1. INTRODUCTION. Hardly anyone would maintain that it is better to begin to learn geography from flat maps than from a globe. But almost all introductions to hyperbolic non-Euclidean geometry, except [6], present plane models, such as the projective and conformal disk models, without even mentioning that there exists a model that has the same relation to plane models that a globe has to flat maps. This model, which is on one sheet of a hyperboloid of two sheets in Minkowski 3-space and which I shall call $H^2$, is over a hundred years old; Killing and Poincaré both described it in the 1880's (see Section 14). It is used by differential geometers [29, p. 4] and physicists (see [21, pp. 724–725] and [23, p. 113]). Nevertheless it is not nearly so well known as it should be, probably because, like a globe, it requires three dimensions.

The main advantages of this model are its naturalness and its symmetry. Being embeddable (distance function and all) in flat space-time, it is close to our picture of physical reality, and all its points are treated alike in this embedding. Once the strangeness of the Minkowski metric is accepted, it has the familiar geometry of a sphere in Euclidean 3-space $E^3$ as a guide to definitions and arguments. For example, the lines of $H^2$ are its non-empty intersections with the planes through the origin of the Minkowski space $M^3$. The length of a line segment of $H^2$ is defined by analogy with arc length in calculus; this leads naturally to the hyperbolic functions. As in the spherical case, every isometry of $H^2$ can be extended to a linear transformation of $M^3$, so that straightforward calculations with matrices can be used to prove theorems and develop the trigonometry of $H^2$. The circles, horocycles, and equidistant curves have a beautiful interpretation: they are precisely the nontrivial intersections of $H^2$ with planes of $M^3$ that do not pass through the origin. An area function for $H^2$ can be constructed from the volume function on $M^3$.

The aim of this article is to show that $H^2$ can be used to give an introduction to hyperbolic geometry to undergraduates who know a little about linear transformations and groups, a bit of special relativity being helpful for motivation. The mixing of rigor and intuition is similar to what is common in calculus courses. The treatment is not axiomatic, since there is no intrinsic reason to stress axiomatics in hyperbolic geometry any more than in, say, spherical trigonometry. For historical and philosophical reasons, however, many treatments are based on axioms; therefore I shall refer to Moise's axioms [22] at the points where they can be verified from my approach. (I have chosen these axioms since, incorporating real-valued functions for distance and measure of angles, they are closer to my analytic approach than Hilbert's [6], [8]; the latter, being weaker, would be easier to verify.) I will treat them not as axioms, but just as properties of $H^2$. I have included something about the hyperbolic analogues of map projections and about the history of the model.
A special feature of my approach, which distinguishes it from Faber's [6, Chapter VII], is its extensive use of orthogonal transformations of $M^3$, the analogues of the rigid motions of $E^3$ that fix the origin, to move subsets of $H^2$ to convenient positions.

I want to thank Alan H. Durfee and Mark E. Kidwell for encouraging me to write up this article. It originated in an undergraduate course at Tufts, and I did some of the work while visiting at Harvard.

2. MINKOWSKI 3-SPACE. By Minkowski 3-space $M^3$ I mean a 3-dimensional real vector space together with a real-valued function $q$ on it such that

$$q\left(\sum_{i=0}^{2} x_i U_i\right) = -x_0^2 + x_1^2 + x_2^2$$

(2.1)

for some basis $\mathcal{B} = (U_0, U_1, U_2)$ of the space. More briefly,

$$q\left(\sum x_i U_i\right) = \sum e_i x_i^2$$

(2.2)

where $e_0 = -1$, $e_1 = e_2 = 1$; then $q(U_i) = e_i$. We can define Minkowski n-space $M^n$ similarly for any $n \geq 2$ (with one minus sign). To relate the cases $n = 2$ and 3, take $M^2$ to be the subspace of $M^3$ with basis $(U_0, U_1)$ together with the restriction of $q$.

If we replace (2.2) by

$$q_E\left(\sum x_i U_i\right) = \sum x_i^2,$$

(2.3)

we get ordinary Euclidean 3-space $E^3$, with $q_E$ giving the square of length. I shall constantly use analogies from $E^3$ to study $M^3$; watch for such analogies when they are not mentioned.

For $X = \sum x_i U_i$ and $Y = \sum y_i U_i$ in $M^3$, define

$$p(X, Y) = \frac{1}{2} [q(X + Y) - q(X) - q(Y)].$$

(2.4)

(This is the bilinear form or pairing corresponding to the quadratic form $q$.) Then

$$q(X) = p(X, X)$$

(2.5)

and

$$p(X, Y) = \sum e_i x_i y_i = -x_0 y_0 + x_1 y_1 + x_2 y_2;$$

(2.6)

in particular $p(U_i, U_j)$ equals $e_i$ if $i = j$ and 0 otherwise. Observe that $p$ is analogous to the ordinary dot product on $E^3$ given by $p_E(X, Y) = X \cdot Y = \sum x_i y_i$ and $\mathcal{B}$ to an orthonormal basis.

3. THE HYPERBOLOIDAL MODEL. In $M^3$, let $H^2$ be the set of all vectors $X = \sum x_i U_i$ for which

$$q(X) = -1,$$

(3.1)

$$x_0 > 0.$$  

(3.2)

These two conditions can be expressed by the equation

$$x_0 = \sqrt{1 + x_1^2 + x_2^2}.$$  

(3.3)

(3.1) describes a hyperboloid of two sheets and (3.2) picks out one sheet. We can define $H^n \subset M^{n+1}$ similarly for arbitrary $n$; for $n = 1$, take $H^1 = H^2 \cap M^2$ with $M^2$ as in the previous section. Figure 1 may help in thinking about $H^2$. In flat
space-time, identified with $M^4$, $H^2$ appears to each observer as a circle whose radius is increasing at slightly greater (!) than the speed of light.

We now begin to construct a model of hyperbolic geometry whose points, or $H$-points, are the elements of $H^2$. We think of them as either points or vectors when considered in $M^3$, and as points when considered in $H^2$. The lines or $H$-lines of the model are defined to be all the nonempty intersections of $H^2$ with 2-dimensional subspaces of $M^3$; for example, $H^1$ is an $H$-line. (The prefix "$H$" will always be optional.)

Incidence of $H$-points and $H$-lines and betweenness for points of an $H$-line are defined in the natural way. Each pair of distinct $H$-points $A$ and $B$ lies on a unique $H$-line $\overline{AB}$, namely the intersection of $H^2$ with the plane $OAB$ of $M^3$, $O$ being the origin of $M^3$. The definitions of the $H$-segment $\overline{AB}$ and the $H$-ray $\overrightarrow{AB}$ are straightforward. We can now check the plane incidence axioms of [22, pp. 37–38].

It should be clear that the complement of each $H$-line in $H^2$ consists of two $H$-half planes, called its sides. This statement can be made precise as the plane-separation axiom or Pasch’s axiom [22, p. 62]. Two distinct $H$-lines have one or zero $H$-points in common according as the line of intersection of the planes of $M^3$ in which they lie intersects $H^2$ or not; this gives the hyperbolic parallel axiom [22, p. 114]. The Archimedean axiom and the axiom of completeness (or continuity) [22, pp. 256 and 265] are also clear.

In the analogous situation in $E^3$, the equation $q_E(X) = 1$ (cf. (2.3)) defines a sphere $S^2$; this leads to a model of spherical (or double elliptic) geometry whose lines or $S$-lines are the great circles of $S^2$.

4. DISTANCE. For distinct $H$-points $A$ and $B$ we want to define the $H$-distance $d(A, B)$, also called the $H$-length of $\overline{AB}$. The natural way to adapt the usual definition of arc length is to partition $\overline{AB}$, as a curve in $M^3$, by suitable points
$P_0 = A, P_1, \ldots, P_m = B$ and to define

$$d(A, B) = \lim_{m \to \infty} \sum_{j=1}^{m} \sqrt{q(P_j - P_{j-1})}; \quad (4.1)$$

but first we must check that $q(P_j - P_{j-1})$, analogous to a squared length, is positive.

The equation $q(X) = 0$ describes a cone in $M^3$ and the vectors with $q(X) > 0$ are the points outside this cone (cf. the spacelike vectors of special relativity). By the mean value theorem there is a point of $P_{j-1}P_j$ at which the tangent vectors in $M^3$ to this $H$-segment are parallel to $P_j - P_{j-1}$, so it suffices to show that all vectors $V \neq O$ tangent to $H^2$ have $q(V) > 0$. We can show this by turning to advantage a limitation of our intuition. (For a different approach, see Section 7.) Since we are used to Euclidean space, any attempt to visualize $M^3$ pictorially as in Figure 1 identifies it with $E^3$, i.e. imposes a Euclidean quadratic form on it. This destroys the symmetry of $M^3$ and of $H^2$, so that different points of $H^2$ do not look alike. Since we cannot avoid this identification, we use it. Suppose then that $M^3 = E^3$ has both forms $q$ and $q_E$ with respect to the basis $\mathcal{B}$. The cone $q(X) = 0$ now consists of all vectors that make (Euclidean) angles of $\pi/4$ with the plane $x_0 = 0$. All tangent vectors $V$ to $H^2$ in $E^3$ make angles less than $\pi/4$ with that plane, so that $q(V) > 0$; this implies that the definition (4.1) of $d(A, B)$ is valid as in the Euclidean case.

To evaluate the limit, parametrize $\overline{AB}$; that is, let $F$ be a smooth one-to-one mapping of an interval $a < t < b$ onto $\overline{AB}$ with $F(a) = A$, $F(b) = B$. $F$ is a vector-valued function of $t$ with derivative $F'$ and $P_j = F(t_j)$ for a partition $a = t_0 < t_1 < \cdots < t_m = b$. The limit is

$$d(A, B) = \int_{a}^{b} \sqrt{q(F'(t))} \, dt. \quad (4.2)$$

Moise’s axioms of distance and definition of betweenness [22, pp. 47–49 and 51] can be obtained now; also see Section 12. For curves in $H^2$, (4.1) and (4.2) give $H$-arc length.

We can use (4.2) to compute $H$-distances along $H^1$; we shall treat $H^2$ in Section 6. Let $\overline{AB}$ be any $H$-line segment of $H^1$, say $A = a_0U_0 + a_1U_1$, $B = b_0U_0 + b_1U_1$, with $a_1 < b_1$. Using (3.3) we can take the parameter $t = x_1$, with $F(t) = \sqrt{1 + t^2} U_0 + tU_1$, $a_1 \leq t \leq b_1$. Then

$$d(A, B) = \int_{a_1}^{b_1} \sqrt{-\frac{t^2}{1 + t^2}} + 1 \, dt = [\ln(t + \sqrt{t^2 + 1})]_{a_1}^{b_1}.$$

Define

$$\text{arcsinh } t = \ln(t + \sqrt{t^2 + 1});$$

arcsinh is monotone increasing and arcsinh $0 = 0$. Then

$$d(A, B) = \text{arcsinh } b_1 - \text{arcsinh } a_1;$$

in particular $d(U_0, B) = \text{arcsinh } b_1$ if $b_1 > 0$, whence the $H$-length of $H^1$ is infinite. If $r = \text{arcsinh } t$, then $t = (e^r - e^{-r})/2 = \sinh r$, the hyperbolic sine of $r$. On $H^1$, $x_0 = \sqrt{1 + t^2} = (e^r + e^{-r})/2 = \cosh r$, the hyperbolic cosine of $r$; so a parametrization of $H^1$ by $H$-length is

$$P(r) = (\cosh r)U_0 + (\sinh r)U_1, \quad -\infty < r < \infty. \quad (4.3)$$
I use the slightly unusual symbol “arcsinh t” since it does represent an (arc) length in our peculiar way of measuring.

5. ORTHOGONAL TRANSFORMATIONS. A linear transformation $T$ of $M^3$ to $M^3$ is called orthogonal (with respect to $q$) if

$$q(T(X)) = q(X), \quad X \in M^3. \quad (5.1)$$

The orthogonal transformations form a group, the orthogonal group $O(M^3)$.

We can express this in terms of matrices as follows. Each vector $X = \sum x_i U_i$ of $M^3$ has its column of coordinates

$$[X] = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} \quad (5.2)$$

with respect to the basis $\mathcal{Q}$. Each linear transformation $T$ has a matrix $[T] = [t_{ij}]$ with respect to $\mathcal{Q}$ such that the matrix equation

$$[T(X)] = [T][X] \quad (5.3)$$

holds for all $X$. By (2.4) and (2.5), $T$ is in $O(M^3)$ if and only if

$$p(T(X), T(Y)) = p(X, Y), \quad X, Y \in M^3, \quad (5.4)$$

or, equivalently,

$$p(T(U_j), T(U_k)) = \sum_{i=0}^2 e_i t_{ij} t_{ik} = \begin{cases} e_j & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases} \quad (5.5)$$

(since $[T(U_j)]$ is the $j$-th column of $[T]$). (5.5) can be written as

$$[T]'[J_0][T] = [J_0], \quad (5.6)$$

where $[T]'$ is the transpose of $[T]$ and

$$[J_0] = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$ 

By (5.6), the determinant of $T$ is $\pm 1$. The elements of $O(M^3)$ with determinant 1 form a subgroup $O^+(M^3)$ of index 2, the special orthogonal group.

The orthogonal group has a subgroup associated to $H^2$ that will be fundamental to us. Let $T \in O(M^3)$; if $X \in H^2$, (3.1) and (5.1) imply that $T(X) \in H^2$ or $-T(X) \in H^2$. By continuity, those $T$ that map $H^2$ onto itself form another subgroup of $O(M^3)$ of index 2, which I shall call $G(M^3)$ or simply $G$. Furthermore $G(M^3)$ has the subgroup $G^+(M^3) = G(M^3) \cap O^+(M^3)$ of index 2.

Let $T \in G(M^3)$; since $T$ maps each 2-dimensional subspace of $M^3$ onto another, it maps each $H$-line onto another. Then, by (4.1) and (5.1), if $A$ and $B$ are points of $H^2$,

$$d(T(A), T(B)) = \lim \sum \sqrt{q(T(P_j) - T(P_{j-1}))}$$

$$= \lim \sum \sqrt{q(T(P_j - P_{j-1}))} = \lim \sum \sqrt{q(P_j - P_{j-1})} = d(A, B). \quad (5.7)$$

This means that the restriction to $H^2$ of every element of $G(M^3)$ is an $H$-isometry or distance-preserving mapping of $H^2$. 

446 HYPERBOLIC GEOMETRY ON A HYPERBOLOID [May
Analogous groups exist for $M^n$; for example, $G^+(M^4)$ is the Lorentz group of special relativity. Similarly, in Euclidean space, $O(E^3)$ consists of the usual orthogonal transformations and $O^+(E^3)$ of the rotations about the origin. (Since there is no analogue of the sheets of $H^2$, $O(E^3)$ corresponds to $G(M^3)$ as well as to $O(M^3)$.)

6. THE FORMULA FOR DISTANCE. The following calculations, first over $M^2$ and $E^2$ and then in $G = G(M^3)$, will motivate equations (6.4) and (6.7); some readers may prefer to skip them or to use the alternative argument indicated at the end of this section. By the two-dimensional analogue of (5.5), it is a simple exercise to show that $T \in O(M^2)$ if and only if

$$[T] = \begin{bmatrix} e \cosh s & f \sinh s \\ e \sinh s & f \cosh s \end{bmatrix}, \quad e = \pm 1, \quad f = \pm 1, \quad (6.1)$$

with $e, f, s$ uniquely determined by $T$. Then $T \in G(M^2)$ if and only if $e = 1$, $T \in O^+(M^2)$ if and only if $e = f$, and $T \in G^+(M^2)$ if and only if $e = f = 1$. Similarly, as is well known, $O(E^2)$ consists of those $T$ such that

$$[T] = \begin{bmatrix} \cos \theta & -h \sin \theta \\ \sin \theta & h \cos \theta \end{bmatrix}, \quad h = \pm 1. \quad (6.2)$$

$T \in O^+(E^2)$ if and only if $h = 1$.

Now let $G_1$ be the subgroup of $G$ consisting of those $T$ that fix the $H$-line $H$ (as a whole), i.e. such that $T(H^1) = H^1$ in the usual notation. If $T \in G_1$, then $T$ also fixes $M^2$, so that $t_{20} = t_{21} = 0$. By (5.5), it is easy to see that $t_{02} = t_{12} = 0$, i.e., $T$ fixes the subspace spanned by $U_2$. (This is a special case of a fact about orthogonal complements [11, p. 364].) On $M^2$, $T$ acts like an element of $G(M^2)$; then, by (6.1),

$$T = L_j J_i J_{j'}, \quad i, j, j' = 0, 1, \quad (6.3)$$

where

$$[L_s] = \begin{bmatrix} \cosh s & \sinh s & 0 \\ \sinh s & \cosh s & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (6.4)$$

for some real $s$,

$$[J_i] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad [J_2] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \quad$$

Thus $G_1$ is the set of all transformations of the form (6.3). $L_s L_t = L_{s+t}$ and $L_0 = I$, the identity transformation.

Similarly, if $T$ is in the subgroup $G_0$ of elements of $G$ that fix $U_0$, $T$ fixes the subspace spanned by $U_1$ and $U_2$. Since

$$q(x_1 U_1 + x_2 U_2) = x_1^2 + x_2^2, \quad (6.5)$$

this subspace with the restriction of $q$ is a Euclidean plane. By (6.2),

$$T = R_\theta J_j, \quad j = 0, 1, \quad (6.6)$$

where

$$[R_\theta] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}. \quad (6.7)$$

1993] HYPERBOLIC GEOMETRY ON A HYPERBOLOID 447

This content downloaded from 128.114.34.22 on Mon, 18 Apr 2016 14:39:25 UTC
All use subject to http://about.jstor.org/terms
thus $G_0$ consists of all $T$ of the form (6.6). Note that in $E^3$, identified with $M^3$ as in Section 4, $R_\theta$ is a rotation through $\theta$ radians. Clearly

$$G_1 \cap G_0 = \{I, R_\pi, J_1, J_2\}. \quad (6.8)$$

Applying $R_\theta$ to the parametrization (4.3) of $H^1$, define $P(r, \theta) = R_\theta(P(r))$. By (5.3),

$$P(r, \theta) = (\cosh r)U_0 + (\sinh r \cos \theta)U_1 + (\sinh r \sin \theta)U_2. \quad (6.9)$$

This is a parametrization of $H^2$. It is true that $r$ and $\theta$ are hyperbolic polar coordinates for $H^2$; the meaning of $r$ follows from (4.3) since $R_\theta$ is an isometry, but that of $\theta$ must wait until we define $H$-angle measure in the next section. By matrix multiplication, $L_s(P(r, 0)) = P(r + s, 0)$ and $R_\theta(P(r, \phi)) = P(r, \phi + \theta)$. Thus we now call $L_s$ the $H$-translation by $s$ along $H^1$ and we will call $R_\theta$ the $H$-rotation by $\theta$ about $U_0$.

At last we can obtain the basic formula for $H$-distance, namely:

$$d(A, B) = \arccosh(-p(A, B)). \quad (6.10)$$

To prove this, note that by (5.7) and (5.4) both sides are unchanged by applying any element of $G$. Applying $L_{-r}R_{-\theta}$ where $A = P(r, \theta)$, we can reduce to the case $A = U_0$; applying another $R_\phi$, we can further assume that $B = P(s, 0)$ with $s > 0$. By (2.6) and (4.3), $p(U_0, B) = -\cosh s$; since $s = d(U_0, B)$, this implies (6.10).

(6.10) may be compared with the fact that in spherical geometry the distance $d_s(A, B)$ between two points of $S^2$ is the radian measure of the Euclidean angle $\angle AOB$, which is $\arccos p(B, A)$.

Most of the above argument was devoted to the reduction to the case $A = U_0$, i.e. the proof of the basic fact that every point $A$ of $H^2$ can be moved to $U_0$ by an element of $G$; this means that the points of $H^2$ are “all alike” as far as $G$ is concerned. At the cost of using more linear algebra, we could have eliminated some calculations by proving this as follows. By a slight variant of the Gram-Schmidt orthogonalization process [11, pp. 356–357] together with Sylvester’s Theorem [11, p. 359], we could show that the coordinate column $[A]$ of $A$ (cf. (5.2)) is the first column of the matrix of some element $T$ of $G$; then $T^{-1}$ is the required element.

7. ANGLES. The $H$-angle $\angle BAC$ is $\overline{AB} \cup \overline{AC}$ where $\overline{AB} \neq \overline{AC}$. Define its ($H$-angle) measure $m(\angle BAC)$ as follows: let $V$ and $W$ be the vectors in $M^3$ tangent to $AB$ and $AC$ respectively at $A$ such that $q(V) = q(W) = 1$; then set

$$m(\angle BAC) = \arccos p(V, W). \quad (7.1)$$

To see that these vectors exist and are unique and that $|p(V, W)| \leq 1$, reduce to the case $A = U_0$ by transforming everything by a suitable element of $G$ as above; then the tangent plane to $H^2$ is $x_0 = 1$, which is Euclidean by (6.5), so that (7.1) makes sense. In fact, the ordinary law of cosines shows that the $H$-angle measure of $\angle BUC$ is the same as its Euclidean measure, so that in (6.9) $\theta$ really measures an $H$-angle. The angle measure between two intersecting curves in $H^2$ is defined similarly. Clearly every element of $G$ preserves angle measure just as it preserves distance. (A similar argument could have been used in Section 4 to show that $q(V)$ is positive.)

8. SUPERPOSITION. Now we have all we need to prove the following main existence and uniqueness theorem for elements of $G$. This theorem tells us exactly how far we can extend the fact that every point of $H^2$ can be moved to $U_0$ by some element of $G$. 

448 HYPERBOLIC GEOMETRY ON A HYPERBOLOID

[May]
Theorem. Let $I_j$ be an $H$-line, $A_1B_j$ a ray on $I_j$, and $S_j$ a side of $I_j$, for $j = 1, 2$. There exists exactly one $T \in G$ such that $T(A_1) = A_2$, $T(I_1) = I_2$, $T(A_1B_1) = A_2B_2$, and $T(S_1) = S_2$.

To see this, note that since $G$ is a group, we can assume, as in the proof of (6.10), that $A_1B_j$ is the ray of $H^1$ for which $x_j \geq 0$ for $j = 1, 2$. Then by (6.8) $T$ is $J_2$ or $I$, depending on whether or not it interchanges the sides of $H^1$; the theorem follows.

Let us call two subsets $K$ and $K'$ of $H^2$ congruent if $T(K) = K'$ for some $T \in G$. Then the theorem implies that two $H$-segments are congruent if and only if they have the same $H$-length, and that two $H$-angles are congruent if and only if they have the same measure. The axioms for angle-construction, angle-addition, and supplements [22, pp. 76–77] and the SAS-axiom [22, p. 84] can now be verified easily. This completes Moise’s list of axioms and shows that $H^2$ is indeed a model for plane hyperbolic geometry.

We could now establish that every isometry of $H^2$ is the restriction of some element of $G$, using the theorem together with the SSS-theorem [22, p. 87]. (Cf. the corresponding fact for $S^2$ and $O(E^3)$.) Thus two subsets of the hyperbolic plane are congruent if and only if one can be mapped on the other by “superposition”, i.e. by applying an isometry. This condition does not depend on the particular model of the hyperbolic plane being used.

9. TRIGONOMETRY. Let $\triangle ABC$ be any triangle in $H^2$; as usual in trigonometry, set $d(B, C) = a$, $d(C, A) = b$, $m(\angle BAC) = \alpha$, etc. Transforming by a suitable element of $G$ we can suppose that $C = U_0$, that $A$ is on the ray of $H^1$ for which $x_1 \geq 0$, and that $B$ is on the side of $H^1$ for which $x_2 > 0$; then $A = P(b, 0)$ and $B = P(a, \gamma)$. $L_{-b}$ carries $\triangle ABC$ to $\triangle U_0B'C'$ where $C' = = P(b, \pi)$ and $B' = P(c, \pi - \alpha)$. Together with (6.4) and (6.9), the equation $[B'] = [L_{-b}[B]$ (cf. (5.3)) yields

$$
\begin{bmatrix}
\cosh c \\
-\sinh c \cos \alpha \\
\sinh c \sin \alpha
\end{bmatrix} =
\begin{bmatrix}
\cosh b & -\sinh b & 0 \\
-\sinh b & \cosh b & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\cosh a \\
\sinh a \cos \gamma \\
\sinh a \sin \gamma
\end{bmatrix}.
$$

After expanding, the top entry gives the hyperbolic law of cosines and the bottom gives the law of sines. These imply the rest of trigonometry, including a formula for the angle of parallelism [6, Chapter VI], [8, Chapter 10].

10. EQUATIONS AND POLES OF LINES. Each $H$-line $l$ is the intersection of $H^2$ with a subspace of $M^3$ with equation

$$
p(V, X) = -v_0x_0 + v_1x_1 + v_2x_2 = 0
$$

for some nonzero $V = \sum v_iU_i \in M^3$. We can regard (10.1) as an equation of $l$; set $l = [V]$. It is not hard to see that $q(V) > 0$; dividing by $\sqrt{q(V)}$, we can assume that $V$ is in the subset

$$
D^2 = \{V \in M^3 | q(V) = -v_0^2 + v_1^2 + v_2^2 = 1\}
$$

of $M^3$, a hyperboloid of one sheet. Each line corresponds to exactly two points $\pm V$ of $D^2$; by the analogy with $S^2$, which is becoming a bit strained, we can call these the poles of the line. $D^2$, or rather the space of which it is a model, has been called the exterior-hyperbolic plane [3]; the symbol comes from the name of the cosmologist de Sitter, who studied its analogue $D^4 \subset M^5$ [9, pp. 124–131], [21, p. 745].
11. CYCLES. In analogy with the “small circles” of $S^2$, consider the curves $C$ on $H^2$ that are the intersections of $H^2$ with planes of $M^3$ that do not contain $O$; call such curves the cycles (or generalized circles) of $H^2$. Any three noncollinear points of $H^2$ are clearly contained in exactly one cycle. These planes have equations of form

$$p(V, X) = -v_0x_0 + v_1x_1 + v_2x_2 = -k$$  \hspace{1cm} (11.1)

for fixed $V \in M^3$ and real $k$, neither $V$ nor $k$ being zero (cf. (10.1)). There are three cases.

Case I. $q(V) < 0$. Dividing (11.1) by $\pm\sqrt{-q(V)}$ gives $V \in H^2$. There exists $T \in G$ with $T(V) = U_0$; the curve $C_k = T(C)$, congruent to $C$, has equation $p(U_0, X) = -k$, or $x_0 = k$. By (3.2), $k > 0$. $C_k$ has polar equation $r = \text{arccosh} \ k$ (cf. (6.9)), so we call $C$ the circle of $H^2$ with center $V$ and radius $s = \text{arccosh} \ k$. In $M^3$, $C$ is an ellipse or circle. The family of all circles of $H^2$ with center $V$ and the family of all (concurrent) $H$-lines through $V$ are orthogonal (cf. Section 7) trajectories, since this is true when $V = U_0$; in this situation, the rotations $R_\theta$ of (6.7) fix each of the circles and the family of lines. Since the plane $x_0 = k$ containing $C_k$ is Euclidean by (6.5), the $H$-arc length of $C_k$ (cf. Section 4) equals its circumference in $E^3$, identified with $M^3$. By (6.9), its radius in $E^3$ is sinh $s$, whence its circumference is $2\pi \sinh s$.

Case II. $q(V) > 0$. We can suppose that $V \in D^2$. Choose $T \in G$ such that $T(V) = U_2$. Then $T(C) = C_k$ has equation $x_2 = k$, $k \neq 0$, whence $C_k$ and $C$ are branches of hyperbolas in $M^3$. Consider the $H$-translations $L_t$ along $H^1$. For the point $P_s = (\cosh s)U_0 + (\sinh s)U_2$ of the $H$-line $x_2 = 0$ coordinatized by $H$-length,

$$L_t(P_s) = (\cosh s \cosh t)U_0 + (\cosh s \sinh t)U_2 + (\sinh s)U_2.$$  \hspace{1cm} (11.2)

Therefore $L_t$ maps $C_k$ on itself and $k = \pm \sinh s$ where $s$ is the perpendicular $H$-distance from any point $L_t(P_s)$ of $C_k$ to $H^1$. Accordingly each curve $x_2 = k$ (or $x_2 = \pm k$) is called an equidistant curve with axis $H^1$. The images under the mappings $L_t$ of the line $x_1 = 0$ form the family of “divergent” or “hyperparallel” $H$-lines perpendicular to $H^1$; this and the family of all curves $C_k$ are orthogonal trajectories. For fixed $s > 0$, the $H$-arc length of $C_k$ from $P_s$ to $L_t(P_s)$ is

$$\int_0^t \sqrt{a(L_t(P_s))} \, dt = t \cosh s, \quad t \text{ being the } H \text{-length of the segment of } H^1 \text{ from } U_0 \text{ to } L_t(U_0).$$

Case III. $q(V) = 0$. $V$ is on the cone $v_0^2 = v_1^2 + v_2^2$. We can suppose that $v_0 = 1$, so that $C$ has equation $-x_0 + (\cos \theta)x_1 + (\sin \theta)x_2 = -k$. Then $R_{-\theta}(C) = C_k$ has equation $-x_0 + x_1 = -k$. Since $C_k$ lies on $H^2$, $k > 0$. The curves $C$ are called horocycles; they are parabolas in $M^3$. A matrix calculation, as in Section 6, shows that the elements of $G$ that map any one of the curves $C_k$ on itself are precisely the transformations $N_tJ_z$ where

$$N_t = \begin{bmatrix} 1 + \frac{t^2}{2} & -\frac{t^2}{2} & t \\ \frac{t^2}{2} & 1 - \frac{t^2}{2} & t \\ t & -t & 1 \end{bmatrix}$$  \hspace{1cm} (11.3)
For fixed $s$ this parametrizes the horocycle $C_k$ for which $k = \cosh s - \sinh s = e^{-s}$. The $H$-arc length on $C_k$ between $L_s(U_0)$ and $N_t(L_s(U_0))$ is $|te^{-s}|$ by a very easy integration.

The $H$-line $N_t(H^1)$ has equation $tx_0 + tx_1 + x_2 = 0$ and is parametrized by $N_t(L_s(U_0))$ with $t$ fixed; these lines form a family of “parallel” (in the asymptotic sense) lines, the orthogonal trajectories of the curves $C_k$. The distance between $C_1$ and $C_k$ along any of these lines, e.g. $H^1$, is $|s|$. In this case there is an extra type of symmetry: $L_s$ maps $C_1$ on $C_k$, so that all horocycles are congruent. Accordingly, the $L_s$ as well as the $N_t$ fix the two orthogonal families; note that $N_tL_s = L_sN_te^{-s}$.

12. AREAS. It is easy to assign an area to each well-behaved region $R$ in $H^2$ in such a way that areas will be invariant under congruence and additive [22, p. 154]: define the $H$-area of $R$ to be any constant times the volume in $M^3$ of the solid consisting of all points of all segments in $M^3$ joining $O$ to points of $R$. By Section 8, congruence amounts to applying some $T \in G$; $T$ preserves volumes since it has determinant $\pm 1$ by Section 5, hence $T$ preserves $H$-areas and is clearly additive. (The natural choice of the constant is 3.) A similar argument interprets $H$-length in terms of the area of a sector of a hyperbola (this goes back to [14, p. 260]) and corresponds to a well-known interpretation of hyperbolic functions (see many calculus books and [4, p. 253].)

13. HYPERBOLIC CARTOGRAPHY. Despite the special place that the hyperboloidal model $H^2$ holds, it is good to have a variety of plane models, in their proper roles as map projections, to bring out different aspects of hyperbolic geometry. The analogy with map projections of the globe [2], [18] is not new, but it unifies many of the known models (see [6, pp. 136–137] and [8, Chapter 7] for example) and encourages the development of new ones. Here are a few examples, without proofs.

An important class consists of the “azimuthal” projections, in which the point $P(r, \theta)$ of $H^2$ (see (6.9)) is mapped on the point of the Euclidean plane with polar coordinates $(f(r), \theta)$ for some function $f$. (Terms used by cartographers for the sphere are in quotation marks.) The “gnomic” projection with $f(r) = \tanh r = (\sinh r / \cosh r)$ gives the projective or Beltrami-Klein model on the unit disk, which represents $H$-lines by line-segments. The “stereographic” projection with $f(r) = 2 \tanh(r/2)$ gives the Poincaré disk model, which represents $H$-lines by circles; this is conformal in the sense that it preserves angles between curves. These two models are among the best-known, and their connection with cartography has been pointed out by Coxeter [4, pp. 255 and 258] and Penrose [23, p. 113]. (Milnor [19] also discusses these models.) Some other choices are $f(r) = 2 \sinh(r/2)$, which preserves areas; $f(r) = \sinh r$, the “orthographic” projection, which projects $H^2$ perpendicularly on the Euclidean plane $x_0 = 0$ (see Gans [7]); and of course the “equidistant” projection $f(r) = r$, whose spherical counterpart appears on the seal of the United Nations.

An example that I have not seen worked out is the analogue of the Mercator projection, a famous conformal map of the sphere with “parallels of latitude” and “meridians” represented by orthogonal families of lines. (This analogue, however,
closely resembles the conformal map of $D^2$ given on [9, pp. 126–127]. With $H^1$ in the role of the "equator", this analogue maps the point $L_t(P_t)$ of (11.2) on the point with rectangular coordinates $(t, f(s))$, where $f$ is chosen so that the map is conformal. Since the shapes of infinitesimal regions are preserved, the formula for arc length on the equidistant curves corresponding to parallels of latitude implies that $f'(s) = 1/\cosh s$, whence $f(s) = \arctan(\sinh s)$, the so-called gudermannian function. Figure 2 shows the grid of this projection, with $H^1$ horizontal and with spacings $\frac{1}{4}$. The horizontal lines (extended infinitely on both sides) represent these equidistant curves and $H^1$; the vertical line-segments, of length $\pi$, represent the meridians, which are divergent $H$-lines. The top and bottom lines do not correspond to points of $H^2$. We cannot enliven the picture by drawing outlines of continents, but Figure 3 shows the image of a polar grid of points $P(r, \theta)$ ("transverse Mercator projection"). Figure 4, rotated 90 degrees, represents the image of a tessellation of $H^2$ by congruent triangles; this is the tessellation used, in the stereographic projection, by M. C. Escher in his picture "Circle Limit IV" [5].

In fact the above is only one of three analogues of the Mercator projection, each conformal with a family of cycles playing the role of parallels and with the orthogonal trajectory lines as meridians. While the projection based on circles seems of little interest, the one based on horocycles gives the Poincaré half-plane model with the horocycles represented by parallel lines and their orthogonal trajectories by half-lines. This model, usually taken on the upper half of the...
complex number plane, is extremely important because of its wide-ranging applications (e.g. [16], [17]).

14. ON THE HISTORY OF THE MODEL. The high points of the prehistory of the model $H^2$ are Lambert’s idea of an “imaginary sphere” in 1766 [1, p. 50] and Taurinus’ calculation of trigonometry on a “sphere of imaginary radius” in 1826 [1, p. 79]. The name “hyperbolic geometry”, introduced by Klein in 1871, does not refer to this model; see [31, p. 63].

Clear statements of $H^2$ seem to have originated twice. According to Killing’s 1885 book [14, pp. 258-259], Weierstrass communicated the coordinates $x_i$ at a seminar in 1872 and “supplied numerous applications”. Accordingly Killing named them “Weierstrass coordinates” and used them extensively. (Killing remarked that Beltrami had previously used some closely related imaginary coordinates, but only in one proof.) There is a distinction, however, between using the coordinates and using the points of a hyperboloid as a model. Killing made it clear that he was doing the latter in a one-paragraph presentation [14, p. 260] of $H^2$ as an “Abbildung” (mapping, image, picture) of the hyperbolic plane in $E^3$, which included a description of $H$-lines and cycles. In 1880, he had described [13, p. 275] the exterior-hyperbolic plane as an ideal region of the hyperbolic plane; although he used Weierstrass coordinates there, he mapped the hyperbolic plane on a hemisphere instead of a hyperboloid [13, pp. 284-287]. See Hawkins [10] for a real historian’s treatment of Killing.

Meanwhile, in 1881, Poincaré [24] took quantities corresponding to our $x_i$ from an 1854 paper of Hermite on quadratic forms, named them “hyperbolic coordi-
nates" without mentioning a hyperboloid, and related them to the projective disk model. (Some unpublished work of Poincaré in 1880 related to our model is discussed by Gray [7a, pp. 271–272].) In 1887, he defined [25] "quadratic geometries" on arbitrary quadric surfaces in 3-space, with distances and angles defined in terms of cross-ratios. These included spherical, hyperbolic (a form of our model), exterior-hyperbolic, and Euclidean geometries, the last on a paraboloid!

The latest part of our model to be developed was the connection with the Minkowski structure of the embedding space. Naturally enough, this seems to have followed the origin of special relativity. Some early references are Minkowski [20, p. 376] (a posthumous paper based on a 1907 lecture), Sommerfeld [28], and Varičak [30]. In 1909, Jansen [12] gave what appears to be the first detailed exposition of $H^2$, referring to Poincaré and Minkowski; he derived most properties by translating them from the half-plane model.

The "imaginary radius" idea survived, mixed with $H^2$, in Klein’s classic posthumous book [15, p. 193] and more recently in [26].

REFERENCES

3. H. S. M. Coxeter, A geometrical background for de Sitter’s world, this MONTHLY, 50 (1943) 217–228.
17. B. Mazur, Number theory as gadfly, this MONTHLY, 98 (1991) 593–610.

454  HYPERBOLIC GEOMETRY ON A HYPERBOLOID [May


Department of Mathematics
Tufts University
Medford, MA 02155

E-mail: reynolds@jade.tufts.edu

---

**Lester R. Ford Award**

The Lester R. Ford Award for an expository article published in the *Monthly* in 1991 was presented at the San Antonio joint meetings to

C. W. H. Lam