Circles and Möbius Geometry

8.1 Möbius geometry

8.1.1 Introduction We need to understand hyperbolic space, the third of the metric geometries in two dimensions (along with plane and sphere geometry). We will however take the scenic route to get there. Once we arrive we will understand some of Escher's pictures better and discover a use for all the more expensive (more than $\$2$) oribifolds.

8.1.2 Set-up We will regard $\mathbb{R}^2$ as the complex plane $\mathbb{C}$. We will add a "point at infinity" (denoted "∞") to form the "extended complex plane"

$$\widetilde{\mathbb{C}} = \mathbb{C} \cup \{\infty\}.$$

A *Möbius transformation* ("M-transformation" for short) is a map $\widetilde{\mathbb{C}} \longrightarrow \widetilde{\mathbb{C}}$ of the form

$$f : z \mapsto \frac{az+b}{cz+d}, \quad a, b, c, d \in \mathbb{C}, \quad ad - bc \neq 0.$$

M-transformations are bijections if one treats the point at infinity in the obvious naive way:

$$f(\infty) = \frac{a}{c}, \quad f\left(-\frac{d}{c}\right) = \infty.$$

8.1.3 Properties It is proved in courses on complex variables that complex analytic functions in general are angle preserving, so that, for instance, the images under a Möbius transformation of two curves that meet orthogonally will also meet orthogonally.

8.1.4 Problem (NM) Show that an M-transformation f such that $f(\infty) = \infty$ is of the form $f(z) = pz + q$ for complex numbers p, q with $p \neq 0$.

8.1.5 Homogeneous coordinates Given complex numbers z_1, z_2 not both zero, we say that they are *homogeneous coordinates* for $z \in \widetilde{\mathbb{C}}$ if

$$\frac{z_1}{z_2} = z, \text{ (where we take } z_2 = 0 \text{ to give } z = \infty).$$

We will often think of our homogeneous coordinates as being a column vector

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.$$

For each z there are thus infinitely many possible homogeneous coordinates, the vectors all non-zero complex multiples of each other.

We will often represent $z \in \widetilde{\mathbb{C}}$ by the choice of homogeneous coordinates

$$\begin{pmatrix} z \\ 1 \end{pmatrix} \text{ or by } \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ if } z = \infty.$$

8.1.6 Matrix representation Given the M-transformation as in §8.1.2, define the matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then the M-transformation can be represented by matrix multiplication on homogeneous coordinates:

$$z \sim \begin{pmatrix} z \\ 1 \end{pmatrix} \mapsto M \begin{pmatrix} z \\ 1 \end{pmatrix} = \begin{pmatrix} az+b \\ cz+d \end{pmatrix} \sim \frac{az+b}{cz+d}.$$

The tildes relate homogeneous coordinates to standard ones.

Composition of M-transformations corresponds to multiplication of the corresponding matrices. We see from this that M-transformations are invertible, with the inverse given by the inverse matrix.

8.1.7 Theorem

1. Let $p, q, r \in \mathbb{C}_\infty$ be distinct points. The M-transformation
$$f(z) = \frac{(p-z)(r-q)}{(r-z)(p-q)}$$
is such that $f(p) = 0,\ f(q) = 1, f(r) = \infty$.

2. An M-transformation that fixes each of p, q, r is the identity.

3. Let p', q', r' also be distinct points. Then there is a unique M-transformation g such that $g(p) = p', g(q) = q', g(r) = r'$.

Proof. The first point is an easy calculation. For the second part, show first that an M-transformation that fixes $0, 1, \infty$ is the identity.

For the last part, existence follows by composing a transformation sending p, q, r to $0, 1, \infty$ with one taking $0, 1, \infty$ to p', q', r'. For uniqueness, suppose one has two transformations, f, g taking p, q, r to p', q', r'. Then $f^{-1} \circ g$ fixes p, q, r and so is the identity. Therefore $f = g$. □

8.1.8 Definition In Möbius geometry, by a *circle* we mean either a normal round circle (hereafter called a "proper circle") in $\mathbb{C}_\infty$ or a straight line in $\mathbb{C}_\infty$. (So a straight line is just a circle that goes through infinity.)

8.1.9 Theorem Through any three distinct points in $\mathbb{C}_\infty$ there is a unique circle.

Proof. If one point is ∞ then it is the line through the other two. Otherwise, if the three points are collinear, then that line is the circle. If not, then the points lie on a "proper" circle by elementary geometry. $\square$

8.1.10 Theorem M-transformations take circles to circles. The group of M-transformations acts transitively on the space of all circles in $\mathbb{C}_\infty$.

Proof. That M-transformations take circles to circles is proved in Honours Complex Variables: it is a straightforward exercise in algebra. The transitivity of the action follows from the fact that there is an M-transformation taking any set of three distinct points to any other three. $\square$

8.1.11 Warning An M-transformation that maps a circle to another circle does NOT usually map the centre to the centre. So, in Möbius geometry a circle does not have a well-defined centre.

8.2 Apollonius's problem and Soddy circles

The "Apollonius tangency problem" is to find all lines or circles tangent to all of three given lines or circles. As stated, this is clearly a problem in Möbius geometry: the fact that it treats lines and circles equally is one clue. But more than that, if f is an M-transformation then f will take a configuration of three circles to another configuration of three circles, and take solutions of the problem to solutions of the problem.

The idea of meeting tangentially is certainly valid in Möbius geometry too: it simply means meeting at angle zero, and M-transformations preserve angles of intersection.

Two straight lines are tangent where they meet at infinity if and only if they are parallel.

8.2.1 Definition Let S, S', S'' be three circles and suppose each one is tangent to the other two and the three resulting points of tangency are all distinct. A *Soddy circle* is another circle T tangent to all three of S, S', S''.

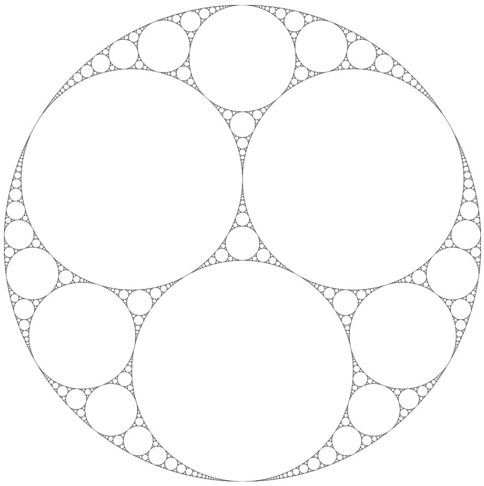
The existence and number of Soddy circles is a particular case of the Apollonian problem and it is not clear at first glance how many solutions it may have.

8.2.2 Theorem Every configuration of three mutually tangent circles (not all tangent at the same point) has precisely two Soddy circles.

8.2.3 Problem (NM) Prove the theorem above by recalling that we can use an M-transformation to take the three points of tangency to any other three distinct points. Choose one that takes them to $0, 1, \infty$ and see what the picture must look like. (Parallel lines in $\mathbb{C}_\infty$ are tangent at infinity.)

See `http://jwilson.coe.uga.edu/EMAT6680Su09/Floer/6690/Soddy%20Circles/Soddy%20Circles.html` for more on this: Soddy wrote a paper on the subject which was written as a poem called "The Kiss precise".

8.2.4 Apollonian gaskets Take three circles S, S', S'' as described above and add the two Soddy circles which we will call T, T'. Now consider triples of circles consisting of two original ones and one of the Soddy circles: for example, choose S, S', T. This triple has two Soddy circles, one of which is S'' and the other of which is new. Generate six "second generation" circles like this and add them to the picture. Now consider all triples of tangent circles involving one of the six added circles and add a third generation to the picture. Continue!



8.2.5 Problem (NM) What does the Apollonian gasket look like if two of the initial circles are parallel lines in the plane?

8.3 Inversive Geometry (in two dimensions)

8.3.1 Motivation In plane and spherical geometry we have indirect (orientation reversing) transformations such as reflections. What is the corresponding thing for Möbius geometry? In $\mathbb{C}_\infty$, reflection in the real axis is given by $z \mapsto \bar{z}$. If we add that to the Möbius group then we need to add all compositions of that with our existing Möbius transformations.

8.3.2 Definition The *inversive group* is the set of all *inversive transformations* of $\mathbb{C}_\infty$, which are those of the form

$$z \mapsto \frac{az+b}{cz+d} \quad \text{or} \quad z \mapsto \frac{a\bar{z}+b}{c\bar{z}+d} \quad \text{where } ad-bc \neq 0.$$

We will refer to those of the second form as *indirect* M-transformations. The set of all direct and indirect M-transformations forms a group, as one can easily check.

8.3.3 Note Since both M-transformations and complex conjugation in $\mathbb{C}_\infty$ preserve angles and sends circles to circles, then so do indirect M-transformations. Note that complex conjugation is orientation reversing, like reflections and glides in the plane whereas (direct) M-transformations preserve orientation as do all complex analytic functions.

8.3.4 Definition Consider a proper circle as in Fig 8.1 centred at the origin. Let X be a point in the circle and consider the construction of the point X' as shown in the figure. (The lines PX' and QX' are tangent to the circle.) We say that X and X' are *inverse points* with respect to the circle.

8.3.5 Problem (NL)

1. Let the circle in the figure have radius R and suppose $|OX| = x$ and $|OX'| = x'$. Show that $xx' = R^2$.

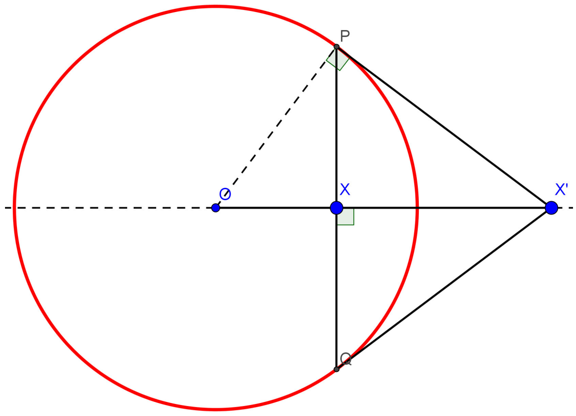


Figure 8.1: Inversion in a circle

2. Deduce that inversion in a circle of radius of radius R centred at the origin is given by

$$z \mapsto \frac{R^2}{\bar{z}} = \frac{R^2 z}{|z|^2}.$$

3. Deduce that inversion in a circle centred at $a \in \mathbb{C}$ of radius R is given by

$$z \mapsto c + \frac{R^2}{\bar{z} - \bar{c}}.$$

   (Hint: conjugate the previous case with translation by c.)

4. Write the general inversion in the previous part as an indirect M-transformation. Hence deduce that inversion sends circles to circles.

5. Show that in the limit as a circle become very large, inversion in it can be regarded as reflection in a line. (You may need to think how to take the limit so that this works.) We therefore say that inversion in a line is simply reflection in it.

6. Show that in the figure any circle passing through both X and X' intersects the original circle orthogonally. (A possible start is to assume the new circle is centred at a point $((x + x')/2, t)$. Note that two circles intersect orthogonally if the sum of the squares of their radii is equal to the square of the distance between their centres.)

The exercise essentially proves the following.

8.3.6 Theorem Let S define a circle and let p and q be distinct points not on S. Then p and q are inverse with respect to S if and only if every circle through p and q meets S orthogonally.

8.3.7 The role of inversion Inversion in inversive geometry plays the same role as reflections in Euclidean and plane geometry.

8.4 Pencils of circles

8.4.1 Problem (NM) Take the equations of two (proper) circles in the familiar real form:

$$x^2 + y^2 - 2fx - 2gy + c = 0, \qquad x^2 + y^2 - 2f'x - 2g'y + c' = 0$$

and subtract one from the other. The result is the equation of a line, known as the "radical axis" of the two circles. If the two circles intersect, which line is it?

8.4.2 Circle and line equations Consider first the equation

$$k(x^2 + y^2) - 2px - 2qy + d = 0$$

where not all the constants $k, p, q, d \in \mathbb{R}$ are zero. The equation can be multiplied through by any non-zero constant without changing the solution set.

Setting $k = 0$ we have the equation of a line (unless $p = q = 0$). Otherwise, we can rescale so that $k = 1$ and we have a circle-like equation of the form

$$S = x^2 + y^2 - 2fx - 2gy + c = 0$$

or equivalently

$$(x - f)^2 + (y - g)^2 = f^2 + g^2 - c.$$

The solution set is an actual circle if $c < f^2 + g^2$ and is a single point if $c = f^2 + g^2$. We refer to the latter case as a *point circle*.

8.4.3 Remark (for interest) We note in passing then that in the above, the set of lines in the plane is identified with all but one point of the set of triples (f, g, c) (not all zero) up to scaling by a non-zero constants. So we see that the set of lines in the plane is topologically a projective plane with a point removed. We know from previously that if we had removed an open disc we would have a Möbius band including boundary and so removing a single point (or a closed disc) leaves a Möbius band with boundary not included. Thus the space of lines in the plane is topologically an (open) Möbius band. You might want to try finding a closed curve in this space (i.e. a 1-parameter family of lines returning to the starting line) that corresponds to a loop going once around the band.

8.4.4 Henceforth in this discussion we will set $k = 1$ so that we are talking about proper, round circles. Let $S = 0, S' = 0$ be equations of circles. Then we could consider circles defined by arbitrary linear combinations $t_1 S + t_2 S' = 0$.

Since non-zero scalar multiples of a matrix determine the same circle, it is enough to consider linear combinations of the form $(1 - t)S + tS' = 0$ for a parameter t in $\mathbb{R}_\infty = \mathbb{R} \cup \{\infty\}$. We define the equation corresponding to $t = \infty$ to be $-S + S' = 0$. Note that our original circles correspond to parameter values $t = 0$ and $t = 1$.

For some values of t the equation may define a point circle or not define a circle at all.

8.4.5 Definition We call a 1-parameter family of circles as above the *pencil of circles* generated by $S = 0$ and $S' = 0$.

8.4.6 Theorem Consider a pencil of circles $(1-t)S + tS'$ in the plane. Every point in the plane is on at least one circle in the pencil. (For the purposes of this statement, we regard a "point circle" in a pencil as being a circle.) Suppose two circles in the pencil pass through a given point. Then every circle in the pencil passes through that point. There are at most two such points.

Proof. Fixing x and y in a pencil one obtains an equation of the form $\alpha t + \beta = 0$ with α, β being constants. Such equations have zero, one or infinitely many solutions for t. No solutions occurs only when $\alpha = 0, \beta \neq 0$ but one can check in that case that $t = \infty$ gives a solution. There cannot be more that two points that lie on all the circles in the pencil because three points determine a unique circle. □

8.4.7 Example Consider all circles passing through the two points $(\pm 1, 0)$. The centre of such a circle must lie at some point $(0, s)$ on the y-axis and it must have squared radius $1+s^2$ and so the equation is

$$x^2 + (y-s)^2 = 1 + s^2.$$

This may not look like a pencil of circles because of the s^2 terms, but if you multiply out these vanish and you get

$$x^2 + y^2 - 2sy - 1 = 0$$

(which we recognise as the pencil generated by $x^2 + y^2 - 1 = 0$ and the line $y = 0$, albeit with a different parameterisation from our default).

A pencil consisting of all circles passing through two fixed points is said to be *elliptic*.

8.4.8 Example Let S and S' be the point circles at $(\pm 1, 0)$:

$$S : (x-1)^2 + y^2 = 0; \qquad S' : (x+1)^2 + y^2 = 0.$$

Then the pencil $(1-t)S + tS' = 0$ is easily checked to be

$$(x - (1-2t))^2 + y^2 = 4t(t-1).$$

The solution set is empty for $0 < t < 1$ but is otherwise a circle centred at $(1-2t), 0)$ of squared radius $4t(t-1)$. (For $t = \infty$ it reduces to the line $x = 0$.) See Figure 8.2 for a picture.

Such a pencil, containing two point circles is said to be *hyperbolic*.

8.4.9 Exercise (Exercise) SN Show that every circle in the pencil in Example 8.4.7 meets every circle in Example 8.4.8 orthogonally. (See Figure **??** for a picture.)

8.4.10 Observations Since we can characterise pencils of circles geometrically (e.g. as all circles through two fixed points), we see that the idea of "pencil of circles" are all invariant under M-transformations. Thus we understand pencils if we understand the cases where the principal features are in standard positions.

8.4.11 Parabolic pencils If one takes an arbitrary pencil and computes the condition for the general circle in the pencil to be a point circle, one arrives at a quadratic in t. Thus a pencil can either have 0,1 or 2 point circles. We have dealt with the case of 0 or 2 point circles.

Pencils with one point circle are called *parabolic*. They consist of a single point circle, a line containing the point circle and every other circle tangent to the line at the point circle. Draw yourself a picture.

As we saw, each elliptic pencil is associated with a hyperbolic one all of whose circles meet those in the original pencil orthogonally. In the parabolic case this “orthogonal” pencil is also parabolic, containing the same point circle but with the line rotated by ninety degrees.

8.4.12 General case So we see that pencils of circles are either all circles through two fixed points (the elliptic case), hyperbolic pencils orthogonal two an elliptic one and containing point circles at the two defining points or they are parabolic.

It is instructive to work out what each case looks like if one of the defining points is at infinity.

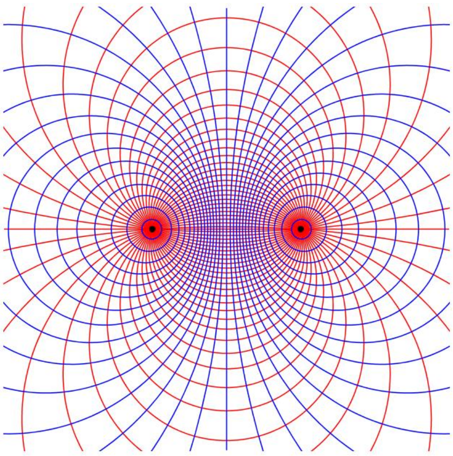


Figure 8.2: Orthogonality of elliptic and hyperbolic pencils defined by the same pair of points.

Table 8.1: Summary of pencils of circles

	Elliptic
General form	All circles through two given points
Special example	All straight lines through a point. (Other defining point is ∞.)
Number of point circles	0
Parameter values	Defined for all
Orthogonal to	Hyperbolic pencil defined by same two points
	Hyperbolic
General form	generated by two point circles.
Special example	All circles with a given centre. (Other defining point is ∞.)
Number of point circles	2
Parameter values	Not defined for all
Orthogonal to	Elliptic pencil defined by same two points
	Parabolic
General form	All circles tangent to a given one at a given point
Special example	All lines with a given direction. (The defining point is ∞.)
Number of point circles	1
Parameter values	Defined for all
Orthogonal to	Parabolic pencil defined by the same point but with perpendicular common tangents.