A similar problem arises when we want to extend the transformation $x \mapsto 1/x$ of $\mathbb{RP}^1$ to the half plane. The appropriate extension is not $z \mapsto 1/z$ because this transformation does not map the upper half plane onto itself. In fact, if we write z in its polar form $z = r(\cos\theta + i\sin\theta)$, then

$$\frac{1}{z} = \frac{1}{r}(\cos\theta - i\sin\theta) = \frac{1}{r}(\cos(-\theta) + i\sin(-\theta))$$

because $\cos(-\theta) = \cos\theta$ and $\sin(-\theta) = -\sin\theta$. Thus, z (at angle θ) and $1/z$ (at angle $-\theta$) have opposite slopes from O. Hence, they lie in different half planes. The appropriate extension of $x \mapsto 1/x$ is $z \mapsto 1/\bar{z}$, which sends

$$z = r(\cos\theta + i\sin\theta) \quad \text{to} \quad 1/\bar{z} = \frac{1}{r}(\cos\theta + i\sin\theta)$$

lying in the same direction θ from O (Figure 8.3). This transformation is called *reflection* (or *inversion*) *in the unit circle*.

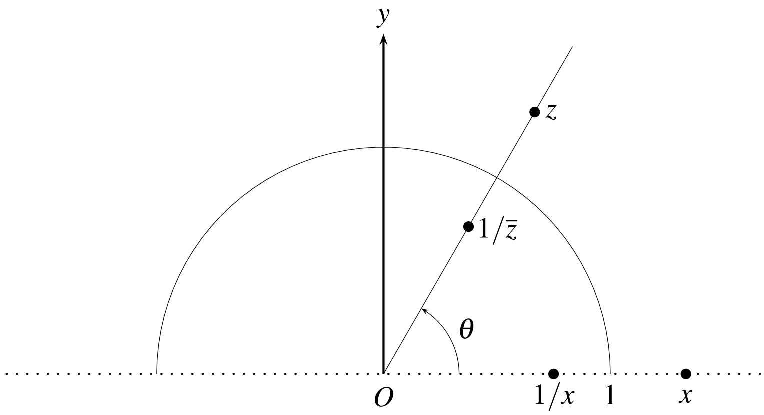

Figure 8.3: Extending inversion from the line to the half plane

Because all transformations $x \mapsto \frac{ax+b}{cx+d}$ of $\mathbb{RP}^1$ are products of $x \mapsto x+l$, $x \mapsto kx$, and $x \mapsto 1/x$, *their extensions to the half plane are products of*

- the horizontal translations $z \mapsto z + l$,
- the dilations $z \mapsto kz$ for $k > 0$,
- reflection in the y-axis $z \mapsto -\bar{z}$,
- reflection in the unit circle $z \mapsto 1/\bar{z}$.

We call these the *generating transformations of the half plane*.

Equations of non-Euclidean "lines"

Complex conjugation not only enables us to express reflection in lines and circles; it also enables us to write the equations of non-Euclidean "lines" very simply as equations in z.

- First consider "lines" that are actual Euclidean lines, namely those of the form $x = a$, where a is a real number. An arbitrary point on this line is of the form $z = a + iy$, so $\bar{z} = a - iy$, and z therefore satisfies the equation
$$z + \bar{z} = 2a. \tag{*}$$

- Next consider "lines" that are semicircles with centers on the x-axis. If the center is c and the radius is r, then any z on the circle satisfies
$$|z - c| = r, \quad \text{or equivalently,} \quad |z - c|^2 = r^2.$$
But now notice that for any complex number $x + iy$ we have
$$|x + iy|^2 = x^2 + y^2 = (x + iy)(x - iy) = (x + iy)\overline{(x + iy)}.$$
Hence,
$$|z - c|^2 = (z - c)\overline{(z - c)} = (z - c)(\bar{z} - \bar{c})$$
and the equation $|z - c|^2 = r^2$ becomes
$$(z - c)(\bar{z} - \bar{c}) = r^2,$$
that is,
$$z\bar{z} - c\bar{z} - \bar{c}z + c\bar{c} = r^2.$$
Finally, because c is a real number, we have $\bar{c} = c$, so the equation can be written as
$$z\bar{z} - c(z + \bar{z}) + c^2 - r^2 = 0. \tag{**}$$

The equations (*) and (**) are both of the form

$$Az\bar{z} + B(z + \bar{z}) + C = 0 \quad \text{for some} \quad A, B, C \in \mathbb{R}. \tag{***}$$

Conversely, if A and B are not both zero, then (***) reduces to one of the equations (*) or (**) above, if it is satisfied by any points z at all.

- If $A = 0$, then (***) becomes $z + \bar{z} + C/B = 0$, which is (*) with $2a = -C/B$.

- If $A \neq 0$, then (***) becomes $z\bar{z} + (z + \bar{z})B/A + C/A = 0$, which is (**) with $c = -B/A$ and $c^2 - r^2 = C/A$ if $r^2 = B^2/C^2 - C/A \geq 0$. If $r^2 < 0$, then no points z satisfy the equation (***), because the equation is equivalent to $|z - c|^2 = r^2$ and $|z - c|^2$ is necessarily > 0.

Thus, *equations of non-Euclidean "lines" are the satisfiable equations*

$$Az\bar{z} + B(z + \bar{z}) + C = 0, \quad \textit{where} \quad A, B, C \in \mathbb{R} \quad \textit{are not all zero.}$$

Exercises

I expect that most readers of this book are familiar with the complex numbers, but it still seems worthwhile to review the properties of the complex conjugate. Its role in geometric transformations may not be familiar, so we develop the basic facts from first principles.

8.2.1 Writing z_1 as $x_1 + iy_1$ and z_2 as $x_2 + iy_2$, show that $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$ and $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$.

8.2.2 Similarly, show that $\overline{1/z} = 1/\bar{z}$.

8.2.3 Deduce from Exercises 8.2.1 and 8.2.2 that, for any $a, b, c, d \in \mathbb{R}$ and $z \in \mathbb{C}$, the complex conjugate of $\frac{az+b}{cz+d}$ is $\frac{a\bar{z}+b}{c\bar{z}+d}$.

With these facts established, we are in a position to determine the extension to the half plane of each linear fractional transformation $x \mapsto \frac{ax+b}{cx+d}$ of $\mathbb{RP}^1$. What we know so far is that the extension of $x \mapsto x + l$ is $z \mapsto z + l$, the extension of $x \mapsto kx$ is $z \mapsto kz$ when $k > 0$ and $z \mapsto k\bar{z}$ when $k < 0$, and that the extension of $x \mapsto 1/x$ is $z \mapsto 1/\bar{z}$. We also know that any transformation $x \mapsto \frac{ax+b}{cx+d}$ is a product of these generating transformations. Hence, the extension of $x \mapsto \frac{ax+b}{cx+d}$ to the half plane is the product of the corresponding extensions. It seems likely that the latter product is either $z \mapsto \frac{az+b}{cz+d}$ or $z \mapsto \frac{a\bar{z}+b}{c\bar{z}+d}$, so the main problem is to decide *when* the product is $z \mapsto \frac{az+b}{cz+d}$ and when it is $z \mapsto \frac{a\bar{z}+b}{c\bar{z}+d}$.

8.2.4 Write each generating transformation of $\mathbb{RP}^1$ in the form $x \mapsto \frac{ax+b}{cx+d}$, and hence, show that those whose extension involves $\bar{z}$ are precisely those for which $ad - bc < 0$.

8.2.5 Deduce from Exercise 8.2.4 and Exercise 5.6.3 that the extension of a product, of transformations $x \mapsto \frac{a_1x+b_1}{c_1x+d_1}$ and $x \mapsto \frac{a_2x+b_2}{c_2x+d_2}$, is the product of their extensions.

8.2.6 Deduce from Exercise 8.2.5, or otherwise, that the extension of $x \mapsto \frac{ax+b}{cx+d}$ is $z \mapsto \frac{az+b}{cz+d}$ when $ad - bc > 0$ and $z \mapsto \frac{a\bar{z}+b}{c\bar{z}+d}$ otherwise.

It may seem unfortunate that the extension of $x \mapsto \frac{ax+b}{cx+d}$ is one of two different types: a function of z or a function of $\bar{z}$. However, these two algebraic types are inevitable because they reflect a geometric distinction: *The functions of z are orientation-preserving, and the functions of $\bar{z}$ are not.* In particular, *the linear fractional transformations $z \mapsto \frac{az+b}{cz+d}$ with $ad - bc > 0$ are precisely the orientation-preserving transformations of the half plane.*

8.3 Reflections and Möbius transformations

The extensions of the transformations $x \mapsto \frac{ax+b}{cx+d}$ from $\mathbb{RP}^1$ to the half plane could be called "linear fractional," but this would be confusing, because one half of them are linear fractional functions of z and the other half are linear fractional functions of $\bar{z}$. Instead they are called *Möbius transformations*, after the German mathematician August Ferdinand Möbius. In 1855, Möbius introduced a theory of transformations generated by reflections in circles, using the obvious generalization from reflection in the unit circle to reflection in an arbitrary circle. We will see below that all Möbius transformations of the half plane are products of reflections.

One advantage of the reflection idea is that it makes sense in three (or more) dimensions, where *reflection in a sphere* is meaningful but "linear fractional transformation" generally is not. It is also revealing to view the transformations of $\mathbb{RP}^1$ as the restrictions of Möbius transformations of the half plane, as this brings to light a concept of "projective reflection".

Reflection in an arbitrary circle is defined by generalizing the relationship between z and $1/\bar{z}$ shown in Figure 8.3. We say that points Q and Q' are *reflections of each other in the circle with center P and radius r* if P, Q, Q' lie in a straight line and $|PQ||PQ'| = r^2$ (Figure 8.4).

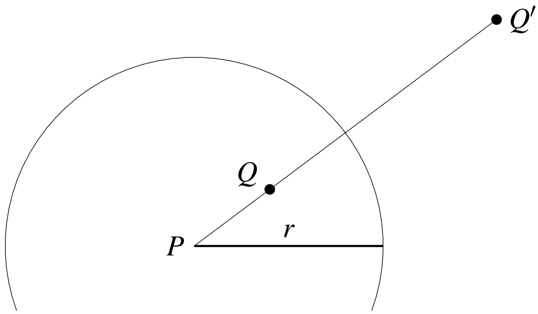

Figure 8.4: Reflection in an arbitrary circle

If the circle (or, rather, its upper half) is a non-Euclidean line, then the center P lies on the x-axis, and reflection in this circle can be composed from generating transformations of the half plane as follows:

- translate P to O,
- reduce the radius to 1 by dilating by $1/r$,
- reflect in the unit circle,
- restore the radius to r by dilating by r,
- translate the center from O back to P.

Likewise, reflection in an arbitrary vertical line, say $x = a$, can be composed from generating transformations of the half plane as follows:

- translate the line $x = a$ to the y-axis,
- reflect in the y-axis,
- translate the y-axis to the line $x = a$.

Thus, *all reflections in non-Euclidean lines are products of generating transformations of the half plane*.

Conversely, we now show that *every generating transformation of the half plane is a product of reflections* (and hence so is every transformation of the half plane). The generating transformations $z \mapsto -\bar{z}$ and $z \mapsto 1/\bar{z}$ are reflections by definition, so it remains to deal with the remaining generating transformations.

- the horizontal translation $z \mapsto z + l$: this is a Euclidean translation, and it is the product of reflections in the lines $x = 0$ and $x = l/2$.
- the dilation $z \mapsto kz$, where $k > 0$: this is the product of the reflection $z \mapsto 1/\bar{z}$ in the unit circle and the map $z \mapsto k/\bar{z}$, which is reflection in the circle with center O and radius $\sqrt{k}$. $\square$

It should be mentioned that ordinary reflection—reflection in a straight line—is the limiting case of reflection in a circle obtained by letting P and r tend to infinity in such a way that the circle tends to a straight line. Because Euclidean lines are the fixed point sets of ordinary reflections, it is natural that the "lines" of the half plane should be the fixed point sets of its "reflections."

Projective reflections

Looking back from the half plane to its boundary line $\mathbb{RP}^1$, we realize that we now know more about projective transformations of the line than we did before. *Any projective transformation of* $\mathbb{RP}^1$ *is a product of projective reflections*, where a *projective reflection* is the restriction, to $\mathbb{RP}^1$, of a reflection of the half plane.

There is a "three reflections theorem" for $\mathbb{RP}^1$, analogous to the three reflections theorem for isometries of the Euclidean plane (Section 3.7). This follows from a three reflections theorem for the half plane, similar to the one for the Euclidean plane, that we will prove in Section 8.8.

Exercises

The simplest reflections of $\mathbb{RP}^1$ are ordinary reflection in O, $x \mapsto -x$, and the restriction of reflection in the unit circle, $x \mapsto 1/x$. The map $x \mapsto 1/x$ might be called "reflection in the point-pair $\{-1, 1\}$," and it generalizes to "reflection in the point-pair $\{a, b\}$." (A point-pair $\{a, b\}$ is a "0-dimensional sphere," because it consists of the points at constant distance $(b-a)/2$ from the "center" $(a+b)/2$.)

8.3.1 Write down the formula for ordinary reflection in the point $x = a$.

8.3.2 Explain why the map $x \mapsto c^2/x$ is reflection in the point-pair $\{-c, c\}$.

8.3.3 Using Exercise 8.3.2, or otherwise, show that reflection in the point-pair $\{a, b\}$ is given by the linear fractional function

$$f(x) = \frac{x(a+b) - 2ab}{2x - (a+b)}.$$

8.3.4 Show that, as $b \to \infty$, the function for reflection in the point-pair $\{a, b\}$ tends to the function for ordinary reflection in the point $x = a$.

8.4 Preserving non-Euclidean lines

We have now extended the projective transformations $x \mapsto \frac{ax+b}{cx+d}$ of $\mathbb{RP}^1$ to Möbius transformations $z \mapsto \frac{az+b}{cz+d}$ or $z \mapsto \frac{a\bar{z}+b}{c\bar{z}+d}$ of the half plane, but are Möbius transformations of the half plane any easier to understand? We intend to show that they are, by showing that they have more easily visible invariants than the transformations of $\mathbb{RP}^1$. First we show the invariance of *non-Euclidean lines*, which we now define officially as the vertical lines $x =$ constant and the semicircles with centers on the x-axis.

Each Möbius transformation of the half plane maps non-Euclidean lines to non-Euclidean lines.

For the generating transformations $z \mapsto z+l$, $z \mapsto kz$ for $k>0$, and $z \mapsto -\bar{z}$, this is easy to see. Each of these transformations sends vertical lines to vertical lines, circles to circles, and the x-axis to the x-axis because:

- $z \mapsto z+l$ is a horizontal translation of the half plane.
- $z \mapsto kz$ with $k>0$ is a dilation of the half plane by k.
- $z \mapsto -\bar{z}$ is the Euclidean reflection of the half plane in the y-axis.

Thus, any product of the three transformations just listed sends vertical lines to vertical lines and semicircles with centers on the x-axis to semicircles with centers on the x-axis. Hence, *all products of the transformations $z \mapsto z+l$, $z \mapsto kz$ for $k>0$, and $z \mapsto -\bar{z}$ preserve non-Euclidean lines.*

To show that all Möbius transformations preserve non-Euclidean lines, it therefore remains to show that *reflection in the unit circle, $z \mapsto 1/\bar{z}$, preserves non-Euclidean lines.* This is less obvious, because reflection in a circle can send a vertical line to a semicircle and vice versa. We prove that non-Euclidean lines are preserved by using their equations (***) derived in Section 8.3.

Given a non-Euclidean line, whose points z satisfy an equation

$$Az\bar{z}+B(z+\bar{z})+C=0 \quad \text{for some} \quad A,B,C\in\mathbb{R}, \tag{***}$$

we wish to find the equation satisfied by the points of its reflection in the unit circle. These are the points $w=1/\bar{z}$, so we seek the equation satisfied by w. The required equation is likely to involve $\overline{w}=1/z$ as well, so we are looking for an equation connecting $1/z$ and $1/\bar{z}$. Such an equation is easy to find: just divide the equation (***) by $z\bar{z}$. Division yields the equation

$$A+B\left(\frac{1}{\bar{z}}+\frac{1}{z}\right)+\frac{C}{z\bar{z}}=0,$$

that is,

$$Cw\overline{w}+B(w+\overline{w})+A=0. \tag{****}$$

Equation (****) is satisfied by the reflections $w=1/\bar{z}$ of the points z satisfying (***), and (****) has the same form as (***), because $A,B,C\in\mathbb{R}$. Hence, (****) also represents a non-Euclidean line. □

Exercises

An example in which reflection in the unit circle sends a vertical line to a semicircle is shown in Figure 8.5.

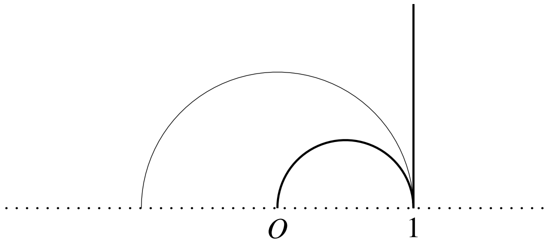

Figure 8.5: Reflection of the line $x = 1$

8.4.1 Give intuitive reasons why the reflection of the line $x = 1$ in the unit circle should have one end at 1 on the x-axis and the other end at O.

8.4.2 Show that the line $x = 1$ has equation $z+\overline{z} = 2$, and that its reflection in the unit circle has equation $w+\overline{w} = 2w\overline{w}$.

8.4.3 Verify that $w+\overline{w} = 2w\overline{w}$ is the equation of the semicircle with ends O and 1 on the x-axis.

8.5 Preserving angle

Next to non-Euclidean lines, the most visible invariant of Möbius transformations is *angle*. Because non-Euclidean lines are not necessarily straight, the angle between two of them is really the angle between their tangents at the point of intersection. Nevertheless, it is easy to see the angle between non-Euclidean lines. Figure 8.6 shows an example.

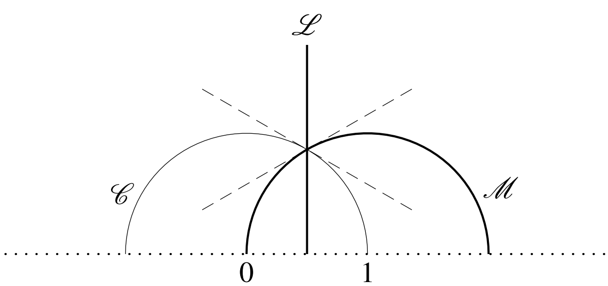

Figure 8.6: Some non-Euclidean lines and the angles between them

The three non-Euclidean lines are the unit circle $\mathscr{C}$, the vertical $\mathscr{L}$ where $x = 1/2$, and its reflection $\mathscr{M}$ in the unit circle, which happens to be the semicircle with endpoints 0 and 2 on the x-axis.

At the point where the three non-Euclidean lines meet, they divide the space around the point into six equal angles, so each angle is $2\pi/6 = \pi/3$. This equality is confirmed by the tangents, which are shown as dashed lines. Notice that any two of $\mathscr{C}$, $\mathscr{L}$, and $\mathscr{M}$ are reflections of each other in the third non-Euclidean line, so the figure shows numerous instances of an angle equal to its reflection. To show that *any* angle is preserved by *any* Möbius transformation, we look once again at the properties of the generating transformations.

The effect of Möbius transformations

The Möbius transformations $z \mapsto z + l$ and $z \mapsto -\bar{z}$ are Euclidean isometries; hence, they certainly preserve angle (along with length, area, and so on). The Möbius transformations $z \mapsto kz$ for $k > 0$ are dilations; hence, they too preserve angle. Thus, *it suffices to prove that angle is preserved by the remaining generator of Möbius transformations: reflection in the unit circle*, $z \mapsto 1/\bar{z}$. The latter transformation is the composite of $z \mapsto -\bar{z}$ and $z \mapsto -1/z$, so it suffices in turn to prove that $z \mapsto -1/z$ preserves angle.

We therefore concentrate our attention on the Möbius transformation $z \mapsto -1/z$. This transformation does not in general preserve Euclidean lines, because it may map them to circles. Thus, we need to be aware that "angle" generally means the angle between curves and hence the angle between the tangents. However, we can avoid computing the position of tangents by taking the *infinitesimal* view of angle. That is, we study what becomes of the direction between two points, z and $z + \Delta z$, when we send them to $-1/z$ and $-1/(z + \Delta z)$, respectively, and let Δz tend to zero.

If Δz is the point at distance ε from O in direction θ, then

$$\Delta z = \varepsilon(\cos\theta + i\sin\theta),$$

because $\cos\theta + i\sin\theta$ is the point at distance 1 from O in direction θ. It follows that the point at distance ε from z in direction θ is

$$z + \Delta z = z + \varepsilon(\cos\theta + i\sin\theta),$$

and that the point $z + \Delta z$ tends to z in the constant direction θ as ε tends to zero.

The difference between the image points $-1/(z+\Delta z)$ and $-1/z$ is therefore

$$\begin{aligned}\frac{1}{z}-\frac{1}{z+\varepsilon(\cos\theta+i\sin\theta)} &= \frac{z+\varepsilon(\cos\theta+i\sin\theta)-z}{z(z+\varepsilon(\cos\theta+i\sin\theta))}\\ &= \frac{\varepsilon(\cos\theta+i\sin\theta)}{z(z+\varepsilon(\cos\theta+i\sin\theta))}.\end{aligned}$$

Now as ε tends to zero, this difference is ever more closely approximated by

$$\frac{\varepsilon(\cos\theta+i\sin\theta)}{z^2}.$$

To be precise, *the direction from* $-1/z$ *to* $-1/(z+\Delta z)$ *tends to the direction of* $\varepsilon(\cos\theta+i\sin\theta)z^{-2}$*, which is* $\theta+$constant. The constant is the argument (angle) of z^{-2}, recalling from Section 4.7 that the argument of a product of complex numbers is the sum of their arguments.

The angle between two smooth curves meeting at z (approximated by the difference in directions from z to points $z+\Delta_1 z$ and $z+\Delta_2 z$ on the respective curves) is therefore the angle between the images of these curves under the map $z\mapsto -1/z$. This is because a smooth curve is one for which the direction from z to $z+\Delta z$ tends to a constant θ as $z+\Delta z$ tends to z along the curve. Non-Euclidean lines are smooth, so the angle between them is preserved by the transformation $z\mapsto -1/z$, as required. □

Tilings of the half plane

If one takes a triangle with angles π/p, π/q, π/r, for some natural numbers p,q,r, then any reflection of that triangle will have angles π/p, π/q, π/r. Reflecting the reflections causes the space around each vertex to be exactly filled with corners of triangles. For example, the space around the vertex of angle π/p becomes filled with $2p$ corners of angle π/p. In fact, the *whole half plane* becomes filled, or *tiled*, by copies of the original triangle. An example is shown in Figure 8.7, where the basic tile has angles $\pi/2$, $\pi/3$, and $\pi/7$.

Notice that the angle sum $\pi/2+\pi/3+\pi/7$ is less than π. In fact, the angle sum of any triangle bounded by non-Euclidean lines is less than π, and *the quantity* ($\pi -$ angle sum) *is proportional to the area of the triangle.* This elegant result is less surprising when one learns that the area of spherical triangle is also proportional to $\pi -$ angle sum (see exercises below).

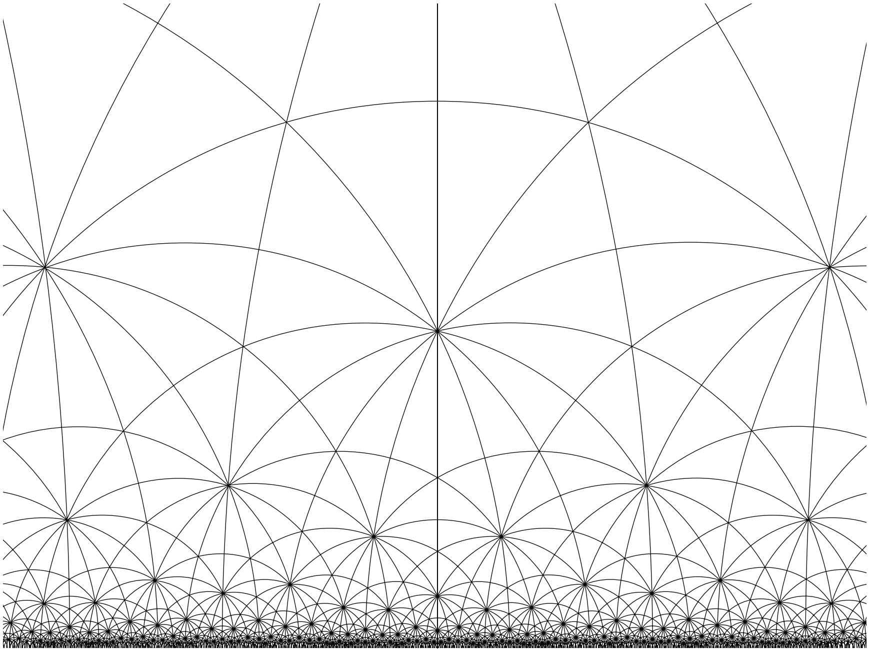

Figure 8.7: Tiling by repeated reflections

However, it does reveal a limitation in the half plane view of non-Euclidean geometry: All the triangles in Figure 8.7 have equal non-Euclidean area, but they certainly do not look equal!

One should think of the half plane as a kind of “perspective view” of the non-Euclidean plane with the x-axis as a horizon. The x-axis is infinitely distant, because there are infinitely many identical triangles between any point of the half plane and the x-axis. In this respect, the half plane is like a perspective view of a Euclidean tiled floor, except that ordinary perspective preserves straightness and distorts angle, whereas this “non-Euclidean perspective” distorts straightness and preserves angle. There are other views of the non-Euclidean plane that make non-Euclidean lines look straight (see Section 8.9), but any such view has a curved horizon!

Another way in which a tiling of the half plane resembles a perspective view is that one can estimate the length of a line by counting the numbers of tiles that lie along it. There is indeed a non-Euclidean measure of distance that is invariant under Möbius transformations, and we will see exactly what it is in Section 8.6.