Chapter 7

Implications of consistency of hyperbolic geometry

1. There are models within $\mathbf{R}^3$ or even $\mathbf{R}^2$ that satisfy the postulates of hyperbolic geometry (Hilbert IBC + HH (p. 259) + Dedekind). (We get to details around pp. 329–330.) The HH hypothesis is actually redundant when we assume Dedekind — see p. 262. The existence of these models shows that hyperbolic geometry is consistent if our theory of $\mathbf{R}^n$ is (the latter needing the real numbers and hence some level of set theory). That is, elementary linear algebra establishes the consistency of hyperbolic geometry just as surely as that of Euclidean geometry — which it does, because $\mathbf{R}^2$ itself is a model of (Hilbert) Euclidean geometry (pp. 139–140 and the Chapter 3 projects), and a similar statement can be made about $\mathbf{R}^3$. In fact, the hyperbolic models can be developed within axiomatic Euclidean geometry, so we don't really need the consistency of the real numbers to reach the conclusion, just consistency of Euclidean geometry.

An ironic consequence of the foregoing is that HE/EV *can't be proved* within the Hilbert axioms + Dedekind (or your favorite continuity axiom), unless Euclidean geometry itself is *inconsistent*. In other words, if Saccheri et al. had *succeeded* in "vindicating" Euclid by proving EV, they would have destroyed Euclidean geometry by proving it inconsistent.

Note that in these models the primitives (especially "straight line" and "congruent") are reinterpreted to mean something rather different than they mean in the embodying Euclidean spaces. This may leave the impression that the "true" geometry of a plane really is Euclidean, and hyperbolic geometry describes something different, so why the big deal? If you *define* Euclidean geometry as the geometry of the vector space $\mathbf{R}^2$ with its standard inner product, no one can argue with you.

2. On the other hand, now that we know that HE/EV is not inevitable, or built into the framework of the human mind so that denying it is inconceivable, as so many mathematicians and philosophers used to say, whether the geometry of *physical space* is Euclidean becomes an experimental question. Logically, it is entirely possible that space is hyperbolic, this time with the primitives meaning (almost) exactly what they always have in Euclidean thinking. In fact, in modern physical theory (general relativity, cosmology) space is something even more general, not homogeneous (i.e., may be different in different regions) and possibly changing in time. It can be "bumpy" as well as curved (see the beginning of Chap. 8 and the end of Appendix A). In other words, nature is not *exactly* described by the $\mathbf{R}^3$ geometry given by linear algebra. On the small scale, planets etc. create bumps in the geometry, and the influence of the bumps on moving bodies constitutes the gravitational force of the planets. Gravity = nonhomogeneous geometry. On the large scale, however, current observational evidence indicates that the average curvature

of the observable universe is very close to flat; but there is no good theoretical reason why that must be true, and elliptic (usually spherical) and hyperbolic 3-spaces are frequently studied by serious physicists. (Because the universe is expanding, the geometry of four-dimensional space-*time* is not flat.)

Beltrami's four models of hyperbolic geometry

1. The pseudosphere or tractrix model (postpone)
2. Klein's disk model
3. Poincaré's disk model
4. Poincaré's upper half plane model

And for completeness, recall

5. Hyperboloid model.

Numbers 1 and 5 in this list are surfaces in a higher-dimensional space. The other three are *maps* of the hyperbolic space onto a part of the flat 2-dimensional plane.

The Klein disk — introduction

In this model the lines are the line segments across the disk (chords). We have already encountered this model in Ex. 2.9(c). There we considered incidence and parallelism only. Betweenness is also elementary, but congruence is nonstandard (i.e., segments of equal length don't appear to be of equal length to our Euclidean eyes, and similarly for angles); it turns out to be most convenient to delay the discussion of congruence until after studying the Poincaré disk. There are many parallel chords to a given chord through a point, so this is a promising model of hyperbolic geometry.

If A and B are points on the bounding circle, the chord between them is denoted A)(B. A and B are *ideal points* or *points at infinity* — not themselves points in the hyperbolic plane. The chord connecting them is called *line of enclosure* of any angle leading to them, hence fulfilling the Warning on p. 115 (see Fig. 7.5, p. 300). The extended sides of the angle are the limiting parallel rays to A)(B from the vertex of that angle.

The proof of I-1 requires line-circle continuity. For purposes of this course we might as well assume that we are in the real Euclidean plane, so that Dedekind's axiom holds (for lines and for chords). The rest of I and B are left for an exercise. The C axioms will be proved indirectly after we study the Poincaré model.

In the Klein model lines are straight, but lengths and angles are nonstandard. In the Poincaré model lines are curved, lengths are nonstandard, but angles are the same as in the Euclidean geometry of the plane containing the disk; such a model is called *conformal*. (A Mercator map has similar properties for a spherical geometry.)

Let us start with the known metric of the hyperboloid,

$$ds^2 = R^2(dr^2 + \sinh^2 r\, d\theta^2) \qquad \text{(take } R = 1\text{)},$$

and perform a rescaling of the radial coordinate (i.e., $r =$ function of ρ and vice versa) to map the hyperboloid to the interior of a disk in such a way that the result is conformal:

$$ds^2 = f(\rho)^2(dx^2 + dy^2) \equiv f(\rho)^2(d\rho^2 + \rho^2 d\theta^2).$$

That is, the coordinate system looks Cartesian at each point but the *scale* can vary from point to point. Remember that such formulas are shorthand for integral formulas for arc length:

$$s = \int_{\tau_0}^{\tau_f} f(\rho(\tau)) \sqrt{\left(\frac{d\rho}{d\tau}\right)^2 + \rho(\tau)^2 \left(\frac{d\theta}{d\tau}\right)^2}\, d\tau.$$

They also tell us how to calculate angles. The cosine of the angle between two vectors $(\Delta x_1, \Delta y_1)$ and $(\Delta x_2, \Delta y_2)$ is, as in ordinary vector calculus, the ratio of the dot product of the vectors as defined by the the metric ds^2 divided by the product of the lengths of the vectors:

$$\frac{f(\rho)^2(\Delta x_1 \Delta x_2 + \Delta y_1 \Delta y_2)}{\sqrt{f(\rho)^2({\Delta x_1}^2 + {\Delta y_1}^2)}\sqrt{f(\rho)^2({\Delta x_2}^2 + {\Delta y_2}^2)}},$$

Since the scale factor cancels out, this angle is the same as in the Euclidean geometry, where $f(\rho) = 1$. This justifies the earlier claim that conformality in the sense of a scale factor implies conformality in the sense of preserving angles.

Comparing the two formulas for ds^2, we see that we must have

$$\sinh^2 r = f(\rho)^2 \rho^2 \quad \text{and} \quad f(\rho)^2 = \frac{dr^2}{d\rho^2} \equiv \left(\frac{dr}{d\rho}\right)^2.$$

Thus

$$\frac{dr}{d\rho} = \pm f(\rho) = \pm\frac{\sinh r}{\rho},$$

or

$$\frac{dr}{\sinh r} = \pm\frac{d\rho}{\rho},$$

or

$$\ln \tanh\left(\frac{r}{2}\right) + c = \pm \ln \rho.$$

Thus either ρ or $1/\rho$ equals $C\tanh(r/2)$. We want $\rho = 0$ when $r = 0$ and $\rho = 1$ when $r \to \infty$, so we choose

$$\rho = \tanh\left(\frac{r}{2}\right).$$

It follows that

$$r = 2\tanh^{-1}\rho = \ln\left(\frac{1+\rho}{1-\rho}\right).$$

Finally,

$$f(\rho) = \frac{\sinh \ln\left(\frac{1+\rho}{1-\rho}\right)}{\rho} = \frac{\frac{1+\rho}{1-\rho} - \frac{1-\rho}{1+\rho}}{2\rho} = \frac{1 + 2\rho + \rho^2 - (1 - 2\rho + \rho^2)}{2\rho(1-\rho^2)} = \frac{2}{1-\rho^2}.$$

Therefore,

$$ds^2 = \frac{4}{(1-\rho^2)^2}(d\rho^2 + \rho^2 d\theta^2).$$

It is standard to drop the factor 4, which amounts to starting with $R = \frac{1}{2}$ instead of 1. Thus we get the standard metric for the Poincaré disk,

$$ds^2 = \frac{dx^2 + dy^2}{[1 - (x^2 + y^2)]^2}.$$

(This equation appears in the middle of p. 565 in a complex-variable notation.)

Lines and lengths in the Poincaré model

The ratio between hyperbolic length and Euclidean length rapidly diminishes as one approaches the bounding circle, γ. We have the situation indicated in the Escher drawing on the cover of the book, with infinite crowding near the edge. (Recall: the angles equal the Euclidean angles.)

Although I can't prove it by elementary means, the geodesics in this geometry are *circular arcs orthogonal to γ at both ends*; these include straight lines that are diameters of the circle. (Later I will prove the corresponding theorem for the Poincaré half-plane model.) There result the Figures on pp. 304–305 of limiting rays, divergent parallels, and Lambert and Saccheri quadrilaterals. Two other Eschers (nos. 22 and 24 in *The Graphic Work of M. C. Escher*) show the circles (lines) better than no. 23.

A geometrical development of the notions of length and congruence in the Poincaré model occupies the (hard!) second half of Chapter 7. The upshot is that the Hilbert congruence axioms can be verified for this model. We shall also see that the Poincaré model can be mapped onto the Klein model or vice versa, so the C axioms transfer from Poincaré to Klein while the I and B axioms transfer from Klein to Poincaré. This will complete the proof that both disk systems are indeed *models* of hyperbolic geometry (as axiomatized by Hilbert). Therefore, hyperbolic geometry is consistent if Euclidean geometry is.

The Poincaré distance is defined on pp. 319–320: Let A and B be points in the interior; let P and Q be the endpoints (on γ) of the corresponding Poincaré line (Fig 7.27). For any pair of these points, define $\overline{\mathrm{AB}}$, etc., to be the ordinary Euclidean distance between them (measured along a Euclidean straight line, not a Poincaré line). For the moment consider the case Q * A * B * P, Then the *cross-ratio*

$$(\mathrm{AB},\mathrm{PQ}) \equiv \frac{\overline{\mathrm{AP}}\,\overline{\mathrm{BQ}}}{\overline{\mathrm{BP}}\,\overline{\mathrm{AQ}}}$$

is greater than 1. The Poincaré distance from A to B is then defined to be the logarithm of (AB,PQ). If we interchange A with B, or P with Q, the cross-ratio is inverted, so it's now less than 1 and its logarithm is the negative of the previous case. So

$$d(\mathrm{AB}) \equiv |\ln(\mathrm{AB},\mathrm{PQ})|$$

is a formula that applies equally in the case P * A * B * Q. A short calculation shows that if A * C * B, then $d(\mathrm{AC}) + d(\mathrm{CB}) = d(\mathrm{AB})$, so distance is additive along a line, as required by Axiom C-3.

Lemma 7.4 on p. 328 shows that if O is the center and $d(\mathrm{OB}) = r$, then r is related to the Euclidean length of OB by

$$\rho \equiv \overline{\mathrm{OB}} = \frac{e^r - 1}{e^r + 1} = \tanh\left(\frac{r}{2}\right) .$$

(Here I take the radius of γ to be 1; Greenberg calls the radius r and the distance d.) The inverse of this formula is

$$r = \ln\frac{1+\rho}{1-\rho} ,$$

a formula we got earlier for the coordinate transformation. This confirms that the two definitions of Poincaré distance are the same, at least for lines through the center. (For the general case, see the paper by Zahar cited below.) In our coordinate calculation, ρ was the radial coordinate in the underlying Euclidean disk, hence indeed equal to the Euclidean length of OB, whereas r was the radial coordinate in the hyperboloid representation,

$$ds^2 = R^2(dr^2 + \sinh^2 r\, d\theta^2),$$

from which it's clear that r is the hyperbolic length of OB. [What happened to the factor $R^2 = \frac{1}{4}$?]

For segments without an endpoint at the center, the cross-ratio gives a coordinate-independent (albeit inscrutable) definition of the length, while any of our formulas for ds^2 gives a coordinate-dependent formula for the same length as an integral along the Poincaré line (circular arc as a geodesic curve). Obviously, arc lengths defined by such integrals are additive:

$$\int_A^B ds = \int_A^C ds + \int_C^B ds.$$

Now consider the six congruence axioms in the Poincaré model. C-5 and C-4 are immediate, because angles are equal to Euclidean angles. (For C-4, Greenberg p. 319 gives a detailed constructive justification of the existence and uniqueness of the geodesic with a given direction at a given point.) Similarly, C-2, C-3, and C-1 are immediate from the definition of length: "Having the same numerical length" is obviously an equivalence relation for segments (C-2). We have already shown additivity (C-3) two ways. C-1 requires us to note that the Poincaré lines are infinite in extent (as measured by arc length), which is obvious from the hyperboloid representation. That leaves C-6, SAS, which is much harder, but can be proved (pp. 327–328) by reducing to a triangle with one vertex at the origin and using the formula relating r to ρ. This also completes the proof that Poincaré length is the same thing as hyperbolic arc length, even for lines that don't go through the origin.

Isomorphisms between the Klein and Poincaré models

My favorite is on pp. 334–335. It shows that a Klein line can be regarded as a Poincaré line just straighened out! The precise construction involves projecting from north pole to southern hemisphere and then vertically back to the equatorial plane.

The alternative construction is on p. 306. It is the same, except that the plane involved is tangent to the south pole instead of through the equator. This makes the Poincaré disk twice as big. [Hmm. Maybe that has something to do with the missing factor of $4 = 2^2$ noticed earlier.] Now projecting from a point on the equator (instead of north pole) yields the next model:

Poincaré's half-plane model

There is a conformal (angle-preserving) mapping of a half plane onto the unit disk (or vice versa). Such mappings are most easily described in terms of complex variables. Let $z = x + iy$, $w = u + iv$. Then

$$z = \frac{w-1}{w+1}$$

maps the right half plane onto the unit disk: If $u = 0$ (i.e., w is on the imaginary axis), then $z = (iv - 1)/(iv + 1)$ which is easily seen to lie on the unit circle. If $u > 0$, then z has absolute value less than 1 — it is inside the disk. Note that

$$w = 0 \;\Rightarrow\; z = -1, \quad w = 1 \;\Rightarrow\; z = 0, \quad w \to \infty \;\Rightarrow\; z \to 1.$$

It is well known (to students in a complex-variables course) that such a fractional-linear (Möbius) transformation is conformal.

In real terms,

$$x + iy = \frac{u - 1 + iv}{u + 1 + iv} = \frac{[u - 1 + iv][u + 1 - iv]}{(u + 1)^2 + v^2} = \frac{u^2 + v^2 - 1 + 2iv}{(u + 1)^2 + v^2},$$

so

$$x = \frac{u^2 + v^2 - 1}{u^2 + v^2 + 2u + 1}, \qquad y = \frac{2v}{u^2 + v^2 + 2u + 1}.$$

Now we can convert

$$ds^2 = \frac{dx^2 + dy^2}{[1 - (x^2 + y^2)]^2}$$

to (after a long calculation)

$$ds^2 = \frac{du^2 + dv^2}{u^2}.$$

The standard convention is to rotate from RHP to UHP, and to rename the variables as x and y:

$$ds^2 = \frac{dx^2 + dy^2}{y^2}.$$

In this new representation (related to the old one somewhat as a Mercator map is related to a polar one), the geodesics are the semicircles with ends orthogonal to the horizontal axis, including the vertical lines as a special case.

The metric (arc length formula) for the Poincaré half plane is algebraically simpler than the one for the Poincaré disk, and its circular "lines" are easier to think about because they are all semicircles.

In this model it is relatively easy to verify that the hyperbolic lines are geodesics — that is, they are stationary points (in fact, local minima) of the arc length with respect to small, local variations.

First let's write down the equations of these lines in analytical geometry. The generic kind is a [semi]circle with center $x = c$ on the horizontal axis and radius a:

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

The vertical lines are a special case: $x =$ constant. To get these as limits of the semicircles, rewrite the latter as

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

divide by $a+c$, and take c and a to infinity with $c-a$ fixed.

Now return to the arc length, and suppose that the curve is parametrized by giving x as a function of y:

$$s = \int ds = \int \frac{1}{y}\sqrt{\left(\frac{dx}{dy}\right)^2 + 1}\, dy.$$

(The integral need not be taken over the entire curve, just over a segment that contains the entire interval where the small variation takes place. There is a slight gap in our logic at the top of each circle, where the curve has a horizontal tangent, but that can be taken care of.)

Suppose we replace $x(y)$ by $x(y) + \delta x(y)$, where $\delta x(y) = 0$ outside a small interval and $|\delta x(y)|$ is always very small. How does s change? Well, $\dot{x} \equiv \frac{dx}{dy}$ changes to $\dot{x} + \delta\dot{x} = \dot{x} + d[\delta x]/dy$. And so the first-order change in $\sqrt{\dot{x}^2+1}$ is

$$\frac{d}{d\dot{x}}\sqrt{\dot{x}^2+1}\,\delta\dot{x} = \frac{\dot{x}}{\sqrt{\dot{x}^2+1}}\,\frac{d\delta x}{dy}\,.$$

So to first order, the change in s is

$$\delta s = \int \frac{1}{y}\,\frac{\dot{x}}{\sqrt{\dot{x}^2+1}}\,\frac{d\delta x}{dy}\, dy.$$

[What we have done here is an infinite-dimensional generalization of the third-semester calculus formula

$$dF = \frac{\partial F}{\partial x}\,dx + \frac{\partial F}{\partial y}\,dy + \frac{\partial F}{\partial z}\,dz = \sum_{j=1}^{3} \frac{\partial F}{\partial x_j}\,dx_j\,.$$

The analogy is $j \mapsto y$, $x_j \mapsto x(y)$, sum $\mapsto$ integral.]

Now the integral can be evaluated by parts; there is no contribution from the endpoints because we assumed that the variation is confined to an interval interior to the integral of integration.

$$\delta s = -\int \frac{d}{dy}\left[\frac{1}{y}\,\frac{\dot{x}}{\sqrt{\dot{x}^2+1}}\right]\delta x(y)\, dy.$$

But δx is arbitrary; if this quantity is to be 0 for *all* variations, then the derivative in the integrand must be identically zero. Thus the function equals a constant:

$$\frac{1}{y}\frac{\dot{x}}{\sqrt{\dot{x}^2+1}} = \frac{1}{a} \quad \text{[or 0]}.$$

With a few steps of algebra you can solve this equation for $\dot{x}$:

$$\dot{x} = \pm\frac{y}{\sqrt{a^2-y^2}}\,.$$

Integrate to get

$$x = \mp\sqrt{a^2-y^2}+c,$$

or $(x-c)^2+y^2=a^2$, as predicted. If the first constant of integration was 0, the equation for $\dot{x}$ was

$$\frac{\dot{x}}{\sqrt{\dot{x}^2+1}} = 0,$$

which is satisfied by (precisely) the vertical lines, $x=c$.

The same kind of argument works for the geodesics in the Poincaré disk model, but the algebra is harder. See E. G. Zahar, *Stud. Hist. Phil. Mod. Phys.* **28** (1997) 183–218.

Local curvature and the pseudosphere model

Geometry	*Curvature*	*Triangle sum*	*Circumference*
Euclidean	zero (flat)	π (no defect)	$2\pi R$
Ellip./sph.	positive	$>\pi$ (neg. def.)	$<2\pi R$
Hyperbolic	negative	$<\pi$ (pos. def.)	$>2\pi R$

[sketches]

Explanation of the *circumference* column: In this discussion R is the radius of a circle, and K (formerly R) will be the "radius" or characteristic length of the whole space. In the case of the Euclidean metric, K is an arbitrary scale factor introduced to maximize the connection with the curved cases. Now we do some trivial calculations:

Euclidean circles: $ds^2 = K^2(dx^2+dy^2) = K^2(dr^2+r^2d\theta^2)$.
Distance from O to r is $\int_0^r K\,d\tilde{r} = Kr \equiv R$.
Circumference at r is $\int_0^{2\pi} Kr\,d\theta = 2\pi Kr$.
Ratio of circ. to radius is 2π (independent of K and of R!).

Spherical circles: $ds^2 = K^2(dr^2 + \sin^2 r\, d\theta^2)$.
Radius from O to r is $\int_0^r K\, d\tilde{r} = Kr \equiv R$.
Circumference at r is $\int_0^{2\pi} K \sin r\, d\theta = 2\pi K \sin r$.
Ratio is $2\pi\, \dfrac{K \sin r}{Kr} = 2\pi\, \dfrac{\sin(R/K)}{R/K} < 2\pi$.

Hyperboloidal circles: $ds^2 = K^2(dr^2 + \sinh^2 r\, d\theta^2)$.
Radius from O to r is $\int_0^r K\, d\tilde{r} = Kr \equiv R$.
Circ. at r is $\int_0^{2\pi} K \sinh r\, d\theta = 2\pi K \sinh r$.
Ratio is $2\pi\, \dfrac{K \sinh r}{Kr} = 2\pi\, \dfrac{\sinh(R/K)}{R/K} > 2\pi$.

Moreover, the ratios increasingly deviate from the flat one as R/K grows, whereas the flat limit is approached as R/K becomes small. The *curvature* is defined so that it turns out to be [proportional to] $1/K^2$. (More details in a moment.)

By considering *small* circles we can define the curvature *at each point* of a two-dimensional space. (That is, given R we can find K from the formulas above. For a general space, K might differ from point to point.) In more detail, if $r \equiv R/K$ is small we can approximate

$$\sin r \approx r - \frac{1}{6}\, r^3, \qquad \sinh r \approx r + \frac{1}{6}\, r^3,$$

and pick off K from the third-order term.

Similarly, one can find the curvature from looking at the defects of small triangles. I omit details.

Now consider a surface embedded in three-dimensional Euclidean space, with the distance function *induced* by the 3D geometry (i.e., distances measured by a tape measure along the shortest path on the surface). It has at each point a curvature defined as we just discussed; this number is called its *intrinsic curvature*. The curvature is determined by how the surface is embedded into 3-space, but *not conversely*; for instance, a cylinder is flat, as far as small circles and triangles are concerned, although it has an *extrinsic curvature* that describes how it is rolled up. For more detail see the end of Appendix A, or Part I of *Differential Geometry and Relativity Theory* by R. L. Faber, or *Elementary Differential Geometry* by Barrett O'Neill. Briefly, the extrinsic curvature is characterized by two numbers associated with the curvatures (in fact, reciprocals of the "radii of curvature") of two orthogonal curves through the point; the intrinsic curvature is their product. (These numbers are the eigenvalues of a particular 2×2 matrix associated with the point, and the intrinsic curvature is its determinant.) For the cylinder one of these *principal curvatures* is 0, so the intrinsic curvature is, too. For a sphere the principal curvatures are equal (to $1/K$), so the intrinsic curvature is positive; but the same product could arise from unequal curvatures of the same sign (think of a football).

Now the big point: Are there surfaces of *constant, negative* curvature embedded in Euclidean $\mathbf{R}^3$? Yes, as mentioned on pp. 295–296, Beltrami's *pseudosphere* is a trumpet-shaped surface obtained by revolving a curve called a *tractrix* about its axis. (The equation of the tractrix is

$$y = a \ln \frac{a + \sqrt{a^2 - x^2}}{x} - \sqrt{a^2 - x^2}\,.$$

See the Wikipedia article "Tractrix" for details.) Notice that the two dimensions of the trumpet "curve in opposite directions", so the principal curvatures are of opposite sign. As you go up the tube of the trumpet, one of the principal curvatures shrinks but the other one grows in an exactly compensating manner, so that the intrinsic curvature stays constant.

The pseudosphere locally models hyperbolic geometry in the same way that the cylinder locally models Euclidean geometry. Even when "unrolled", it does not cover the whole hyperbolic space. See Greenberg, Figs. 10.11 and 10.12, pp. 483–487 (in Chapter 10).

It is important to understand the differences between the pseudosphere model and the hyperboloid model. The hyperboloid corresponds to the *entire* hyperbolic space constructed axiomatically by Bolyai, Lobachevsky, Gauss, Beltrami, Klein, Poincaré, and Hilbert. But it is embedded in the 3-dimensional space-time with indefinite metric, so lengths, angles, and curvature upon it are not accurate when looked at in 3-dimensional *Euclidean* terms. The pseudosphere model is merely *local*: it represents only a fragment of the hyperbolic space (but a completely representative one, since all points are geometrically equivalent). But its ambient 3-dimensional space is Euclidean, so a piece of sheet metal, say, in the shape of a pseudosphere does have exactly the local geometry of a hyperbolic space.

Axiomatization of elliptic geometry

From the point of view of Riemannian manifolds of constant curvature, the sphere is just as valid as the hyperboloid (in fact, more elementary and visualizable). Our axiomatic development excluded spherical/elliptic geometry at a surprisingly early stage. Recall that the AIA theorem is inconsistent with "elliptical parallelism" and an examination of its proof revealed that it depended on Axioms I-1 and B-4. Spherical geometry violates I-1, and its cured version, elliptic geometry, violates B-4. Both versions violate B-3, as well.

In Appendix A Greenberg states *separation axioms* to replace the betweenness axioms that become untenable in elliptic geometry. He does not make clear whether these axioms can also be used to formulate Euclidean and hyperbolic geometry, thereby creating a unified treatment. (I don't think so.)

He states that the congruence and continuity axioms "all make sense when rephrased," but this is far from clear for continuity. Aristotle's axiom is blatantly inconsistent with elliptic geometry. (Its proof from the more powerful continuity axioms requires the Hilbertian betweenness concept – see pp. 135–136.) Dedekind's axiom as stated in the book (p. 134) does not carry over without repair. One has to say that a *segment* (defined on p. 543) has the topological structure of an interval of the real line, and the details would require defining a *ray* — or some replacement for that concept — without using betweenness. Better references would have been appreciated.

A good thing in this appendix is the demonstration on pp. 546–547 of why the (negative) angle defect of a spherical triangle is proportional to the area.

As the example spherical $\rightarrow$ elliptic illustrates, spaces can have the same local geometry and different global structures. In the flat case we can have cylinders and tori as well as the whole plane — not to mention Möbius strips and Klein bottles. The analogs of tori in hyperbolic space are even more numerous and are still the subject of research. References:

1. Janna Levin, "Topology and the cosmic microwave background", Physics Reports **365** (2002) 271–333.

2. Jeffrey Weeks, *The Shape of Space*.

3. David Borthwick, "Spectral theory on hyperbolic surfaces", `http://www.math.dartmouth.edu/~specgeom/Borthwick_slides.pdf`.