5 What is geometry?

5.1 The Erlanger Programm

Felix Klein moved on from his work on the Klein quadric and moved on physically from Bonn where he had been a student. Dodging the Franco-Prussian war in 1870 he passed through Paris, went briefly to Göttingen and then to Erlangen, near Nuremberg, in the south of Germany.

He prepared for his inaugural address in 1872 in Erlangen a paper which gave a very general view on what geometry should be regarded as. It was somewhat controversial at the time and in fact he spoke on something different for his lecture, but the point of view is still called the *Erlanger Programm*. Klein saw geometry as:

the study of invariants under a group of transformations.

This throws the emphasis on the group rather than the space, and was highly influential in a number of ways. Recall

Definition 18 *An action of a group G on a set Ω is a homomorphism $f : G \to \mathrm{Sym}(\Omega)$ to the group of all bijections of Ω to itself.*

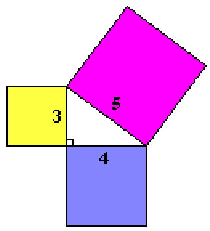
So for Klein, we have a set, say $\mathbf{R}^2$, and a group acting on it. For Euclidean geometry the group is the group of all transformations of the form

$$x \mapsto Ax + b$$

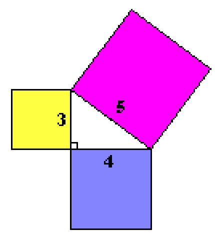
where A is an orthogonal transformation and $b \in \mathbf{R}^2$. This does form a group since

$$A(A'x + b') + b = AA'x + Ab' + b.$$

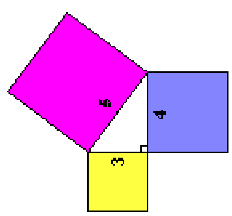
Every element in the group is the composition of a rotation or reflection $x \mapsto Ax$ and a translation $x \mapsto x + b$. So Pythagoras' theorem is a theorem of Euclidean geometry because if this is true:



so is this:



and this:



There are other groups that act on $\mathbf{R}^2$, though, and each of these gives rise to a different geometry. For example if we look at transformations

$$x \mapsto Ax + b$$

where A is just invertible and not orthogonal (this is called an *affine transformation*) then even the statement of Pythagoras's theorem is not invariant – a right angled triangle can be taken to any other triangle by an affine transformation. Even more interesting is the case of the hyperbolic plane, which we shall look at later. In fact it was the study of this, and Klein's realization that both Euclidean and hyperbolic geometry are special cases of projective geometry, which led him to formulate his proposal. We have to realize that in the early 19th century, as for most preceding centuries, Euclidean geometry had a special status, bound up with the logical structure of deductions from a set of axioms. Seeing it as just one of many geometries was as radical as Copernicus saying that the earth goes round the sun and not vice-versa.

The group that plays Klein's role for projective geometry is the group of projective transformations $\tau : P^n(F) \to P^n(F)$. Such a τ is defined by an invertible linear transformation $T : F^{n+1} \to F^{n+1}$, where we identify T and λT. Some names now:

Definition 19 *The group of all invertible linear transformations* $T : F^n \to F^n$ *is called the general linear group* $GL(n, F)$.

Definition 20 *The group of all invertible linear transformations* $T : F^n \to F^n$ *whose determinant is* 1 *is called the special linear group* $SL(n, F)$.

Definition 21 *The projective linear group* $PGL(n, F)$ *is the quotient of* $GL(n, F)$ *by the normal subgroup of non-zero multiples of the identity, and is the group of all projective transformations of* $P^{n-1}(F)$ *to itself.*

All the theorems we have proved so far are invariant under the projective linear group, but they have all been about lines, planes intersections etc. – just set-theoretical properties about linear subspaces. When Klein spoke about invariants he also meant numerical quantities, such as in Euclidean geometry the distance between two points $x, y \in \mathbf{R}^2$:

$$|(Ax + b) - (Ay + b)| = |A(x - y)| = |x - y|.$$

In projective geometry there is no invariant distance. Indeed Theorem 3 tells us that there is a projective transformation that takes any two distinct points to any other two. There are nevertheless more complicated invariants as we shall see next.

5.2 The cross-ratio

We shall deal with the geometry of the projective line $P^1(F)$ and the action of the group $PGL(2, F)$. We know from Theorem 3 that any three points go into any other three, but this is not the case for four points. There is a numerical invariant called the cross-ratio:

Definition 22 *Let* P_1, P_2, P_3, P_4 *be four distinct points on the projective line* $P^1(F)$, *with homogeneous coordinates* $P_i = [x_i, y_i]$. *The cross-ratio is defined by*

$$(P_1P_2; P_3P_4) = \frac{(x_1y_3 - x_3y_1)(x_2y_4 - x_4y_2)}{(x_1y_4 - x_4y_1)(x_2y_3 - x_3y_2)}.$$

First note that multiplying (x_i, y_i) by $\lambda_i \neq 0$ introduces the same factor in the numerator and denominator and leaves the formula for the cross-ratio unchanged. This shows it is well-defined. Similarly note that the linear transformation

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

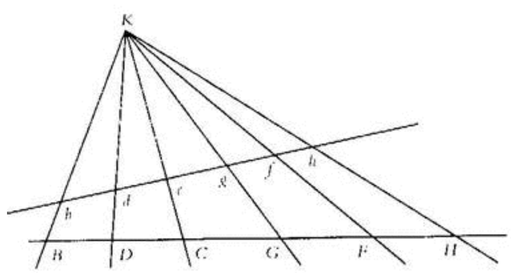
transforms $x_1y_2 - x_2y_1$ to

$$(ax_1 + by_1)(cx_2 + dy_2) - (ax_2 + by_2)(cx_1 + dy_1) = (ad - bc)(x_1y_2 - x_2y_1).$$

Again the factors in numerator and denominator cancel. We deduce that

- the definition of cross-ratio is independent of the choice of basis in which to write the homogeneous coordinates
- if $\tau : P(U) \to P(V)$ is a projective transformation between two projective lines then
$$(\tau(P_1)\tau(P_2); \tau(P_3)\tau(P_4)) = (P_1P_2; P_3P_4).$$

A particular case is projection from a point in the plane:



In the diagram, $(b\,d; c\,g) = (B\,D; C\,G)$.

If none of the y_i in the definition of cross-ratio is zero, then we can write it in terms of the scalars $z_i = x_i/y_i$ and obtain

$$(z_1\,z_2; z_3\,z_4) = \frac{(z_1 - z_3)(z_2 - z_4)}{(z_1 - z_4)(z_2 - z_3)} \tag{9}$$

Even when one of y_i is zero, an intelligent use of ∞ gives the correct results, for example if $y_1 = 0$, then the definition (22) provides

$$(\infty\,z_2; z_3\,z_4) = \frac{y_3(x_2y_4 - x_4y_2)}{y_4(x_2y_3 - x_3y_2)} = \frac{z_2 - z_4}{z_2 - z_3}$$

whereas the formula (9) gives

$$(\infty\, z_2; z_3\, z_4) = \frac{(\infty - z_3)(z_2 - z_4)}{(\infty - z_4)(z_2 - z_3)}.$$

On feature of the cross-ratio is the fact that you need to get the order right! Here (check for yourself the formulae) is what happens when you permute the variables:

- $(P_1P_2; P_3P_4) = (P_2P_1; P_4P_3) = (P_3P_4; P_1P_2)$
- $(P_2P_1; P_3P_4) = (P_1P_2; P_3P_4)^{-1}$
- $(P_1P_3; P_2P_4) = 1 - (P_1P_2; P_3P_4)$

Now, in the terminology of (9),

$$(\infty\, 0; 1\, x) = x$$

and so, since we assumed the four points were distinct, we see that $x \neq 0, 1$. But we have seen that there is a unique projective transformation τ that takes P_1, P_2, P_3 to $\infty, 0, 1$ so

$$(P_1P_2; P_3P_4) = (\tau(P_1)\tau(P_2); \tau(P_3)\tau(P_4)) = (\infty\, 0; 1\, \tau(P_4))$$

and thus any cross-ratio avoids the values $0, 1$. If we go through all the permutations of the four variables then in general we find the following six possible values of the cross-ratio:

$$x, \quad \frac{1}{x}, \quad 1 - x, \quad \frac{1}{1-x}, \quad 1 - \frac{1}{x}, \quad \frac{x}{x-1}.$$

There is a special case when these coincide in pairs: if $x = -1, 2$ or $1/2$ and this particular configuration of points has a name:

Definition 23 *The points P_1, P_2, P_3, P_4 are harmonically separated if the cross-ratio $(P_1P_2; P_3P_4) = -1$.*

Remark: There is some value in seeing the cross-ratio as a point in the projective line $P^1(F)\backslash\{0, 1, \infty\}$. We have seen that the symmetric group S_3 of order 6 permuting $\{0, 1, \infty\}$ acts as projective transformations of $P^1(F)$ to itself. The six values of the cross-ratio then constitute an orbit of this action.

5.3 Affine geometry

One of the features which motivated Klein was the fact that different types of geometry were special cases of projective geometry, and simply involved restricting the group. Affine geometry is a case in point. Consider a hyperplane $P(U)$ in a projective space $P(V)$ and the subgroup of the group of projective transformations of $P(V)$ to itself which takes $P(U)$ to itself. The complement $P(V)\backslash P(U)$ is acted on by this group and is called an *affine space*. Basically it is just a vector space without a distinguished origin, but it is the *group* which determines the way we look at it.

To get a hold on this, consider choosing a basis such that in homogeneous coordinates $[x_0, \ldots, x_n]$ the hyperplane in $P^n(F)$ is defined by $x_0 = 0$. Then a projective transformation τ which maps this to itself comes from a linear transformation T with matrix of the form

$$\begin{pmatrix} * & 0 & 0 & \ldots & 0 \\ * & * & * & \ldots & * \\ * & * & * & \ldots & * \\ * & * & * & \ldots & * \\ * & * & * & \ldots & * \end{pmatrix}$$

and we can unambiguously choose a representative for T by taking $T_{00} = 1$. The action of the subgroup of $GL(n+1, F)$ defined by these matrices on $P^n(F)\backslash P^{n-1}(F)$ which we identify with F^n as usual by

$$(x_1, x_2, \ldots, x_n) \mapsto [1, x_1, \ldots, x_n]$$

is of the form

$$x \mapsto Ax + b$$

where $A \in GL(n, F)$ and $b \in F^n$. The group of such transformations is called the *affine group* $A(n, F)$.

Example: The simplest affine situation is the real affine line, which is just $\mathbf{R}$ with the group of transformations

$$x \mapsto ax + b$$

with $a \neq 0$. The cross-ratio, which is a projective invariant of four points in $P^1(\mathbf{R})$ give an affine invariant of three points, since $A(1, \mathbf{R})$ is the subgroup of $PGL(2, \mathbf{R})$ which preserves the single point $\infty \in P^1(\mathbf{R})$. We have

$$(\infty\, x_1; x_2\, x_3) = \frac{x_1 - x_3}{x_1 - x_2}$$

so if $x_1 < x_2 < x_3$ the proportion in which x_2 divides the segment $[x_1, x_3]$ is the affine invariant $(x_1\,\infty; x_2\,x_3)$. The special value 1/2 (the concept of "harmonically separated" in projective geometry corresponds to the midpoint of the segment. So although there is no invariant notion of distance, we retain the notion of proportionality.

In higher dimensions, the centre of mass of m points $x_1, \ldots, x_m \in \mathbf{R}^n$

$$\bar{x} = \frac{1}{m}(x_1 + \ldots + x_m)$$

is a well-defined affine concept since

$$\frac{1}{m}(Ax_1 + b + \ldots + Ax_m + b) = A\bar{x} + b.$$

5.4 Gaussian optics

One application of this group-theoretical way of looking at geometry is in the simplest model of optical systems – those which have an axis of symmetry. If we track the effect on light rays which lie in a plane through that axis, then we have a 2-dimensional problem but with an extra symmetry – if the system was originally axially symmetric, and in $\mathbf{R}^2$ the x-axis is the axis of symmetry, then the effect on light rays must be symmetric with respect to the reflection in the x-axis

$$(x, y) \mapsto (x, -y).$$

To model optics this way we map $\mathbf{R}^2$ into $P^2(\mathbf{R})$ by

$$(x, y) \mapsto [1, x, y]$$

and the we consider first the subgroup of projective transformations which commutes with the reflection in the x-axis:

$$(x_0, x_1, x_2) \mapsto (x_0, x_1, -x_2)$$

These are given by linear transformations T whose matrix is of the form

$$\begin{pmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & * \end{pmatrix}.$$

We can choose $T_{22} = 1$, but now we need also the input of the physics. This tells us that the 2×2 matrix T_{ij}, $0 \leq i, j \leq 1$ must have determinant 1. This is part of what is called symplectic geometry which underlies many physical theories. So Gaussian optics can be put into geometrical form by considering an action of the group $SL(2,\mathbf{R})$.

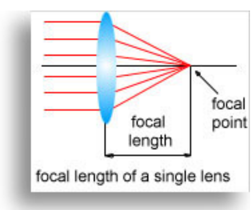
Note how the matrix in $SL(2,\mathbf{R})$

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

acts:

$$(x,y) \mapsto (\frac{ax+b}{cx+d}, \frac{1}{cx+d})$$

so the axis of the system is treated as a projective line and the group acts as projective transformations. This is necessary because the image of a point on the axis through the action of a lens may be "at infinity": that point is called the focal point. From the point of view of projective geometry all those parallel light rays from infinity are the lines passing through the intersection of the x-axis with the line at infinity, the point $[0,1,0]$, and the projective transformation takes them to the lines passing through the focal point.



So let us consider the effect of a thin symmetric lens at the origin on objects on the optical axis. It is given by an element $T \in SL(2,\mathbf{R})$ which acts on the x axis as the projective transformation τ. If the focal point is $(x,y) = (-f,0)$ then

$$\tau(\infty) = -f.$$

Because the lens is symmetric

$$\tau^{-1}(\infty) = f.$$

Light rays at the origin go straight through so

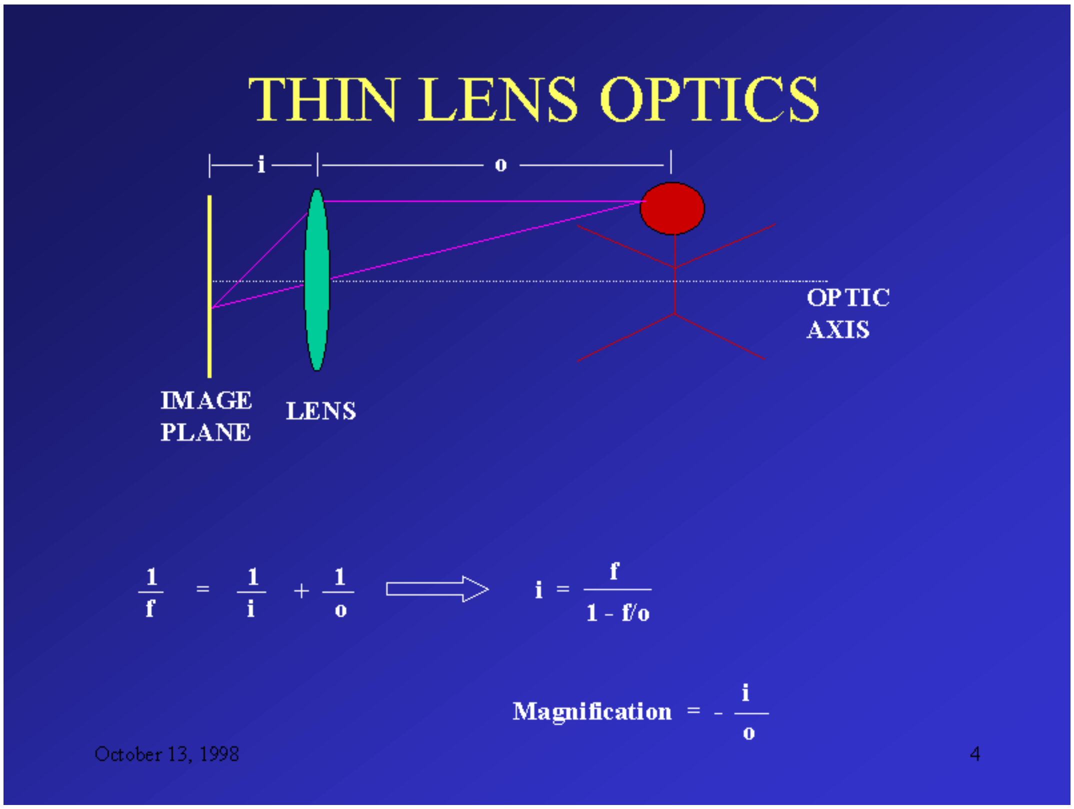
$$\tau(0) = 0.$$

The projective transformation of the line is is uniquely determined by this action on three points, and we easily find

$$\tau(x) = \frac{-fx}{x-f}.$$

An object at $x = \mathbf{o}$ then has image at $x = -\mathbf{i}$ where

$$\frac{1}{f} = \frac{1}{\mathbf{i}} + \frac{1}{\mathbf{o}}.$$



The group of projective transformations of the line $PGL(2, \mathbf{R})$ is obtained by the quotient of $GL(2, \mathbf{R})$ by the scalar matrices. In $SL(2, \mathbf{R})$ those scalars are just $\{\pm 1\}$, so we just need to choose a sign to get T:

$$T = \pm \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix}.$$

Find out which it is!

5.5 Non-Euclidean geometry

The group $SL(2, \mathbf{R})$ acts on other spaces than $P^1(\mathbf{R})$ or the optical space above, and if we study its invariants then we return to something close to the Euclidean geometry of the plane but also something essentially different, which is nowadays called *hyperbolic geometry*, or when it was first discovered *Non-Euclidean geometry.* A real matrix is a special type of complex matrix, so the group $PGL(2, \mathbf{R})$ is a subgroup of $PGL(2, \mathbf{C})$ and so acts on the complex projective line $P^1(\mathbf{C})$. It preserves the copy of $P^1(\mathbf{R}) \subset P^1(\mathbf{C})$ given by points with real homogeneous coordinates $[x_0, x_1]$, but we want to consider instead the complement of this subset. Since $\infty = [1, 0]$ is real, then

$$P^1(\mathbf{C}) \backslash P^1(\mathbf{R}) \subset P^1(\mathbf{C}) \backslash \{\infty\}$$

and is $\mathbf{C} \backslash \mathbf{R}$. This has two components, the upper and lower half-planes. An element of $PGL(2, \mathbf{R})$ may interchange these two (e.g. $z \mapsto -z$) but the subgroup $PSL(2, \mathbf{R}) = SL(2, \mathbf{R})/\{\pm 1\}$ preserves each one. We choose the upper one H. In fact the group acts transitively on H for the Möbius transformation

$$z \mapsto az + b$$

takes i to $ai + b$, and if $a > 0$, the transformation is defined by the 2×2 matrix

$$\frac{1}{\sqrt{a}} \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$$

which has determinant 1.

We can define an invariant function of a pair of distinct points in H using the cross-ratio:

$$(z\, w; \bar{w}\, \bar{z}).$$

Since z and w are in the upper half-plane $\bar{z}$ and $\bar{w}$ are in the lower half-plane so if $z \neq w$ the four points $z, w, \bar{w}, \bar{z}$ are distinct and the cross-ratio is well defined. Spelling it out we get

$$(z\, w; \bar{w}\, \bar{z}) = \frac{(z - \bar{w})(w - \bar{z})}{(z - \bar{z})(w - \bar{w})} = \frac{|z - \bar{w}|^2}{4 \Im z \Im w}$$

which is positive, since $\Im z$ and $\Im w$ are positive. It is also symmetric in w and z. We shall use this to define a notion of the distance between z and w, but first note that we can allow $w = z$ in the cross-ratio $(z\, w; \bar{w}\, \bar{z})$ because

$$(z\, z; \bar{z}\, \bar{z}) = \frac{|z - \bar{z}|^2}{4 \Im z \Im z} = 1.$$

Definition 24 *Let z, w be two points in the upper half plane. The hyperbolic distance between z and w is*

$$d(z, w) = 2 \sinh^{-1} \sqrt{(z\, w; \bar{w}\, \bar{z})}.$$

This is clearly invariant under the action of $PSL(2, \mathbf{R})$. The particular analytical formula shows that if $z = ix, w = iy$ are imaginary points in H with $x > y$ then

$$e^{d/2} - e^{d/2} = 2\sqrt{\frac{(x-y)^2}{4xy}} = \left(\frac{x}{y}\right)^{1/2} - \left(\frac{y}{x}\right)^{1/2}$$

so that

$$d(ix, iy) = \log x - \log y \tag{10}$$

This notion of distance has the following property:

Proposition 23 *Let z_1, z_2 and w_1, w_2 be pairs of distinct points in the upper half plane and suppose that $d(z_1, z_2) = d(w_1, w_2)$. The there is a unique element $\tau \in PSL(2, \mathbf{R})$ such that $\tau(z_i) = w_i$ for $i = 1, 2$.*

Proof: Consider $z_1, z_2, \bar{z}_2$. These are distinct because $z_1 \neq z_2$ and $\bar{z}_2$ is in the lower half plane. By Theorem 3 there is a unique *complex* projective transformation τ such that

$$\tau(z_1) = w_1, \quad \tau(z_2) = w_2, \quad \tau(\bar{z}_2) = \bar{w}_2.$$

By projective invariance

$$(z_1\, z_2; \bar{z}_2\, \bar{z}_1) = (\tau(z_1)\, \tau(z_2); \tau(\bar{z}_2)\, \tau(\bar{z}_1)) = (w_1\, w_2; \bar{w}_2\, \tau(\bar{z}_1))$$

However, $d(z_1, z_2) = d(w_1, w_2)$ so

$$(w_1\, w_2; \bar{w}_2\, \tau(\bar{z}_1)) = (z_1\, z_2; \bar{z}_2\, \bar{z}_1) = (w_1\, w_2; \bar{w}_2\, \bar{w}_1)$$

and thus $\tau(\bar{z}_1) = \bar{w}_1$. We then have

$$\tau(\bar{z}_1) = \overline{\tau(z_1)}, \quad \tau(\bar{z}_2) = \overline{\tau(z_2)}$$

i.e. two solutions in the lower half-plane to the equation

$$\frac{az+b}{cz+d} = \frac{\bar{a}z+\bar{b}}{\bar{c}z+\bar{d}} \tag{11}$$

This is quadratic with imaginary coefficients:

$$(a\bar{c} - \bar{a}c)z^2 + (b\bar{c} - \bar{b}c + a\bar{d} - \bar{a}d)z + (b\bar{d} - \bar{b}d) = 0$$

which either vanishes identically or has conjugate roots. But since the imaginary part of both roots is negative, it must vanish. Thus (11) is identically true. This means that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \lambda \begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix}$$

and $|\lambda| = 1$, so $\lambda = e^{i\theta}$. It follows that

$$e^{-i\theta} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is real. Its determinant must be positive since it maps z_1 in the upper half plane to w_1 in the upper half plane. Thus rescaling we get an element in $SL(2, \mathbf{R})$. □

We noted above that $PSL(2, \mathbf{R})$ acts transitively on the upper half plane H. The theory of group actions tells us that there is a bijection

$$G/K \leftrightarrow H$$

from the space of cosets G/K of the subgroup K which fixes a point $x_0 \in H$. The map is just $gK \mapsto g \cdot x_0$. In our case, if we take $x_0 = i$, then K is the group of transformations

$$z \mapsto \frac{\cos\theta z + \sin\theta}{-\sin\theta z + \cos\theta}$$

which is isomorphic by $\theta \mapsto e^{2i\theta}$ to the group of unit complex numbers.

Klein's point of view then says that the geometry of the upper half plane is basically the study of the invariants of this space of cosets. It offers the opportunity to view this geometry in a different way – to find another space on which $PSL(2, \mathbf{R})$ acts transitively with the same stabilizer K. We look at such a space next.

We consider the space of quadrics in the projective line $P^1(\mathbf{R})$. By this, we mean the real projective plane defined by the 3-dimensional real vector space of quadratic forms

$$B(v, v) = ax_0^2 + bx_0x_1 + cx_1^2$$

on $\mathbf{R}^2$. The singular ones are given by the equation

$$b^2 - 4ac = 0$$

which, when we write it as

$$b^2 - (a + c)^2 + (a - c)^2$$

is clearly a non-empty non-singular conic in $P^2(\mathbf{R})$. We consider the subset $D \subset P^2(\mathbf{R})$ defined by

$$D = \{[a, b, c] \in P^2(\mathbf{R}) : b^2 - 4ac < 0\}.$$

If $a + c = 0$, $b^2 - 4ac = b^2 + (a - c)^2 > 0$ so D lies in the complement of the line $a + c = 0$ in $P^2(\mathbf{R})$ and we can regard D as the interior of the circle $(b/(a + c))^2 + ((a - c)/(a + c))^2 = 1$ in $\mathbf{R}^2$. The group $PGL(2, \mathbf{R})$, and its subgroup $PSL(2, \mathbf{R})$, acts naturally on this space of quadrics.

Now when $b^2 - 4ac < 0$, $a \neq 0$ and the quadratic equation $ax^2 + bx + c = 0$ has complex conjugate roots $z, \bar{z}$, one of which, say z, must be in the upper half plane. The quadratic form factorizes as $a(x_0 - zx_1)(x_0 - \bar{z}x_1)$. Conversely any $z \in H$ defines a point in D by $(x_0 - zx_1)(x_0 - \bar{z}x_1)$. We therefore have a bijection

$$F : H \to D$$

commuting with the action of $PSL(2, \mathbf{R})$. This means that the geometry of D, the interior of a conic in $P^2(\mathbf{R})$, is identical to that of H, the upper half plane. Historically, H is called the *Poincaré model* and D the *(Klein)-Beltrami model*. Each one has its distinct advantages.

How do we define distance in the Beltrami model? Two points $X, Y \in D$ define two quadrics and we take the pencil of quadrics generated by them – the unique projective line in $P^2(\mathbf{R})$ which joins them. This meets the conic in two points A and B – the two singular conics in the pencil. So we have four points on a line and we have an invariant, the cross-ratio.

To see how this relates to the hyperbolic distance, we need only consider z, w as imaginary because of Proposition 23 and (10) which shows that any distance can be realized by imaginary points in H. So if $z = ix, w = iy$, the two points $X = F(ix), Y = F(iy)$ in D are the quadrics

$$x_0^2 + x^2x_1^2, \qquad x_0^2 + y^2x_1^2$$

and the pencil generated by them is

$$\lambda(x_0^2 + x^2x_1^2) + \mu(x_0^2 + y^2x_1^2) = (\lambda + \mu)x_0^2 + (\lambda x^2 + \mu y^2)x_1^2.$$

With $[\lambda, \mu]$ as homogeneous coordinates on this line in $P^2(\mathbf{R})$ we have $X = [1, 0], Y = [0, 1]$ and the two singular quadrics occur when $\lambda + \mu = 0$ and $\lambda x^2 + \mu y^2 = 0$. So $A = [1, -1], B = [y^2, -x^2]$. Thus the cross-ratio

$$(YX; AB) = x^2/y^2.$$

From (10) we see that in this model

$$d(X,Y) = \frac{1}{2} \log |(XY;AB)| \tag{12}$$

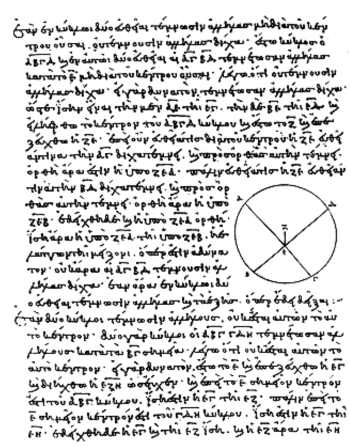
5.6 The parallel postulate

For centuries, Euclid's deduction of geometrical theorems from self-evident common notions and postulates was thought not only to represent a model of the physical space in which we live, but also some absolute logical structure. One postulate caused some problems though – was it really self-evident? Did it follow from the other axioms? This is how Euclid phrased it:

"That if a straight line falling on two straight lines makes the interior angle on the same side less than two right angles, the two straight lines if produced indefinitely, meet on that side on which the angles are less than two right angles".

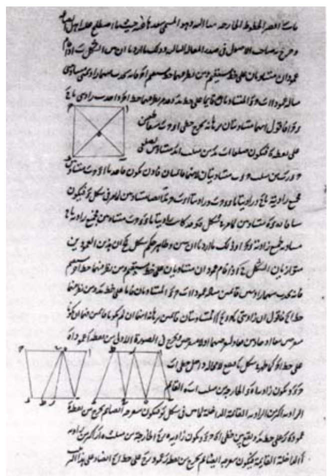
Some early commentators of Euclid's *Elements*, like Posidonius (1st Century BC), Geminus (1st Century BC), Ptolemy (2nd Century AD), Proclus (410 - 485) all felt that the parallel postulate was not sufficiently evident to accept without proof.

Here is a page from a medieval edition of Euclid dating from the year 888. It is handwritten in Greek. The manuscript, contained in the Bodleian Library, is one of the earliest surviving editions of Euclid.



The controversy went on and on; here is the Islamic mathematician Nasir al-Din al-Tusi (1201-1274)

struggling with the parallel postulate:



Finally Janos Bolyai (1802-1860) and Nikolai Lobachevsky (1793-1856)

discovered non-Euclidean geometry simultaneously. It satisfies all of Euclid's axioms except the parallel postulate, and we shall see that it is the geometry of H or D that we have started to look at.

Bolyai became interested in the theory of parallel lines under the influence of his father Farkas, who devoted considerable energy towards finding a proof of the parallel postulate without success. He even wrote to his son:

"I entreat you, leave the doctrine of parallel lines alone; you should fear it like a sensual passion; it will deprive you of health, leisure and peace – it will destroy all joy in your life."

Another relevant figure in the discovery was Carl Friedrich Gauss (1777-1855)

who was the first to consider the possibility of a geometry denying the parallel postulate. However, for fear of being ridiculed he kept his work unpublished. When he read Janos Bolyai's work he wrote to Janos's father:

"If I commenced by saying that I must not praise this work you would certainly be surprised for a moment. But I cannot say otherwise. To praise it, would be to praise myself. Indeed the whole contents of the work, the path taken by your son, the results to which he is led, coincide almost entirely with my meditations, which have occupied my mind partly for the last thirty or thirty-five years."

Yes, well..... he would say that wouldn't he?

Euclid's axioms were made rigorous by Hilbert. They begin with undefined concepts of

- "point"
- "line"
- "lie on" (a point lies on a line)
- "betweenness"
- "congruence of pairs of points"
- "congruence of pairs of angles".

Euclidean geometry is then determined by logical deduction from the following axioms:

EUCLID'S AXIOMS

I. AXIOMS OF INCIDENCE

1. Two points have one and only one straight line in common.

2. Every straight line contains a least two points.

3. There are at least three points not lying on the same straight line.

II. AXIOMS OF ORDER

1. Of any three points on a straight line, one and only one lies between the other two.

2. If A and B are two points there is at least one point C such that B lies between A and C.

3. Any straight line intersecting a side of a triangle either passes through the opposite vertex or intersects a second side.

III. AXIOMS OF CONGRUENCE

1. On a straightline a given segment can be laid off on either side of a given point (the segment thus constructed is congruent to the give segment).

2. If two segments are congruent to a third segment, then they are congruent to each other.

3. If AB and $A'B'$ are two congruent segments and if the points C and C' lying on AB and $A'B'$ respectively are such that one of the segments into which AB is divided by C is congruent to one of the segments into which $A'B'$ is divided by C', then the other segment of AB is also congruent to the other segment of $A'B'$.

4. A given angle can be laid off in one and only one way on either side of a given half-line; (the angle thus drawn is congruent to the given angle).

5. If two sides of a triangle are equal respectively to two sides of another triangle, and if the included angles are equal, the triangles are congruent.

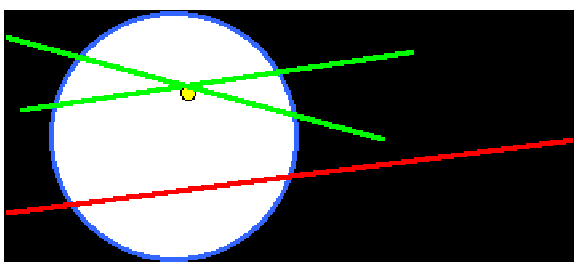
IV. AXIOM OF PARALLELS

Through any point not lying on a straight line there passes one and only one straight line that does not intersect the given line.

V. AXIOM OF CONTINUITY

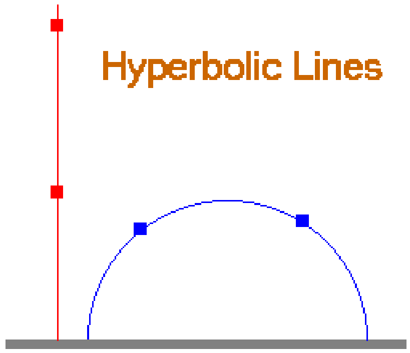
1. If AB and CD are any two segments, then there exists on the line AB a number of points $A_1, \ldots, A_n$ such that the segments $AA_1, A_1A_2, \ldots, A_{n-1}A_n$ are congruent to CD and such that B lies between A and A_n

The main property of hyperbolic geometry we have not introduced is the notion of a line. This is easiest in the Beltrami model – it is just the intersection of a projective line in $P^2(\mathbf{R})$ with the interior D of the conic. This obviously does *not* satisfy the parallel postulate:



In the Poincaré model we find:

Proposition 24 *In the upper half plane, a line is either a half-line parallel to the imaginary axis, or a semi-circle which intersects the real axis orthogonally.*



Proof: The map $F : H \to D$ is defined in homogeneous coordinates by

$$z \mapsto [1, -(z+\bar{z}), z\bar{z}]$$

so since a line in the Beltrami model is given by a projective line $\alpha_0 x_0 + \alpha_1 x_1 + \alpha_2 x_2 = 0$, in the upper half plane this is

$$\alpha_0 - \alpha_1(z+\bar{z}) + \alpha_2 z\bar{z} = 0.$$

If $\alpha_2 = 0$ this is a line of the form $ax + b = 0$, i.e. $x = constant$, which for $y > 0$ is a half-line parallel to the y-axis. Otherwise it is the intersection of the circle

$$\alpha_2(x^2 + y^2) - 2\alpha_1 x + \alpha_0 = 0$$

with D, and this is of the form

$$(x-a)^2 + y^2 = c$$

which is a circle centred at the point $(a, 0)$ on the x-axis, as required. $\square$

Remark: Note that from (10) that ix, iy, iz lie on a hyperbolic line and if $x > y > z$ then

$$d(x,z) = \log x - \log z = \log x - \log y + \log y - \log z = d(x,y) + d(y,z) \qquad (13)$$

But there is a unique line through any two points, from the Beltrami model. It follows from Proposition 23 that there is an element of $PSL(2, \mathbf{R})$ which takes any line to any other and so (13) holds for any line.

The fact that hyperbolic geometry satisfies all the axioms except the parallel postulate is now only of historic significance and the reader is invited to do all the checking. Often one model is easier than another. Congruence should be defined through the action of the group $PGL(2, \mathbf{R})$. We mentioned that this has a natural action on D and not just its index 2 subgroup $PSL(2, \mathbf{R})$. On H we have to add in transformations of the form

$$z \mapsto \frac{a\bar{z} + b}{c\bar{z} + d}$$

with $ad - bc = -1$ to get the action of $PGL(2, \mathbf{R})$.

Finally we should point out the link with the *Geometry of Surfaces* course. The upper half plane has a Riemannian metric, a first fundamental form, given by

$$\frac{dx^2 + dy^2}{y^2} \tag{14}$$

This is just a rescaling of the flat space fundamental form $dx^2 + dy^2$ by the positive function y^{-2}. This affects the lengths of curves but not the angles between them. In the coordinates x, y of H, angles between curves as measured by the metric (14) are just Euclidean angles. We say then that the metric is *conformally flat*. In the Beltrami model, we have affine coordinates in which the lines of the hyperbolic geometry become straight lines. In fact, the *geodesics* of the metric (14) are either half-lines parallel to the imaginary axis, or semi-circles which intersect the real axis orthogonally. In this case we say that the metric is *projectively flat*. The hyperbolic distance between two points is the metric distance, and then it follows that the basic axioms of a metric space hold for $d(z, w)$.