Other statements of Axiom P_1 are less obviously equivalent to it. For example,

(i) The angle sum of a triangle $= \pi$ (Euclid).

(ii) The locus of points equidistant from a straight line is a straight line. (al-Haytham, around 1000 CE).

(iii) Similar triangles of different sizes exist [Wallis (1663); see Fauvel and Gray (1988), p. 510].

Thus a denial of the parallel axiom entails denial of (i), (ii), and (iii). A denial of (iii) means in particular that scale models are impossible, since three points in the original object and the three corresponding points of a scale model would define similar triangles of different sizes.

Such unlikely consequences convinced many people that the parallel axiom was a logically necessary property of straight lines, already implied by the other axioms of Euclid, and so efforts were made to prove it outright.

The most tenacious attempt, entitled *Euclides ab omni naevo vindicatus* (Euclid cleared of every flaw), was made by Saccheri (1733). Saccheri's plan of attack began by subdividing the denial of the parallel axiom into two alternatives:

Axiom P_0. There is no line through P that does not meet L.

Axiom P_2. There are at least two lines through P that do not meet L.

The next step was to destroy each alternative by deducing a contradiction from it. He succeeded in deducing a contradiction from Axiom P_0, using other axioms of Euclid, such as the axiom that a straight line can be prolonged indefinitely. (Such additional assumptions are certainly necessary, since great circles on the sphere have some properties of straight lines, except that they are finite in length.)

Saccheri was less successful with Axiom P_2. The consequences he derived from it, hoping to obtain a contradiction, were as follows. Among the lines M through P that do not meet L are two extremes, M^+ or M^-, called *parallels* or *asymptotic lines* (Figure 18.1); any of these lines M strictly between M^+ and M^- has a common perpendicular with L and, moreover, the position of this perpendicular tends to infinity as M tends to M^+ or M^-. Although curious, these consequences of Axiom P_2 were not contradictory and Saccheri, sensing that the contradiction was slipping away from him, tried to overtake it by proceeding to infinity.

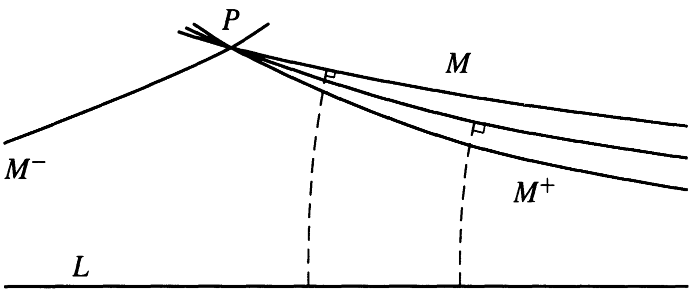

Figure 18.1: Asymptotic lines

He claimed that the asymptotic line M^+ would meet L at infinity and have a common perpendicular with it there. This was perhaps plausible, given similar arguments in projective geometry, though Euclid certainly would not have accepted it. But it *still* was not contradiction. Saccheri merely claimed that such a conclusion was "repugnant to the nature of the straight line" [Saccheri (1733), p. 173], perhaps visualizing an intersection like Figure 18.2. But why should asymptotic lines not be tangential at infinity? History was to show that this was an appropriate resolution of Saccheri's "contradiction" (see Section 18.5). Thus Saccheri's results were not, as he thought, steps toward a proof of the parallel axiom; they were the first theorems of a *noneuclidean* geometry in which Axiom P_2 replaces the parallel axiom.

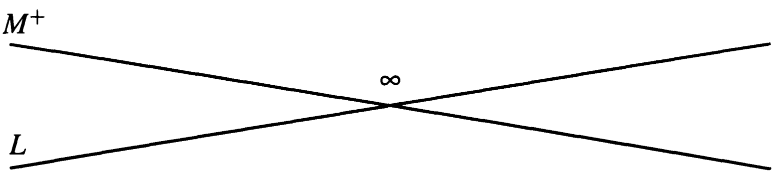

Figure 18.2: Hypothetical intersection at infinity

EXERCISES

The connection between the parallel axiom and the angle sum of a triangle is very direct and elegant.

18.1.1 Deduce, from Euclid's version of the parallel axiom, that a line falling on two *parallel* lines makes the interior angles sum to π.

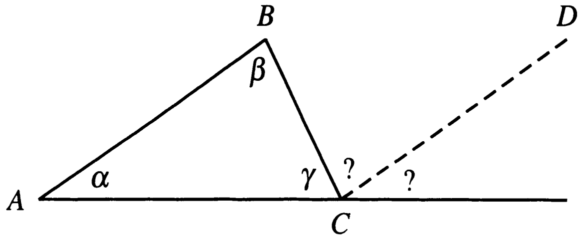

Figure 18.3: The angle sum of a triangle

18.1.2 Use Exercise 18.1.1 and the construction in Figure 18.3 (in which CD is parallel to AB) to show that $\alpha+\beta+\gamma=\pi$.

18.1.3 Deduce from Exercise 18.1.2 that the angle sum of any quadrilateral is 2π and, in particular, that squares exist.

Thus theorems mentioning squares, such as Pythagoras' theorem, can only hold when Euclid's parallel axiom is assumed.

18.2 Spherical Geometry

In rejecting P_0 because of its incompatibility with infinite lines, Saccheri avoided having to consider the most natural geometry in which P_0 holds, that of the sphere with great circles as "lines." Spherical geometry had been cultivated since ancient times to meet the needs of astronomers and navigators, and formulas for the side lengths and areas of spherical triangles were well known. But the sphere was considered part of Euclid's spatial geometry, so the axiomatic significance of spherical geometry was at first ignored. What did happen, though, was that the first explorations of Axiom P_2 were guided by the analogy of the sphere.

Lambert (1766) made the striking discovery that Axiom P_2 implies that the area of a triangle with angles α,β,γ is proportional to $\pi-(\alpha+\beta+\gamma)$, its angular defect. In other words,

$$\text{area} = -R^2(\alpha+\beta+\gamma-\pi)$$

for some positive constant R^2. Having rediscovered Harriot's theorem that

$$\text{area} = R^2(\alpha+\beta+\gamma-\pi)$$

for a triangle on the sphere of radius R, Lambert mused that one "could almost conclude that the new geometry would be true on a sphere of imaginary radius." What a sphere of radius iR might be was never explained, but the idea of using complex numbers to generate the formulas of a hypothetical geometry proved fruitful.

It was found that formulas derived from Axiom P_2 could also be obtained by replacing R by iR in corresponding formulas of spherical geometry. For example, Gauss (1831) deduced from Axiom P_2 that the circumference of a circle of radius r is $2\pi R \sinh r/R$. The same result follows by replacing R by iR in $2\pi R \sin r/R$, which is the circumference of a circle of radius r on the sphere (where, of course, r is measured *on* the spherical surface. See Exercise 18.2.1).

The geometry of Axiom P_2 was called *hyperbolic* by Klein (1871). One reason for this is that its formulas involve hyperbolic functions, whereas those of spherical geometry involve circular functions. Lambert (1766) introduced the hyperbolic functions and noted their analogy with the circular functions, but he did not follow through with a complete translation of spherical formulas into hyperbolic formulas. This was first done by Taurinus (1826), one of a small circle who corresponded with Gauss on geometric questions.

This gave hyperbolic geometry a second leg to stand on, but there was still nothing solid under its feet. Neither Gauss nor Taurinus seemed confident of finding a convincing interpretation of hyperbolic geometry, even though Gauss (1827) came remarkably close with the "Gauss–Bonnet" theorem. As mentioned in Section 17.6, this theorem shows that surfaces of constant negative curvature give a geometry in which angular defect is proportional to area, and Gauss knew that the pseudosphere was such a surface. Gauss' student Minding (1840) even showed that the hyperbolic formulas for triangles hold on the pseudosphere, but no one at that time commented on the likely importance of this result for hyperbolic geometry. Perhaps it was clear that the pseudosphere cannot serve as a "plane," because it is infinite in only one direction. Only in 1868, when Beltrami extended the pseudosphere to a true *hyperbolic plane*—a surface locally like the pseudosphere but infinite in all directions—was hyperbolic geometry finally placed on a firm foundation.

Exercises

18.2.1 Prove that the circumference of the circle C of radius r on the sphere of radius R (Figure 18.4) is $2\pi R \sin(r/R)$.

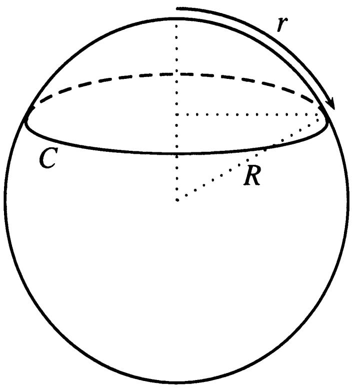

Figure 18.4: Radius and circumference on the sphere

18.2.2 Show that both $2\pi R \sin(r/R)$ and $2\pi R \sinh(r/R)$ tend to $2\pi r$ as $R \to \infty$.

These results illustrate that even noneuclidean geometry is "Euclidean in the small"—its formulas tend to the Euclidean formulas as size (in this case, size relative to the radius of curvature) tends to zero.

The same is true of the angle sum formula for a triangle.

18.2.3 Deduce from Harriot's area formula that the angle sum of a spherical triangle tends to π as its size tends to zero.

18.3 Geometry of Bolyai and Lobachevsky

The most important contributors to hyperbolic geometry between Gauss and Beltrami were Lobachevsky and Bolyai, who published independent discoveries of the subject: Lobachevsky (1829) and János Bolyai (1832b). Because of their courage in advocating an unconventional geometry, Bolyai and Lobachevsky have won the admiration of many historians. Nevertheless, the historical significance of their work is debatable. The bulk of their results were already known to Gauss and his circle and could have been picked up, at least in hazy form, from existing publications and personal contacts. Lambert (1766) and Taurinus (1826) were in print, and Bolyai's father, F. Bolyai, was a lifelong friend of Gauss, as was Lobachevsky's teacher Bartels. In any case their work, though more systematic than previous attempts and expressed with a lot more conviction, attracted very little attention at first. We have seen how the possibility of using differential geometry to justify hyperbolic geometry was overlooked until 1868. Up to that time, there seemed no reason to take hyperbolic geometry seriously.

In retrospect, of course, the theorems of Bolyai and Lobachevsky can be seen to unify the fragmentary results of their predecessors very nicely. They cover the basic relations between sides and angles of triangles (hyperbolic trigonometry), the measure of polygonal areas by angular defect, and formulas for circumference and area of circles. Lobachevsky (1836) broke new ground by finding volumes of polyhedra, which turn out to be far from elementary, involving the function $\int_0^\theta \log 2|\sin t|\, dt$.

Both Bolyai and Lobachevsky considered a three-dimensional space satisfying Axiom P_2 and made extensive use of a surface peculiar to this space, the *horosphere*. A horosphere is a "sphere with center at infinity," and it is *not* a hyperbolic plane. Wachter, a student of Gauss, observed in a letter of 1816 [published in Stäckel (1901)] that the geometry of the horosphere is in fact Euclidean. This astonishing result was rediscovered by Bolyai and Lobachevsky, and they anticipated that it would make Euclidean geometry subordinate to hyperbolic. We shall see in Section 18.5 how this view was vindicated in the work of Beltrami.

18.4 Beltrami's Projective Model

Interest in hyperbolic geometry was rekindled in the 1860s when unpublished work of Gauss, who had died in 1855, came to light. Learning that Gauss had taken hyperbolic geometry seriously, mathematicians became more receptive to noneuclidean ideas. The works of Bolyai and Lobachevsky were rescued from obscurity and, approaching them from the viewpoint of differential geometry, Beltrami (1868a) was able to give them the concrete explanation that had eluded all his predecessors.

Beltrami was interested in the geometry of surfaces and he had found the surfaces that could be mapped onto the plane in such a way that their geodesics went to straight lines [Beltrami (1865)]. They turned out to be just the surfaces of constant curvature. In the case of positive curvature, the sphere, such a mapping is central projection onto a tangent plane (Figure 18.5), though of course this maps only half the sphere onto the whole plane.

The mappings of surfaces of constant negative curvature, on the other hand, take the *whole* surface onto only *part* of the plane. Figure 18.6, adapted from Klein (1928), shows some of these mappings (the middle one being of the pseudosphere).

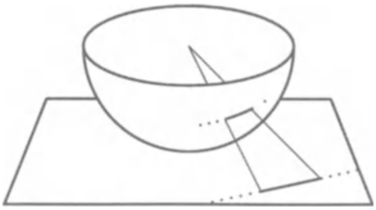

Figure 18.5: Central projection

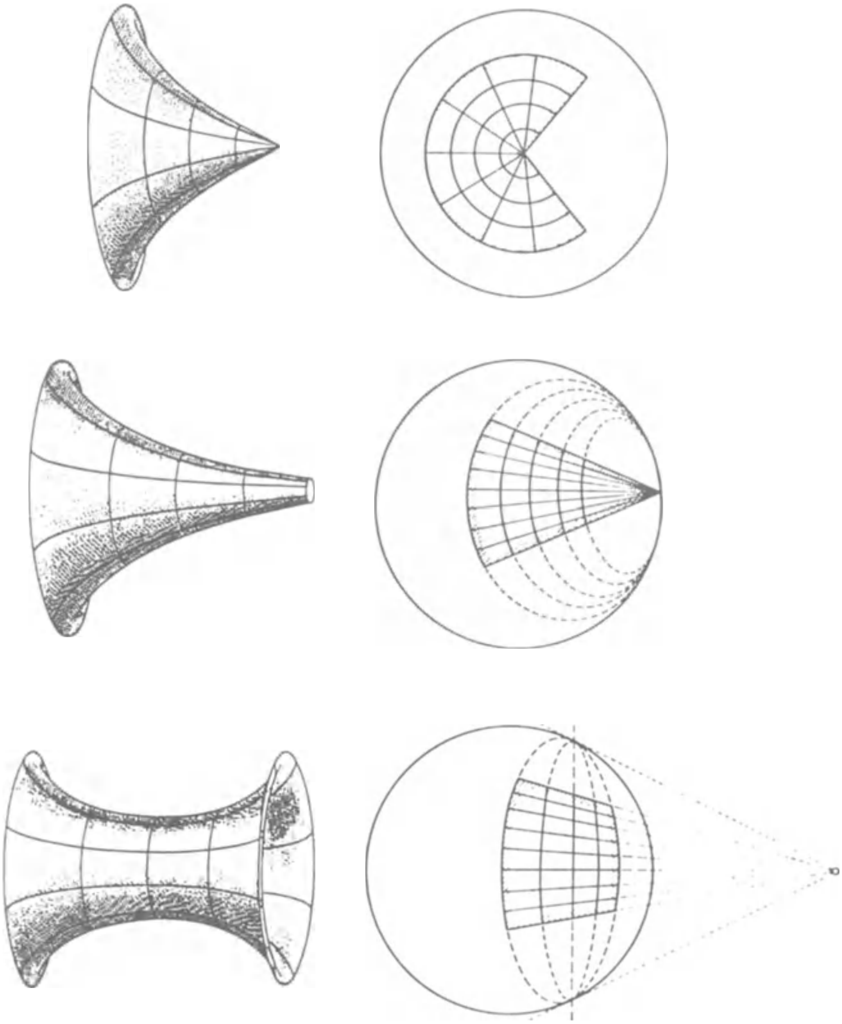

Figure 18.6: Geodesic-preserving mappings

Each negatively curved surface S is mapped onto a portion of the unit disk. Beltrami (1868a) realized that the disk can then be viewed as a natural extension of S to an "infinite plane," thus bypassing the problem of constructing "planelike" surfaces of constant negative curvature in ordinary space. Instead one takes the disk as the "plane," line segments within it as "lines," and "distance" between two points of the disk as the distance between their preimage points on the surface S. The function $d(P,Q)$, giving "distance" between points P, Q of the disk in this way, turns out to be meaningful for all points inside the unit circle, so the notion of "distance" extends to the whole open disk. As Q approaches the unit circle, $d(P,Q)$ tends to infinity, so the "plane," and hence the "lines" in it, are indeed infinite with respect to this nonstandard "distance."

It follows that all the axioms of Euclid, except the parallel axiom, are satisfied with the new interpretation of "plane," "line," and "distance." Instead of the parallel axiom, one has of course Axiom P_2, since there is more than one "line" through a point P outside a given "line" L (Figure 18.7).

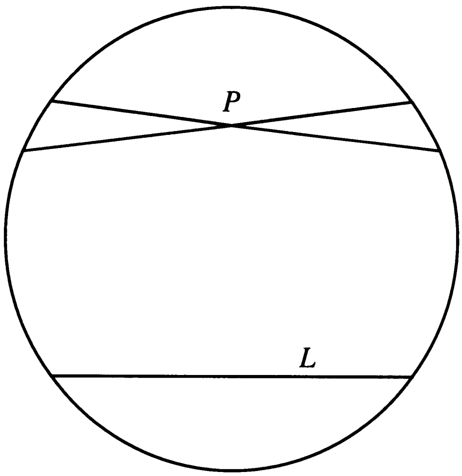

Figure 18.7: Failure of the parallel axiom

Beltrami also observed that the rigid motions of the "plane," since they preserve straight lines, are necessarily projective transformations. They are precisely those projective transformations of the plane that map the unit circle onto itself. Consequently, this model of the hyperbolic plane is often called the *projective model.* Cayley (1859) had already observed that these