation

$$2\pi = n\pi + \sum_{i=1}^{n} (\psi_i - \varphi_i - \alpha_i).$$

Rearranging, we get

$$\sum_{i=1}^{n} (\varphi_i - \psi_i) = (n-2)\pi - \sum_{i=1}^{n} \alpha_i.$$

By Eq. 11.4, this is the desired area. $\square$

Note that the area of a hyperbolic triangle, i.e., a triangle whose sides are hyperbolic lines, depends only on its angles. We found in Proposition 6.5.8 that the same result holds in spherical geometry, but as we noted there this is completely different to the Euclidean situation, where we can change the size of a triangle (and hence its area) without changing its angles. In fact, we shall show in the next section that, as in spherical geometry (Exercise 6.5.2), two hyperbolic triangles with the same angles are *congruent*. But first we must discuss what congruence means in hyperbolic geometry.

EXERCISES

11.1.1 Show that, if l is a half-line geodesic in $\mathcal{H}$ and a is a point not on l, there are infinitely many hyperbolic lines passing through a that do not intersect l.

11.1.2 Complete the proof of Proposition 11.1.4 by dealing with the case in which the hyperbolic line passing through a and b is a half-line.

11.1.3 Show that for any $a \in \mathcal{H}$ there is a unique hyperbolic line passing through a that intersects the hyperbolic line l given by $v = 0$ perpendicularly. If b is the point of intersection, one calls $d_{\mathcal{H}}(a, b)$ the *hyperbolic distance of a from l.*

11.1.4 The *hyperbolic circle* $\mathcal{C}_{a,R}$ with centre $a \in \mathcal{H}$ and radius $R > 0$ is the set of points of $\mathcal{H}$ which are a hyperbolic distance R from a:

$$\mathcal{C}_{a,R} = \{z \in \mathcal{H} \mid d_{\mathcal{H}}(z, a) = R\}.$$

Show that $\mathcal{C}_{a,R}$ is a Euclidean circle.

Show that the Euclidean centre of $\mathcal{C}_{ic,R}$, where $c > 0$, is ib and that its Euclidean radius is r, where

$$c = \sqrt{b^2 - r^2}, \quad R = \frac{1}{2} \ln \frac{b+r}{b-r}.$$

Deduce that the hyperbolic length of the circumference of $\mathcal{C}_{ic,R}$ is $2\pi \sinh R$ and that the hyperbolic area inside it is $2\pi(\cosh R - 1)$. Note that these do not depend on c; in fact, it follows from the results of the next section that the circumference and area of $\mathcal{C}_{a,R}$ depend only on R and not on a (see the remarks preceding Theorem 11.2.4).

Compare these formulas with the case of a spherical circle in Exercise 6.5.3, and verify that they are consistent with Exercise 10.2.3.

11.2 Isometries of $\mathcal{H}$

In Euclidean plane geometry, two triangles are said to be *congruent* if one triangle can be moved until it coincides with the other. The types of motion that are allowed are combinations of rotations, translations, and reflections, i.e., the *isometries* of the plane (see Appendix 1). Similarly, to discuss congruence in spherical geometry, it was necessary in Section 6.5 to determine the isometries of the sphere.

It is easy to identify some isometries of $\mathcal{H}$:

(i) *Translations parallel to the real axis*, given by

$$T_a(z) = z + a, \qquad a \in \mathbb{R}.$$

(ii) *Reflections in lines parallel to the imaginary axis*, given by

$$R_a(z) = 2a - \bar{z}, \qquad a \in \mathbb{R}.$$

$R_a(z)$ is the 'reflection' of z in the line $\mathfrak{Re}(z) = a$, thought of as a mirror; each point of this line is fixed by R_a.

(iii) *Dilations by a factor* $a > 0$, given by

$$D_a(z) = az.$$

In terms of the parameters (v, w), these maps are given by $(v, w) \mapsto (v + a, w)$, $(v, w) \mapsto (2a - v, w)$ and $(v, w) \mapsto (av, aw)$, respectively, each of which obviously takes $\mathcal{H}$ to $\mathcal{H}$ and preserves the first fundamental form (11.1). But there is also a fourth type of isometry that is not quite as obvious:

(iv) *Inversions in circles with centres on the real axis.* The inversion in the circle with centre $a \in \mathbb{R}$ and radius $r > 0$ is

$$\mathcal{I}_{a,r}(z) = a + \frac{r^2}{\bar{z} - a}$$

(see Appendix 2).

To see that $\mathcal{I}_{a,r}$ is an isometry of $\mathcal{H}$, we consider first the case $a = 0$, $r = 1$, and denote $\mathcal{I}_{0,1}$ by $\mathcal{I}$. Then,

$$\mathcal{I}(v + iw) = \frac{v + iw}{v^2 + w^2},$$

which makes it clear that $\mathcal{I}$ takes any point in $\mathcal{H}$ to another point of $\mathcal{H}$ and any point on the real axis to another point on the real axis. To see that $\mathcal{I}$ is indeed an isometry of $\mathcal{H}$, we use the result of Exercise 6.1.4: if $\tilde{v} = \frac{v}{v^2+w^2}$ and $\tilde{w} = \frac{w}{v^2+w^2}$, then

$$d\tilde{v} = \frac{(w^2 - v^2)dv - 2vwdw}{(v^2 + w^2)^2}, \quad d\tilde{w} = \frac{-2vwdv + (v^2 - w^2)dw}{(v^2 + w^2)^2},$$

and hence

$$\begin{aligned}\frac{d\tilde{v}^2 + d\tilde{w}^2}{\tilde{w}^2} &= \frac{1}{w^2(v^2 + w^2)^2}\Big\{\big((w^2 - v^2)dv - 2vwdw\big)^2 \\ &\quad + \big(-2vwdv + (v^2 - w^2)dw\big)^2\Big\} \\ &= \frac{(w^2 - v^2)^2 + 4v^2w^2}{w^2(v^2 + w^2)^2}(dv^2 + dw^2) = \frac{dv^2 + dw^2}{w^2}.\end{aligned}$$

Returning to the general case, we note that

$$\begin{aligned}\mathcal{I}_{a,r}(z) &= T_a\left(\frac{r^2}{\bar{z}-a}\right) = T_a D_{r^2}\left(\frac{1}{\bar{z}-a}\right)\\ &= T_a D_{r^2}\mathcal{I}(z-a) = T_a D_{r^2}\mathcal{I}T_{-a}(z),\end{aligned}$$

so $\mathcal{I}_{a,r}$ is a composite $T_a \circ D_{r^2} \circ \mathcal{I} \circ T_{-a}$ of maps that are already known to be isometries of $\mathcal{H}$. Since any composite of isometries is an isometry, it follows that $\mathcal{I}_{a,r}$ is an isometry of $\mathcal{H}$.

We summarize our observations as follows:

Proposition 11.2.1

Any composite of a finite number of maps of the types (i)–(iv) defined above is an isometry of $\mathcal{H}$.

We shall call an isometry of one of the types (i)–(iv) an *elementary isometry* of $\mathcal{H}$. In fact, every isometry of $\mathcal{H}$ is a composite of a finite number of elementary isometries, but since we shall not make use of this result we leave its proof to the exercises.

Since isometries take geodesics to geodesics (Corollary 9.2.7), we know that the elementary isometries take half-lines and semicircles perpendicular to the real axis to other half-lines and semicircles perpendicular to the real axis. In fact, it is clear that translations, dilations and reflections take half-lines to half-lines and semicircles to semicircles, but the situation for inversions is a little more complicated:

Proposition 11.2.2

The inversion $\mathcal{I}_{a,r}$ in the circle with centre $a \in \mathbb{R}$ and radius $r > 0$ takes hyperbolic lines that intersect the real axis perpendicularly at a to half-lines, and all other hyperbolic lines to semicircles.

See Appendix 2 for the proof. The result is intuitively clear, since if a point of $\mathcal{H}$ "tends to" a point a its image under $\mathcal{I}_{a,r}$ "tends to infinity" (both limits in the Euclidean sense) and so cannot lie on a semicircle geodesic.

Isometries can be used to simplify the solution of many problems in hyperbolic geometry, by reducing the problem to a 'standard' situation. The basic result needed for this is

Proposition 11.2.3

Let l_1 and l_2 be hyperbolic lines in $\mathcal{H}$, and let z_1 and z_2 be points on l_1 and l_2, respectively. Then, there is an isometry of $\mathcal{H}$ that takes l_1 to l_2 and z_1 to z_2.

Proof

We observe first that it is enough to prove this result in the special case in which l_2 is the half-line l passing through the origin and $z_2 = i$. For if the proposition has been proved in this case, there is an isometry F_1 that takes l_1 to l and z_1 to i, and an isometry F_2 that takes l_2 to l and z_2 to i. Then, $F_2^{-1} \circ F_1$ is an isometry that takes l_1 to l_2 and z_1 to z_2.

There are now two cases depending on whether l_1 is a half-line or a semicircle. If l_1 is the half-line $v = a$, say, the translation T_{-a} takes l_1 to l and z_1 to some point ib, say, on l, where $b > 0$. Then, the dilation $D_{b^{-1}}$ takes l to itself and ib to i, and so the isometry we want is $D_{b^{-1}} \circ T_{-a}$.

Finally, suppose that l_1 is a semicircle, and let a be one of the two points in which it intersects the real axis. By Proposition 11.2.2, the inversion $\mathcal{I}_{a,1}$ takes l to a half-line geodesic l', say, and z_1 to some point z' on l'. By the preceding case, there is an isometry F that takes l' to l and z' to i, so the isometry we want is $F \circ \mathcal{I}_{a,1}$. □

As a simple application, we can now complete Exercise 11.1.4. If $a, b \in \mathcal{H}$, there is an isometry F of $\mathcal{H}$ that takes a to b. Then, F will clearly take the hyperbolic circle $\mathcal{C}_{a,R}$ to $\mathcal{C}_{b,R}$ for all $R > 0$. It follows that these hyperbolic circles have the same circumference and area.

Here is a more important application.

Theorem 11.2.4

In hyperbolic geometry, similar triangles are congruent.

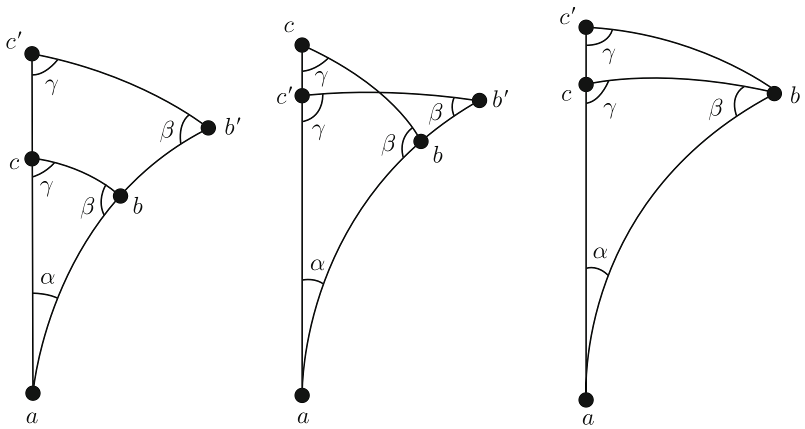
Proof

We have to prove that if we have two triangles T and T' with vertices a, b, c and a', b', c', and if the angle α of T at a is equal to that of T' at a', and similarly for the angles β at b and b' and for the angles γ at c and c', then there is an isometry F of $\mathcal{H}$ such that $F(a) = a'$, $F(b) = b'$ and $F(c) = c'$.

Let l, m, n and l', m', n' be the sides of T and T' (so that l is the side opposite the vertex a, etc.). It is enough to prove the theorem in the special case in which $a = a' = i$ and $m = m'$ is the imaginary axis. For by Proposition 11.2.3, there is an isometry G that takes a to i and m to the imaginary axis, and an isometry G' that takes a' to i and m' to the imaginary axis. If F is the desired isometry in the special case, then $(G')^{-1} \circ F \circ G$ is the desired isometry in the general case.

Assume then that $a = a' = i$ and $m = m'$ is the imaginary axis. By applying the reflection in the imaginary axis if necessary, we can further assume that b and b' are on the same side of the imaginary axis. Then either the hyperbolic lines n and n' coincide, or one is obtained from the other by applying the inversion $\mathcal{I}_{0,1}$ (which fixes m and the vertex i). Hence, we can assume that $n = n'$.

If now $b = b'$ and $c = c'$ the theorem is proved. If not, then we must be in one of the three situations shown below. By making use of Theorem 11.1.5, we shall prove that each of these situations is impossible.



In the first case, the angle sum of the quadrilateral with vertices b, c, c', b' is

$$(\pi - \beta) + (\pi - \gamma) + \gamma + \beta = 2\pi,$$

whereas by Theorem 11.1.5 the angle sum must be $< 2\pi$.

In the second case, the angle sum of the triangle with vertices d, b', b is

$$\delta + (\pi - \beta) + \beta,$$

where δ is the angle between l and l' at their intersection point d. This is $> \pi$, again contradicting Theorem 11.1.5.

Finally, in the third case the triangle with vertices b, c, c' has angle sum

$$\delta + (\pi - \gamma) + \gamma > \pi,$$

where δ is as in the preceding case (if c and c' are interchanged the argument is the same). $\square$

It follows from this theorem that there must be a formula for the lengths of the sides of a triangle in $\mathcal{H}$ in terms of its angles. Although we could prove such a formula now, it is slightly easier to establish it in a different model of hyperbolic geometry, and this is what we consider next.

EXERCISES

11.2.1 Show that if $a, b \in \mathcal{H}$, the hyperbolic distance $d_{\mathcal{H}}(a, b)$ is the length of the shortest curve in $\mathcal{H}$ joining a and b.

11.2.2 Show that, if l is any hyperbolic line in $\mathcal{H}$ and a is a point not on l, there are infinitely many hyperbolic lines passing through a that do not intersect l.

11.2.3 Let a be a point of $\mathcal{H}$ that is not on a hyperbolic line l. Show that there is a unique hyperbolic line m passing through a that intersects l perpendicularly. If b is the point of intersection of l and m, and c is any other point of l, prove that

$$d_{\mathcal{H}}(a, b) < d_{\mathcal{H}}(a, c).$$

Thus, b is the unique point of l that is closest to a.

11.2.4 This exercise and the next determine all the isometries of $\mathcal{H}$.

(i) Let F be an isometry of $\mathcal{H}$ that fixes each point of the imaginary axis l and each point of the semicircle geodesic m at the centre of the origin and radius 1. Show that F is the identity map.

(ii) Let F be an isometry of $\mathcal{H}$ such that $F(l) = l$ and $F(m) = m$, where l and m are as in (i). Prove that F is the identity map, the reflection R_0, the inversion $\mathcal{I}_{0,1}$ or the composite $\mathcal{I}_{0,1} \circ R_0$ (in the notation at the beginning of this section).

(iii) Show that every isometry of $\mathcal{H}$ is a composite of elementary isometries.

(iv) Show that every isometry of $\mathcal{H}$ is a composite of reflections and inversions in lines and circles perpendicular to the real axis.

11.2.5 A Möbius transformation (see Appendix 2) is said to be *real* if it is of the form

$$M(z) = \frac{az + b}{cz + d},$$

where $a, b, c, d \in \mathbb{R}$. Show that:

(i) Any composite of real Möbius transformations is a real Möbius transformation, and the inverse of any real Möbius transformation is a real Möbius transformation.

(ii) The Möbius transformations that take $\mathcal{H}$ to itself are exactly the real Möbius transformations such that $ad - bc > 0$.

(iii) Every real Möbius transformation is a composite of elementary isometries of $\mathcal{H}$, and hence is an isometry of $\mathcal{H}$.

(iv) If $J(z) = -\bar{z}$ and M is a real Möbius transformation, $M \circ J$ is an isometry of $\mathcal{H}$.

(v) If we call an isometry of type (iii) or (iv) a *Möbius isometry*, any composite of Möbius isometries is a Möbius isometry;

(vi) Every isometry of $\mathcal{H}$ is a Möbius isometry.

11.3 Poincaré disc model

We now consider a model of hyperbolic geometry based on the unit disc in the complex plane. Poincaré used this model to bring hyperbolic geometry into the mainstream of mathematics by establishing its connections with other areas, notably complex analysis and number theory.

We consider the transformation

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

It defines a bijection between the complex plane with the point $-i$ removed and the complex plane with the point 1 removed, its inverse being

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

In particular, $\mathcal{P}$ is well defined at all points of $\mathcal{H}$ and its boundary the real axis.

Let us determine the image of $\mathcal{H}$ under $\mathcal{P}$. We have,

$$\mathcal{P}(v + iw) = \frac{v + i(w - 1)}{v + i(w + 1)},$$

so

$$|\mathcal{P}(v+iw)| = \left(\frac{v^2+w^2+1-2w}{v^2+w^2+1+2w}\right)^{1/2}.$$

Hence, $|\mathcal{P}(v+iw)|$ is <1 if $w>0$, is $=1$ if $w=0$ and is >1 if $w<0$. Thus, $\mathcal{P}$ takes $\mathcal{H}$ to the unit disc

$$\mathcal{D} = \{z \in \mathbb{C} \,|\, |z| < 1\},$$

and the real axis to the boundary of $\mathcal{D}$, i.e., the unit circle $\mathcal{C}$ given by $|z| = 1$.

Definition 11.3.1

The Poincaré disc model $\mathcal{D}_P$ of hyperbolic geometry is the disc $\mathcal{D}$ equipped with the first fundamental form for which $\mathcal{P} : \mathcal{H} \to \mathcal{D}_P$ is an isometry.

Proposition 11.3.2

The first fundamental form of $\mathcal{D}_P$ is

$$\frac{4(dv^2+dw^2)}{(1-v^2-w^2)^2}.$$

In particular, $\mathcal{D}_P$ is a conformal model of hyperbolic geometry.

Proof

If $\tilde{v}+i\tilde{w} = \mathcal{P}^{-1}(v+iw)$, we find that

$$\tilde{v} = \frac{-2w}{(v-1)^2+w^2}, \quad \tilde{w} = \frac{1-v^2-w^2}{(v-1)^2+w^2},$$

which gives

$$d\tilde{v} = \frac{4(v-1)wdv - 2((v-1)^2-w^2)dw}{((v-1)^2+w^2)^2},$$
$$d\tilde{w} = \frac{2((v-1)^2-w^2)dv + 4(v-1)wdw}{((v-1)^2+w^2)^2}.$$

Hence,

$$d\tilde{v}^2 + d\tilde{w}^2 = \frac{16(v-1)^2w^2 + 4((v-1)^2-w^2)^2}{((v-1)^2+w^2)^4}(dv^2+dw^2) = \frac{4(dv^2+dw^2)}{((v-1)^2+w^2)^2}$$

and so

$$\frac{d\tilde{v}^2 + d\tilde{w}^2}{\tilde{w}^2} = \frac{4(dv^2 + dw^2)}{(1 - v^2 - w^2)^2}.$$

Since the first fundamental form of $\mathcal{D}_P$ is a multiple of $du^2 + dv^2$, $\mathcal{D}_P$ is a conformal model. □

Since $\mathcal{P} : \mathcal{H} \to \mathcal{D}_P$ is an isometry, it follows that the isometries of $\mathcal{D}_P$ are exactly the maps

$$\mathcal{P} \circ F \circ \mathcal{P}^{-1},$$

where F is any isometry of $\mathcal{H}$. Indeed, since any composite of isometries is an isometry, $\mathcal{P} \circ F \circ \mathcal{P}^{-1}$ is an isometry of $\mathcal{D}_P$ if F is an isometry of $\mathcal{H}$; conversely, if G is any isometry of $\mathcal{D}_P$, then $F = \mathcal{P}^{-1} \circ G \circ \mathcal{P}$ is an isometry of $\mathcal{H}$, and $G = \mathcal{P} \circ F \circ \mathcal{P}^{-1}$.

Here is a simple application of this observation:

Proposition 11.3.3

(i) Let Γ be a circle that intersects $\mathcal{C}$ perpendicularly. Then, inversion in Γ is an isometry of $\mathcal{D}_P$.

(ii) Let l be a line passing through the origin (and so perpendicular to $\mathcal{C}$). Then, (Euclidean) reflection in l is an isometry of $\mathcal{D}_P$.

Proof

For (i), let Γ have centre $a \in \mathbb{C}$ and radius $r > 0$; then, the inversion in Γ is given by

$$\mathcal{I}_{a,r} = a + \frac{r^2}{\bar{z} - \bar{a}}.$$

By Proposition A.2.8, $\mathcal{I}_{a,r}$ takes $\mathcal{D}_P$ to itself. We have to show that $\mathcal{P}^{-1} \circ \mathcal{I}_{a,r} \circ \mathcal{P}$ is an isometry of $\mathcal{H}$. We find that

$$\mathcal{I}_{a,r}(\mathcal{P}(z)) = \frac{(a - |a|^2 + r^2)\bar{z} + i(a + |a|^2 - r^2)}{(1 - \bar{a})\bar{z} + i(1 + \bar{a})}.$$

Now, since Γ intersects $\mathcal{C}$ at right angles, $|a|^2 = r^2 + 1$, so

$$\mathcal{I}_{a,r}(\mathcal{P}(z)) = \frac{(a - 1)\bar{z} + i(a + 1)}{(1 - \bar{a})\bar{z} + i(1 + \bar{a})}.$$

This leads to

$$\mathcal{P}^{-1}(\mathcal{I}_{a,r}(\mathcal{P}(z))) = \frac{i(a-\bar{a})\bar{z} - (2+a+\bar{a})}{(2-a-\bar{a})\bar{z} - i(a-\bar{a})}.$$

This is a real Möbius transformation (Exercise 11.2.5) and so is an isometry of $\mathcal{H}$.

For (ii), let l make an angle θ with the real axis, so that reflection in l is the map $\mathcal{R}(z) = e^{2i\theta}\bar{z}$. We find that

$$\mathcal{P}^{-1}(\mathcal{R}(\mathcal{P}(z))) = \frac{z\cos\theta + \sin\theta}{-z\sin\theta + \cos\theta},$$

which is again a real Möbius transformation. □

Note that simple isometries in one model may not correspond to simple isometries in the other. For example, it is clear from Proposition 11.3.2 that any rotation about the origin is an isometry of $\mathcal{D}_P$ (because such a rotation is an isometry of the Euclidean plane, and hence preserves $dv^2 + dw^2$ and $v^2 + w^2$), but the corresponding isometry of $\mathcal{H}$ is quite complicated (it is not an elementary isometry, for example).

Since $\mathcal{P}$ is an isometry, the geodesics (i.e., the hyperbolic lines) in $\mathcal{D}_P$ are the images under $\mathcal{P}$ of the geodesics in $\mathcal{H}$. Hence, the properties of the hyperbolic lines in $\mathcal{H}$ can be transferred to $\mathcal{D}_P$. For example, if a and b are two distinct points of $\mathcal{D}_P$, then by Proposition 11.1.3, there is a unique hyperbolic line l in $\mathcal{H}$ passing through the distinct points $\mathcal{P}^{-1}(a)$ and $\mathcal{P}^{-1}(b)$, so $\mathcal{P}(l)$ is the unique hyperbolic line in $\mathcal{D}_P$ passing through a and b. Similarly, Proposition 11.2.3 holds as stated with $\mathcal{H}$ replaced by $\mathcal{D}_P$.

The distance between two points of $\mathcal{D}_P$ is given by

$$d_{\mathcal{D}_P}(a,b) = d_{\mathcal{H}}(\mathcal{P}^{-1}(a), \mathcal{P}^{-1}(b)), \qquad a, b \in \mathcal{D}_P.$$

Using the formula in Proposition 11.1.4, it is straightforward (see Exercise 11.3.1) to prove

Proposition 11.3.4

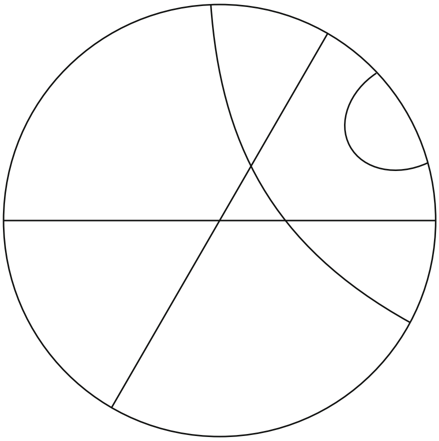
For $a, b \in \mathcal{D}_P$, we have

$$d_{\mathcal{D}_P}(a,b) = 2\tanh^{-1}\frac{|b-a|}{|1-\bar{a}b|}.$$

The explicit form of the hyperbolic lines in $\mathcal{D}_P$ can, of course, be determined from the first fundamental form in Proposition 11.3.2. But it is easier to make use of some simple properties of the map $\mathcal{P}$.

Proposition 11.3.5

The hyperbolic lines in $\mathcal{D}_P$ are the lines and circles that intersect $\mathcal{C}$ perpendicularly (see the diagram below).



Proof

This follows from Proposition 11.1.2 and the fact that $\mathcal{P}$ takes the boundary of $\mathcal{H}$ to that of $\mathcal{D}_P$ and, being a Möbius transformation, preserves (Euclidean) angles and takes lines and circles to lines and circles (see Appendix 2). $\square$

Note that Proposition 11.3.3 tells us that 'reflection' in any hyperbolic line in $\mathcal{D}_P$ is an isometry of $\mathcal{D}_P$ – 'reflection' in a circle being interpreted as inversion (and Exercise 11.3.5 shows that every isometry of $\mathcal{D}_P$ is a composite of such reflections).

We shall now establish some new properties of hyperbolic geometry to which the Poincaré model is particularly well suited, starting with the basic result in hyperbolic trigonometry.

Theorem 11.3.6

Consider a hyperbolic triangle with angles α, β, γ and sides of length A, B, C (so that A is the length of the side opposite α, etc.). Then,

$$\cosh C = \cosh A \cosh B - \sinh A \sinh B \cos \gamma,$$

and two analogous formulas can be obtained by applying the cyclic permutations $\alpha \to \beta \to \gamma \to \alpha$ and $A \to B \to C \to A$.

This formula is called the 'hyperbolic cosine rule' because it becomes the usual cosine rule when A, B, and C are small: using the approximations $\cosh A = 1 + \frac{1}{2}A^2$ and $\sinh A = A$, etc. we get

$$C^2 = A^2 + B^2 - 2AB\cos\gamma$$

(compare the spherical case treated in Proposition 6.5.3(i)).

Proof

Let a, b, and c be the vertices of the triangle, so that α is the angle at a, etc. By applying an isometry of $\mathcal{D}_P$ that takes c to the origin followed by a suitable rotation about the origin (i.e. another isometry), we can assume that $c = 0 \in \mathcal{D}_P$ and that $a > 0$. By Proposition 11.3.4,

$$a = \tanh\frac{1}{2}B, \quad b = e^{i\gamma}\tanh\frac{1}{2}A.$$

Now

$$\cosh A = \cosh^2\frac{1}{2}A + \sinh^2\frac{1}{2}A = \frac{1+\tanh^2\frac{1}{2}A}{\operatorname{sech}^2\frac{1}{2}A} = \frac{1+\tanh^2\frac{1}{2}A}{1-\tanh^2\frac{1}{2}A} = \frac{1+|a|^2}{1-|a|^2}$$

and by Proposition 11.3.4 again

$$\tanh\frac{1}{2}C = \frac{|b-a|}{|1-\bar{a}b|},$$

so

$$\begin{aligned}
\cosh C &= \frac{1+\tanh^2\frac{1}{2}C}{1-\tanh^2\frac{1}{2}C} \\
&= \frac{|1-\bar{a}b|^2 + |b-a|^2}{|1-\bar{a}b|^2 - |b-a|^2} \\
&= \frac{(1-\bar{a}b)(1-a\bar{b}) + (b-a)(\bar{b}-\bar{a})}{(1-\bar{a}b)(1-a\bar{b}) - (b-a)(\bar{b}-\bar{a})} \\
&= \frac{1+|a|^2+|b|^2+|a|^2|b|^2 - 2(\bar{a}b + a\bar{b})}{1-|a|^2-|b|^2+|a|^2|b|^2} \\
&= \frac{(1+|a|^2)(1+|b|^2) - 2(\bar{a}b+a\bar{b})}{(1-|a|^2)(1-|b|^2)} \\
&= \cosh A\cosh B - 4\cos\gamma\frac{\tanh\frac{1}{2}A\tanh\frac{1}{2}B}{(1-\tanh^2\frac{1}{2}A)(1-\tanh^2\frac{1}{2}B)} \\
&= \cosh A\cosh B - \sinh A\sinh B\cos\gamma,
\end{aligned}$$

using $\sinh A = 2\sinh\frac{1}{2}A\cosh\frac{1}{2}A$. $\square$

In particular, we have the hyperbolic analogue of Pythagoras' theorem:

Corollary 11.3.7

Suppose that a hyperbolic triangle has sides of lengths A, B, and C and that the angle opposite the side of length C is a right angle. Then,

$$\cosh C = \cosh A \cosh B.$$

Further results in hyperbolic trigonometry can be found in the exercises.

EXERCISES

11.3.1 Prove Proposition 11.3.4.

11.3.2 Let l and m be hyperbolic lines in $\mathcal{D}_P$ that intersect at right angles. Prove that there is an isometry of $\mathcal{D}_P$ that takes l to the real axis and m to the imaginary axis. How many such isometries are there?

11.3.3 Show that the Möbius transformations that take $\mathcal{D}_P$ to itself are those of the form

$$z \mapsto \frac{az+b}{\bar{b}z+\bar{a}}, \qquad |a| > |b|.$$

Recall (Exercise 6.5.4) that these are unitary Möbius transformations.

11.3.4 Show that the isometries of $\mathcal{D}_P$ are the transformations of the following two types:

$$z \mapsto \frac{az+b}{\bar{b}z+\bar{a}}, \qquad z \mapsto \frac{a\bar{z}+b}{\bar{b}\bar{z}+\bar{a}},$$

where a and b are complex numbers such that $|a| > |b|$. Note that this and the preceding exercise show that the isometries of $\mathcal{D}_P$ are exactly the Möbius and conjugate-Möbius transformations that take $\mathcal{D}_P$ to itself.

11.3.5 Prove that every isometry of $\mathcal{D}_P$ is the composite of finitely many isometries of the two types in Proposition 11.3.3.

11.3.6 Consider a hyperbolic triangle with vertices a, b, and c, sides of length A, B, and C and angles α, β, and γ (so that A is the length of the side opposite a and α is the angle at a, etc.). Prove the *hyperbolic sine rule*

$$\frac{\sin\alpha}{\sinh A} = \frac{\sin\beta}{\sinh B} = \frac{\sin\gamma}{\sinh C}.$$

11.3.7 With the notation in the preceding exercise, suppose that $\gamma = \pi/2$. Prove that:

(i) $\cos\alpha = \frac{\sinh B \cosh A}{\sinh C}$.

(ii) $\cosh A = \frac{\cos\alpha}{\sin\beta}$.

(iii) $\sinh A = \frac{\tanh B}{\tan\beta}$.

11.3.8 With the notation in Exercise 11.3.6, prove that

$$\cosh A = \frac{\cos\alpha + \cos\beta\cos\gamma}{\sin\beta\sin\gamma}.$$

This is the formula we promised at the end of Section 11.2 for the lengths of the sides of a hyperbolic triangle in terms of its angles.

11.3.9 Show that if $\mathbb{R}^2$ is provided with the first fundamental form

$$\frac{4(du^2 + dv^2)}{(1 + u^2 + v^2)^2},$$

the stereographic projection map $\Pi : S^2 \backslash \{\text{north pole}\} \to \mathbb{R}^2$ defined in Example 6.3.5 is an isometry. Note the similarity between this formula and that in Proposition 11.3.2: the plane with this first fundamental form provides a 'model' for the sphere in the same way as the half-plane with the first fundamental form in Proposition 11.3.2 is a 'model' for the pseudosphere.

11.4 Hyperbolic parallels

In Euclidean plane geometry, there are many equivalent criteria for two lines l and m to be parallel. For example:

(i) l and m do not intersect.

(ii) l and m have a common perpendicular line.

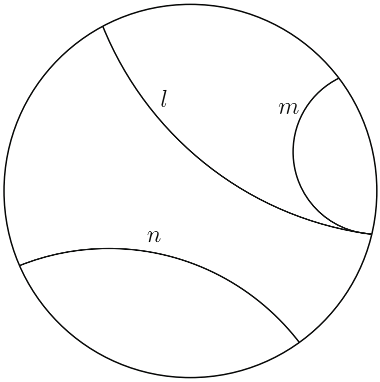
(iii) l and m are a constant distance apart.

(A fourth criterion is considered in Exercise 11.4.3.) In hyperbolic geometry, these conditions are *not* equivalent. In fact, two distinct hyperbolic lines are *never* a constant distance apart (see Exercise 11.4.2), so (iii) is not relevant to the discussion of parallels in hyperbolic geometry. Further, it is clear that in hyperbolic geometry (i) does not imply (ii) (consider two half-line geodesics in $\mathcal{H}$, for example), so we must distinguish two cases:

Definition 11.4.1

Let l and m be hyperbolic lines in $\mathcal{D}_P$ that do not intersect at any point of $\mathcal{D}_P$. If l and m intersect at a point of the boundary of $\mathcal{D}_P$ they are said to be *parallel*; otherwise they are said to be *ultra-parallel*.

In the diagram below, l and m are parallel, and l and n are ultra-parallel.



We have already noted (Proposition 11.1.3(ii)) that the parallel axiom does not hold in hyperbolic geometry. In fact, if a is a point that is not on a hyperbolic line l, there are infinitely many hyperbolic lines through a that do not intersect l (see Exercise 11.1.1). The following result shows that exactly two of these hyperbolic lines are parallel to l.

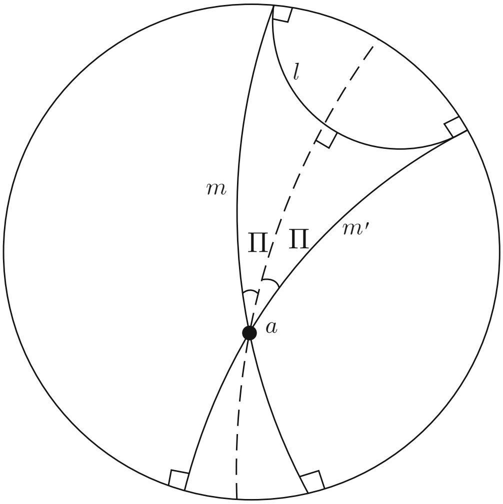
Proposition 11.4.2

Suppose that $a \in \mathcal{D}_P$ is a point not on a hyperbolic line l. Then, there are exactly two hyperbolic lines, say m and m', passing through a that are parallel to l. The angle between m and m' at a is 2Π, where

$$\sin \Pi = \operatorname{sech} d,$$

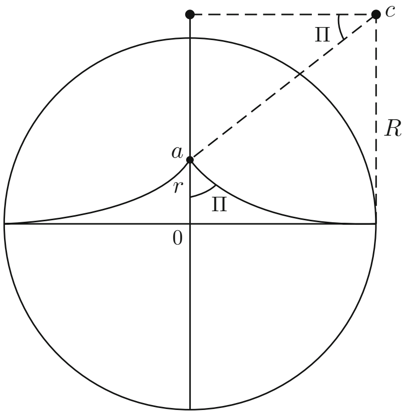
and d is the hyperbolic distance of a from l (Exercise 11.1.3). Moreover, a hyperbolic line through a intersects l if and only if it lies between m and m' on the same side of a as l, and the hyperbolic line through a perpendicular to l bisects the angle between m and m'.

The angle Π is called the *angle of parallelism*.



Proof

We first show that there is an isometry of $\mathcal{D}_P$ that takes l to the real axis and a to a point on the imaginary axis. In that case, all the assertions made in the proposition are clear, except for the formula for Π.



Let $\tilde{a} = \mathcal{P}^{-1}(a)$, $\tilde{l} = \mathcal{P}^{-1}(l)$. There is an isometry F of $\mathcal{H}$ that takes $\tilde{l}$ to the imaginary axis; let $b = F(\tilde{a})$. The isometry $D_{1/|b|}$ takes b to a point on the unit circle $v^2 + w^2 = 1$ and fixes the imaginary axis. Now note that $\mathcal{P}$ takes the imaginary axis in $\mathcal{H}$ to the real axis in $\mathcal{D}_P$ and the unit circle in $\mathcal{H}$ to the imaginary axis in $\mathcal{D}_P$.

We can therefore assume that l is the real axis and that $a = ir$ where $r = \tanh\frac{1}{2}d$ by Proposition 11.3.4. The circle m through a that touches the real axis at 1 has centre $c = 1 + iR$ and radius R for some $R > 0$, and so has equation

$$|z - 1 - iR| = R.$$

Since m passes through ir, we have $|-1 + i(r - R)| = R$, which gives

$$R = \frac{1 + r^2}{2r}.$$

In the right-angled (Euclidean) triangle with vertices a, iR and c, the hypotenuse is perpendicular to m, so the angle of the triangle at a is $\pi/2 - \Pi$ (see the diagram above). Hence, by Euclidean trigonometry,

$$R \sin \Pi = R - r$$

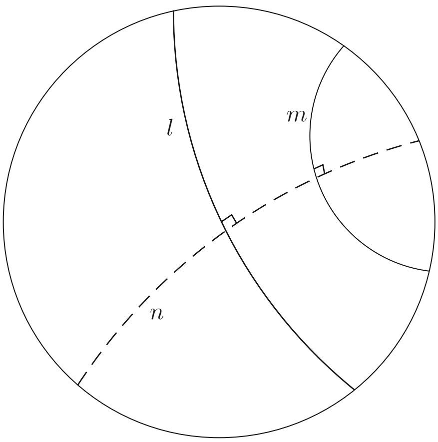
and we get

$$\sin \Pi = 1 - \frac{r}{R} = 1 - \frac{2r^2}{1 + r^2} = \frac{1 - r^2}{1 + r^2} = \frac{1 - \tanh^2 \frac{1}{2}d}{1 + \tanh^2 \frac{1}{2}d} = \frac{1}{\cosh d}. \qquad \square$$

As we mentioned above, one characterization of parallel lines in Euclidean plane geometry is that such lines have a common perpendicular. In hyperbolic geometry, this property characterizes ultra-parallels:

Proposition 11.4.3

Two hyperbolic lines in $\mathcal{D}_P$ are ultra-parallel if and only if they have a common perpendicular (i.e., a hyperbolic line that intersects them both at right-angles). In that case the common perpendicular is unique.



In Euclidean plane geometry, of course, two parallel lines have infinitely many common perpendiculars.

Proof

Suppose first that l and m are hyperbolic lines in $\mathcal{D}_P$ that have a common perpendicular n which intersects them at the points a and b. We can assume that l and n are the real and imaginary axes, respectively, and that a is the origin (see Exercise 11.3.2). Then m is part of a circle with centre at some point iR on the imaginary axis, where $|R| > 1$. Since m intersects $\mathcal{C}$ at right angles, the radius r of $\mathcal{C}$ satisfies

$$R^2 = r^2 + 1.$$

In particular, $|R| > r$, so m does not intersect the real axis. Hence, l and m are ultra-parallel.

Conversely, suppose that l and m are ultra-parallel. As before, we can assume that l is the real axis. Suppose that m is the circle with centre a and radius r; then, as above,

$$|a|^2 = r^2 + 1. \tag{11.5}$$

We claim that

$$-1 < \mathfrak{Re}(a) < 1. \tag{11.6}$$

Indeed, m intersects the real axis at a point v if and only if

$$|v - a| = r.$$

In view of (11.5), this gives

$$v^2 - 2v\,\mathfrak{Re}(a) + 1 = 0. \tag{11.7}$$

If $|\mathfrak{Re}(a)| > 1$, Eq. 11.7 has two distinct real roots whose product is equal to 1, hence one root v satisfies $-1 < v < 1$. This means that l and m intersect in $\mathcal{D}_P$, contrary to assumption. Similarly, if $|\mathfrak{Re}(a)| = 1$, Eq. 11.10 has ± 1 as a repeated root, so l touches m at 1 or -1 on the boundary of $\mathcal{D}_P$, again contrary to assumption. Hence, (11.6) must hold.

We now consider a circle with centre b on the real axis and radius s. This intersects both m and $\mathcal{C}$ at right angles if and only if

$$b^2 = s^2 + 1 \quad \text{and} \quad |b - a|^2 = r^2 + s^2.$$

If $\mathfrak{Re}(a) \neq 0$, these equations have the unique solution

$$b = \frac{1}{\mathfrak{Re}(a)}, \quad s = \sqrt{(\mathfrak{Re}(a)^{-2} - 1},$$

and the corresponding circle n is the unique common perpendicular to l and m. If $\mathfrak{Re}(a) = 0$, it is clear that the imaginary axis is the unique common perpendicular. □

EXERCISES

11.4.1 Which pairs of hyperbolic lines in $\mathcal{H}$ are parallel? Ultra-parallel?

11.4.2 Let l be the imaginary axis in $\mathcal{H}$. Show that, for any $R > 0$, the set of points that are a distance R from l is the union of two half-lines passing through the origin, but that these half-lines are *not* hyperbolic lines. This contrasts with the situation in Euclidean geometry, in which the set of points at a fixed distance from a line is a pair of lines.

11.4.3 Let a and b be two distinct points in $\mathcal{D}_P$, and let $0 < \mathcal{A} < \pi$. Show that the set of points $c \in \mathcal{D}_P$ such that the hyperbolic triangle with vertices a, b and c has area $\mathcal{A}$ is the union of two segments of lines or circles, but that these are not hyperbolic lines. Note that this equal-area property could be used to characterize lines in Euclidean geometry.

11.5 Beltrami–Klein model

The final model of non-Euclidean geometry that we shall discuss was actually the first to be introduced by Beltrami, but it was Klein who realised that this model could be used to unify non-Euclidean geometry with *projective geometry*, a subject that has been studied since antiquity. (We do not assume that the reader is familiar with projective geometry.)

The model is constructed by using two projections of the unit sphere S^2. We recall the stereographic projection map Π (Example 6.3.5) from S^2 to the xy-plane. This map defines a diffeomorphism from the 'southern hemisphere'

$$S^2_- = \{(x, y, z) \in S^2 \mid z < 0\}$$

to the unit disc

$$\mathcal{D} = \{(x, y, 0) \in \mathbb{R}^3 \mid x^2 + y^2 < 1\}.$$

We shall also need the 'vertical' projection of $\mathbb{R}^3$ onto the xy-plane:

$$\mathrm{pr}(x, y, z) = (x, y, 0).$$

This also defines a diffeomorphism from S^2_- to $\mathcal{D}$. Hence, the composite map

$$\mathcal{K} = \mathrm{pr} \circ \Pi^{-1} : \mathcal{D} \to \mathcal{D}$$

is a diffeomorphism. It is easy to see (Exercise 11.5.1) that, if we identify the xy-plane with $\mathbb{C}$ by $(x, y, 0) \mapsto x + iy$ as usual, then

$$\mathcal{K}(z) = \frac{2z}{|z|^2 + 1}, \qquad z \in \mathcal{D}. \tag{11.8}$$

Definition 11.5.1

The Beltrami-Klein model $\mathcal{D}_K$ of non-Euclidean geometry is the disc $\mathcal{D}$ equipped with the first fundamental form for which the diffeomorphism

$$\mathcal{K} : \mathcal{D}_P \to \mathcal{D}_K$$

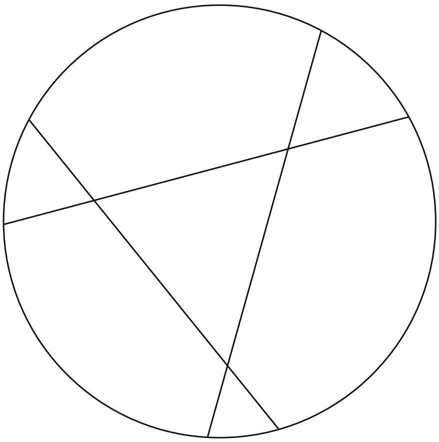
is an isometry.

We shall not need to know the first fundamental form of $\mathcal{D}_K$ explicitly (it was actually computed in Exercise 8.3.1(iii)).

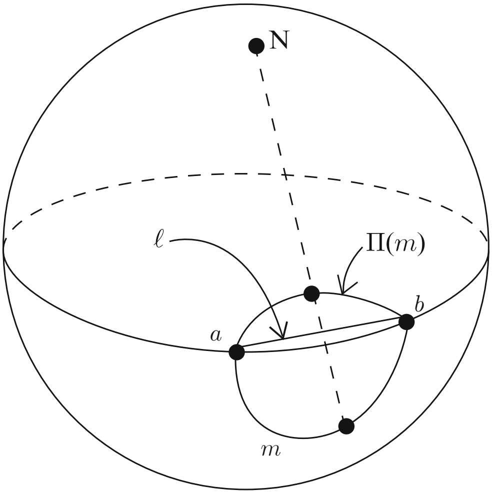
The Beltrami-Klein model has the following remarkable property.

Proposition 11.5.2

The hyperbolic lines in the Beltrami-Klein model are the (Euclidean) straight line segments contained in the disc $\mathcal{D}_K$.



Proof 11.5.2 Let l be the line segment joining points a and b on $\mathcal{C}$. The curve on S^2_- that corresponds to ℓ under the projection pr is the intersection of S^2 with the plane perpendicular to the xy-plane that contains ℓ. This is a semicircle m, say, that intersects $\mathcal{C}$ at right angles at a and b.



Since Π is a conformal map that takes circles on S^2 to lines and circles in the xy-plane (see Example 6.3.5 and Exercise 6.3.7), $\Pi\,(m)$ is an arc of a circle in $\mathcal{D}$ that intersects the boundary of $\mathcal{D}$ at right angles, in other words a hyperbolic line in $\mathcal{D}_P$. It follows that every line segment in $\mathcal{D}_K$ is a hyperbolic line. Since there is a line segment passing through any given point of $\mathcal{D}_K$ in any given direction, these must be all of the hyperbolic lines in $\mathcal{D}_K$ (see Proposition 9.2.4). □

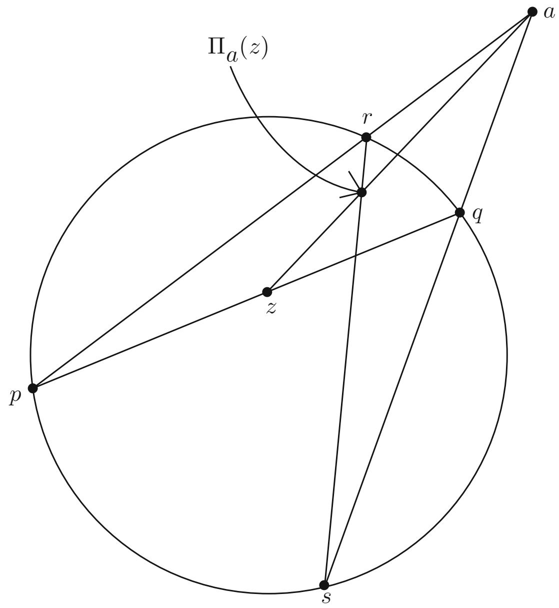
Corollary 11.5.3

$\mathcal{D}_K$ is not a conformal model of hyperbolic geometry.

Proof

Consider a hyperbolic triangle in $\mathcal{D}_K$. By Proposition 11.5.2 this is also a Euclidean triangle, so the sum of its internal Euclidean angles is π. But, by Theorem 11.1.5, the sum of its internal hyperbolic angles is $< \pi$. □

The isometries of $\mathcal{D}_K$ can, of course, be deduced from those of $\mathcal{D}_P$ by using the isometry $\mathcal{K}$. For example, any rotation about the origin is an isometry of $\mathcal{D}_K$. For, if ρ_θ is such a rotation by an angle θ, so that $\rho_\theta(z) = e^{i\theta}z$, it is clear from Eq. 11.8 that $\mathcal{K} \circ \rho_\theta \circ \mathcal{K}^{-1} = \rho_\theta$ and we know that ρ_θ is an isometry of $\mathcal{D}_P$ (see the remarks following the proof of Proposition 11.3.3). But to proceed further, it is more instructive to take a different, and more geometric, approach.



If $a \in \mathbb{C}$ and $|a| > 1$, define the *perspectivity*

$$\Pi_a : \mathcal{D}_K \to \mathcal{D}_K$$

with centre a as follows. Let $z \in \mathcal{D}_K$ and let l be any hyperbolic line in $\mathcal{D}_K$ passing through z. Thus, l is a (Euclidean) line segment that intersects $\mathcal{C}$ at two points, say p and q. Let the lines through a and p and through a and q intersect $\mathcal{C}$ again at r and s, respectively (if the line through a and p happens to be tangent to $\mathcal{C}$ at p, then $r = p$; and similarly for the line through a and q). Then, $\Pi_a(z)$ is defined to be the point of intersection of the line through a and z with the line through r and s (see the diagram above).

Of course, it is not obvious that this definition makes sense, i.e., that $\Pi_a(z)$ depends only on z (and a) and not on the choice of the line l, but this follows from

Proposition 11.5.4

With the above notation,

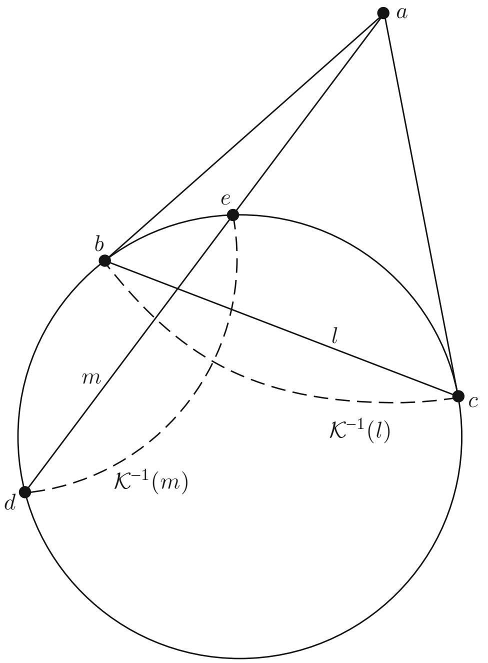
$$\Pi_a = \mathcal{K} \circ \mathcal{I}_{a,r} \circ \mathcal{K}^{-1},$$

where $r = \sqrt{|a|^2 - 1}$. In particular, Π_a is an isometry of $\mathcal{D}_K$.

To prove this we need

Lemma 11.5.5

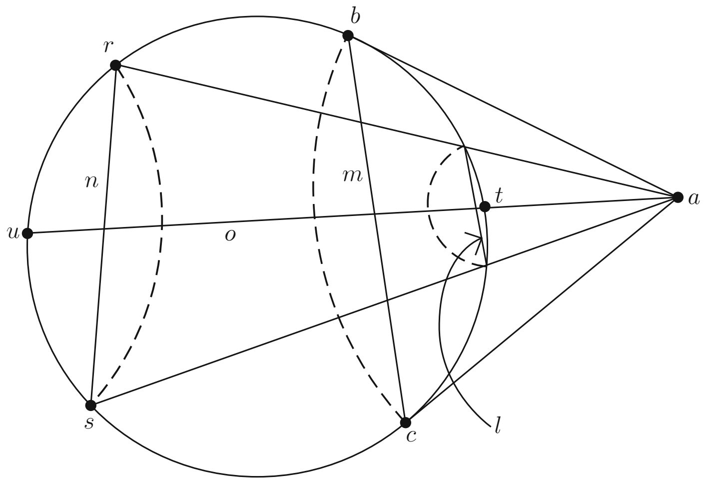
Let l and m be hyperbolic lines in $\mathcal{D}_K$ and suppose that these lines intersect $\mathcal{C}$ at the points b, c and d, e, respectively. Suppose that the tangents to $\mathcal{C}$ at b and c intersect at a, and that the extension of m passes through a. Then, l and m intersect at right angles in the hyperbolic sense.



Proof 11.5.5 The hyperbolic lines $\mathcal{K}^{-1}(l)$ and $\mathcal{K}^{-1}(m)$ in $\mathcal{D}_P$ corresponding to l and m are circular arcs that intersect $\mathcal{C}$ at right angles at the points b, c and d, e, respectively. Let $\mathcal{I}$ be the inversion in the circle of which $\mathcal{K}^{-1}(l)$ is an arc, so that $\mathcal{I}$ is an isometry of $\mathcal{D}_P$ (see Appendix 2, especially Proposition A.2.8). Now $\mathcal{I}$ takes $\mathcal{K}^{-1}(m)$ to a circular arc that intersects $\mathcal{C}$ at right angles (Corollary A.2.7), and it obviously interchanges the points d and e. It follows that $\mathcal{I}$ preserves $\mathcal{K}^{-1}(m)$. This implies that $\mathcal{K}^{-1}(l)$ and $\mathcal{K}^{-1}(m)$ are perpendicular in the Euclidean sense (Proposition A.2.8), and hence in the hyperbolic sense since $\mathcal{D}_P$ is a conformal model. Since $\mathcal{K} : \mathcal{D}_P \to \mathcal{D}_K$ is an isometry, l and m are perpendicular in the hyperbolic sense. □

Proof 11.5.4 Let the tangents from a to $\mathcal{C}$ touch it at b and c, let m be the line segment with endpoints b, c and let the line through a and z intersect $\mathcal{C}$ at t and u. Let l be any line passing through z and let l intersect $\mathcal{C}$ at p, q. Let r, s

be the points of $\mathcal{C}$ such that the lines through p and r and through q and s pass through a, let n be the line segment with endpoints r, s and let u be the point of intersection of $\mathcal{C}$ with the line o passing through a and z. Since z is the intersection of l and o, $\mathcal{K}^{-1}(z)$ is the intersection of $\mathcal{K}^{-1}(l)$ and $\mathcal{K}^{-1}(o)$; similarly, $\mathcal{K}^{-1}(\Pi_a(z))$ is the intersection of $\mathcal{K}^{-1}(n)$ and $\mathcal{K}^{-1}(o)$.



By Lemma 11.5.5, m and o are perpendicular in $\mathcal{D}_K$, so $\mathcal{K}^{-1}(m)$ and $\mathcal{K}^{-1}(o)$ are perpendicular in $\mathcal{D}_P$. It follows that $\mathcal{I}_{a,r}$ fixes $\mathcal{K}^{-1}(o)$. Since $\mathcal{I}_{a,r}$ takes p to r and q to s, it takes $\mathcal{K}^{-1}(l)$ to $\mathcal{K}^{-1}(n)$. Hence, $\mathcal{I}_{a,r}$ takes $\mathcal{K}^{-1}(z)$ to $\mathcal{K}^{-1}(\Pi_a(z))$:

$$\mathcal{I}_{a,r}(\mathcal{K}^{-1}(z)) = \mathcal{K}^{-1}(\Pi_a(z)).$$

This is what we wanted to prove. □

Now that we have the isometries Π_a at our disposal, we can prove a beautiful formula for the distance between two points of $\mathcal{D}_K$. For this, we shall need the following concept from projective geometry.

Definition 11.5.6

If a, b, c, and d are distinct complex numbers, their *cross-ratio* is

$$(a, b; c, d) = \frac{(a-c)(b-d)}{(a-d)(b-c)}.$$

Proposition 11.5.7

Suppose that the points a, b, c, and d lie on a line and that a and b are between c and d. Then, $(a, b; c, d) > 0$. Moreover, if p is a point distinct from a, b, c, and d and if the lines through p and each of the points a, b, c, and d intersect another line at a', b', c', and d', then

$$(a, b; c, d) = (a', b'; c', d').$$

This result is expressed by saying that the cross-ratio is a 'projective invariant': the cross-ratio of four points on a line is unchanged when they are 'projected' from some point p onto another line.

Proof

Let l be the line containing a, b, c, and d. Since a and b are on the 'same side' of l relative to c, $\arg(a - c) = \arg(b - c)$, so

$$\frac{a-c}{b-c} = \frac{|a-c|}{|b-c|}.$$

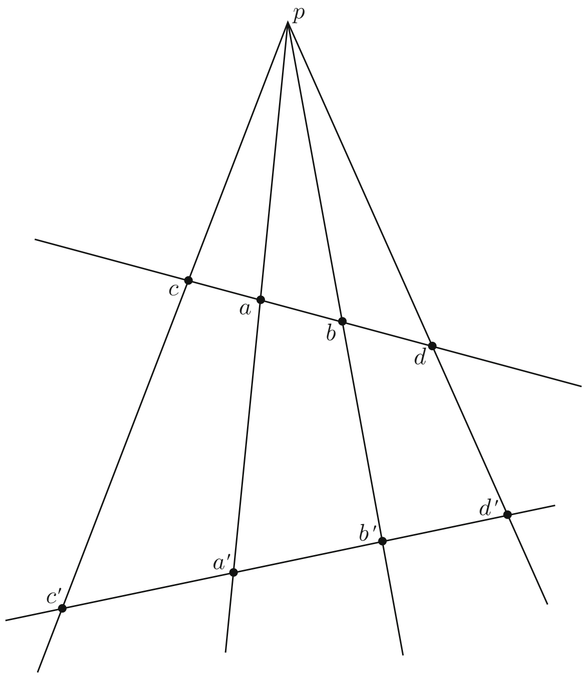
Similarly,

$$\frac{b-d}{a-d} = \frac{|b-d|}{|a-d|}.$$

Hence,

$$(a, b; c, d) = \frac{|a-c||b-d|}{|a-d||b-c|}.$$

In particular, this cross-ratio is a positive number.



Let $\angle apb$ be the angle between the lines through p and a and through p and b, etc. By the Euclidean sine rule,

$$\frac{|a-c|}{\sin\angle apc} = \frac{|p-c|}{\sin\angle pac}, \quad \frac{|a-d|}{\sin\angle apd} = \frac{|p-d|}{\sin\angle pad},$$
$$\frac{|b-c|}{\sin\angle bpc} = \frac{|p-c|}{\sin\angle pbc}, \quad \frac{|b-d|}{\sin\angle bpd} = \frac{|p-d|}{\sin\angle pbd}.$$

Hence,

$$(a,b;c,d) = \frac{\sin\angle apc \sin\angle bpd}{\sin\angle apd \sin\angle bpc}.$$

But obviously $\angle a'p'c' = \angle apc$, etc., hence the result. $\square$

In particular, the cross-ratio $(a,b;c,d)$, with $a,b,c,d \in \mathcal{D}_K$, is unchanged if a,b,c, and d are subjected to any perspectivity. Note that the cross-ratio is also unchanged if a,b,c, and d are subjected to any rotation about the origin, since this amounts to multiplying each of a,b,c, and d by a non-zero complex number.

Theorem 11.5.8

Let $a,b \in \mathcal{D}_K$ and let c,d be the points of intersection of the line through a,b with $\mathcal{C}$. Then, the Beltrami-Klein distance between a and b is

$$d_{\mathcal{D}_K}(a,b) = \frac{1}{2}|\ln(a,b;c,d)|.$$

Proof

We use a suitable isometry of $\mathcal{D}_K$ to reduce to the case in which a and b are real. Let l be the line through c and 1, and m the line through d and -1. We consider two cases, according to whether l and m are parallel (in the Euclidean sense) or not.

If l and m are parallel, the line joining c and d passes through the origin, and a suitable rotation about the origin will take c to 1, d to -1 and a,b to points a',b' on the real axis. Such a rotation is an isometry of $\mathcal{D}_K$ by the remarks following Corollary 11.5.3.

Suppose, on the other hand, that l and m intersect at a point p, say. If $|p| > 1$, the perspectivity Π_p takes c to 1, d to -1 and a,b to points a',b' on the line joining -1 and 1, i.e., the real axis. If $|p| < 1$, the lines l' joining c and -1 and m' joining d and 1 intersect at a point p' with $|p'| > 1$ and the perspectivity $\Pi_{p'}$ takes c to -1, d to 1 and a,b to points a',b' on the real axis.

We compute the distance $d_{\mathcal{D}_K}(a,b) = d_{\mathcal{D}_K}(a',b')$ by transferring to $\mathcal{D}_P$ using the isometry $\mathcal{K} : \mathcal{D}_P \to \mathcal{D}_K$, so that $d_{\mathcal{D}_K}(a',b') = d_{\mathcal{D}_P}(\mathcal{K}^{-1}(a'), \mathcal{K}^{-1}(b'))$. Using Proposition 11.3.2, this gives

$$d_{\mathcal{D}_K}(a',b') = \int_{\mathcal{K}^{-1}(a')}^{\mathcal{K}^{-1}(b')} \frac{2dv}{1-v^2} = \ln \frac{(1+\mathcal{K}^{-1}(b'))(1-\mathcal{K}^{-1}(a'))}{(1+\mathcal{K}^{-1}(a'))(1-\mathcal{K}^{-1}(a'))}. \tag{11.9}$$

Using the formula (11.8) for $\mathcal{K}$, we find that

$$\mathcal{K}^{-1}(\lambda) = \frac{1}{\lambda}(1-\sqrt{1-\lambda^2}), \quad \lambda \in \mathcal{D},$$

which implies that

$$\frac{1+\mathcal{K}^{-1}(\lambda)}{1-\mathcal{K}^{-1}(\lambda)} = \sqrt{\frac{1+\lambda}{1-\lambda}}.$$

Using this, (11.9) becomes

$$d_{\mathcal{D}_K}(a',b') = \frac{1}{2}\ln\frac{(1+b')(1-a')}{(1-b')(1+a')}. \tag{11.10}$$

On the other hand, we have seen that there is a perspectivity or a rotation about the origin that takes (a,b,c,d) to $(a',b',1,-1)$ or $(a',b',-1,1)$ with $a',b' \in \mathbb{R}$, and that these transformations of $\mathcal{D}_K$ leave the cross-ratio unchanged (see the remarks following the proof of Proposition 11.5.7). In the first case,

$$(a,b;c,d) = (a',b';1,-1) = \frac{(1-a')(1+b')}{(1+a')(1-b')},$$

and in the second case,

$$(a,b;c,d) = (a',b';-1,1) = \frac{(1+a')(1-b')}{(1-a')(1+b')},$$

so in both cases

$$d_{\mathcal{D}_K}(a,b) = d_{\mathcal{D}_K}(a',b') = \frac{1}{2}|\ln(a,b;c,d)|. \qquad \square$$

EXERCISES

11.5.1 Prove Eq. 11.8.

11.5.2 Extend the definition of cross-ratio in the obvious way to include the possibility that one of the points is equal to ∞, e.g., $(\infty,b;c,d) = (b-d)/(b-c)$. Show that, if $M : \mathbb{C}_\infty \to \mathbb{C}_\infty$ is a Möbius transformation, then

$$(M(a), M(b); M(c), M(d)) = (a,b;c,d) \text{ for all distinct points } a,b,c,d \in \mathbb{C}_\infty. \tag{11.11}$$

Show, conversely, that if $M : \mathbb{C}_\infty \to \mathbb{C}_\infty$ is a bijection satisfying this condition, then M is a Möbius transformation.

11.5.3 Use the preceding exercise to show that, if (a, b, c) and (a', b', c') are two triples of distinct points of $\mathbb{C}_\infty$, there is a unique Möbius transformation M such that $M(a) = a'$, $M(b) = b'$ and $M(c) = c'$.

11.5.4 Let $a, b \in \mathbb{C}_\infty$ and let d be the spherical distance between the points of S^2 that correspond to a, b under the stereographic projection map Π (Example 6.3.5). Show that

$$-\tan^2 \frac{1}{2}d = \left(a, -\frac{1}{\bar{a}}; b, -\frac{1}{\bar{b}}\right).$$

11.5.5 Show that, if $\mathcal{R}$ is the reflection in a line passing through the origin, then $\mathcal{K}\mathcal{R} = \mathcal{R}\mathcal{K}$. Deduce that $\mathcal{R}$ is an isometry of $\mathcal{D}_K$.

11.5.6 Show that the isometries of $\mathcal{D}_K$ are precisely the composites of (finitely many) perspectivities and reflections in lines passing through the origin.