William P. Thurston

The Geometry and Topology of Three-Manifolds

Electronic version 1.0 - October 1997

http://www.msri.org/gt3m/

This is an electronic edition of the 1980 notes distributed by Princeton University. The text was typed in TeX by Sheila Newbery, who also scanned the figures. Typos have been corrected (and probably others introduced), but otherwise no attempt has been made to update the contents.

Numbers on the right margin correspond to the original edition's page numbers.

Thurston's *Three-Dimensional Geometry and Topology*, Vol. 1 (Princeton University Press, 1997) is a considerable expansion of the first few chapters of these notes. Later chapters have not yet appeared in book form.

Please send corrections to Silvio Levy at levy@msri.org.

CHAPTER 2

Elliptic and hyperbolic geometry

There are three kinds of geometry which possess a notion of distance, and which look the same from any viewpoint with your head turned in any orientation: these are elliptic geometry (or spherical geometry), Euclidean or parabolic geometry, and hyperbolic or Lobachevskiian geometry. The underlying spaces of these three geometries are naturally Riemannian manifolds of constant sectional curvature $+1$, 0, and -1, respectively.

Elliptic n-space is the n-sphere, with antipodal points identified. Topologically it is projective n-space, with geometry inherited from the sphere. The geometry of elliptic space is nicer than that of the sphere because of the elimination of identical, antipodal figures which always pop up in spherical geometry. Thus, *any* two points in elliptic space determine a unique line, for instance.

In the sphere, an object moving away from you appears smaller and smaller, until it reaches a distance of $\pi/2$. Then, it starts looking larger and larger and optically, it is in focus behind you. Finally, when it reaches a distance of π, it appears so large that it would seem to surround you entirely.
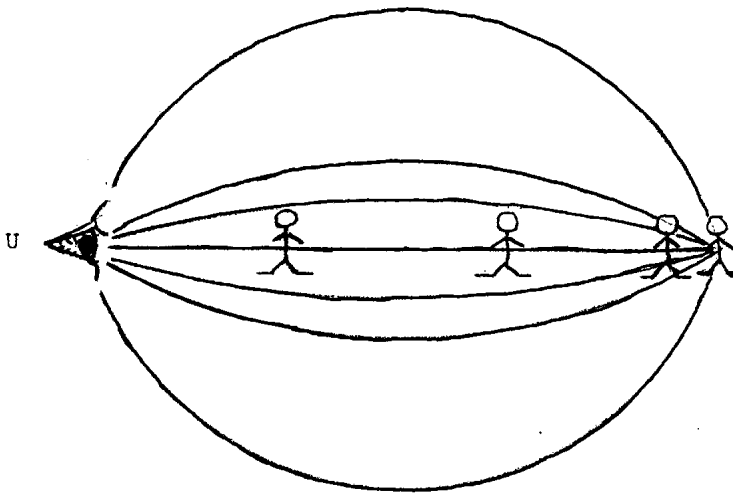

2.2

In elliptic space, on the other hand, the maximum distance is $\pi/2$, so that apparent size is a monotone decreasing function of distance. It would nonetheless be

distressing to live in elliptic space, since you would always be confronted with an image of yourself, turned inside out, upside down and filling out the entire background of your field of view.

Euclidean space is familiar to all of us, since it very closely approximates the geometry of the space in which we live, up to moderate distances. Hyperbolic space is the least familiar to most people. Certain surfaces of revolution in $\mathbb{R}^3$ have constant curvature -1 and so give an idea of the local picture of the hyperbolic plane.

2.3

The simplest of these is the pseudosphere, the surface of revolution generated by a tractrix. A tractrix is the track of a box of stones which starts at $(0,1)$ and is dragged by a team of oxen walking along the x-axis and pulling the box by a chain of unit length. Equivalently, this curve is determined up to translation by the property that its tangent lines meet the x-axis a unit distance from the point of tangency. The pseudosphere is not complete, however—it has an edge, beyond which it cannot be extended. Hilbert proved the remarkable theorem that *no* complete C^2 surface with curvature -1 can exist in $\mathbb{R}^3$. In spite of this, convincing physical models can be constructed.
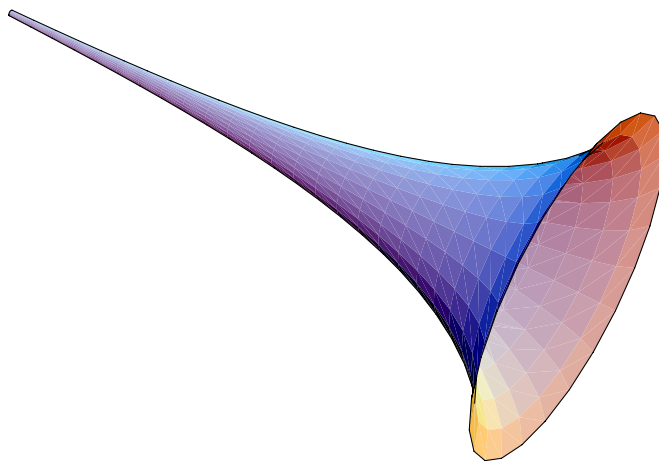
We must therefore resort to distorted pictures of hyperbolic space. Just as it is convenient to have different maps of the earth for understanding various aspects of its geometry: for seeing shapes, for comparing areas, for plotting geodesics in navigation; so it is useful to have several maps of hyperbolic space at our disposal.

2.1. The Poincaré disk model.

Let D^n denote the disk of unit radius in Euclidean n-space. The interior of D^n can be taken as a map of hyperbolic space H^n. A hyperbolic line in the model is any Euclidean circle which is orthogonal to ∂D^n; a hyperbolic two-plane is a Euclidean sphere orthogonal to ∂D^n; etc. The words "circle" and "sphere" are here used in

the extended sense, to include the limiting case of a line or plane. This model is conformally correct, that is, hyperbolic angles agree with Euclidean angles, but distances are greatly distorted. Hyperbolic arc length $\sqrt{ds^2}$ is given by the formula   2.4

$$ds^2 = \left(\frac{1}{1 - r^2}\right)^2 dx^2,$$

where $\sqrt{dx^2}$ is Euclidean arc length and r is distance from the origin. Thus, the Euclidean image of a hyperbolic object, as it moves away from the origin, shrinks in size roughly in proportion to the Euclidean distance from ∂D^n (when this distance is small). The object never actually arrives at ∂D^n, if it moves with a bounded hyperbolic velocity.
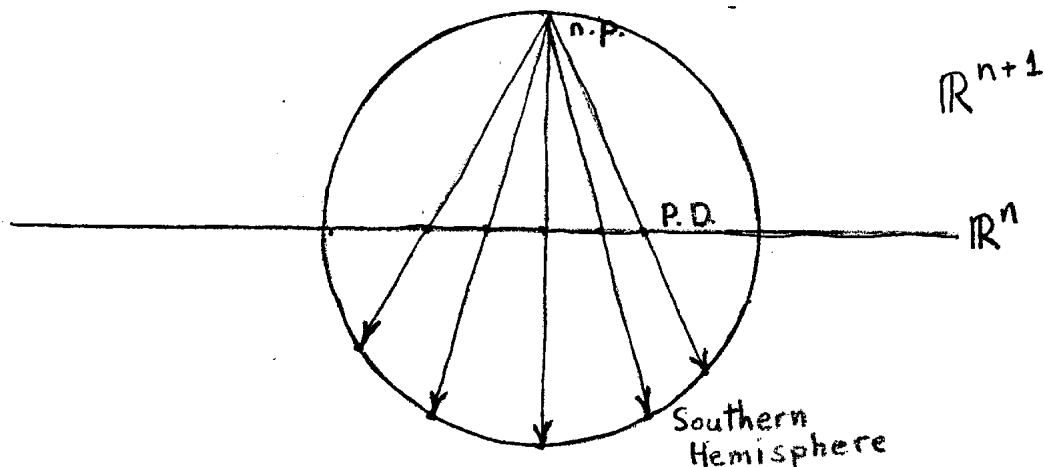

Lines

People
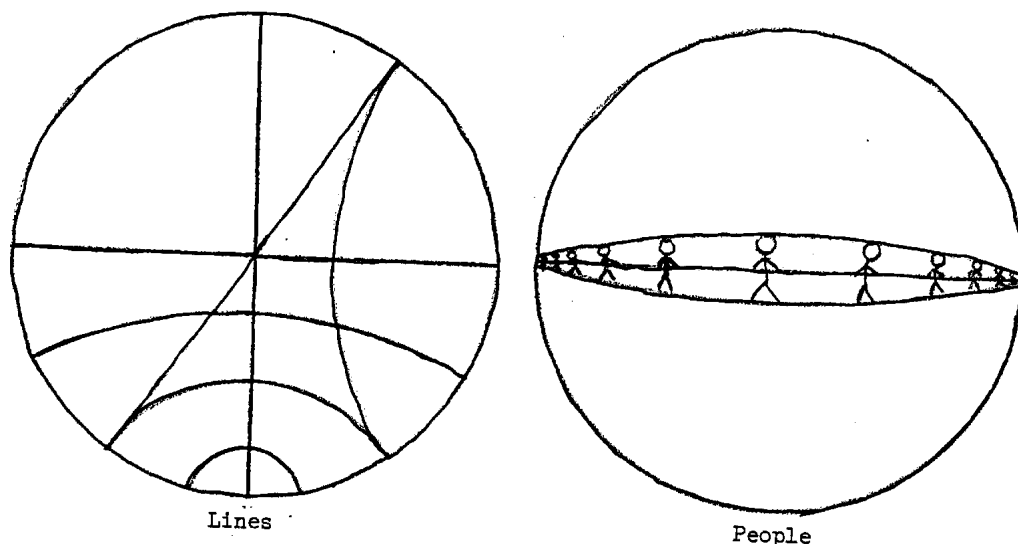
The sphere ∂D^n is called the *sphere at infinity*. It is not actually in hyperbolic space, but it can be given an interpretation purely in terms of hyperbolic geometry, as follows. Choose any base point p_0 in H^n. Consider any geodesic ray R, as seen from p_0. R traces out a segment of a great circle in the visual sphere at p_0 (since p_0 and R determine a two-plane). This visual segment converges to a point in the visual sphere. If we translate H^n so that p_0 is at the origin of the Poincaré disk model, we   2.5 see that the points in the visual sphere correspond precisely to points in the sphere at infinity, and that the end of a ray in this visual sphere corresponds to its Euclidean endpoint in the Poincaré disk model.

2.2. The southern hemisphere.

The Poincaré disk $D^n \subset \mathbb{R}^n$ is contained in the Poincaré disk $D^{n+1} \subset \mathbb{R}^{n+1}$, as a hyperbolic n-plane in hyperbolic $(n + 1)$-space.

Stereographic projection (Euclidean) from the north pole of ∂D^{n+1} sends the Poincaré disk D^n to the southern hemisphere of D^{n+1}.

Thus hyperbolic lines in the Poincaré disk go to circles on S^n orthogonal to the equal S^{n-1}.

There is a more natural construction for this map, using only hyperbolic geometry. For each point p in $H^n \subset H^{n+1}$, consider the hyperbolic ray perpendicular to H^n at p, and downward normal. This ray converges to a point on the sphere at infinity, which is the same as the Euclidean stereographic image of p.

2.6
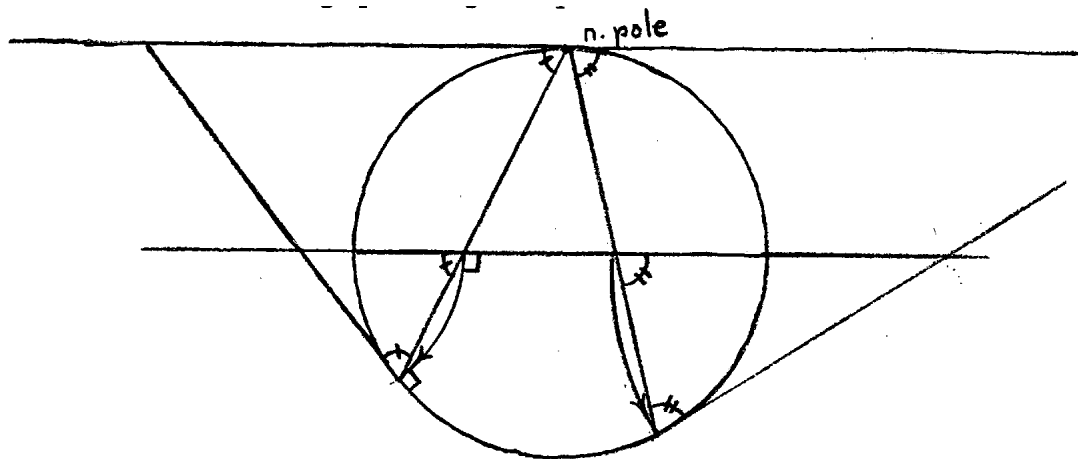
2.3. The upper half-space model.

This is closely related to the previous two, but it is often more convenient for computation or for constructing pictures. To obtain it, rotate the sphere S^n in $\mathbb{R}^{n+1}$ so that the southern hemisphere lies in the half-space $x_n \geq 0$ is $\mathbb{R}^{n+1}$. Now

stereographic projection from the top of S^n (which is now on the equator) sends the
southern hemisphere to the upper half-space $x_n > 0$ in $\mathbb{R}^{n+1}$.   2.7
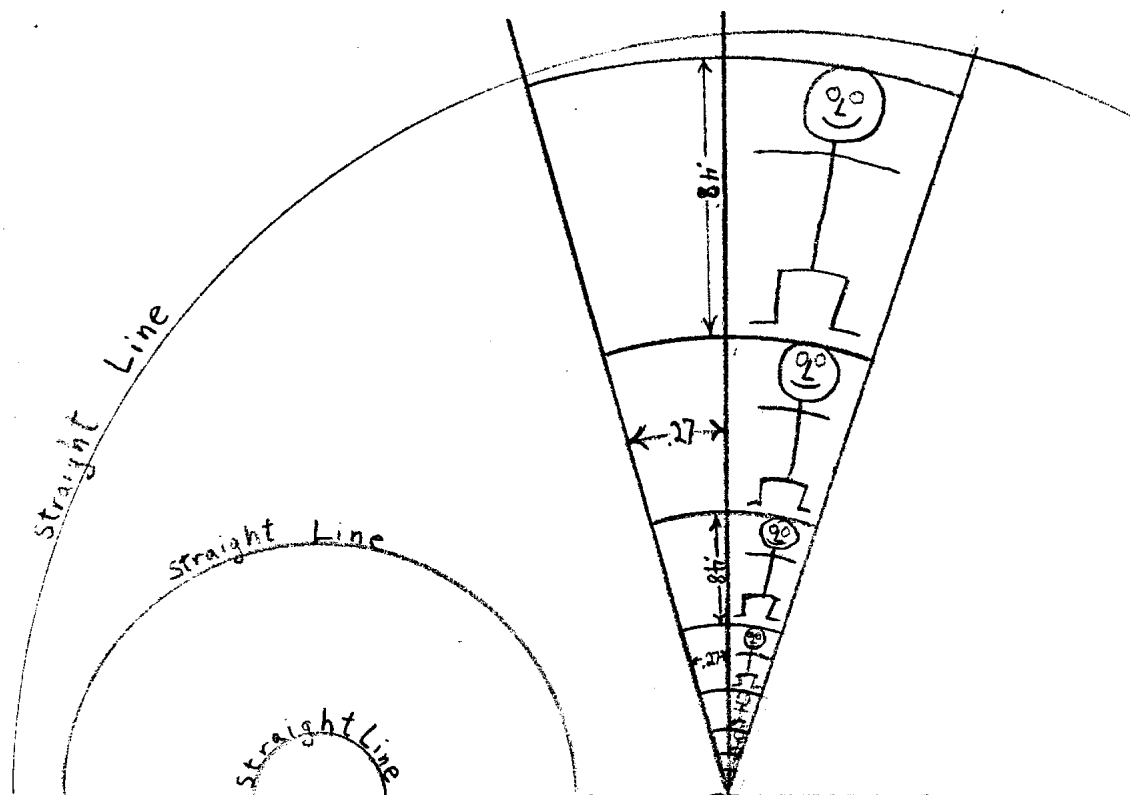
A hyperbolic line, in the upper half-space, is a circle perpendicular to the bounding
plane $\mathbb{R}^{n-1} \subset \mathbb{R}^n$. The hyperbolic metric is $ds^2 = (1/x_n)^2 \, dx^2$. Thus, the Euclidean
image of a hyperbolic object moving toward $\mathbb{R}^{n-1}$ has size precisely proportional to
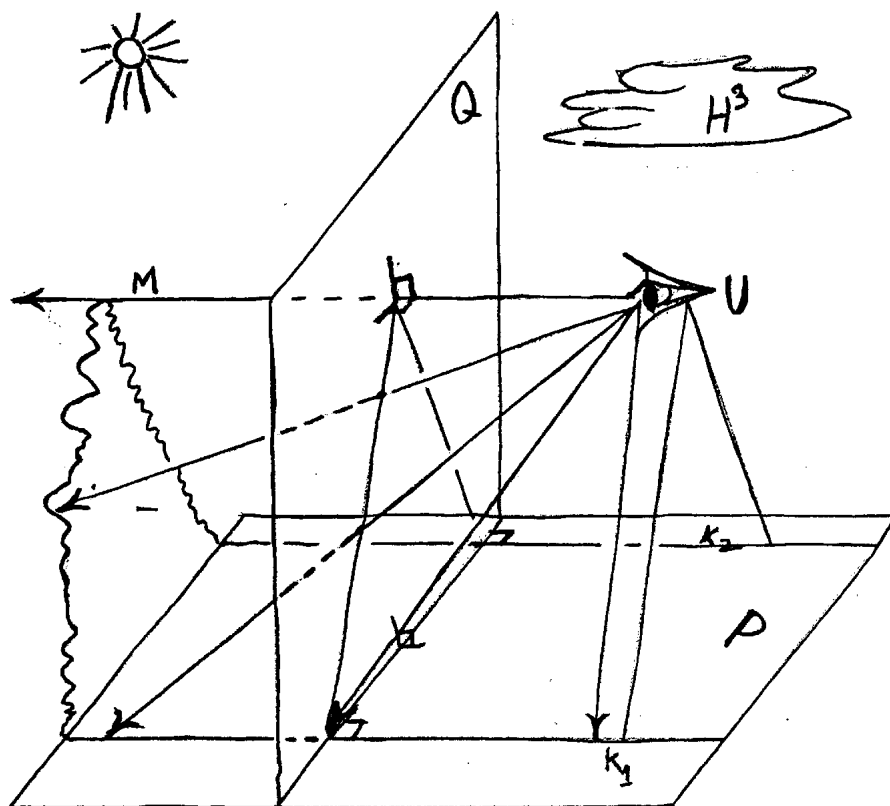the Euclidean distance from R^{n-1}.

2.4. The projective model.

This is obtained by Euclidean orthogonal projection of the southern hemisphere
of S^n back to the disk D^n. Hyperbolic lines become Euclidean line segments. This
model is useful for understanding incidence in a configuration of lines and planes.
Unlike the previous three models, it fails to be conformal, so that angles and shapes
are distorted.

It is better to regard this projective model to be contained not in Euclidean
space, but in projective space. The projective model is very natural from a point of   2.8
view inside hyperbolic $(n + 1)$-space: it gives a picture of a hyperplane, H^n, in true
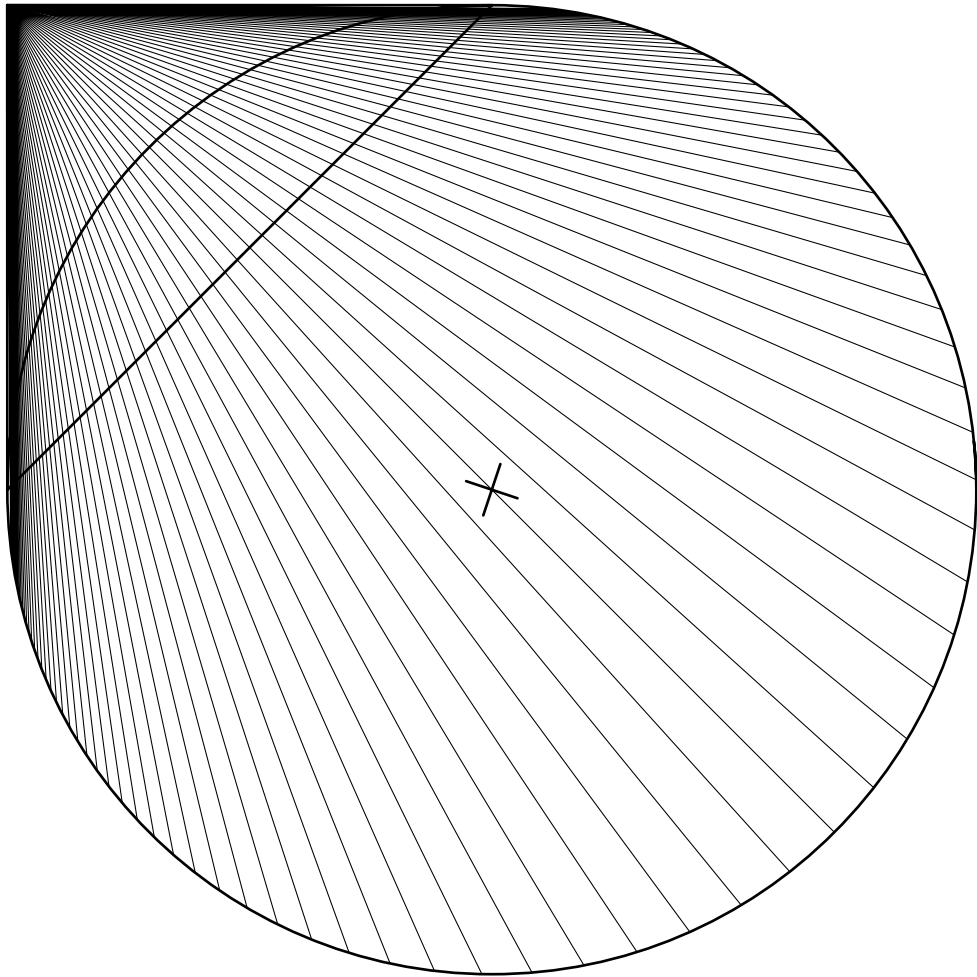perspective. Thus, an observer hovering above H^n in H^{n+1}, looking down, sees H^n

as the interior of a disk in his visual sphere. As he moves farther up, this visual disk shrinks; as he moves down, it expands; but (unlike in Euclidean space), the visual radius of this disk is always strictly less than $\pi/2$. A line on H^2 appears visually straight.

It is possible to give an intrinsic meaning within hyperbolic geometry for the points outside the sphere at infinity in the projective model. For instance, in the two-dimensional projective model, any two lines meet somewhere. The conventional sense of meeting means to meet inside the sphere at infinity (at a *finite* point). If the two lines converge in the visual circle, this means that they meet on the circle at infinity, and they are called *parallels*. Otherwise, the two lines are called *ultraparallels*; they have a unique common perpendicular L and they meet in some point x in the Möbius band outside the circle at infinity. *Any other line perpendicular to L passes through x, and any line through x is perpendicular to L.*   2.9

To prove this, consider hyperbolic two-space as a plane $P \subset H^3$. Construct the plane Q through L perpendicular to P. Let U be an observer in H^3. Drop a perpendicular M from U to the plane Q. Now if K is any line in P perpendicular

Evenly spaced lines. The region inside the circle is a plane, with a base line and a family of its perpendiculars, spaced at a distance of .051 fundamental units, as measured along the base line shown in perspective in hyperbolic 3-space (or in the projective model). The lines have been extended to their imaginary meeting point beyond the horizon. U, the observer, is directly above the X (which is .881 fundamental units away from the base line). To see the view from different heights, use the following table (which assumes that the Euclidean diameter of the circle in your printout is about 5.25 inches or 13.3cm):

To see the view of U at a height of	hold the picture a distance of	To see the view of U at a height of	hold the picture a distance of
2 units	11″ (28 cm)	5 units	17′ (519 cm)
3 units	27″ (69 cm)	10 units	2523′ (771 m)
4 units	6′ (191 cm)	20 units	10528.75 miles (16981 km)

For instance, you may imagine that the fundamental distance is 10 meters. Then the lines are spaced about like railroad ties. Twenty units is 200 meters: U is in a hot air balloon.

to L, the plane determined by U and K is perpendicular to Q, hence contains M; hence the visual line determined by K in the visual sphere of U passes through the visual point determined by K. The converse is similar.

2.10
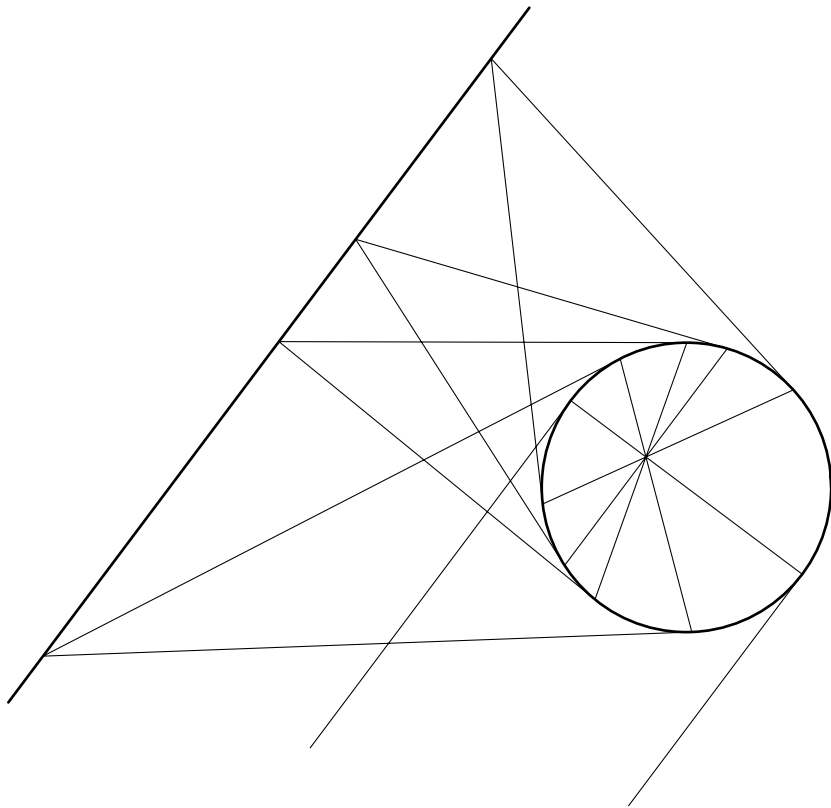
This gives a one-to-one correspondence between the set of points x outside the sphere at infinity, and (in general) the set of hyperplanes L in H^n. L corresponds to the common intersection point of all its perpendiculars. Similarly, there is a correspondence between points in H^n and hyperplanes outside the sphere at infinity: a point p corresponds to the union of all points determined by hyperplanes through p.

2.5. The sphere of imaginary radius.

A sphere in Euclidean space with radius r has constant curvature $1/r^2$. Thus, hyperbolic space should be a sphere of radius i. To give this a reasonable interpretation, we use an indefinite metric $dx^2 = dx_1^2 + \cdots + dx_n^2 - dx_{n+1}^2$ in R^{n+1}. The sphere of radius i about the origin in this metric is the hyperboloid
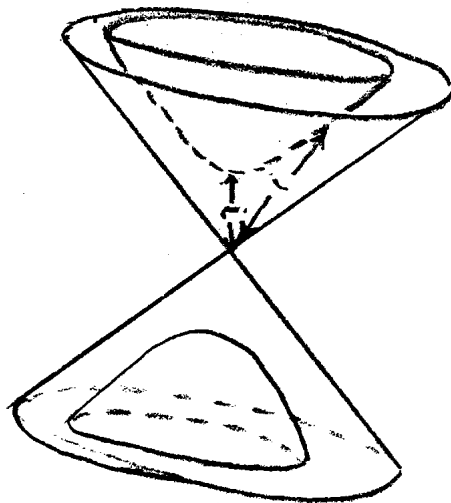
$$x_1^2 + \cdots + x_n^2 - x_{n+1}^2 = -1.$$

The metric dx^2 restricted to this hyperboloid is positive definite, and it is not hard to check that it has constant curvature -1.

Any plane through the origin is dx^2-orthogonal to the hyperboloid, so it follows from elementary Riemannian geometry that it meets the hyperboloid in a geodesic. The projective model for hyperbolic space is reconstructed by projection of the hyperboloid from the origin to a hyperplane in $\mathbb{R}^n$. Conversely, the quadratic form $x_1^2 + \cdots + x_n^2 - x_{n+1}^2$ can be reconstructed from the projective model. To do this, note that there is a unique quadratic equation of the form

$$\sum_{i,j=1}^{n} a_{ij} x_i x_j = 1$$

defining the sphere at infinity in the projective model. Homogenization of this equation gives a quadratic form of type $(n, 1)$ in $\mathbb{R}^{n+1}$, as desired. Any isometry of the quadratic form $x_1^2 + \cdots + x_n^2 - x_{n+1}^2$ induces an isometry of the hyperboloid, and hence any projective transformation of $\mathbb{P}^n$ that preserves the sphere at infinity induces an isometry of hyperbolic space. This contrasts with the situation in Euclidean geometry, where there are many projective self-homeomorphisms: the affine transformations. In particular, hyperbolic space has no similarity transformations except isometries. This is true also for elliptic space. This means that there is a well-defined unit of measurement of distances in hyperbolic geometry. We shall later see how this is related to three-dimensional topology, giving a measure of the "size" of manifolds.

2.6. Trigonometry.

Sometimes it is important to have formulas for hyperbolic geometry, and not just pictures. For this purpose, it is convenient to work with the description of hyperbolic

space as one sheet of the "sphere" of radius i with respect to the quadratic form

$$Q(X) = X_1^2 + \cdots + X_n^2 - X_{n+1}^2$$

in $\mathbb{R}^{n+1}$. The set $\mathbb{R}^{n+1}$, equipped with this quadratic form and the associated inner product

$$X \cdot Y = \sum_{i=1}^{n} X_1 Y_1 - X_{n+1} Y_{n+1},$$

is called $E^{n,1}$.

First we will describe the geodesics on level sets $S_r = \{X : Q(X) = r^2\}$ of Q. Suppose that X_t is such a geodesic, with speed

$$s = \sqrt{Q(\dot{X}_t)}.$$

We may differentiate the equations

$$X_t \cdot X_t = r^2, \qquad \dot{X}_t \cdot \dot{X}_t = s^2,$$

to obtain

$$X_t \cdot \dot{X}_t = 0, \qquad \dot{X}_t \cdot \ddot{X}_t = 0,$$

and

$$X_t \cdot \ddot{X}_t = -\dot{X}_t \cdot \dot{X}_t = -s^2.$$

Since any geodesic must lie in a two-dimensional subspace, $\ddot{X}_t$ must be a linear combination of X_t and $\dot{X}_t$, and we have

2.6.1.
$$\ddot{X}_t = -\left(\frac{s}{r}\right)^2 X_t.$$

This differential equation, together with the initial conditions

2.13

$$X_0 \cdot X_0 = r^2, \qquad \dot{X}_0 \cdot \dot{X}_0 = s^2, \qquad X_0 \cdot \dot{X}_0 = 0,$$

determines the geodesics.

Given two vectors X and Y in $E^{n,1}$, if X and Y have nonzero length we define the quantity

$$c(X, Y) = \frac{X \cdot Y}{\|X\| \cdot \|Y\|},$$

where $\|X\| = \sqrt{X \cdot X}$ is positive real or positive imaginary. Note that

$$c(X, Y) = c(\lambda X, \mu Y),$$

where λ and μ are positive constants, that $c(-X, Y) = -c(X, Y)$, and that $c(X, X) = 1$. In Euclidean space E^{n+1}, $c(X, Y)$ is the cosine of the angle between X and Y. In $E^{n,1}$ there are several cases.

We identify vectors on the positive sheet of S_i ($X_{n+1} > 0$) with hyperbolic space. If Y is any vector of real length, then Q restricted to the subspace $Y^\perp$ is indefinite

of type $(n-1, 1)$. This means that $Y^{\perp}$ intersects H^n and determines a hyperplane. We will use the notation $Y^{\perp}$ to denote this hyperplane, *with the normal orientation determined by Y*. (We have seen this correspondence before, in 2.4.)

2.6.2.   If X and $Y \in H^n$, then $c(X,Y) = \cosh d(X,Y)$,

where $d(X,Y)$ denotes the hyperbolic distance between X and Y.

To prove this formula, join X to Y by a geodesic X_t of unit speed. From 2.6.1 we   2.14
have

$$\ddot{X}_t = X_t, \qquad X_t \cdot \dot{X}_0 = 0,$$

so we get $c(\ddot{X}_t, X_t) = c(X_t, X_t)$, $c(\dot{X}_0, X_0) = 0$, $c(X, X_0) = 1$; thus $c(X, X_t) = \cosh t$. When $t = d(X,Y)$, then $X_t = Y$, giving 2.6.2.

If $X^{\perp}$ and $Y^{\perp}$ are distinct hyperplanes, then

2.6.3.

$X^{\perp}$ and $Y^{\perp}$ intersect

$\Longleftrightarrow$ Q is positive definite on the subspace $\langle X, Y \rangle$ spanned by X and Y

$\Longleftrightarrow$ $c(X,Y)^2 < 1$

$\Longrightarrow$ $c(X,Y) = \cos \angle(X,Y) = -\cos \angle(X^{\perp}, Y^{\perp})$.
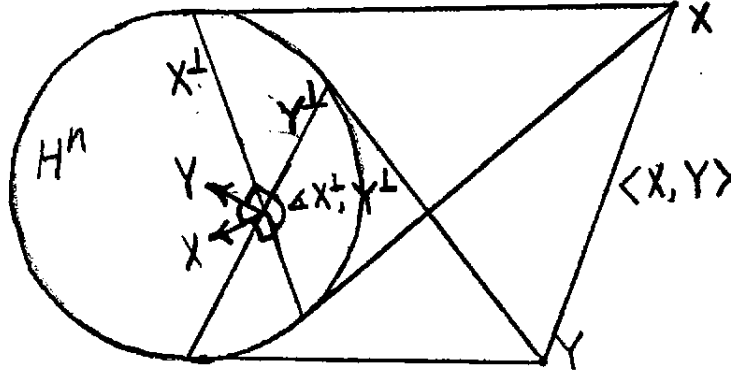
To see this, note that X and Y intersect in $H^n \iff Q$ restricted to $X^{\perp} \cap Y^{\perp}$ is indefinite of type $(n-2, 1) \iff Q$ restricted to $\langle X, Y \rangle$ is positive definite. ($\langle X, Y \rangle$ is the normal subspace to the $(n-2)$ plane $X^{\perp} \cap Y^{\perp}$).   2.15

There is a general elementary formula for the area of a parallelogram of sides X and Y with respect to an inner product:

$$\text{area} = \sqrt{X \cdot X\, Y \cdot Y - (X \cdot Y)^2} = \|X\| \cdot \|Y\| \cdot \sqrt{1 - c(X,Y)^2}.$$

This area is positive real if X and Y span a positive definite subspace, and positive imaginary if the subspace has type $(1,1)$. This shows, finally, that $X^{\perp}$ and
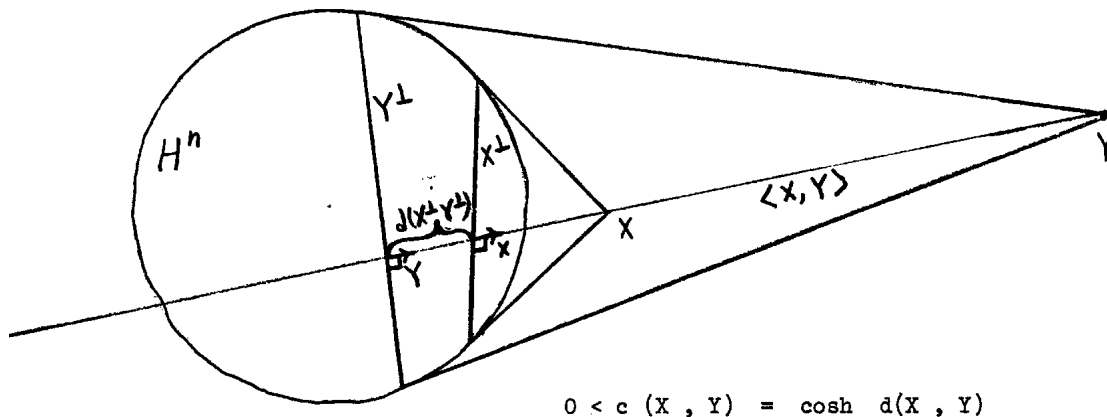
$Y^\perp$ intersect $\iff$ $c\,(X,Y)^2 < 1$. The formula for $c(X,Y)$ comes from ordinary trigonometry.

2.6.4.

$X^\perp$ and $Y^\perp$ have a common perpendicular $\iff$ Q has type $(1,1)$ on $\langle X, Y \rangle$

$$\iff c(X,Y)^2 > 1$$
$$\implies c(X,Y) = \pm \cosh\big(d\,(X^\perp, Y^\perp)\big).$$

The sign is positive if the normal orientations of the common perpendiculars coincide, and negative otherwise.

$$0 < c\ (X\ ,\ Y)\ =\ \cosh\ d(X\ ,\ Y)$$

2.16

The proof is similar to 2.6.2. We may assume X and Y have unit length. Since $\langle X, Y \rangle$ intersects H^n as the common perpendicular to $X^\perp$ and $Y^\perp$, Q restricted to $\langle X, Y \rangle$ has type $(1,1)$. Replace X by $-X$ if necessary so that X and Y lie in the same component of $S_1 \cap \langle X, Y \rangle$. Join X to Y by a geodesic X_t of speed i. From 2.6.1, $\ddot{X}_t = X_t$. There is a dual geodesic Z_t of unit speed, satisfying $Z_t \cdot X_t = 0$, joining $X^\perp$ to $Y^\perp$ along their common perpendicular, so one may deduce that

$$c,(X,Y) = \pm\tfrac{d\,(X,Y)}{i} = \pm d\,(X^\perp, Y^\perp).$$

There is a limiting case, intermediate between 2.6.3 and 2.6.4:

2.6.5. $X^\perp$ and $Y^\perp$ are parallel

$$\iff Q \text{ restricted to } \langle X, Y \rangle \text{ is degenerate}$$
$$\iff c(X,Y)^2 = 1.$$

In this case, we say that $X^\perp$ and $Y^\perp$ form an angle of 0 or π. $X^\perp$ and $Y^\perp$ actually have a distance of 0, where the distance of two sets U and V is defined to be the infimum of the distance between points $u \in U$ and $v \in V$.
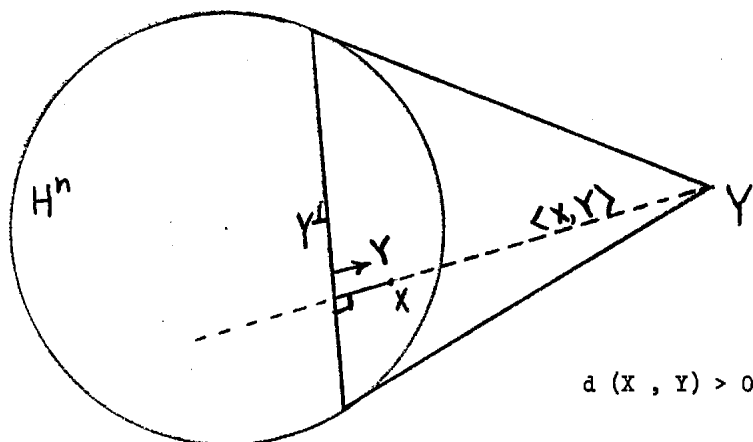
There is one more case in which to interpret $c(X, Y)$:

2.6.6. If X is a point in H^n and $Y^\perp$ a hyperplane, then
$$c(X, Y) = \frac{\sinh\big(d\,(X, Y^\perp)\big)}{i},$$

where $d\,(X, Y^\perp)$ is the *oriented* distance. 2.17

d (X , Y) > 0

The proof is left to the untiring reader.

With our dictionary now complete, it is easy to derive hyperbolic trigonometric formulae from linear algebra. To solve triangles, note that the edges of a triangle with vertices u, v and w in H^2 are $U^\perp$, $V^\perp$ and $W^\perp$, where U is a vector orthogonal to v and w, etc. To find the angles of a triangle from the lengths, one can find three vectors u, v, and w with the appropriate inner products, find a dual basis, and calculate the angles from the inner products of the dual basis.

Here is the general formula. We consider triangles in the projective model, with vertices inside or outside the sphere at infinity. Choose vectors v_1, v_2 and v_3 of length i or 1 representing these points. Let $\epsilon_i = v_i \cdot v_i$, $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$ and $c_{ij} = c(v_i, v_j)$. Then the matrix of inner products of the v_i is

$$C = \begin{bmatrix} \epsilon_1 & \epsilon_{12} c_{12} & \epsilon_{13} c_{13} \\ \epsilon_{12} c_{12} & \epsilon_2 & \epsilon_{23} c_{23} \\ \epsilon_{13} c_{13} & \epsilon_{23} c_{23} & \epsilon_3 \end{bmatrix}.$$

The matrix of inner products of the dual basis $\{v^1, v^2, v^3\}$ is C^{-1}. For our purposes, though, it is simpler to compute the matrix of inner products of the basis 2.18

$$\{\sqrt{-\det C^i}\},$$

$$-\operatorname{adj} C = (-\det C) \cdot C^{-1} =$$

$$\begin{bmatrix} -\epsilon_2\epsilon_3(1 - c_{23}^2) & -\epsilon_{12}\epsilon_3(c_{13}c_{23} - c_{12}) & -\epsilon_{13}\epsilon_2(c_{12}c_{23} - c_{13}) \\ -\epsilon_{12}\epsilon_3(c_{13}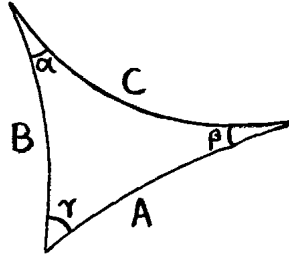c_{23} - c_{12}) & -\epsilon_1\epsilon_3(1 - c_{13}^2) & -\epsilon_{23}\epsilon_1(c_{12}c_{13} - c_{23}) \\ -\epsilon_{13}\epsilon_2(c_{12}c_{23} - c_{13}) & -\epsilon_{23}\epsilon_1(c_{12}c_{13} - c_{23}) & -\epsilon_1\epsilon_2(1 - c_{12}^2) \end{bmatrix}.$$

If v^1, v^2, v^3 is the dual basis, and $c^{ij} = c(v^i, v^j)$, we can compute

2.6.7.
$$c^{12} = \epsilon \cdot \frac{c_{13}c_{23} - c_{12}}{\sqrt{1 - c_{23}^2}\sqrt{1 - c_{13}^2}},$$

where it is easy to deduce the sign

$$\epsilon = \frac{-\epsilon_{12}\epsilon_3}{\sqrt{-\epsilon_2\epsilon_3}\sqrt{-\epsilon_1\epsilon_3}}$$

directly. This specializes to give a number of formulas, in geometrically distinct cases. In a real triangle,

2.6.8.
$$\cosh C = \frac{\cos\alpha\cos\beta + \cos\gamma}{\sin\alpha\sin\beta},$$

2.19

2.6.9.
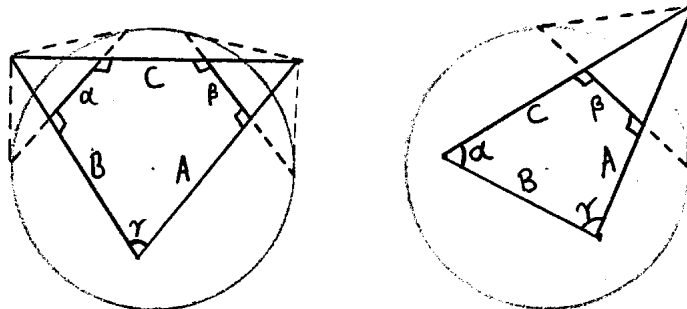$$\cos\gamma = \frac{\cosh A \cosh B - \cosh C}{\sinh A \sinh B},$$

or $\cosh C = \cosh A \cosh B - \sinh A \sinh B \cos c$. (See also 2.6.16.) In an all right hexagon,
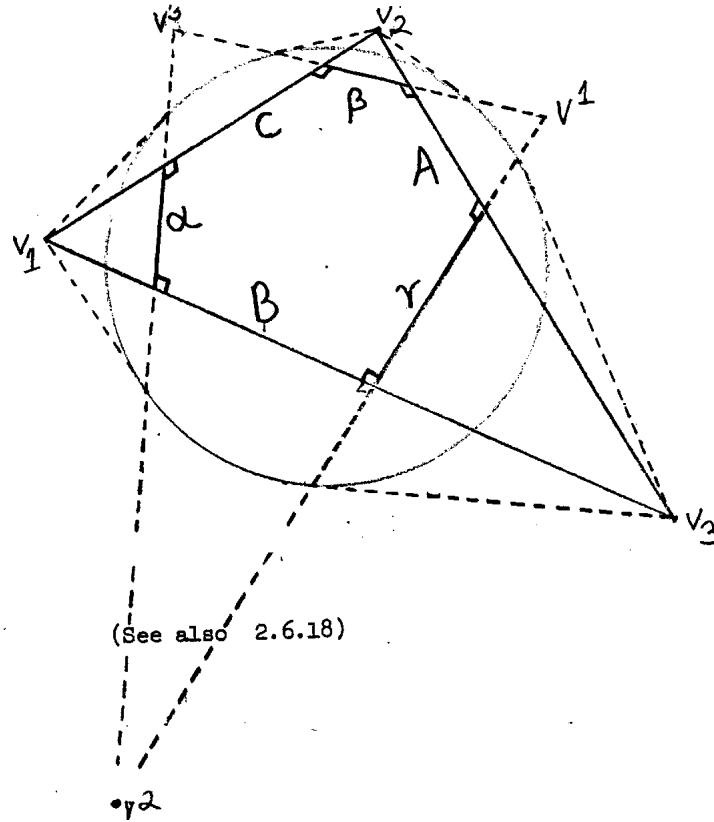
(See also 2.6.18)

2.6.10.
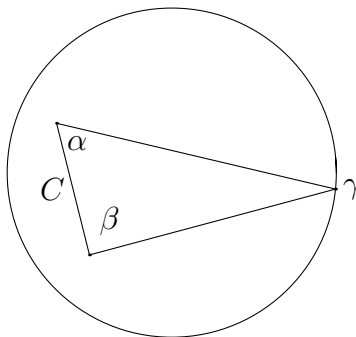$$\cosh C = \frac{\cosh \alpha \cosh \beta + \cosh \gamma}{\sinh \alpha \sinh \beta}.$$

(See also 2.6.18.) Such hexagons are useful in the study of hyperbolic structures on surfaces. Similar formulas can be obtained for pentagons with four right angles, or quadrilaterals with two adjacent right angles:
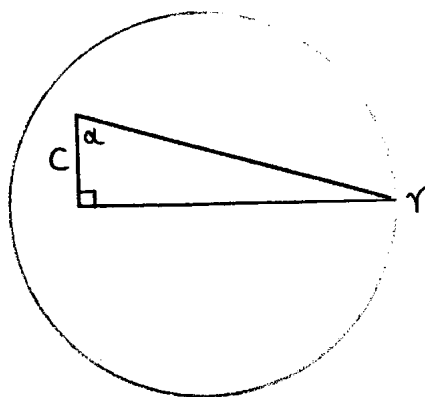

2.20

By taking the limit of 2.6.8 as the vertex with angle γ tends to the circle at infinity, we obtain useful formulas:

2.6.11.
$$\cosh C = \frac{\cos \alpha \cos \beta + 1}{\sin \alpha \sin \beta},$$

and in particular

2.6.12.
$$\cosh C = \frac{1}{\sin \alpha}.$$

These formulas for a right triangle are worth mentioning separately, since they are particularly simple.

2.21
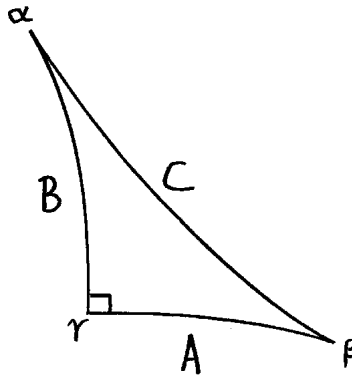
From the formula for $\cos \gamma$ we obtain the hyperbolic Pythagorean theorem:

2.6.13. $$\cosh C = \cosh A \cosh B.$$

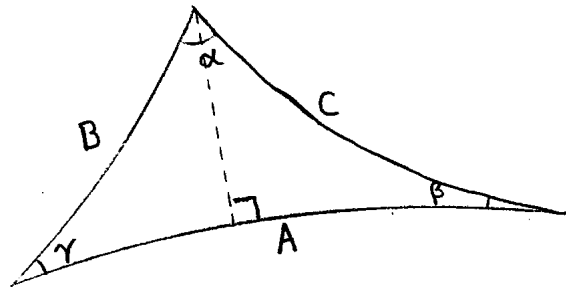Also,

2.6.14. $$\cosh A = \frac{\cos \alpha}{\sin \beta}.$$

(Note that $(\cos \alpha)/(\sin \beta) = 1$ in a Euclidean right triangle.) By substituting

$$\frac{(\cosh C)}{(\cosh A)}$$

for $\cosh B$ in the formula 2.6.9 for $\cos \alpha$, one finds:

2.6.15. In a right triangle, $\sin \alpha = \dfrac{\sinh A}{\sinh C}.$

This follows from the general law of sines,

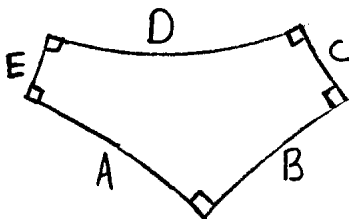2.6.16. In any triangle, $\dfrac{\sinh A}{\sin \alpha} = \dfrac{\sinh B}{\sin \beta} = \dfrac{\sinh C}{\sin \gamma}.$

2.22

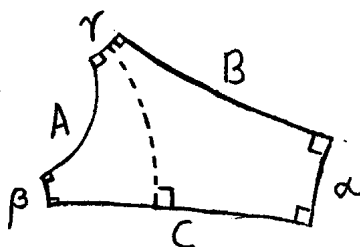Similarly, in an all right pentagon,

one has

2.6.17. $$\sinh A \sinh B = \cosh D.$$

It follows that in any all right hexagon,

there is a law of sines:

2.6.18. $$\frac{\sinh A}{\sinh \alpha} = \frac{\sinh B}{\sinh \beta} = \frac{\sinh C}{\sinh \gamma}.$$