4. Hyperbolic Geometry

4.1 The three geometries

Here we will look at the basic ideas of hyperbolic geometry including the ideas of lines, distance, angle, angle sum, area and the isometry group and finally the construction of Schwartz triangles. We develop enough formulas for the disc model to be able to understand and calculate in the isometry group and to work with the isometries arising from Schwartz triangles. Some of the derivations are complicated or just brute force symbolic computations, so we illustrate the basic idea with hand calculation and relegate the drudgery to Maple worksheets. None of the major results are proven but rather are given as statements of fact. Refer to the Beardon [10] and Magnus [21] texts for more background on hyperbolic geometry.

4.2 Synthetic and analytic geometry similarities

To put hyperbolic geometry in context we compare the three basic geometries. There are three two dimensional geometries classified on the basis of the parallel postulate, or alternatively the angle sum theorem for triangles.

Geometry	**Parallel Postulate**	**Angle sum**	**Models**
spherical	no parallels	$> 180°$	sphere
euclidean	unique parallels	$= 180°$	standard plane
hyperbolic	infinitely many parallels	$< 180°$	disc, upper half plane

The three geometries share a lot of common properties familiar from synthetic geometry. Each has a collection of lines which are certain curves in the given model as detailed in the table below.

Geometry	**Symbol**	**Model**	**Lines**
spherical	S^2, $\widehat{\mathbb{C}}$	sphere	great circles
euclidean	$\mathbb{C}$	standard plane	standard lines
unit disc	$\mathbb{D}$	$\{z \in \mathbb{C} : \lvert z \rvert < 1\}$	lines and circles perpendicular to the boundary
upper half plane	$\mathbb{U}$	$\{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$	lines and circles perpendicular to the real axis

Remark 4.1 If we just want to refer to a model of hyperbolic geometry we will use $\mathbb{H}$ to denote it.

The lines of a geometry give rise to the following ideas and constructs: incidence, betweenness, line segments, distance, angles, triangles, polygons, interiors and exteriors of polygons, and congruence. The following properties hold for these constructs in all three geometries. Note that to make the spherical model work globally as well as locally it is necessary to identify antipodal points. For instance, the intersection of two great circles is a pair of antipodal points. Also, antipodal points do not determine a unique great circle.

Incidence

- Two points determine a unique line.
- If two lines intersect, their intersection is a unique point.

Betweenness, line segments

- There is a notion of betweenness on lines (at least for points which are close).
- Line segments are defined as the set of all points between two points on a line.

Distance and angle

- There is a way of measuring distance and lengths of line segments which satisfies the standard rules.
- There is a way of measuring angles between lines which satisfies the standard rules.
- Circles are defined in the standard centre-radius way.

Triangles, polygons, interiors

- Triangles and polygons are defined.
- There is an interior and exterior to each polygon (more work required for a good definition in the sphere model).

Congruence

- Line segments are congruent if they have the same length.
- Angles are congruent if they have the same measure.
- Standard congruence theorems hold for segments and angles.
- Two polygons are congruent if the vertices of one can be associated to the vertices of the other such that all corresponding edge and (interior) angles are congruent.
- For triangles the SSS, SAS and ASA congruence theorems hold.

Parallel Postulate Two lines are called parallel if they do not intersect. We have:

- *Spherical geometry:* Every two lines intersect and hence there are never any parallels.
- *Euclidean geometry:* For a point P not on a given line ℓ there is a unique line m parallel to ℓ passing through P.
- *Hyperbolic geometry:* For a point P not on a given line ℓ there are at least two (and hence infinitely many) lines m parallel to ℓ passing through P.

Remark 4.2 In euclidean geometry triangles with congruent angles are similar though they need not be congruent. In the non-euclidean models the triangles must be congruent because their areas are determined by the angles.

4.3 Riemann sphere models

We may close up or compactify the complex plane by adding a point at infinity which we denote by ∞. The compactified plane $\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ is called the Riemann sphere. To topologize $\widehat{\mathbb{C}}$ we say that a sequence $\{z_n\}$ converges to infinity, $\lim_{n\to\infty} z_n = \infty$, if $\lim_{n\to\infty} |z_n| = \infty$ in the usual sense. The Riemann sphere $\widehat{\mathbb{C}}$ is homeomorphic to the standard sphere S^2 via stereographic projection as follows (see Figure 4.1). For S^2 we pick the standard sphere equation $s^2 + t^2 + u^2 = 1$, and for $\mathbb{C}$ we pick the s, t plane defined by $u = 0$. Let $N = (0, 0, 1)$ be the north pole of the sphere and $z = s + it = (s, t)$ a point of the complex plane. The ray from N to z intersects the sphere at N and at one other point $z' = \Phi(z)$ as in Figure 4.1.

Extend this definition of Φ to all of $\widehat{\mathbb{C}}$ by $\Phi(\infty) = N$, the resulting map $\Phi : \widehat{\mathbb{C}} \to S^2$ is a homeomorphism. The formulas for Φ and its inverse $\Psi = \Phi^{-1}$ are:

$$\begin{aligned} \Phi(z) &= (rz, 1-r) = (rs, rt, 1-r), \text{ where } r = \frac{2}{1+|z|^2}, \\ \Psi(s,t,u) &= \frac{s+it}{1-u}. \end{aligned}$$

Stereographic projection has the following properties:

- The Φ-image of a circle in the plane is a circle on the sphere. The Φ-image of a line in the plane is a circle on the sphere passing through the north pole.
- The Ψ-image of a circle on the sphere not passing through the north pole is a circle in the plane. The Ψ-image of a circle on the sphere passing through the north pole is a line in the plane.
- Both of Φ and Ψ are conformal maps, namely the Φ or Ψ-images of two curves intersect at the same angle as the original curves.

Now, via stereographic projection, all three geometries can be viewed as subsets of $\widehat{\mathbb{C}}$. Clearly the spherical geometry is all of $\widehat{\mathbb{C}}$. Euclidean geometry is the sphere punctured at the north pole. The disc model $\mathbb{D}$ of the hyperbolic geometry corresponds to the southern hemisphere defined by $u < 0$ and the upper half plane $\mathbb{U}$ corresponds to the "eastern hemisphere" of points on the sphere with $t > 0$.

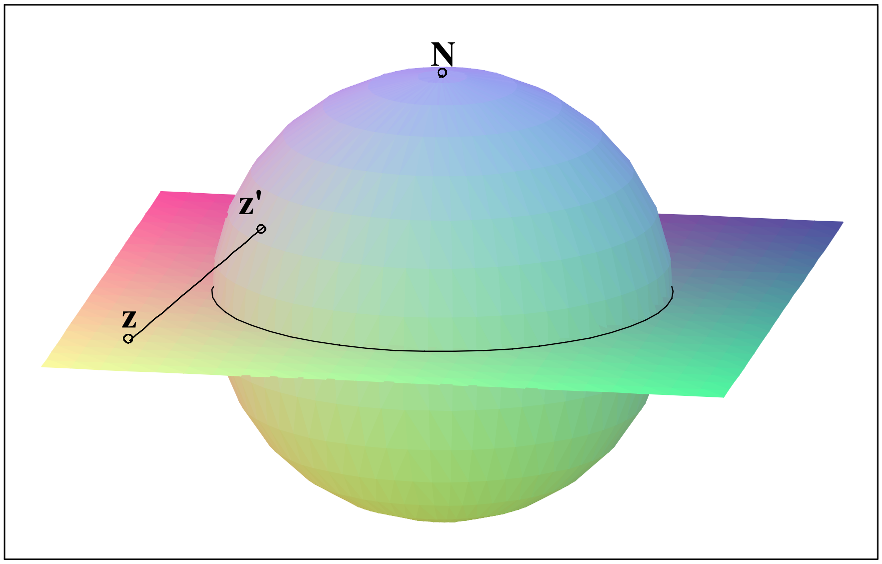

Figure 4.1 Stereographic Projection

4.4 Hyperbolic lines, angle and distance

Many of the constructions here are in the Maple worksheet `Hgeom.mws` which may be found on the website [30].

Lines Lets develop the equations of the hyperbolic lines in the disc. The lines which are diameters are easy, so lets concentrate on the lines which are parts of circles. To this end let (a, b) be the centre of the circle C of radius r which meets the unit circle at a right angle as in Figure 4.2. Since the boundary of the unit circle is not included in $\mathbb{D}$ we have drawn it as a dotted circle. The hyperbolic line is the portion of the circle C in the interior of $\mathbb{D}$ as shown in the figure. Since the distance of the centre of the unit disc to (a, b) is $\sqrt{a^2 + b^2}$ then by the Pythagorean theorem we have:

$$a^2 + b^2 = 1 + r^2, \text{or } r = \sqrt{a^2 + b^2 - 1}. \tag{1}$$

From this it is obvious that the centre of the circle is outside the unit disc. The equation of the circle is

$$\begin{aligned} (x-a)^2 + (y-b)^2 &= r^2, \text{ or} \\ 2xa + 2yb &= x^2 + y^2 + a^2 + b^2 - r^2, \text{ or} \\ 2xa + 2yb &= x^2 + y^2 + 1. \end{aligned} \tag{4.2}$$

From this we are immediately able to deduce that the centre of the $\mathbb{D}$-line connecting the two points (x_1, y_1) and (x_2, y_2) is given by solving the system of equations:

$$\begin{aligned} 2x_1a + 2y_1b &= x_1^2 + y_1^2 + 1 \\ 2x_2a + 2y_2b &= x_2^2 + y_2^2 + 1. \end{aligned}$$

This has a unique solution unless $\det \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} = 0$, i.e., if (x_1, y_1) and (x_2, y_2) lie on the diameter of the unit circle. The point of intersection of two $\mathbb{D}$-lines is given by solving the two simultaneous equations:

$$\begin{aligned} 2xa_1 + 2yb_1 &= x^2 + y^2 + 1, \\ 2xa_2 + 2yb_2 &= x^2 + y^2 + 1. \end{aligned}$$

The equations may be solved by subtracting one from the other, eliminating a variable and then back substitution. There are two solutions but only one is in the unit disc.

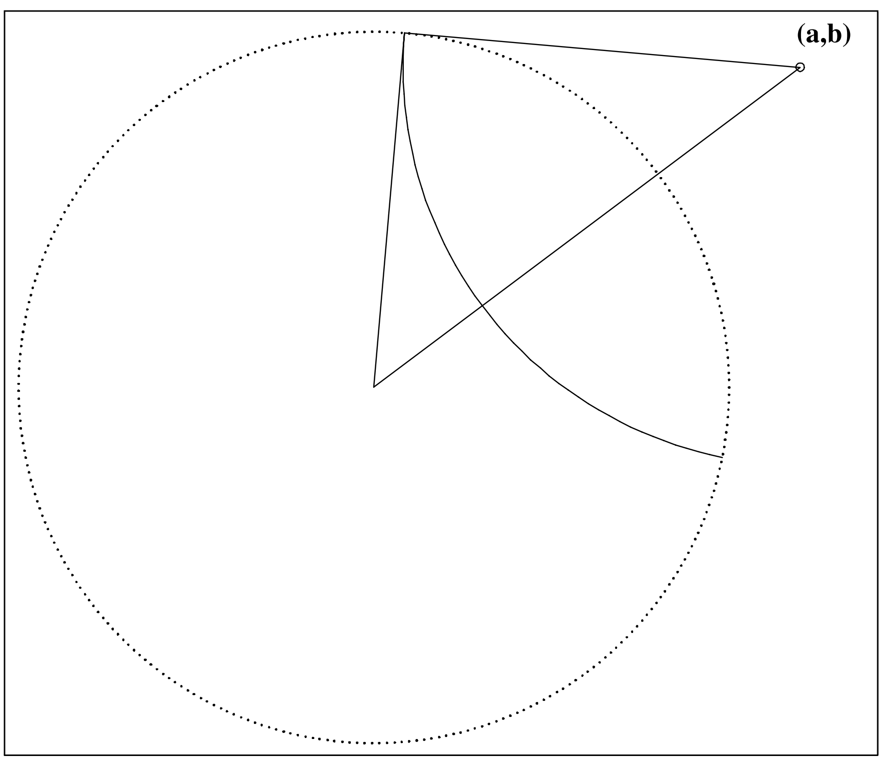

Figure 4.2 Hyperbolic line with circle centre

The other model of the hyperbolic plane is the upper half plane model $\{z \in \mathbb{C} :$

$\mathrm{Im}(z) > 0\}$. A typical non-vertical line is pictured in Figure 4.3.

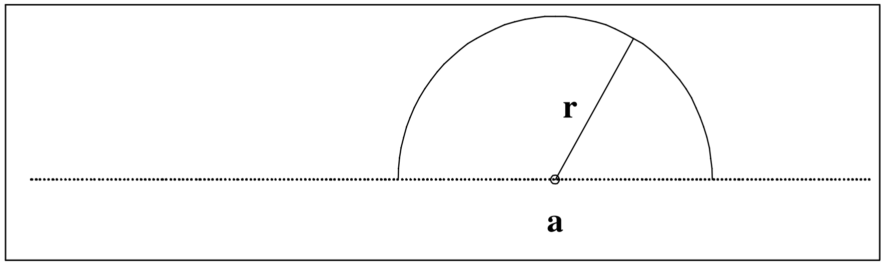

Figure 4.3 Hyperbolic line in upper half plane

The equation of the typical lines are:

$$\begin{aligned}(x-a)^2 + y^2 &= r^2, \text{ or} \\ 2xa &= x^2 + y^2 + a^2 - r^2\end{aligned}$$

where a is the centre of the circle on the x -axis and r is the radius. The r and a for the line determined by two points (x_1, y_1) and (x_2, y_2) is obtained by simultaneously solving the two equations.

$$\begin{aligned}2x_1a &= x_1^2 + y_1^2 + a^2 - r^2, \\ 2x_2a &= x_2^2 + y_2^2 + a^2 - r^2.\end{aligned}$$

If $x_1 \neq x_2$ then subtracting both equations gives us a value for a, from which r may be found. Similarly for finding the intersection of two lines we look at the equations

$$\begin{aligned}2xa_1 &= x^2 + y^2 + a_1^2 - r_1^2, \\ 2xa_2 &= x^2 + y^2 + a_2^2 - r_2^2.\end{aligned}$$

Subtracting the two equations we may find x and then y.

Angles and distance The angle between two lines is the angle between the tangents to the curves at the point as we travel along the lines in prescribed directions. The formula for distance is more complex so we defer its discussion to a later section and in the Maple worksheets `DHgeom.mws` and `UHgeom.mws` [30].

4.5 Law of sines and cosines

Let ΔABC be a non-degenerate triangle in any one of the geometries, with vertices A, B, and C and sides labeled a, b, and c as in Figure 4.4. Let α, β, and γ be the interior angles at A, B, and C respectively and let a, b, and c also denote the respective sides. The standard euclidean law of sines and law of cosines theorems have non-euclidean analogues and a second cosine theorem with non-hyperbolic analogue. Though we are mainly interested in the hyperbolic formulas, we include the formulas for spherical trigonometry for completeness. Proofs for the hyperbolic case are given in Beardon.[10]

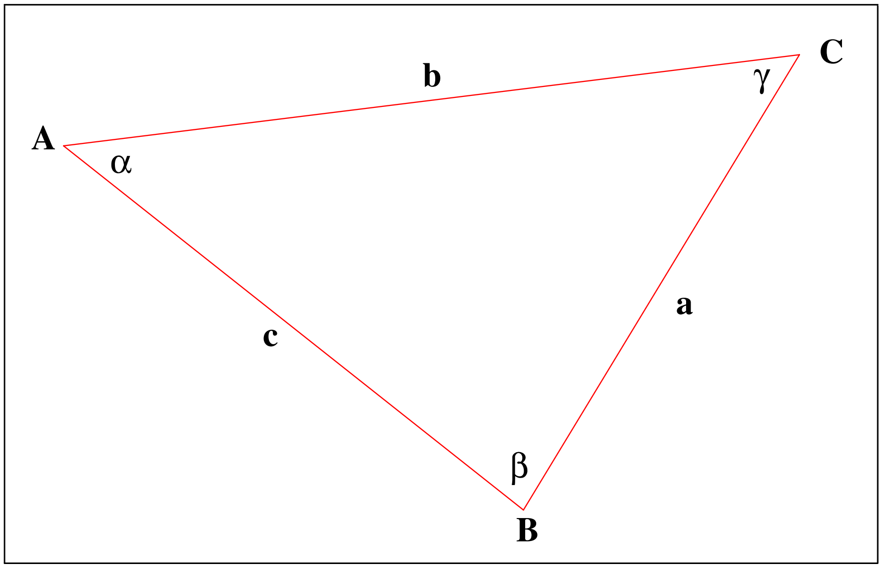

Figure 4.4- Setup for Laws of Sines and Cosines

Theorem 4.1 *Let ΔABC be a non-degenerate triangle in any one of the geometries, as in Figure 4.4. Then with the interpretation of α, β, γ and a, b, c, as above we have (measure in radians on teh sphere)f:*

Law of Sines

Spherical	$\frac{\sin\alpha}{\sin a} = \frac{\sin\beta}{\sin b} = \frac{\sin\gamma}{\sin c}$
Euclidean	$\frac{\sin\alpha}{a} = \frac{\sin\beta}{b} = \frac{\sin\gamma}{c}$
Hyperbolic	$\frac{\sin\alpha}{\sinh a} = \frac{\sin\beta}{\sinh b} = \frac{\sin\gamma}{\sinh c}$

,

Law of Cosines I

Spherical	??????????
Euclidean	$c^2 = a^2 + b^2 - 2ab\cos\gamma$
Hyperbolic	$\cosh c = \cosh a \cosh b - \sinh a \sinh b \cos\gamma$

,

and

Law of Cosines II

Spherical	none
Euclidean	none
Hyperbolic	$\cosh c = \frac{\cos\alpha\cos\beta + \cos\gamma}{\sin\alpha\sin\beta}$

,

4.6 Transforms of the sphere

4.6.1 Linear fractional transformations

Since each of the geometries is a subset of the Riemann sphere we may give a unified treatment of the isometries of the geometry from the conformal mappings of the sphere. For any matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with complex entries and $\det(A) = ad - bc \neq 0$ we define the associated linear fractional transformation $T_A : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ by:

$$\begin{aligned} T_A(z) &= \frac{az+b}{cz+d},\ z \neq \infty, \frac{-d}{c} \\ T_A(\infty) &= \frac{a}{c}, \\ T_A(\frac{-d}{c}) &= \infty. \end{aligned}$$

The last two are defined by taking limits for the values $T_A(z)$ in $\widehat{\mathbb{C}} - \{\infty, \frac{-d}{c}\}$. If $c = 0$ then $d \neq 0$ and $\frac{-d}{c}$ is interpreted as ∞. The mapping T_A has two main properties:

- the image of a line or circle is a line or circle.
- the magnitude and sense of angle of the intersection of two curves is preserved by T_A.

To illustrate these two properties let us examine the linear fractional transformation $z \to T(z) = -\frac{z-i}{z+1}$ which maps the upper half plane onto the unit disc. The before and after pictures for the map T in Figures 4.5 and 4.6 illustrate the bulleted properties above. Also note that T maps the upper half plane model of hyperbolic geometry to disc model of the plane taking hyperbolic lines to hyperbolic lines. In fact, T is an isometry between the two models.

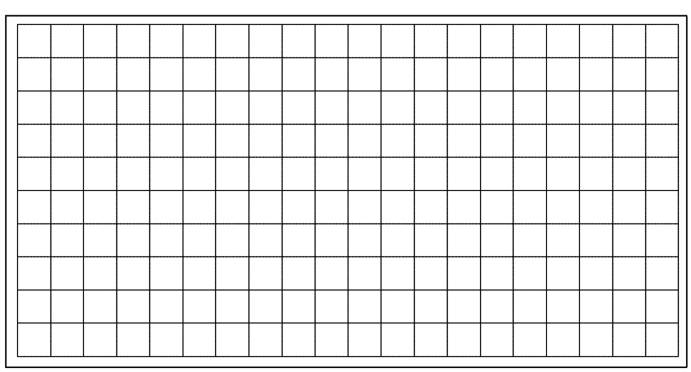

Figure 4.5 Grid in upper half plane

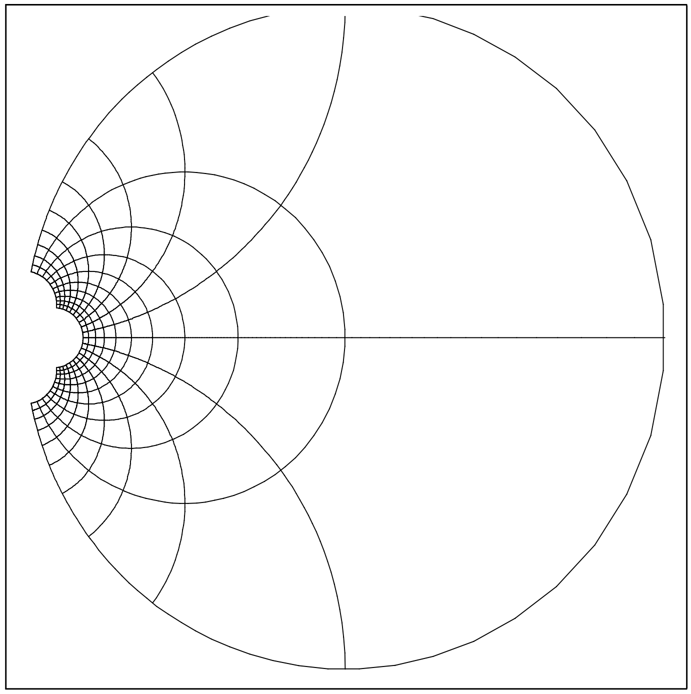

Figure 4.6 The T_A-image of grid lies in the disc.

Here are some homomorphism properties about linear fraction transformations which we leave as exercises.

$$T_{AB} = T_A \circ T_B$$
$$T_{A^{-1}} = T_A^{-1}$$
$$T_{rA} = T_A, r \neq 0.$$

From these properties we get the following:

Proposition 4.2 *Let* $GL_2(\mathbb{C}) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d, \in \mathbb{C}, ad - bc \neq 0 \right\}$ *denote the group of* 2×2 *invertible matrices with complex entries The map*

$$T : GL_2(\mathbb{C}) \to Aut(\widehat{\mathbb{C}}),\ A \to T_A$$

is a homomorphism with kernel $\left\{ \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} : a \in \mathbb{C}^* \right\}$. *Furthermore, when restricted to the subgroup* $SL_2(\mathbb{C})$ =$\{A \in GL_2(\mathbb{C}) : \det(A) = 1\}$ *the mapping has the same image and the kernel is now* $\left\{ \pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$.

The last statement of the propositions follows from the fact that for $A \in GL_2(\mathbb{C})$ the matrix $\frac{1}{\sqrt{\det(A)}} A \in SL_2(\mathbb{C})$ and both matrices have the same image under T.

Certain special matrices correspond to-certain special transformations

Matrix	Transformation	Description
$\begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$	$z \longrightarrow z+b$	translation
$\begin{bmatrix} \sqrt{r} & 0 \\ 0 & \frac{1}{\sqrt{r}} \end{bmatrix}$	$z \longrightarrow rz$	scaling transformation
$\begin{bmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{bmatrix}$	$z \longrightarrow e^{i\theta}z$	rotation
$\begin{bmatrix} \sqrt{a} & 0 \\ 0 & \frac{1}{\sqrt{a}} \end{bmatrix}$	$z \longrightarrow az$	rotation and scaling, $a = re^{i\theta}$
$\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix}$	$z \longrightarrow az+b$	affine transformation
$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$	$z \longrightarrow \frac{-1}{z}$	conformal inversion in unit circle

4.6.2 Inversion in a circle

Let C be a circle with centre z_0 and radius r. Inversion in the circle c takes the point z and maps it to a point $z^{'}$ such that $\overrightarrow{zz_0}$ and $\overrightarrow{z'z_0}$ are the same ray and the distance formula

$$|z' - z_0|\,|z - z_0| = r^2,$$

holds - see Figure 4.7. Since $z - z_0$ and $z' - z_0$ have the same complex argument it follows that

$$\begin{aligned} (z' - z_0)(\overline{z - z_0}) &= |z' - z_0|\,|z - z_0| = r^2, \text{ or} \\ z' - z_0 &= \frac{r^2}{\overline{z - z_0}}, \text{ or} \\ z^{'} &= z_0 + \frac{r^2}{\overline{z - z_0}} = \frac{z_0\overline{z} + (r^2 - z_0\overline{z_0})}{\overline{z} - \overline{z_0}}. \end{aligned} \tag{4.3}$$

If we denote complex conjugation by θ then inversion in a circle has the form: $T_A \circ \theta$, where

$$A = \begin{bmatrix} z_0 & r^2 - z_0\overline{z_0} \\ 1 & -\overline{z_0} \end{bmatrix}.$$

It follows then that the composition of any two inversions in circle is a linear fractional transformation:

$$(T_A \circ \theta) \circ (T_B \circ \theta) = T_A \circ (\theta \circ T_B \circ \theta) = T_A \circ T_{\overline{B}}, \tag{4}$$

where $\overline{B}$ is the matrix obtained by taking the complex conjugates of the entries of B.

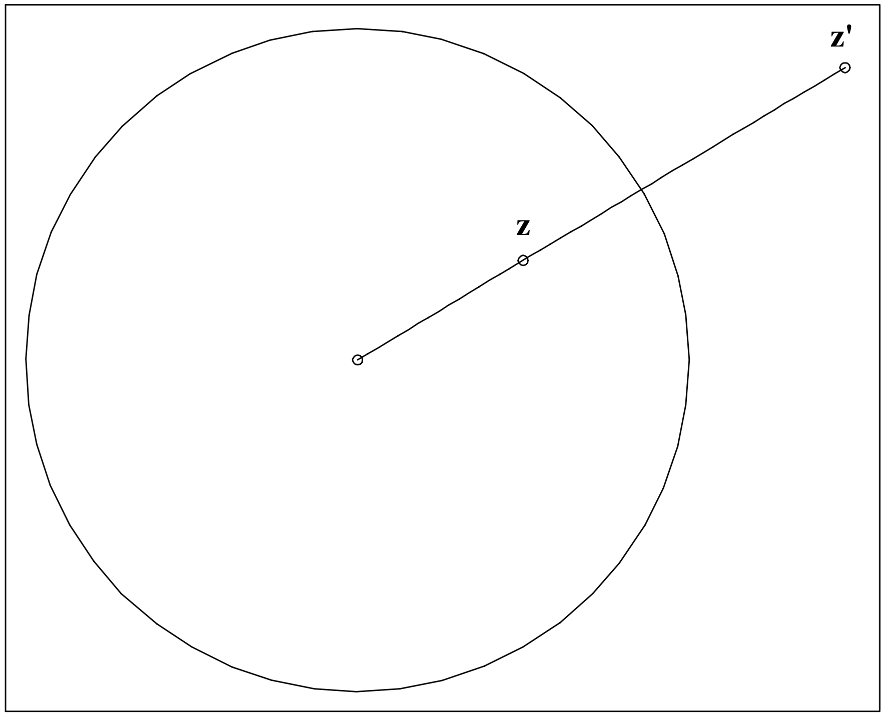

Figure 4.7 Inversion in a circle

Inversions in circles have the following properties:

- Inversion in a circle is a transformation of order 2.
- A point is fixed by inversion in C if and only if it lies on C.
- Every circle or line is mapped to a circle or line. The circle or line is mapped to a line if and only if the circle or line passes through z_0, the centre of C.
- A reflection in a line is the same as "inversion in a circle of infinite radius tangent to the reflecting line".
- Inversion is anti-conformal, i.e., angle preserving and orientation-reversing.
- If a circle or line meets C at a right angle then the circle or line is mapped back to itself.
- If C is a circle perpendicular to the boundary of the unit disc $\mathbb{D}$ then $z_0\overline{z_0} = r^2+1$ as in 4.1 and hence the formula is given by:

$$z' = \frac{z_0\overline{z} - 1}{\overline{z} - \overline{z_0}} = T_A(\overline{z}),\ A = \begin{bmatrix} z_0 & -1 \\ 1 & -\overline{z_0} \end{bmatrix}. \tag{5}$$

4.7 Isometries

4.7.1 Isometries of the models

Definition 4.1 An *isometry* of a geometry is a distance preserving, transformation. Thus if T is an isometry of X then

$$d(T(x), T(y) = d(x, y),\ x, y \in X,$$

where $d(x, y)$ is the distance between the points x and y.

Isometries automatically have inverses and are continuous. It follows from the properties of synthetic geometry discussed earlier that isometries take lines to lines and preserve angles because of the SSS theorem. Thus isometries are preserve the measure of angles. If the isometry preserves the sense angles it is called a direct or orientation preserving isometry, otherwise, it is called an indirect or orientation-reversing isometry. As previously noted each of the geometries may be considered as subsets of the Riemann sphere. Each linear fractional transformation that maps these sets to themselves are conformal (angle preserving maps) though they may not be isometries. It turns out that all the direct isometries can be realized by fractional linear transformations. Here are the groups of isometries.

Table of Isometry Groups

Geometry	**Symbol**	**Isometry group**
spherical	S^2, $\widehat{\mathbb{C}}$	$PSU_2(\mathbb{C}) = \left\{ \begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix} : a, b \in \mathbb{C},\ \|a\|^2 + \|b\|^2 = 1 \right\} / \{\pm I\}$
euclidean	$\mathbb{C}$	$\left\{ \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix} : a, b \in \mathbb{C}, \|a\| = 1 \right\} / \{\pm I\}$
unit disc	$\mathbb{D}$	$PSU_2^{1,1}(\mathbb{C}) = \left\{ \begin{bmatrix} a & b \\ \overline{b} & \overline{a} \end{bmatrix} : a, b \in \mathbb{C},\ \|a\|^2 - \|b\|^2 = 1 \right\} / \{\pm I\}$
upper half plane	$\mathbb{U}$	$PSL_2(\mathbb{R}) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in \mathbb{R},\ ad - bc = 1 \right\} / \{\pm I\}$

Notes on the table of isometry groups:

- In each of the cases above the matrices have been chosen so that the determinants equal 1. Note that both A and $-A$ yield the same transformation.
- If we adopt the first three models, then the diagonal portion of A is of the form $\begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix}$ where $|a| \leq 1$, $|a| = 1$ or $|a| \geq 1$ in the spherical, euclidean and hyperbolic cases, respectively.
- The trace of A, $\mathrm{tr}(A)$ is real in all cases.

- When the transformations of $PSU_2(\mathbb{C})$ are realized on the sphere via stereographic projection they simply become rotations of the sphere. The PSU stands for *projective special unitary.* Unitary means $A^{-1} = A^*$, special means $\det(A) = 1$, and projective means that we are "mod"-ing out by the scalar matrices.

- The PS of $PSU_2^{1,1}(\mathbb{C})$ has the same meaning as in $PSU_2(\mathbb{C})$. The $U_2^{1,1}$ means $(1,1)$ unitary, i.e., $A^{-1} = JA^*J$ where $J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

Orientation reversing isometries are obtained by composing the isometries above with an appropriate "complex conjugation" θ leaving mapping the given domain back to itself. We record these in a table:

Geometry	**Symbol**	**Complex Conjugation**
spherical	S^2, $\widehat{\mathbb{C}}$	$\theta : z \to \overline{z}$
euclidean	$\mathbb{C}$	$\theta : z \to \overline{z}$
hyperbolic unit disc	$\mathbb{D}$	$\theta : z \to \overline{z}$
hyperbolic upper half plane	$\mathbb{U}$	$\theta : z \to -\overline{z}$

Note that equation 2.9 holds for all of the first three geometries.

4.7.2 Properties of isometries

Fixed points Let A be the matrix of a linear factional transformation If z is fixed point of T_A then we have: $\frac{az+b}{cz+d} = z$. The solutions are:

$$\begin{aligned} z &= \frac{1}{2c}\left(a - d + \sqrt{(a^2 - 2ad + d^2 + 4cb)}\right), \\ z &= \frac{1}{2c}\left(a - d - \sqrt{(a^2 - 2ad + d^2 + 4cb)}\right), \end{aligned}$$

or upon noting that $ad - bc = 1$ we get:

$$\begin{aligned} z &= \frac{1}{2c}\left(a - d + \sqrt{(a+d)^2 - 4}\right), \\ z &= \frac{1}{2c}\left(a - d - \sqrt{(a+d)^2 - 4}\right). \end{aligned}$$

If $\mathrm{tr}(A) \neq \pm 2$, then there are exactly 2 fixed points, otherwise there is one unless the transformation is the identity. If $c = 0$ the fixed points are ∞ and $\frac{b}{a-d}$.

Next lets look at the fixed points of an anti-conformal map of $\widehat{\mathbb{C}}$. We have $T_A(\overline{z}) = z$ or

$$c\overline{z}z + dz - a\overline{z} - b = 0.$$

If there are solutions to this equation then it a line or a circle depending on whether $c = 0$ or $c \neq 0$. If the fixed point subset is non-empty then transformation must be inversion in a circle.

4.7.3 Classification of isometries

Linear fractional transformations can be classified according to the value of $a + d = \mathrm{tr}(A)$

$tr(A)$	**type**	**description**
real, $\lvert\mathrm{tr}(A)\rvert < 4$	elliptic	rotation about two fixed points
$\mathrm{tr}(A) = \pm 2$	parabolic	one fixed point, euclidean translation
real, $\lvert\mathrm{tr}(A)\rvert > 4$	hyperbolic	two fixed points, hyperbolic translation
non-real	loxodromic	spiraling trajectories from one fixed point to another

Since the isometries of the geometries all have real traces the loxodromic elements are not of interest. In the table below we have a classification of the various type of isometries that occur in the three geometries.

Geometry	**Orientation preserving isometries**
spherical	rotations
euclidean	rotations euclidean translations
hyperbolic	rotations, euclidean translations, hyperbolic translations

In the figures below the trajectories of various types of isometries in the hyperbolic plane have been produced. In each of the three pictures the transformation moves points along the curves for a certain distance.

Notes on classification of isometries

- The rotations have one fixed point inside the unit disc, the other is outside. The trajectories are all circles centred at the fixed point. The rotation moves a point along the circle it lies on through the angle of rotation ϕ. Note that the angle is only known up to sign since it depends on which fixed point we measure it from. In the case of the unit disc it can be measured from the fixed point inside the unit disc. The angle of rotation satisfies the following:

$$\cos(\phi) = \pm\frac{\mathrm{tr}(A)}{2}$$

- In the hyperbolic translation of both fixed points are on the boundary. Only one of the curves is a hyperbolic line namely the circle perpendicular to the boundary. This line is called the *axis of the hyperbolic translation.* The other curves are called equidistant curves since they consist of points whose distance from the line is a fixed value. In euclidean geometry such a curve is a line but in hyperbolic geometry it is not! In spherical geometry the equidistant curves are circles. The distance moved along the axis is given by:

$$2\ln\left(\frac{1}{2}\left(tr(A) + \sqrt{tr(A)^2 - 4}\right)\right)$$

  Details are given in the Maple worksheets `DHgeom.mws` and `UHgeom.mws` [30]..

- In the euclidean translations the trajectories are called *horocycles* and though they look like hyperbolic circles, they are not. Notice that as we move along the horocycle in the positive or negative direction the limit point is always the fixed point on the boundary. In the upper half plane when the fixed point is infinity the horocycles are simply horizontal lines and the translations are simply translations in the horizontal direction.

Finally we will want to know the type of orientation reversing isometries with fixed points. We call such transformations reflections. Here they classified in table form. Observe that the fixed point set is line in the geometry.

Geometry	**Reflections**
spherical	reflection in a great circle
euclidean	relection in a line
hyperbolic	inversion in a circle

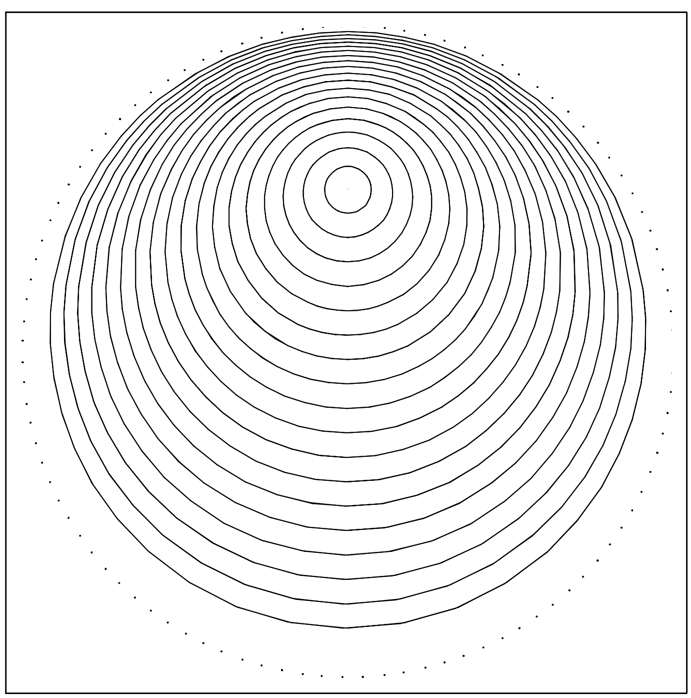

Figure 4.8 Elliptic rotation in $\mathbb{D}$

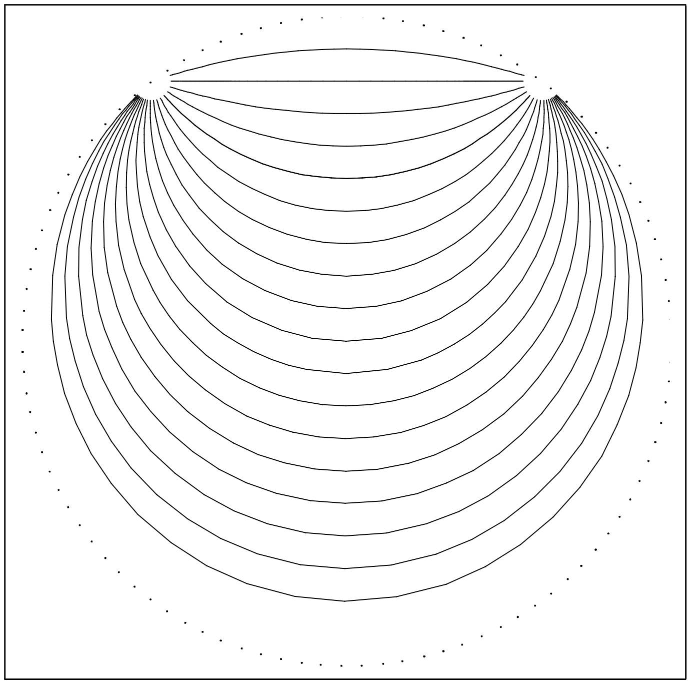

Figure 4.9 Hyperbolic translation in $\mathbb{D}$

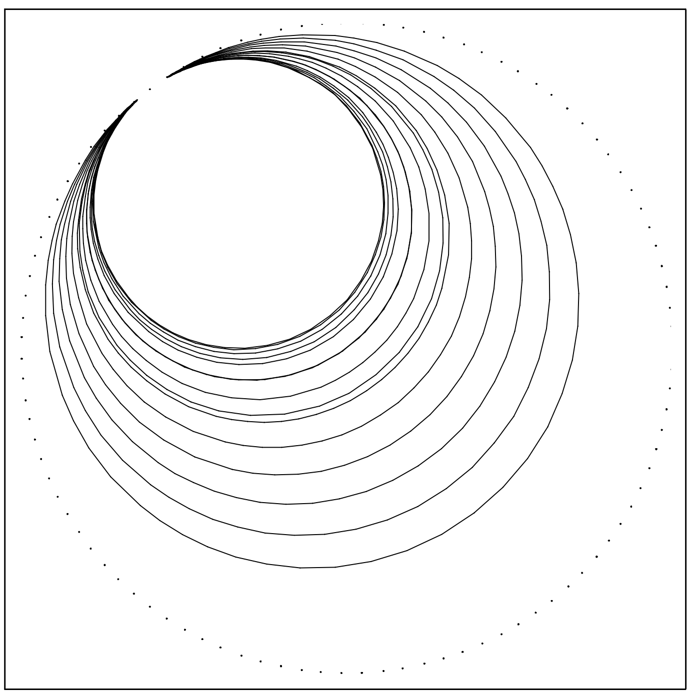

Figure 4.10 Euclidean translation in $\mathbb{D}$

4.8 Discrete groups of isometries

Most of the subgroups of isometries we are interested in are discrete groups of isometries. Here is a rough and ready definition.

Definition 4.2 Let Λ be a group of isometries of the geometry X. The orbit Λx of a point x is the set $\{gx : g \in \Lambda\}$. A group is called discrete if for each $x \in X$ the set Λx has no limit points in the interior of the geometry.

Example 4.1 Let $\Lambda = \{\gamma_{m,n} : m, n \in \mathbb{Z}\}$, where $\gamma : (x, y) \to (x + m, y + n)$. The orbits are all translates of the lattice of points with integer coordinates.

Example 4.2 Let Δ be a polygon such that all the angles are integer submultiples of 2π Then the (kaleidoscopic) group generated by the reflections in the sides of Δ is a discrete subgroup. If Δ is a Schwartz (l, m, n)-triangle then the kaleidoscopic group Λ^* has the following presentation:

$$\Lambda^* = \left\langle p, q, r : p^2 = q^2 = r^2 = (pq)^l = (qr)^m = (rp)^n = 1 \right\rangle .$$

The p, q and r are the reflection in the side of the master tile as discussed in section 2. The subgroup, Λ, of orientation preserving symmetries in Λ^* is generated by $a = pq$, $b = qr$ and $c = rp$. and has the following presentation

$$\Lambda = \left\langle a, b, c : a^l = b^m = c^n = abc = 1 \right\rangle .$$

There is a tiling on the universal cover, examples of which have been given previously.

If X is the geometry then the orbit space of X X/Λ is defined by $X/\Lambda = \{\Lambda x : x \in X\}$. The topology of X/Λ is defined by the limit condition $\Lambda x_n \to \Lambda x_0$ if and only if there are $g_n \in \Lambda$ such that $g_n x_n \to x_0$. A subgroup is called co-compact if X/Λ is compact. The discrete co-compact subgroups have the following types of isometries:

Geometry	**Possible direct isometries**
spherical	rotation of finite order, Λ is a finite group
euclidean	rotations of finite order, eulidean translations
hyperbolic	rotations of finite order, hyperbolic translations

If Λ is only discrete then we may also have euclidean translations in the group in the hyperbolic plane. Discrete groups will be discussed later in the section on the topology of group actions.

4.9 The differential geometry point of view

In this section we give a quick development of the differential geometric notion of length on a surface so that we can verify what the invariant metric is for hyperbolic space and develop a distance formula for it.

Metrics, distance and angle A metric for the geometry X is a distance function $d(x, y)$, $x, y \in X$ satisfying the following properties:

$$\begin{aligned} d(x,y) &\geq 0,\ d(x,y) = 0 \text{ if and only if } x = y, \\ d(x,y) &= d(y,x), \text{ (symmetry) ,and} \\ d(x,z) &\leq d(x,y) + d(y,z) \text{ (triangle inequality).} \end{aligned}$$

The metric is invariant under the group Λ if each $\gamma \in \Lambda$ is an isometry for the metric:

$$d(gx, gy) = d(x, y), \; x, y \in X, g \in \Lambda.$$

The invariant metrics are defined for the sphere and euclidean plane via the table.

Geometry	**Distance formula**
sphere	$\arccos(x \cdot y)$
plane	$\sqrt{(x-y)\cdot(x-y))}$

For the sphere we simply measure the angle subtended at the centre by the two points, the euclidean distance is simply the standard distance. For the hyperbolic distance we can develop a formula by assuming that the isometry group is known as above and working from there. First we need digress a little on Riemannian metrics on a region Ω in the plane.

Recall that the length of a path γ parametrized by $\mathbf{p}(t) = (x(t), y(t))$ $t_0 \leq t \leq t_1$ is given by the integral:

$$\int_{t_0}^{t_1} ds = \int_{t_0}^{t_1} \sqrt{\mathbf{p}'(t) \cdot \mathbf{p}'(t)} dt$$

and the acute angle θ between two curves $\mathbf{p}(t)$ and $\mathbf{q}(t)$ which intersect at $t = t_0$ is given by:

$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{\sqrt{\mathbf{u} \cdot \mathbf{u}}\sqrt{\mathbf{u} \cdot \mathbf{v}}}, \; \text{where } \mathbf{u} = \mathbf{p}'(\mathbf{t}_0), \; \mathbf{v} = \mathbf{q}'(\mathbf{t}_0)$$

We can generalize this notion by taking a more general definition of dot product, one which will vary as we move around Ω:

$$\mu_{(x,y)}(\mathbf{u}, \mathbf{v}) = \mathbf{u}^t \begin{bmatrix} E(x,y) & F(x,y) \\ F(x,y) & G(x,y) \end{bmatrix} \mathbf{v} \tag{6}$$

where we consider $\mathbf{u}$ and $\mathbf{v}$ as a column matrices. The inner product matrix

$$\begin{bmatrix} E(x,y) & F(x,y) \\ F(x,y) & G(x,y) \end{bmatrix}$$

is supposed to positive definite.

Remark 4.3 One way these other metrics arise is parametrize a surface embedded three space or a higher-dimensional space. For example the sphere of radius R can be modeled by the rectangle $-\pi \leq u \leq \pi$ and $-\frac{\pi}{2} \leq v \leq \frac{\pi}{2}$ with the parametrization:

$$\begin{aligned} x(u,v) &= R\cos(u)\cos(v), \\ y(u,v) &= R\sin u)\cos(v), \\ z(u,v) &= R\sin(v), \end{aligned}$$

The E, F and G matrix in u, v coordinates is then defined as

$$\begin{aligned} \begin{bmatrix} E(u,v) & F(u,v) \\ F(u,v) & f(u,v) \end{bmatrix} &= J^t(u,v)J(u,v), \; \text{where} \\ J(u,v) &= \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{bmatrix} \end{aligned}$$

is the differential of the parametrization.

We often write out the metric definition in terms of the differential of arclength ds :

$$ds^2 = Edx^2 + 2Fdxdy + Gdy^2$$

The formula for the area becomes:

$$dA = (EG - F^2)dxdy.$$

Of course the E, F and G vary around the region. Arclength is now defined as:

$$\begin{aligned}\int_{t_0}^{t_1} ds &= \int_{t_0}^{t_1} \sqrt{\mu(\mathbf{p}'(t)\cdot\mathbf{p}'(t))}dt \\ &= \int_{t_0}^{t_1} \sqrt{E\cdot(x'(t))^2 + 2F\cdot x'(t)y'(t) + G\cdot(y'(t))^2}dt.\end{aligned}$$

The terms E, F, G in the above integral need to be replaced by $E(x(t), y(t))$ etc., we have not written this out this explicit dependence to keep the notation simple.

Having defined arclength we can then define the distance between two points as the length of the shortest curve joining them. A *geodesic* is a curve which is the curve of shortest $\mu-$length between any two nearby points on the geodesic. Though the geodesic is locally the shortest curve between any of its two points it may not be the shortest curve between its end points, e.g., think of a helices on a cylinder. We can now define a geometry on a surface by taking the geodesics as the straight lines, at least locally. Now we can do all the standard constructions in geometry on this surface -at least in the small - using small segments of geodesics to be line segments, and distance and angles are defined by the metric.

Though the metric in 4.6 allows for very general surfaces we are interested in conformal geometries where $E = G$ and $F = 0$. This mean that the scalar product $\mu_{(x,y)}(\mathbf{u}, \mathbf{v}) = E(x, y)(\mathbf{u}\cdot\mathbf{v})$ and hence immediately implies that angle measurements are the same as the standard measurements. For the hyperbolic plane the distance and area differentials are:

Model	**distance**	**area**
disc	$ds^2 = 2\frac{dx^2+dy^2}{(1-r^2)^2}$	$dA = \frac{4dxdy}{(1-r^2)^2}$
upper half plane	$ds^2 = \frac{dx^2+dy^2}{y^2}$	$dA = \frac{dxdy}{y^2}$

Now, we have a distance on our abstract surface we would like to determine the isometry group. Pick a differentiable transformation $T : (x, y) \to (f(x, y), g(x, y))$ of the surface. We can guarantee that T is an isometry if for every path

$$\int_{t_0}^{t_1} ds = \int_{t_0}^{t_1} \sqrt{\mu(\mathbf{p}'(t)\cdot\mathbf{p}'(t))}dt = \int_{t_0}^{t_1} \sqrt{\mu(T(\mathbf{p})^{'}(t)\cdot T(\mathbf{p})^{'}(t))}dt.$$

To see if this will be true we compute ds along the path $T(\mathbf{p}(t))$ any verify that we get the same ds. The new ds, say ds^*, will be given by:

$$(ds^*)^2 = E^*(dx^*)^2 + 2F^*dx^*dy^* + G^*(dy)^{*2},$$

where

$$\begin{aligned} dx &= df(x,y) = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \\ dy &= dg(x,y) = \frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial y}dy \\ E^* &= E(f(x,y), g(x,y)), \end{aligned}$$

and analogous formulas for F^* and G^*. For an example lets show that $z \to rz + b$ is an isometry of the upper half plane when r is real and positive and b is real. We have $f(x,y) = rx + b$ and $g(x,y) = ry$. We compute:

$$\begin{aligned} (ds^*)^2 &= E^*((dx^*)^2 + (dy^*)^2) \\ &= \frac{(d(rx+b))^2 + (d(ry))^2}{(ry)^2} \\ &= \frac{r^2(dx^2 + dy^2)}{r^2y^2} = ds^2. \end{aligned}$$

The remainder of the proof of invariance for both models is given in the Maple worksheets `UHgeom.mws` and `DHgeom.mws` [30]. Now let us use invariance to compute distances. In the disc model, if z_0 is a point in the disc then the fractional linear transformation $z \to \frac{z+z_0}{\overline{z_0}z+1}$ maps 0 to z_0. Combining the inverse of this transformation with a rotation at 0 we can isometrically move any pair of points (z_0, z_1) to the pair $(0, r_0)$, where r_0 is on the positive real axis. The parametrization of the geodesic from 0 to r_0 is given by $x(t) = t,\ y(t) = 0$, so

$$ds^2 = 2\frac{dt^2 + d0^2}{(1-(t^2+0^2))^2} = \frac{2dt}{(1-t^2)}.$$

The length is then equal to:

$$\int_0^{r_0} \frac{2dt}{(1-t^2)} = 2\operatorname{arctanh} r_0 = \ln(\frac{1+r_0}{1-r_0}).$$

By combining transformations we may derive the distance function for two arbitrary points in the disc. The formula in all its glory in given in the in the Maple worksheets `UHgeom.mws` and `DHgeom.mws` [30].

The actual history Our presentation does not give the real historical development of the differential-geometric study of non-euclidean geometry. Prior to the development of Riemannian geometry the non-euclidean models were constructed and studied synthetically mainly as a outgrowth of why the parallel postulate was not derivable from the other axioms of geometry. In early differential geometry the notions of distance, area and curvature developed for surfaces in 3 space and then abstracted as above to regions in the plane. Of special interest were the surfaces of constant curvature - the sphere (constant positive curvature) the plane, and cylinder (constant zero curvature) and the pseudosphere (constant negative curvature). A surface of constant curvature in three space in which all geodesics can be indefinitely continued is called a complete surface of constant curvature. The sphere, plane and cylinder are complete surfaces of constant curvature but there is no complete surface of constant negative curvature in three space. To do this we need to resort to abstract surfaces with conformal μ-metrics as above. Once we have created them we can determine the distance formula and the isometry group. Though the concept of curvature would take us too far afield here the Gauss-Bonnet theorem gives a very nice relationship among the curvature, the area of a triangle, and the deficiency of the angle sum of a triangle. If α, β, and γ are the angles of the triangle Δ, and κ its curvature then :

$$\alpha + \beta + \gamma - \pi = \iint_{\Delta} \kappa dA = \kappa \cdot \text{area}(\Delta).$$

4.10 Schwartz triangles

A Schwartz triangle is a triangle in a geometry with angles are $\frac{\pi}{l}$, $\frac{\pi}{m}$, and $\frac{\pi}{n}$. As previously discussed, a Schwartz triangle generates a kaleidoscopic, geodesic tiling of each of the three geometries discussed above. In this section we derive the formulae for the reflections in the sides of a hyperbolic Schwartz triangle. Our triangle will have one vertex at the origin and one edge on the x-axis. A second edge will be a ray in the first quadrant meeting the x -axis at an angle of $\frac{\pi}{l}$. This will make two of the reflections in the sides trivial to find we will only have to work for the third side. The third side will be a circle orthogonal to the unit circle and meeting the x -axis at an acute angle of $\frac{\pi}{m}$ and meeting the second side at an acute angle of $\frac{\pi}{n}$ (Figure 4.11). To find this circle we will determine the locus of centres of circle that meet the x-axis at an acute angle of $\frac{\pi}{m}$, but such that the angle measured from the x-axis counter-clockwise to the circle is $\pi - \frac{\pi}{m}$. We then intersect this locus with the locus of centres meeting the second edge at an acute angle of $\frac{\pi}{n}$. The various sides and curves are shown in Figure 4.11 where $(l, m, n) = (3, 4, 8)$.

Let us first find the locus of centres of hyperbolic lines meeting the x-axis at an angle of $\frac{\pi}{m}$. The equation of this circle is $2xa + 2yb = x^2 + y^2 + 1$ where (a, b) denotes the centre of the circle. If we let x_0 denote the intersection with the x -axis then since the radial line segment from $(x_0, 0)$ to (a, b) is perpendicular to the circle and the circle cuts at an angle of $\frac{\pi}{m}$ then the slope of the line segment is $\cot(\frac{\pi}{m})$. It follows that $\frac{b}{a-x0} = \cot(\frac{\pi}{m})$ hence $x_0 = b + a\tan(\frac{\pi}{m})$. Substituting $x = b + a\tan(\frac{\pi}{m})$ and $y = 0$ in the equation for the line we get $2a(b + a\tan(\frac{\pi}{m})) - (b + a\tan(\frac{\pi}{m}))^2 = 1$ or

$$\cos^2(\frac{\pi}{m})a^2 - \sin^2(\frac{\pi}{m})b^2 = \cos^2(\frac{\pi}{m}). \tag{7}$$

The graph of equation 4.7 is a hyperbola tangent to the unit circle at $(1, 0)$. The portion in the first quadrant is graphed in Figure 4.11. The portion of the hyperbola in the fourth quadrant is the locus of centres of lines that meet the x-axis at angle of $\frac{\pi}{m}$ measure counterclockwise from the x-axis. The left half of the hyperbola corresponds to intersections with the negative x-axis. The locus of the equation for the circle crossing the second edge may be obtained by first determining the locus of centres intersecting the x-axis at an angle of $\frac{\pi}{n}$ (measured counterclockwise from the x-axis and then rotating the hyperbola through $\frac{\pi}{l}$. The portion of the hyperbola in the first quadrant is shown in Figure 4.11. Note that it too is tangent to the unit circle where the second diameter meets the unit circle.

The centre of our desired circle is the intersection of the two hyperbolas in the first quadrant. The Maple worksheet `NSchwartz.mws` [30], shows how the centre may be calculated numerically. It may also be calculated as a formula in the sines an cosines of the three angles but the formula is not instructive. Finally, the matrices of the hyperbolic rotations at the vertices may now be calculated as in equation 4.5 above. The details are given in the Maple worksheet `NSchwartz.mws`. See the paper [5] for further detail on the construction of kaleidoscopic polygons.

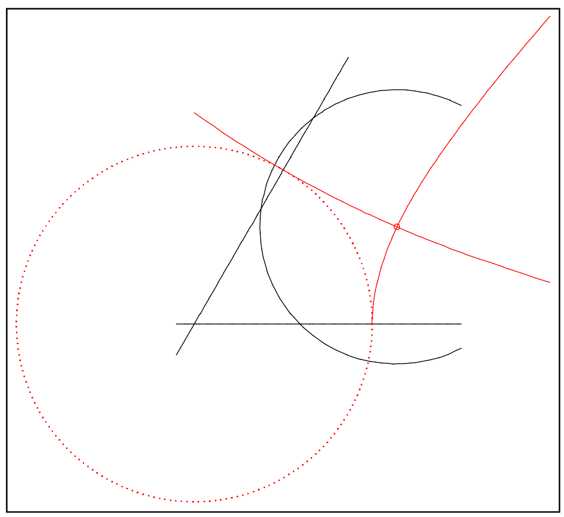

Figure 4.11 Schwartz triangle diagram.