Non-Euclidean graphics

Bill Casselman
University of British Columbia
cass@math.ubc.ca

This note is concerned with practical algorithms to carry out basic drawing exercises in non-Euclidean geometry.

There are several possible models of non-Euclidean geometry to choose from. The one I shall mostly work with is that in which the non-Euclidean plane is identified with the interior of the unit disk $\mathbb{D}$ embedded in the complex numbers $\mathbb{C}$. The metric expressed in Riemann's terms is

$$\frac{dx^2+dy^2}{(1-r^2)^2} \quad (r^2 = x^2+y^2)\,.$$

The geodesics are either (a) the arcs of circles inside the disk orthogonal to its boundary or (b) diameters. In this model, the non-Euclidean isometries are the Möbius transformations

$$z \longmapsto \frac{\alpha z+\beta}{\overline{\beta} z+\overline{\alpha}}$$

associated to matrices

$$\begin{bmatrix} \alpha & \beta \\ \overline{\beta} & \overline{\alpha} \end{bmatrix}$$

with $\det = |\alpha|^2 - |\beta|^2 = 1$. This is the special unitary group of the Hermitian form $|w|^2 - |z|^2$, and it preserves the unit disk precisely because it preserves this form. Möbius transformations take circles and lines to circles and lines, and preserve angles. So any Möbius transformation that takes the closed unit circle to itself takes the geodesic arcs defined here to other geodesic arcs.

Another model is that in which the plane is the Poincaré upper half plane $\mathfrak{H}$, and the geodesics are either vertical lines or semi-circles perpendicular to the real axis. Here the group of isometries is the projective image of $\mathrm{SL}_2(\mathbb{R})$. Although the disk model is the one we shall be primarily interested in, it is sometimes convenient to do things on $\mathfrak{H}$ and then transform results into the unit disk. This is done by means of the **Cayley transform**

$$C\colon\ z \longmapsto \frac{z-i}{z+i}\,,$$

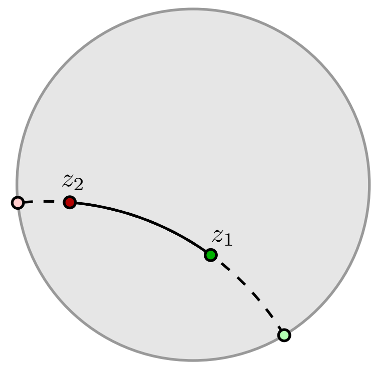
which maps $\mathfrak{H}$ isometrically (in the sense of non-Euclidean geometry) onto the interior of $\mathbb{D}$. Thus the real matrix g in $\mathrm{SL}_2(\mathbb{R})$ acts as CgC^{-1} on $\mathbb{D}$.

A third model is the Kleinian one where the plane is still the interior of the unit disk, but the geodesics are straight line segments. Here the isometries are the projective transformations generated by planar slices of the cone of 2×2 matrices with trace 0, on which $\mathrm{SL}_2(\mathbb{R})$ acts by conjugation. I shall discuss it in one of the last sections.

Contents

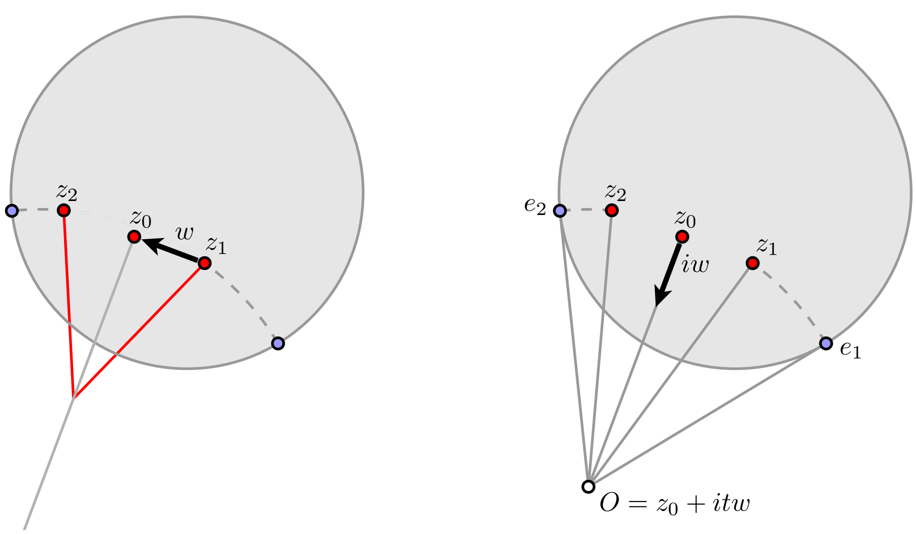
1. Geodesic segments in the disk: circular arcs

The basic problem in non-Euclidean graphics is to draw the non-Euclidean geodesic segment between two points z_1 and z_2 inside the unit disk.



This is generically an arc of a circle, but it also might be a diameter. We must decide which case occurs, and in the case of a circular arc determine the center O of the arc. Solving this problem involves only elementary geometry, but limitations of machine floating point computation makes it a bit more interesting. The reason for this is essentially mathematical: a small Euclidean length ds at a point at radius r corresponds to a non-Euclidean length $ds/(1-r^2)$, which certainly blows up approaching $r = 1$. Points that are close in the Euclidean metric may be far apart in the non-Euclidean one, if they are near the boundary of $\mathbb{D}$. A small error in Euclidean coordinates can translate to a large error in non-Euclidean geometry.

I'll begin by assuming that the geodesic on which z_1 and z_2 lie is the arc of a circle (instead of on a diameter), and find a formula for the centre O of the circle. The distance from O to z_1 is the same as that from O to z_2, which means neither more nor less than that it lies on the perpendicular bisector of the segment between them. To find this, let w be the vector halfway from z_1 to z_0, and let z_0 be the midpoint of the Euclidean segment from z_1 to z_2, which is the vector sum $z_1 + w$. The center O we are looking for will be somewhere on the line through z_0 perpendicular to w.

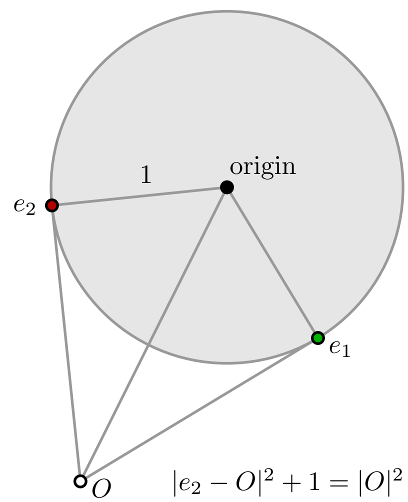


Since O is at equal distance from z_1 and z_2, we have

$$O = z_0 + itw$$

for some real number t. The exact location of the centre (i.e. the value of t) is determined by the condition that the tangent line from O to the unit circle grazes the circle at the end points e_i of the complete arc through the z_i. This gives the condition

$$|e_1 - O|^2 = |e_2 - O|^2 = |O|^2 - 1\,.$$



But the distance from O to e_2 is the same as that from O to z_2, so we obtain the equation

$$|z_2 - O|^2 = |O|^2 - 1\,.$$

In this we set $z_2 = z_0 + w$, $O = z_0 + itw$ and get equations to be solved for t.

Since

$$z_2 - O = (z_0 + w) - (z_0 + itw) = w(1 - it)\,,$$

we solve:

$$\begin{aligned} |w(1-it)|^2 &= |w|^2(1+t^2) \\ &= |z_0|^2 + t^2|w|^2 + 2(z_0 \bullet itw) - 1 \\ &= |z_0|^2 + t^2|w|^2 + 2t(z_0 \bullet iw) - 1 \\ t &= \frac{|w|^2 + 1 - |z_0|^2}{2(z_0 \bullet iw)}\,. \end{aligned}$$

Here $w \bullet z$ is the Euclidean dot product of w and z considered as 2D vcetors.

To summarize the sequence of calculations:

$$\begin{aligned} z_0 &= 0.5 * (z_1 + z_2) \\ w &= 0.5 * (z_2 - z_1) \\ t &= \frac{|w|^2 + 1 - |z_0|^2}{2(z_0 \bullet iw)} \\ O &= z_0 + itw\,. \end{aligned}$$

Let's look more closely at the formula for t.

1.1. Lemma. *Given z_1, z_2 in the closed unit disk $\mathbb{D}$.*

- *The numerator in the formula above for t vanishes only if $z = z_1 = z_2$ lies on the unit circle.*
- *The denominator vanishes only if z_1 and z_2 lie on a diameter of the circle.*

Proof. Since the disk is strictly convex, $|z_0|$ will be less than 1 if $z_1 \neq z_2$. If $z_1 = z_2$, then $z_0 = z_1$ and $|z_0| < 1$ unless $|z_1| = 1$. But $z_1 = z_2$ and $|z_0| = 1$ is equivalent to the condition that the numerator equals 1.

If z_1 and z_2 lie on a diameter, $z_2 = cz_1$ for sone real c and

$$2w = z_2 - z_1 = (c-1)z_1, \quad z_0 \bullet iw = (c-1)(w \bullet iw) = 0\,.$$

If z_1 and z_2 do not lie on a diameter, the Euclidean line through them passes at a distance $\delta > 0$ from the origin. Let u be the vector of length δ from the origin to this line and perpendicular to it. It will be a non-zero multiple of iw. Let v be a unit vector parallel to it. The line consists of all points of the form $u + sv$. In particular, z_0 is on this line, and hence $z_0 \bullet iw \neq 00$.

So the simple mathematical rule is this:

If $z_0 \bullet iw = 0$, draw the straight line from z_1 to z_2. Otherwise find O and draw the arc from z_1 to z_2 with center O.

For this, we need to know how to draw the circular arc from z_1 to z_2 centred at O. This is simple—if α_i is the argument of $z_i - O$ then we draw the arc from α_1 to α_2, in the positive direction if $t > 0$ and in the negative direction if $t < 0$.

The problem with this rule is that computer calculations with real numbers are not exact, and in practice is not possible to tell if $z_0 \bullet iw$ is exactly equal to 0. We want a test that is a bit more stable under perturbations. Also, we don't want to find ourselves drawing arcs of circles of very large radii (and large t).

The trick is to introduce the variable

(1.2) $$T = \frac{1}{t} = \frac{2(z_0 \bullet iw)}{|w|^2 + 1 - |z_0|^2}\,.$$

This will be 0 if and only if $t = \infty$, which means that the geodesic arc is a straight line. Also, we have just seen that the denominator bever vanishes, so T is well defined. A closer inspection will disclose that the denominator is close to 0 only when z_1 and z_2 are close to each other and to the boundary of the unit disk. But as I have already remarked, troubles in that region are to be expected.

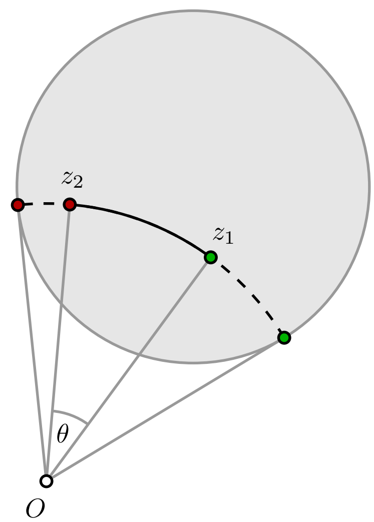
We shall see in the next section that the number T is in fact an accurate measure of the straightness of the geodesic from z_1 to z_2. If T is small, this geodesic will be very close to a straight line, and it turns out that we can find easily a Bézier curve that approximates it well. If T is not small, we can use it to locate O as well as calculate accurately the angle θ between the rays Oz_i. The new rule is therefore:

Calculate T from (1.2) . If it is suitably small, draw the arc from z_1 to z_2 as a Bézier curve. Otherwise find O and draw the arc from z_1 to z_2 with center O.

It remains to explain how to find that Bézier arc in a stable way. Also, again because of machine treatment of floating point numbers, evaluating the denominator in the formula for T is not straightforward.

2. Interpreting T

In this section I'll justify the use of T in a test for straightness of the non-Euclidean path from z_1 to z_2. The obvious measure of straightness is the angle θ spanned by the arc.



But θ and T are closely related:

2.1. Proposition. *In these circumstances,* $T = \arctan(\theta/2)$.

Proof. We know that

$$\begin{aligned} z_2 - O &= z_2 - (z_0 + itw) \\ &= (z_2 - z_0) - itw \\ &= w - itw \\ &= w(1 - it) \\ z_1 - O &= z_1 - (z_0 + itw) \\ &= -w - itw \\ &= -w(1 + it)\,. \end{aligned}$$

So we get

$$\begin{aligned} \frac{z_2 - O}{z_1 - O} &= -\frac{1 - it}{1 + it} \\ &= \frac{it - 1}{it + 1} \\ &= \frac{1 - 1/it}{1 + 1/it} \\ &= \frac{1 + iT}{1 - iT}\,. \end{aligned}$$

But $(z_2 - O)/(z_1 - O) = e^{i\theta}$, so this implies that if θ is the angular span of the arc then $\theta/2$ is the argument of $1 + iT$.

This gives us

$$\theta/2 = \arctan(T)$$

Thus T is nearly the same as $\theta/2$ if it is small. So our rule for drawing the segment is justified.

3. Geodesic segments in the disk: using Bézier curves

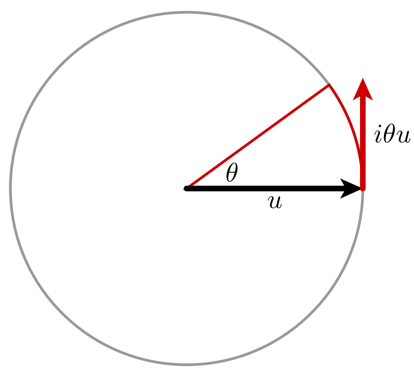
If T is small, we are going to draw a Bézier curve from z_1 to z_2. For this purpose, we must calculate the velocity vectors for the normalized path from one to the other. As explained in Chapter 6 of **Mathematical illustrations**, these determine the control points P_i of the curve:

Given the normalized tangent vectors v_i, we can find the intermediate control points of the approximating Bézier curve (as explained in §6.5 of Mathematical Illustrations):

$$\begin{aligned} P_1 &= z_1 + (1/3)v_1 \\ P_2 &= z_2 - (1/3)v_2 \,. \end{aligned}$$

This is just the same as the very general problem of finding velocity vectors at the start and end of a trajectory with uniform speed along a circular arc of angular span θ whose radius vector u at the start of the arc is given.

Let $r = \|u\|$. The unit tangent vector (in the positive direction of travel) at the end of u is iu/r. If we traveled along the circle for time 1 at constant speed, starting out at velocity iu/r, we would travel a unit length of arc. We want to travel an angle θ along the arc. Since θ is measured in radians, and a radian of angle θ means arc distance θr. So we start out at velocity $i(\theta r)(u/r) = i\theta u$.



In our case, the radius vector is $-w(1 + it)$, so we get

$$\begin{aligned} v_1 &= i \cdot -w(1 + it) \cdot \theta \\ &= -2T \cdot i \cdot w(1 + it) \cdot \frac{\theta}{2T} \\ &= -2Tw(i - t) \cdot \frac{\theta}{2T} \\ &= -2w(i/t - 1) \cdot \frac{\theta}{2T} \\ &= 2w(1 - iT) \cdot \frac{\theta}{2T} \end{aligned}$$

and similarly

$$v_2 = 2w(1 + iT) \cdot \frac{\theta}{2T}$$

Because $\theta/2T \sim 1$ as $T \to 0$, these formulas make sense even in the case where the geodesic line is a diameter.

Notice that according to Proposition 2.1,

$$\frac{\theta}{2T} = \arctan(T)/T = (1/T)\int_0^T \frac{ds}{1+s^2} = 1 - \frac{T^2}{3} + \frac{T^4}{5} - \cdots$$

For T small, this gives an efficient way to compute $\theta/2T$.

4. Technical problems

Because of the way computers deal with real numbers (as opposed to integers), the input data, which is in this case the two points z_1 and z_2, will usually only approximate in the sense that they will be the result of computations which are only approximate. I ask a question one should always ask in dealing with real numbers in computers: Do small variations in the input data affect the output? Do small changes in z_1 and z_2 affect the arc that's drawn? Is this an artefact of the particular technique used to do the drawing, or can we do better with a different one?

First of all, I recall that computers handle real numbers in floating point format, which means that it stores significant digits. Thus a machine with 10 figure accuracy would store $\pi/1000$ as 3.14159265410^{-3} instead of 0.003141593. On the whole this scheme is very efficient, but it also causes misunderstanding if one is not careful.

Let me illustrate what I mean by a simple example from high school algebra. The solutions to the quadratic equation

$$x^2 + ax + b = 0$$

are given by the formula

$$x = -a/2 \pm \sqrt{(a/2^)2 - b}\,.$$

Most of the time this formula can be used without trouble, but occasionally a difficulty arises. Consider

$$x^2 - 1634x + 2 = 0.$$

On my calculator, which handles real numbers with 10 significant figures, I get solutions

$$x = 817 \pm 667,487x_1 = 817 + 816.9987760 = 1633.998776x_2 = 817 - 816.9987760 = 0.0012240.$$

What's going on here is that the original data is specified exactly, and one of the roots is given correctly to 10 significant decimals, but the accuracy of the second root is only 4 significant figures. The problem arises in the subtraction $817 - 816.9987760$. Both figures are accurate to 10 significant figures, and this cannot be improved. But the subtraction involves a loss of accuracy. This is not necessary. We know that the product $x_1x_2 = 2$, so we can set

$$x_2 = 2/x_1 = 0.001223991125\,,$$

which is correct to 10 significant figures. This leads to a modification of the usual rule for solving quadratic equations:

One root is calculated by the formula

$$x^2 + ax + b = 0x_1 = -(a/2) - \varepsilon p(a/2) - b$$

where (a) the square root here always means the positive root, and (b) ε is chosen to be ± 1, so that the second term has the same sign as a. The other root is given by

$$x_2 = b/x_1.$$

In other words, it is the careless use of the quadratic formula that causes difficulty here, intrinsic to the problem. This phenomenon is called cancellation error, and is discussed elegantly in Chapter 1 of [Henrici:1982], from which this example was taken.

In our case, if $|z_0|$ is near 1 we get cancellation problems in calculating $1 - |z_0|^2$, which occurs in the denominator of the formula for T. In some sense, this phenomenon is unavoidable, if we are given only the points z_1 and z_2 as data. If we are placed at a point at radius r from the origin in the unit disk and move a small Euclidean distance dr, then the non-Euclidean distance moved is $dr/(1-r^2)$. (This can be justified by using the fact that along the imaginary axis in the upper half plane $\mathfrak{H}$ the metric is dy/y.) Hence, as I have already said, perturbations in the data z_1, z_2 are magnified.

But in practice the points z_1 and z_2 are found as the non-Euclidean transformations of other points, and this gives us a little more information. There is a very useful formula in these circumstanmces.

Remark. If z is any point of $\mathbb{C}$ with $|z| \neq 1$ and

$$g = \begin{bmatrix} \alpha & \beta \\ \overline{\beta} & \overline{\alpha} \end{bmatrix}$$

then

$$1 - |g(z)|^2 = \frac{1-|z|^2}{|\overline{\beta}z + \overline{\alpha}|^2} \,.$$

This is easy to verify, and explains among other things why the matrices in SU take $\mathbb{D}$ into itself.

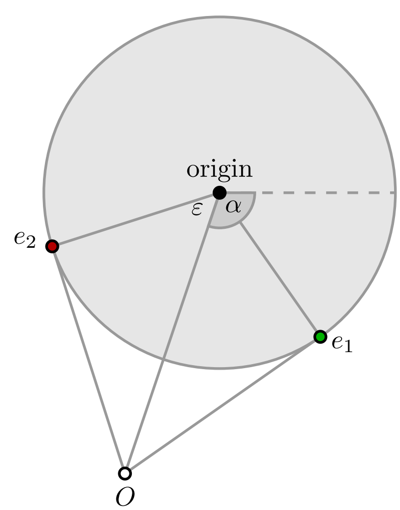
The quantity $1-|z|^2$, for a point z in $\mathbb{D}$, occurs in many computations. If z is close to 1, computing it from representating z as $x+iy$ will involve cancellation error. I have been tempted to incorporate this quantity as part of the data of a non-Euclidean point, but have not seen so far a pressing need to do so.

5. Endpoints of geodesics in the disk

In the rest of this essay, I'll discuss how to solve other problems involving non-Euclidean computation, particularly involving graphics.

The first is, given points z_1 and z_2, how does one find the endpoints e_1, e_2 on the unit circle of the arc through them?

Suppose first that we are looking at the case in which we plot the circular arc between the two points. Let O be the centre of the arc, with argument α. Then the arguments of the endpoints are $\alpha \pm \varepsilon$ where $\cos\varepsilon = 1/|O|$.



But now we receive a small surprise—this makes sense even when O is very large, since O is never 0, and

$$\frac{1}{|O|} = \frac{1}{z_0 + iwt} = \frac{1}{t}\,\frac{1}{z_0/t + iw} = \frac{T}{z_0T + iw} \,.$$

This offers trouble only when z_1 and z_2 are close to each other and to the origin. But in that case even small changes in their locations can seriously affect the location of the end-points. Nothing we can do about it.

No matter which of these methods we use, we have to decide which endpoint goes with which z_i. The correct assignment is that making $(e_2 - e_1) \bullet (z_2 - z_1)$ positive.

6. Triangles in the disk

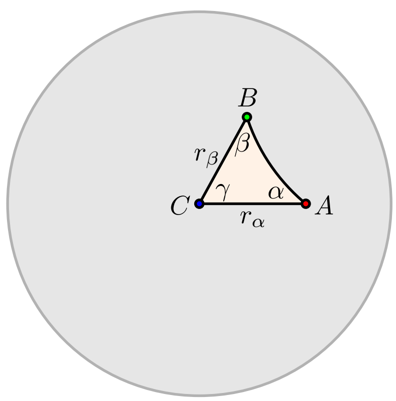
The next problem is one of drawing triangles. In non-Eucliean geometry, unlike Euclidean geometry, the congruence class of a triangle is completely determined by its interior angles, say α, β, γ with $\alpha + \beta + \gamma < \pi$. That is to say, any two triangles with the same angles may be transformed into one another by a Möbius transformation in SU. For example, the non-Euclidean area of the triangle is completely determined:

$$\text{area} = \pi - (\alpha + \beta + \gamma).$$

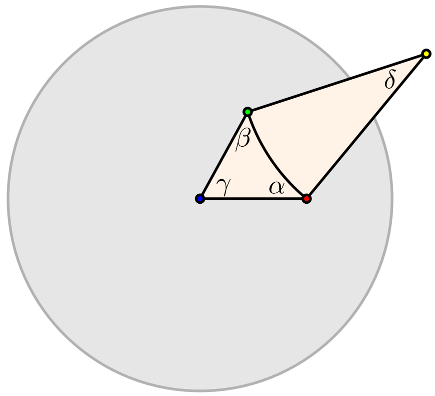
There is no notion of similarity in non-Euclidean geometry.

The problem then arises, is we are given the interior angles, one of the vertices, say C, and the ray from A along one side, how do we find the other vertices A and B?

We may assume that C is the origin, that A is on the positive x-axis, and that B is above A. The triangle is now uniquely determined, as in this figure:



We must find the numbers $r_\alpha = |A|$, $r_\beta = |B|$ (Euclidean lengths). This can be done by following a recipe explained to me by Robert Bédard. First, extend the diagram to include the center D of the non-Euclidean geodesic through A and B. This is possible, because in these circumstances the side through A, B is certainly not a diameter.

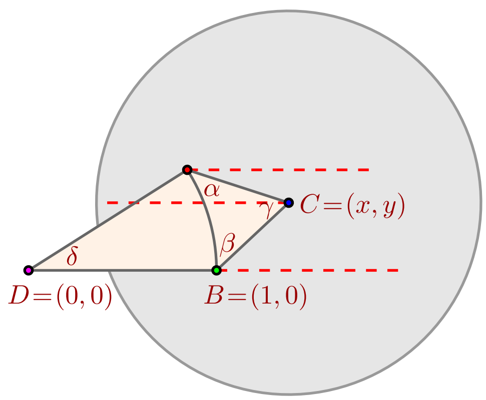


The angle δ satisfies

$$(\alpha + \pi/2) + (\beta + \pi/2) + \gamma + \delta = 2\pi, \quad \delta = \pi - (\alpha + \beta + \gamma)\,,$$

confirming that $\pi - (\alpha + \beta + \gamma)$ is positive.

Then, the figure above is rotated and shifted, so that D becomes the origin, and the line through the origin and B becomes the positive x-axis. We are going to calculate a ratio, so we may scale the diagram, too, making $B = (1, 0)$.



We must now find the coordinates (x, y) of C in the new coordinate system. A simple calculation of angles shows that the top line AC now has (signed) slope τ and the bottom one BC now has slope σ, where

$$\begin{aligned} \sigma &= \tan(\pi/2 - \beta) \\ \tau &= \tan(\pi - \delta - \pi/2 - \alpha) \\ &= \tan(\pi/2 - \beta - \gamma)\,. \end{aligned}$$

We now have

$$\begin{aligned} B &= (x_B, y_B) = (1, 0) \\ A &= (x_A, y_A) = (\cos\delta, \sin\delta) \end{aligned}$$

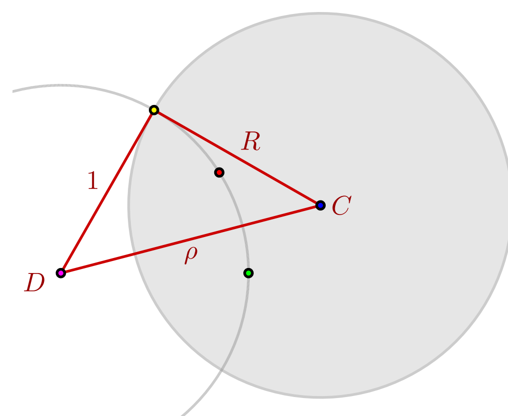
The point C is the intersection of the two lines. Hence to find (x, y) we have to solve for x, y in the pair of equations

$$\begin{aligned} y - y_A &= \tau(x - x_A) \\ y - y_B &= \sigma(x - x_B) \\ -\tau x + y &= -\tau x_A + y_A \\ -\sigma x + y &= -\sigma x_B + y_B \,. \end{aligned}$$

so we get

$$\begin{aligned} x &= \frac{\sigma + \sin\delta - \tau\cos\delta}{\sigma - \tau} \\ y &= \sigma(x - 1) \\ C &= x + iy \,. \end{aligned}$$

Finally, we want to know in these units the radius of the circle centred at C.



$$\begin{aligned} \rho &= |C - D| \\ R &= \sqrt{\rho^2 - 1}\,, \end{aligned}$$

and finally get

$$\begin{aligned} r_\alpha &= |A - C|/R \\ r_\beta &= |B - C|/R \,. \end{aligned}$$

7. Shifts

In this and the next section I'll explain how to implement the effects of certain non-Euclidean isometries. In this one, I'll deal with shifts.

Among the isometries of the upper half plane $\mathfrak{H}$ are those associated to matrices like

$$\begin{bmatrix} t & 0 \\ 0 & 1/t \end{bmatrix},$$

which takes z to $t^2 z$. All these fix both 0 and ∞. Conjugated by the Cayley transform, this becomes

$$\begin{bmatrix} \dfrac{t + 1/t}{2} & \dfrac{t - 1/t}{2} \\ \dfrac{t - 1/t}{2} & \dfrac{t + 1/t}{2} \end{bmatrix}$$

which is

$$\begin{bmatrix} \cosh(x) & \sinh(x) \\ \sinh(x) & \cosh(x) \end{bmatrix}$$

if $t = e^x$. All these fix the endpoints ± 1. A matrix in this form is equivalent to (induces the same Möbius transformation as)

$$\begin{bmatrix} 1 & \tau \\ \tau & 1 \end{bmatrix} \quad (\tau = \sinh(x)/\cosh(x))$$

with τ real. This is very convenient to work with. I call the associated transformation a **shift**. In fact, there is no reason to require τ to be real, any complex τ with $|\tau| < 1$ will determine the Möbius transformation corresponmding to the matrix

$$\begin{bmatrix} 1 & \tau \\ \overline{\tau} & 1 \end{bmatrix}.$$

that takes $\mathbb{D}$ and its boundary to themselves and $-\tau$ to 0.

Why does the shift for such a τ take $\mathbb{D}$ to itself? Because of:

7.1. Lemma. *If*

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

and g is a complex matrix such that

$${}^t\overline{g} H g = \mu H$$

with $\det(g) > 0$ then the Möbius transformation associated to g takes the unit circle as well as the unit disk to themselves.

Proof. Because for such a g

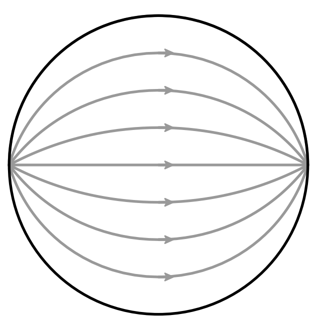
$$1 - |g(z)|^2 = \frac{\det(g)(1 - |z|^2)}{|cz + d|^2}.$$

The explicit formula for a shift is

$$z \longmapsto \frac{z + \tau}{\overline{\tau} z + 1} = \frac{(z + \tau)(\tau \overline{z} + 1)}{|\overline{\tau} z + 1|^2} = \frac{\tau |z|^2 + \tau^2 \overline{z} + z + \tau}{|\overline{\tau} z + 1|^2}.$$

The important thing about shifts is that they take some diameter, called its axis, to itself. The shift parametrized by τ takes 0 to τ, and the line through 0 and τ into itself.

The orbits of the shifts are arcs of circles passing through a pair of opposite points in the unit circle.



8. Reflections

The data determining a non-Euclidean reflection r is a geodesic. In practice it will be specified by a pair of points z_1, z_2 on it. We want a formula for the image of a point z with respect to this reflection.

There is one very simple case, in which the geodesic through z_1 and z_2 lie on a diameter. The non-Euclidean and Euclidean reflections then agree being just the reflection in the diameter. If that diameter is the line at angle θ to the x-axis, the equation of the diameter is $x \sin\theta - y\cos\theta$, and the reflection is:

$$z \longmapsto z - 2(z \bullet (\sin\theta, -\cos\theta)) \cdot (\sin\theta, -\cos\theta)\,.$$

The steps below reduce to that case.

(1) Apply a shift that takes z_1 to 0.
(2) Reflect in the line through 0 and the shifted z_2.
(3) Apply the inverse of the shift in step (1).

9. The Klein model

The group $\mathrm{SL}_2(\mathbb{R})$ acts on the three-dimensional space of symmetric real 2×2 matrices with trace 0 by conjugation:

$$x \longmapsto gxg^{-1}\,.$$

This action preserves the negative determinant, the quadratic form

$$-\det \begin{bmatrix} a & b \\ c & -a \end{bmatrix} = a^2 + bc$$

and of course each â^Ÿsphere’ is taken into itself. This signed determinant can also be written

$$a^2 + \left(\frac{b+c}{2}\right)^2 - \left(\frac{b-c}{2}\right)^2 = x^2 + y^2 - z^2$$

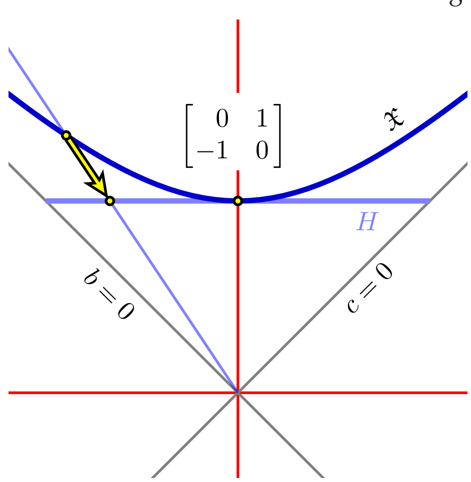
where

$$x = a, \quad y = \frac{b+c}{2}, \quad z = \frac{b-c}{2}\,.$$

The point

$$\kappa = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

is fixed by the subgroup $\mathrm{SO}(2)$ of rotations, and its SL_2-orbit $\mathfrak{X}$, which is one connected component of the sphere $\det = 1$, may therefore be identified with $\mathrm{SL}_2(\mathbb{R})/\mathrm{SO}(2)$. The metric induced on $\mathfrak{X}$ by the signed determinant is also $\mathrm{SL}_2(\mathbb{R})$-invariant, so in other words we have another model of the non-Euclidean plane. If $\mathfrak{H}$ is the slice $z = (b-c)/2 = 1$ through the cone $x^2 + y^2 < z^2$, then there is a canonical projection of $\mathfrak{X}$ onto $\mathfrak{H}$. This defines the Klein model of non-Euclidean geometry.



The advantage of this model over the others is that the action of $\mathrm{SL}_2(\mathbb{R})$ is essentially linear. In this model, the geodesics are the intersections of $\mathfrak{H}$ with planes through the origin.

Reflections in this model are by reflections that are orthogonal with respect to the quadratic form $x^2+y^2-z^2$. For the moment, let the dot product be with respect to it. Thus reflection in the linear plane

$$\alpha \bullet v = (a,b,c) \bullet (x,y,z) = ax+by-cz = 0$$

is

$$v \longmapsto v - 2\left(\frac{v \bullet \alpha}{\alpha \bullet \alpha}\right)\alpha\,.$$

The correspondence between the Klein model and the Poincaré model is not too complicated. One erects a hemisphere over the slice H, projects points vertically on $\mathfrak{H}$ to those on the hemisphere, and then maps this hemisphere in turn onto the unit disk by stereographic projection.

10. Coxeter tilings

A Coxeter matrix is an integral matrix $(m_{i,j})$ with $m_{i,i}=1$, $m_{i,j}=m_{j,i}\ge 2$. To every Coxeter matrix is associated a **Coxeter group** W defined by generators s_i in a set S and relations $(s_is_j)^{m_{i,j}}=1$. In particular $s_i^2=1$ for all i.

Every Coxeter group possesses a standard representation on a real vector space V of dimension $|S|$. If e_i is a basis of the dual $\widehat{V}$, we first define an inner product:

$$e_i \bullet e_j = -\cos(\pi/m_{i,j})\,.$$

so that in particular $e_i \bullet e_i = 1$. Then, in V itself we define vectors $\widehat{e}_i$ by the formula

$$\langle e_i, \widehat{e}_j\rangle = 2(e_i \bullet e_j)\,.$$

These will be linearly dependent if the inner product has non-trivial radical. This happens, for example, if the matrix $(e_i \bullet e_j)$ is

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

when $\widehat{e}_1 = -\widehat{e}_2$.

If the inner product is non-degenerate, we may use it to identify V with its dual.

We then associate to s_i the reflection

$$r_i\colon\ v \longmapsto v - \langle e_i, v\rangle \widehat{e}_i,\,.$$

Let C be the region

$$\langle e_i, v\rangle > 0 \text{ for all } i$$

in V . Its walls are the hyperplanes of reflection for the generators in S. Let $\mathcal{C}$ be the interior of the union of the transforms of C by products of the reflections r_i, which is of course stable under W.

• The region $\mathcal{C}$ is an open convex cone in V;
• the region $\mathcal{C}$ is a fundamental domain for W acting on C.

Exactly what $\mathcal{C}$ looks like depends on the group W. If we can write S as the union of two pieces S' and S'' where the e_i for distinct components are orthogonal, then we can factor W as $W' \times W''$. From now on, I'll assume that S cannot be expressed as such a union—that it is **irreducible**.

The cone $\mathcal{C}$ is all of V if and only if the group is finite and the inner product is positive definite. It will be a half-space if and only if the inner product is semi-definite, and in this case it is called an affine Coxeter group, because it may also be realized as a group of affine transformations in one dimension less. In all other cases, it will be properly contained in a half-space. It is this last case that we shall be eventually interested in.

The real point of looking at the geometry of the standard representation of a Coxeter group is that there is a strong relationship between this geometry and the combinatorics of the group. An expression for w in terms of the generators s_i is reduced if there is no shorter expression for it. The length $\ell(w)$ is the length of a reduced word for w.

We have $\ell(s_i w) > \ell(w)$ if and only if wC lies on the same side of $\langle e_i, v\rangle = 0$ as C,or equivalently if and only if $hei, wCi > 0$.

This leads to a simple way to do computations in Coxeter groups. It will not in practice be infallible, but it will work well most of the time. Suppose ρ to be the unique vector in V such that $\langle e_i, \rho\rangle = 1$ for all i. It lies in the interior of C, and since C is a fundamental domain for W an element w in W is determined uniquely by the vector $w\rho$. Thus $\langle e_i, wC\rangle > 0$ if and only if $\langle e_i, w\rho\rangle > 0$.

We shall use this observation in order to figure out how to make a list of all elements of W up to some fixed length. We represent each w by the array $(\langle e_i, w\rho\rangle)$, so that the identity, for example, is the array $(1, 1, \dots, 1)$. Reflection in these terms is simple:

$$\langle e_i, s_j w\rho\rangle = \langle e_i, w\rho\rangle - \langle e_j, w\rangle\langle e_i, \widehat{e}_j\rangle \,.$$

This means that we can certainly calculate all elements up to any given length. But how do we assure listing each w only once? We shall count a possibly new element $v = s_i w$ only if i is least such that $\langle e_i, v\rangle < 0$. In this way, every w in W is effectively assigned the unique expression in terms of the s_i which is lexicographically minimal.

The problem with this technique is that on a computer real numbers possess only limited accuracy, so that the calculation of the $w\rho$ is not exact. There are ways to get around this, very elegant and sophisticated ones, but for our purpose we shall not need them.

There is just one case we are going to deal with, where $|S| = 3$. The nature of the group depends on the sum

$$\frac{1}{m_{1,2}} + \frac{1}{m_{1,3}} + \frac{1}{m_{2,3}} \,.$$

If it is > 1, the group W is finite; if it is exactly equal to 1 the region C is half of $R3$ and W is the symmetry group of a planar tiling; and if is < 1 then W may be realized as a group of non-Euclidean tiling symmetries. These assertions can be verified by figuring out the signature of the matrix of the inner product

$$\begin{bmatrix} 1 & -\cos\pi/m_{1,2} & -\cos\pi/m_{1,3} \\ -\cos\pi/m_{1,2} & 1 & -\cos\pi/m_{2,3} \\ -\cos\pi/m_{1,3} & -\cos\pi/m_{2,3} & 1 \end{bmatrix}$$

We are interested only in the third case. For the moment, let $u \bullet v$ be the inner product with respect to the metric $x^2 + y^2 - z^2$. Under the assumption that $\sum 1/m_{i,j} < 1$, we deduce from the assertion above that we can find vectors e_i with

$$e_i \bullet e_j = -\cos\pi/m_{i,j}.$$

Normalizing suitably, we may assume that

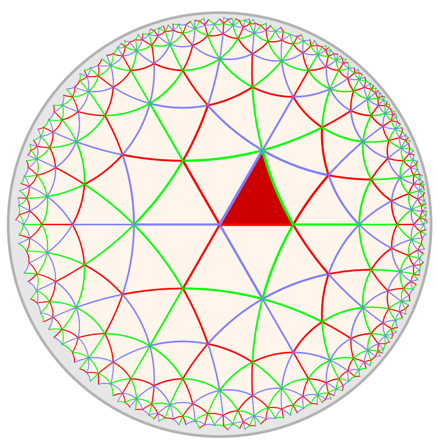
$$e_1 = (x_1, 0, 0)e_2 = (x_2, y_2, 0)e_3 = (x_3, y_3, z_3)$$

with all last coordinates positive. This gives us formulas

$$
\begin{aligned}
&e_1 \bullet e_1 = x_1^2 = 1 \\
&\quad \circ\ x_1 = 1 \\
&e_2 \bullet e_1 = x_2 x_1 = -\cos \pi/m_{1,2} \\
&\quad \circ\ x_2 = -\cos \pi/m_{1,2} \\
&e_2 \bullet e_2 = x_2^2 + y_2^2 \\
&\quad \circ\ y_2 = \sqrt{1 - x_2^2} \\
&e_3 \bullet e_1 = x_3 x_1 = -\cos \pi/m_{1,3} \\
&\quad \circ\ x_3 = -\cos \pi/m_{1,3} \\
&e_3 \bullet e_2 = x_3 x_2 + y_3 y_2 = -\cos \pi/m_{2,3} \\
&\quad \circ\ y_3 = \frac{-\cos \pi/m_{2,3} - x_2 x_3}{y_2} \\
&e_3 \bullet e_3 = x_3^2 + y_3^2 - z_3^2 = 1 \\
&\quad \circ\ z_3 = \sqrt{x_3^2 - y_3^2 - 1}
\end{aligned}
$$

Now we are ready to describe the process of constructing a non-Euclidean tiling of the plane by triangles with angle $\pi/mi, j$ where $\sum 1/m_{i,j} < 1$. We proceed by recursion. We start with the normalized triangle T found by Bédard's technique. The state at any moment (the collection of arguments to the recursive procedure) after that is characterized by (1) the triangle wT to draw, (2) the vector $w\rho$, (3) the length left to go. When the recursive procedure is called, the triangle is drawn, and if $n > 0$ then for each $s_i w > w$ the procedure is called for $s_i wT$, $s_i w\rho$, and $n - 1$.

The following picture took less than a second to draw.



Another, very interesting, approach to drawing Coxeter tilings can be found in [Gagern & Richter-Gebert:2009]. Ruler and compass constructions in non-Euclidean geometry can be found in [Goodman-Strauss:2001].

11. References

1. Bill Casselman, **Mathematical illustrations**, Cambridge University Press, 2004.

2. Martin Gagern and Jürgen Richter-Bebert, 'Hyperbolization of Euclidean ornaments', *The Electronic Journal of Combinatorics* **17** (2009), #R12.

3. Chaim Goodman-Strauss, 'Compass and straightedge in the Poincaré disk', *American Mathematical Monthly* **128** (2001), 38–49.

4. Peter Henrici, **Essentials of numerical analysis**, Wiley, 1982.