From: Whittlesey, M.A., Spherical Geometry and its Applications
(CRC Press: Boca Raton).

Chapter 2

The sphere in space

We begin our study of spherical geometry by noting two different ways of presenting the subject, both of which we consider in this book.

We may first consider a sphere the way most people understand it: as a round object that lives in three-dimensional space. By doing so, we may use the tools of three-dimensional geometry, trigonometry, and vectors to learn about the sphere. This we will call an *extrinsic* approach to spherical geometry. An understanding of it is particularly important in applications. On the other hand, we may study spherical geometry via the *intrinsic* properties of the sphere — that is, properties of the sphere that can be thought of without reference to the larger three-dimensional space in which a sphere sits. It will be our main goal to do this in Chapter 3.

We begin by looking at the extrinsic properties of the sphere in this chapter. As we do so, it will be important to understand how to see these properties as intrinsic, and hence motivate the axiom system that we set up in Chapter 3.

5 Great circles

Definition 5.1 *Let O be a point in space and r a positive real number. The* sphere *with* center O *and* radius r *is the set of all points in space which lie at distance r from O.*

The following proposition summarizes the properties of intersections between spheres and planes in space.

Proposition 5.2 *The intersection between a sphere and a plane in space is a circle, a point, or the empty set.*

In Figure 2.1 we may see these possibilities geometrically by considering the line in space through the center O of the sphere perpendicular to the plane

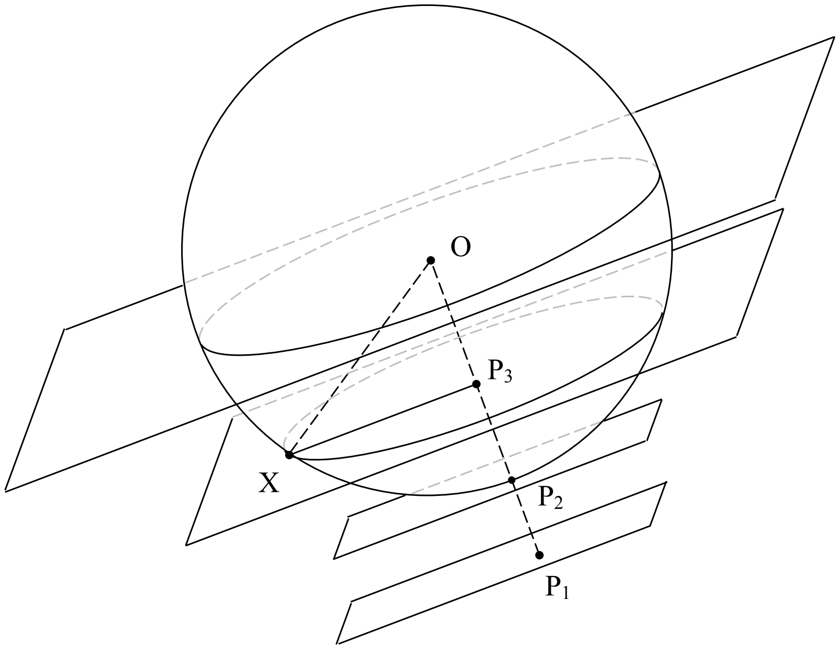

Figure 2.1: The intersection of a plane with a sphere.

at point P. We distinguish four cases: P is outside the sphere, on the sphere, interior to the sphere but different from O, or equal to O. In the first case, every point in the plane is outside the sphere because $P = P_1$ is the closest point on the plane to O — hence every point on the plane is at a greater distance from O than the radius. If $P = P_2$ is on the sphere, then the plane is tangent to the sphere, hence meets it in a single point. In the case of $P = O$, the intersection of the sphere with the plane is the set of all points in the plane at a fixed distance from O. This is a circle. Lastly, suppose $P = P_3$ is inside the sphere, but different from O. Then let X be any point on the intersection of the plane and the sphere, let r be the radius of the sphere, and let d be the distance between P_3 and O. Then $\triangle OP_3X$ has a right angle at P_3, so the length of $\overline{P_3X}$ must be $\sqrt{r^2 - d^2}$. Thus the points on both the plane and sphere all lie at the same distance from P_3, so lie on a circle. To be complete, we must show that all points in the plane at distance $\sqrt{r^2 - d^2}$ from P_3 are also on the sphere; this is left as an exercise. So the intersection of the plane and the sphere is a circle of radius $\sqrt{r^2 - d^2}$.

A similar analysis works for the intersection of a line with a sphere: it consists of zero, one, or two points.

Definition 5.3 *(See Figure 2.2.) Let s be a sphere with center O, p a plane in space and ℓ the line passing through O perpendicular to p.*

If $p = p_1$ *passes through* O *then the intersection of* p_1 *and* s *is called a* great circle *of the sphere. The two points where* ℓ *meets the sphere are called the* poles *of the great circle. We will say that* p_1 *is the plane of the great circle.*

If $p = p_2$ *lies at distance* $0 < d < r$ *from* O*, then the intersection of* p_2 *and* s *is called a* small circle*. The two points where* ℓ *meets the sphere are called the* poles *of the small circle. We will say that* p_2 *is the plane of the small circle.*

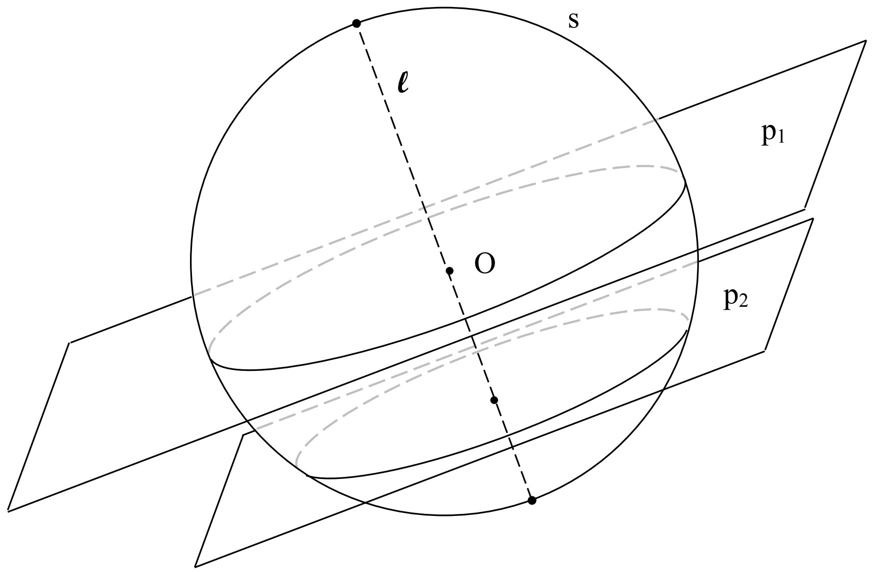

Figure 2.2: Great circles, small circles, and poles.

Definition 5.4 *Two distinct points on a sphere are said to be* antipodal *if the line passing through them passes through the center of the sphere. Each point is said to be the* antipode *of the other.*

If points A and B are antipodal then we write $B = A^a$ (and $A = B^a$). Note that the poles of a great circle are antipodal (as are the poles of a small circle).

Two points on a great circle which are antipodal divide a great circle into two semicircles which are called *great semicircles*.

As an example, the surface of the earth is (approximately) a sphere. The equator of the earth is a great circle, and the north and south poles of the earth are its poles. The circles of constant latitude strictly between zero and ninety degrees are small circles; again, the poles are the north and south poles of the earth. Each meridian forms a semicircle which is a great semicircle.

Great circles are significant because they are the "lines" of spherical geometry. That is, they serve many of the same functions in spherical geometry that lines do in plane geometry. We will see that the shortest path between two points on a sphere follows the route of a great circle, and that most of the time, two points determine a unique great circle passing through them.

Two great circles arise by intersecting the sphere with two planes through the center of the sphere. These two planes intersect in a line through the center of the sphere, which meets the sphere in two points. Thus we have the fundamental result:

Proposition 5.5 *Two distinct great circles meet in exactly two points, and these two points are antipodal. These two points divide each great circle into two great semicircles.*

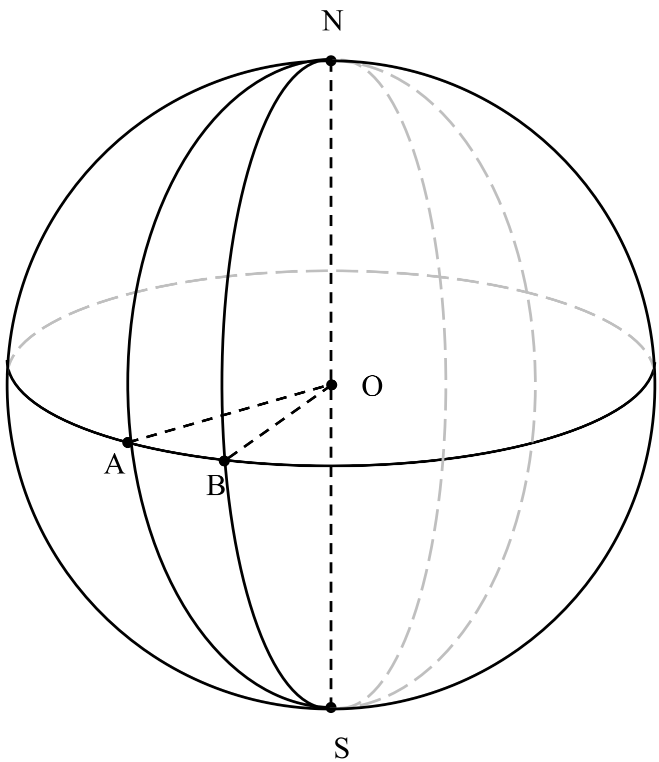

Figure 2.3: Pairs of great circles meeting at antipodal points N and S. Measure of arc. Great semicircle NAS, arc $\widehat{AB}$, and lune $NASBN$.

Suppose we are given two great semicircles with common antipodal end points N and S. If the semicircles are not on the same great circle, they form a two-sided object called a *lune*. The semicircles are said to be the *sides* of the lune, and N and S are the *vertices* of the lune.

The reader will recall that two points in a plane determine a unique line containing them. What about great circles?

Considering Figure 2.3, it is clear that this is not always the case on the sphere. Two antipodal points N and S lie on a line through the center of the sphere. There are many planes containing that line, each of which determines a great circle containing the antipodal points. But it turns out that this is the only exception:

Proposition 5.6 *If two distinct points on a sphere are not antipodal then there exists a unique great circle passing through them.*

We may see why by thinking about what properties the plane containing the great circle would have to satisfy. Such a plane would have to contain the two distinct non-antipodal points (A and B) on the sphere along with the center O of the sphere. Three points determine a unique plane in space if they do not all lie on the same line. If these three points lay on a line, then because one of the points is the center of the sphere, the two points on the sphere would have to be antipodal by definition. Thus the plane containing the three points is unique, and hence the great circle containing them is also unique.

In plane geometry, two points determine a unique "line segment" consisting of points "between" the given points, together with the two original points. On the sphere, the situation is again somewhat different. Given two points on a great circle, there are two circular arcs between them on that great circle. Which one shall be "the" segment between the two points?

Again, there are two cases to consider. The points are either antipodal or not. If they are antipodal (e.g., N and S in Figure 2.3), there are many great circles to choose from which contain the two points. In this case, there is no natural choice to make of a great circle arc between the points. We will have to specify a third point to determine a semicircle. That is, there is only one great semicircle from N to S passing through A (denoted NAS), and only one great semicircle from N to S passing through B, etc. If two given points are not antipodal, then there is a unique great circle through them and two arcs which together make up that great circle, one shorter and one longer. Typically we shall choose the shorter of the two to denote "the" arc between the given points. The points on that arc are then said to be "between" the given points (except that the original two points, the "endpoints" of the arc, are not said to be "between" the endpoints). If A and B are the endpoints then we denote the arc as $\overset{\frown}{AB}$.

In a plane a line segment between two points has a length which is the distance between the two points. On the sphere we have a similar phenomenon. We define spherical distance between two points to be the measure of the shortest arc of a great circle between those points. Typically by "measure" we mean the measure of the central angle of that arc (e.g., in Figure 2.3, $\angle AOB$ is the central angle of the great circle arc from A to B). The "measure" of the arc is closely related to the length of the arc in space. If an arc of a circle of radius r has central angle of radian measure θ then its length is $r\theta$. Thus on a sphere of radius 1, the length of the arc is the same as its measure.

A great semicircle will have a measure of π radians (180°). For units of arc measure we use either degrees or radians depending on the situation. In order to avoid speaking of radians and degrees at the same time, we often shall speak of a pair of points as being a "semicircle" apart (π radians, 180°) or a "quarter circle" apart ($\frac{\pi}{2}$ radians, 90°).

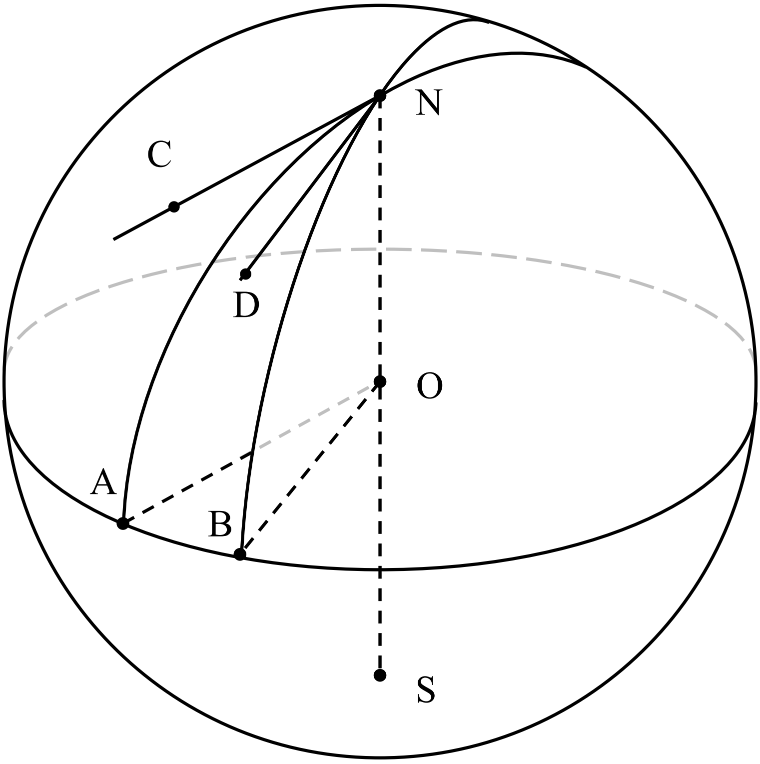

Figure 2.4: Measure of a spherical angle: $m \prec ANB = m \overset{\frown}{AB} = m\angle AOB = m\angle CND = m\angle A - ON - B$.

Figure 2.4 illustrates how we will understand the measure of angles between great circles on the sphere. The great circle arcs $\overset{\frown}{NA}$ and $\overset{\frown}{NB}$ lie on the sides of what we will understand to be a spherical angle. The rays $\overrightarrow{NC}$ and $\overrightarrow{ND}$ are tangent to these two arcs at N. The measure of the angle between these two rays ($\angle CND$) will be the measure of the angle between the arcs. But since ray $\overrightarrow{OA}$ is parallel to the ray $\overrightarrow{NC}$ and $\overrightarrow{OB}$ is parallel to $\overrightarrow{ND}$, the measures of $\angle CND$ and $\angle AOB$ are the same. We above noted that the measure of arc $\overset{\frown}{AB}$ equals the measure of $\angle AOB$. Then $m \overset{\frown}{AB} = m\angle AOB = m\angle CND$ is the measure of what we will call the "spherical angle" $\prec ANB$. Note also that since $\angle AOB$ is a plane angle of the dihedral angle $\angle A-ON-B$, the measure of spherical angle $\prec ANB$ is the measure of the dihedral angle $\angle A-ON-B$. Thus the measure of the angle between two great circles is the measure of the angle between the planes containing them in space. (Figure 2.5 illustrates this situation with the view from above N.)

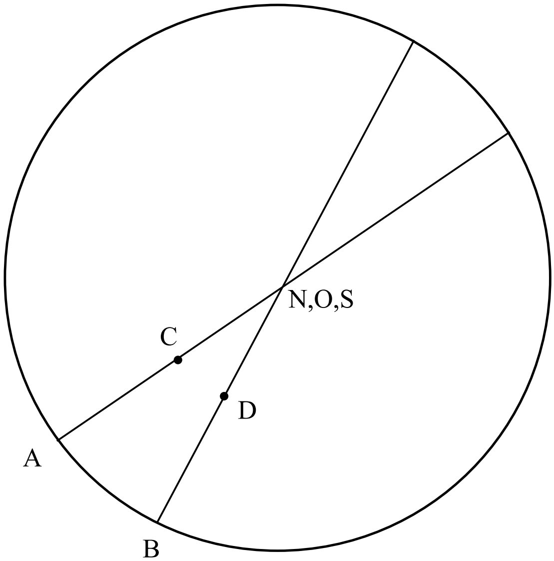

Figure 2.5: Measure of a spherical angle, top view.

A lune has two angles — one based at each of its vertices. These two angles have the same measure. This can be seen in Figure 2.3: lune $NASBN$ has two angles $\prec ANB$ and $\prec ASB$, both of which have measure equal to the measure of arc $\overset{\frown}{AB}$. The *measure* of a lune is defined to be the same as the measure of its angles.

Our definition of the notion of poles of a great circle arose from the fact that a line perpendicular to the plane of a great circle at its center meets the sphere in two points. We shall want to have an understanding of properties of the poles which depends solely on intrinsic properties of the sphere:

Proposition 5.7 *A great circle of a sphere s in space is the set of all points of s which lie a spherical distance of a quarter circle from a pole of the great circle.*

Proof. The great circle is the intersection of a plane p with sphere s and the poles of the great circle are the intersection of a line ℓ with s, where ℓ is perpendicular to p at O. First we show that any point of the great circle is at distance $\frac{\pi}{2}$ from a pole. Any point A of the great circle is a point of p and a pole N of the great circle lies on ℓ. Since ℓ is perpendicular to p, $\overrightarrow{ON}$ is perpendicular to $\overrightarrow{OA}$. By definition the spherical distance between A and N is the measure of $\angle AON$, which is $\frac{\pi}{2}$. Conversely, suppose point A is at spherical distance $\frac{\pi}{2}$ from pole N. By definition, $\angle AON$ has measure $\frac{\pi}{2}$, so $\overrightarrow{OA}$ is perpendicular to $\overrightarrow{ON}$. By Proposition 2.9, we conclude that A is a point of p, so belongs to the great circle. $\diamondsuit$

Proposition 5.7 is useful in that it provides a description of the relationship between a great circle and its poles that depends only on an intrinsic property (distance) on the sphere, and not on lines and planes in space. A similar characterization is desired for small circles.

Proposition 5.8 *A small circle on a sphere is the set of all points on the sphere at a fixed spherical distance $\rho < \frac{\pi}{2}$ from one of its poles P.*

See Exercise 4 to justify this via three-dimensional geometry.

Definition 5.9 *The point P found in Proposition 5.8 is called the* (spherical) center *of the small circle, and the quantity ρ is called the* (spherical) radius *of the small circle.*

It is not hard to see that the points on the small circle are also equidistant from the antipode of P; see Exercise 5.

In plane geometry we learn that a line in a plane splits the plane into two sets of points off the line called half-planes. These sets are convex in the sense that given two points in a half-plane, the line segment between them also lies in the half-plane.

To see that this idea carries over to the sphere, one simply notes that a plane in space splits space into two convex sets called half-spaces. Given a great circle, we consider the two half-spaces associated with its plane. If we intersect these half-spaces with the sphere, we obtain two so-called hemispheres into which the sphere is divided by the great circle. These hemispheres are the "sides" of the great circle.

Are these hemispheres "convex"? What does "convex" mean on the sphere, if anything? We take as our understanding of convexity on the sphere that the hemisphere is "spherically convex" if the shortest great circle arc between any two (non-antipodal) points also lies in that hemisphere. That this is the case may be seen as suggested in Figure 2.6.

Given two points in a hemisphere, we take a great circle containing them, which meets the edge of the hemisphere in two antipodal points. These antipodal points split the original great circle into two great semicircles, one of which is in our hemisphere, and the other of which is in the opposite hemisphere. Our two points must lie in the semicircle in our hemisphere, but not at the endpoints. Thus the shorter arc between them lies in that semicircle, as desired.

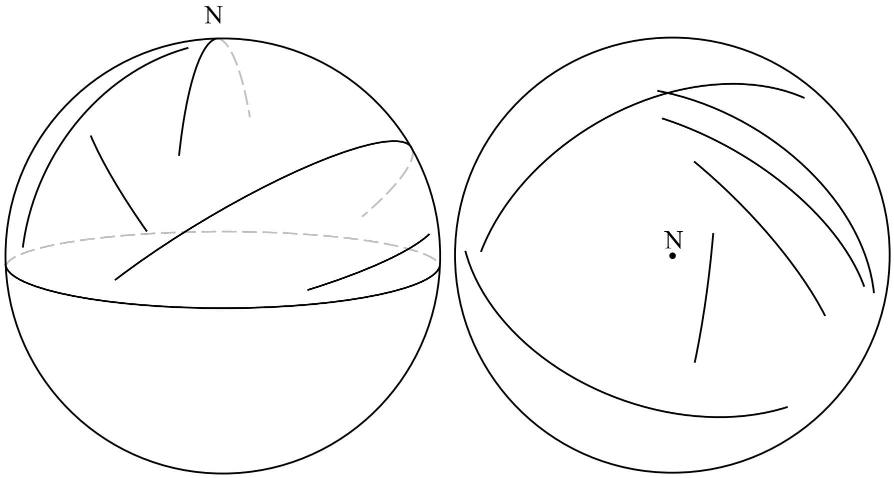

Figure 2.6: A hemisphere is spherically convex — side and top views. The shorter great circle arc between two points of the hemisphere lies in the hemisphere.

A great circle is spherically convex since a spherical arc between two of its non-antipodal points lies on the given great circle. Furthermore, a spherical arc is also spherically convex.

We conclude with one more observation about hemispheres as suggested by Figure 2.7. There we have a hemisphere, a great circle which is the edge of that hemisphere, and a pole N of the great circle contained in the hemisphere. Then the hemisphere consists of all points which are within a quarter circle of the pole N. Any great circle passing through N meets the edge of the hemisphere in two points which cut the great circle into two great semicircles, one of which contains N in the middle. Thus the points of the semicircle lie within a quarter circle of N (except for the endpoints).

Exercises §5

For each of these exercises, use whatever knowledge and technique from three-dimensional geometry is necessary.

1. For each of the following values of r and θ, determine the length of an arc of a great circle whose measure is θ on a sphere of radius r.

   (a) $r = 50$, $\theta = \frac{\pi}{4}$ radians

   (b) $r = 2000$, $\theta = 50°$

   (c) $r = 100$, $\theta = 140°$

2. Suppose that point A is on one side of a great circle. Show that its antipode is on the other side of the great circle.

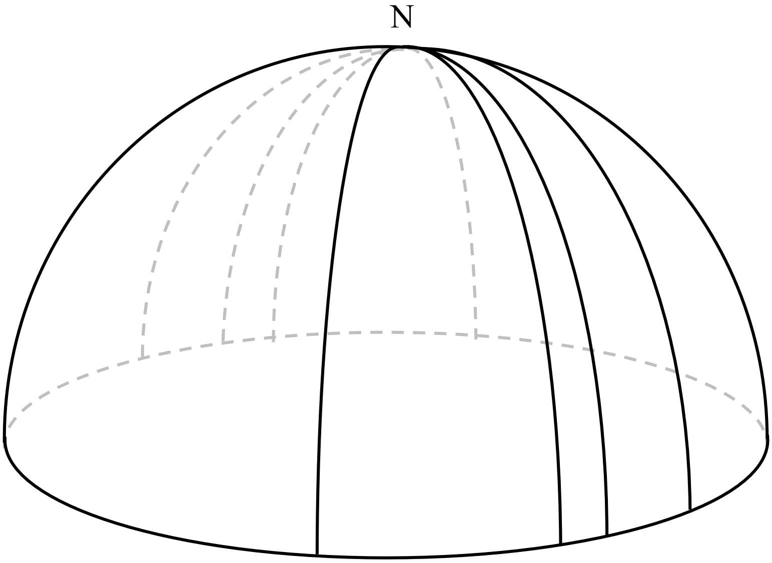

Figure 2.7: A hemisphere is the set of points less than a quarter circle from a pole N of its edge.

3. Suppose that A is a point of a sphere with antipode A^a and B is any other point of the sphere. Explain why $m\,\overset{\frown}{BA} + m\,\overset{\frown}{BA^a} = \pi$.
4. Use three-dimensional geometry to justify Proposition 5.8.
5. Let c be a small circle with center P and spherical radius ρ. Show that c is the set of all points on the sphere which are at spherical distance $\pi - \rho$ from P^a, the antipode of P.
6. Given three points on a sphere which do not lie on a single great circle, argue that they lie on a unique small circle.
7. Let A and B be two points on a sphere. Using a similar theorem in space, explain why the set of all points Γ on the sphere which are at the same spherical distance from A and B must be a great circle perpendicular to any great circle passing through A and B. If A and B are not antipodal, then Γ is called the *perpendicular bisector* of $\overset{\frown}{AB}$). If A and B are antipodal, the great circle is called the *polar circle* or *polar* of A and B. (Compare this extrinsic approach to the problem to an intrinsic approach in §12, Exercise 10.)
8. Suppose that a small circle has two points on opposite sides of a great circle. Argue that the small circle and great circle intersect in two distinct points.

9. Suppose that a sphere has radius r and a small circle on it has spherical radius ρ. Show that the perimeter of the small circle in space is $2\pi r \sin(\rho)$. Conclude that if a small circle arc from point A to point B has arc measure ϕ radians, then the length (in space) of that arc is $r\phi \sin(\rho)$.

10. Referring to problem 9, we define the *spherical length* of the small circle arc given to B to be $\phi \sin(\rho)$. Show that the spherical distance from A to B is less than $\phi \sin(\rho)$.

6 Distance and angles

Plane geometry has the good fortune that the most important theorem regarding distances (the Pythagorean theorem) can be stated without the use of trigonometric functions. Spherical geometry is not so fortunate. Even the most basic propositions about relationships in right triangles require the use of trigonometric functions. We illustrate with an example.

Proposition 6.1 *Suppose two spherical arcs form an angle with measure θ. Let x be the spherical distance between two points (one on each arc) at spherical distance ϕ from the vertex. Then* $\sin(\frac{x}{2}) = \sin(\phi)\sin(\frac{\theta}{2})$.

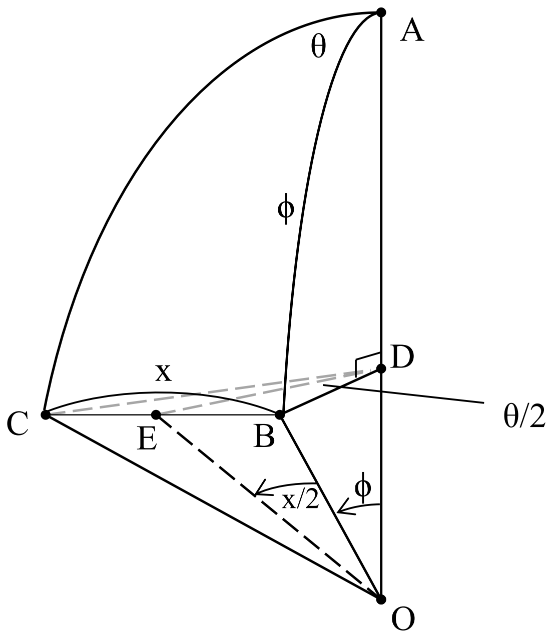

Figure 2.8: Proposition 6.1.

Proof. Suppose the sphere has center O and radius r. Since $\overset{\frown}{AB}$ and $\overset{\frown}{AC}$ have the same measure ϕ, there is a plane perpendicular to $\overleftrightarrow{OA}$ (at D) passing through both B and C. By symmetry, we also have $DB = DC$ and $OB = OC = r$. Let E be the midpoint of the line segment $\overline{BC}$. Then $\triangle DEB$ and $\triangle OEB$ both have a right angle at E. We get $m\angle EOB = \frac{1}{2}m\angle COB = \frac{x}{2}$ by the definition of the measure of $\overset{\frown}{BC}$. Also, $m\angle EDB = \frac{1}{2}m\angle BDC$. But $\angle BDC$ is a plane angle of $\angle B - OA - C$, so has measure θ by definition of the measure of spherical angle $\prec BAC$. So $m\angle EDB = \theta/2$. By definition of arc measure, $m\angle BOD = m\angle BOA = \phi$. By plane right triangle trigonometry applied to $\triangle OBD$, $BD = OB\sin(\phi) = r\sin(\phi)$. Applying plane trignometry to $\triangle BED$, $EB = BD\sin(\frac{\theta}{2}) = r\sin(\phi)\sin(\frac{\theta}{2})$. Applying plane trigonometry to $\triangle OEB$, we find $EB = r\sin(\frac{x}{2})$. Thus $r\sin(\frac{x}{2}) = EB = r\sin(\phi)\sin(\frac{\theta}{2})$ and the conclusion follows. (This argument assumes ϕ is acute; a similar argument holds in the cases where it is right or obtuse.) $\diamondsuit$

Proposition 6.1 can be thought of as our first result about what we will call spherical triangles. We have three points A, B, and C with three spherical arcs $\overset{\frown}{AB}$, $\overset{\frown}{AC}$, and $\overset{\frown}{BC}$. The union of three such arcs is a spherical triangle. The proposition assumes the measures of $\overset{\frown}{AB}$ and $\overset{\frown}{AC}$ are known, as is the measure of the angle between them, and relates them to the measure of the third arc. In the exercises, the interested reader can see how to use techniques of three-dimensional geometry to prove a number of other properties of spherical triangles in space.

Exercises §6.

1. Let O be the center of a sphere and let A, B, and C be distinct points of the sphere, no pair of which is antipodal. Thus we have three arcs $\overset{\frown}{AB}$, $\overset{\frown}{AC}$, and $\overset{\frown}{BC}$ whose measures we denote by c, b, and a, respectively. We let A, B, and C denote the measures of the spherical angles $\prec BAC$, $\prec ABC$, and $\prec ACB$, respectively. Suppose that the spherical angle $\prec ACB$ is right. (That is, the plane containing O, A, and C is perpendicular to the plane containing O, B, and C.) Then we will think of A, B, and C as forming a "spherical triangle" $\triangle^s ABC$ with a right angle at C. Suppose that D is between O and A, E is between O and B, and F is between O and C. (See Figure 2.9.) Also assume that $\overline{DE}$ and $\overline{DF}$ are perpendicular to $\overline{OD}$.

   (a) Explain why $\angle DFE$ and $\angle OFE$ are right angles.

   (b) Explain why the measure of $\angle FDE$ is the same as the measure of the spherical angle $\prec BAC$.

   (c) Using the setting with the structure $ODEF$ introduced at the beginning of this section, prove that $\cos(c) = \cos(a)\cos(b)$, $\sin(A) = \sin(a)/\sin(c)$, $\cos(A) = \tan(b)/\tan(c)$, and $\tan(A) = \tan(a)/\sin(b)$.

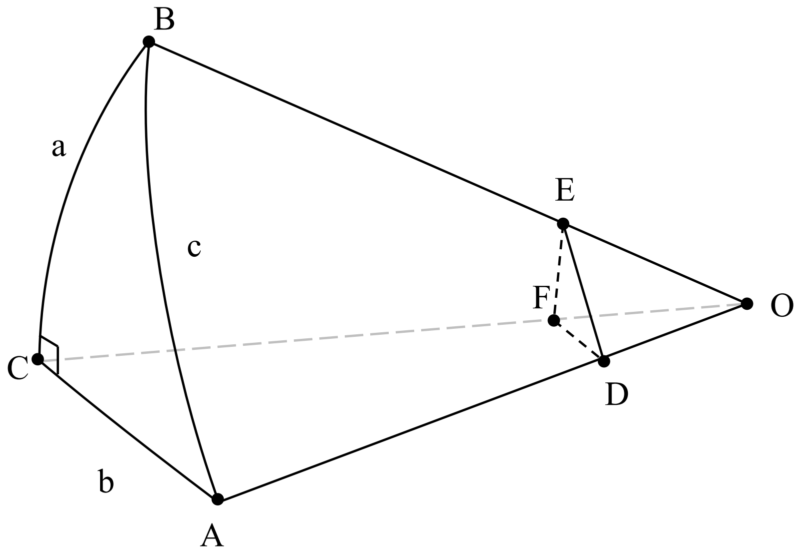

Figure 2.9: Exercise 1.

(d) Prove that $\sin(A) = \cos(B)/\cos(b)$ and $\cos(c) = \cot(A)\cot(B)$.

7 Area

We shall determine areas of a curved surface by approximating it with pieces of surfaces for which the area is known. The idea is that if a sequence of surfaces approaches the surface with unknown area, then the area of the unknown surface is the limiting value of the surfaces in the sequence. An entirely rigorous discussion of this usually involves integral calculus (and real analysis). However, for readers who are interested in an understanding of area that does not require these nonelementary methods, we here provide a discussion which will hopefully be persuasive. We begin by providing an approach to determining the area of the whole sphere.

In space, suppose we are given a planar set of points b and a point P not in the plane of the circle. (See Figure 2.10.) Then we recall that the *cone* with *base* b and *vertex* P is the union of all line segments with one endpoint at P and one endpoint in b.

Suppose that the base is a circle. Then we say that the cone is *circular*. We say that the cone is a *right circular cone* if the segment between P and the center of the circle is perpendicular to the plane of the circle. For a right circular cone, the length of the segment between P and a point in the circle is always the same length called the *slant height*. The radius of the circle is called the *radius* of the cone.

Suppose that the base of the cone is a polygon. Then we say that the cone is *polygonal*. A polygonal cone is *regular* if the base is a regular polygon. We say that a regular polygonal cone is *right* if the segment between the vertex and the center of the polygon is perpendicular to the plane of the polygon.

A *frustum* of a cone is that portion of the cone obtained as follows: Let p be a plane parallel to the base between P and the base. Then the frustum obtained is the set of points of the cone on p, or on the side of p opposite from the vertex. The *bases* of the frustum are (1) the base of the original cone, and (2) the intersection of p with the cone.

A frustum of a right regular polygonal cone is the union of n isosceles trapezoids and their interiors, where n is the number of sides in the polygon. The *slant height* ℓ of the frustum is the height of the trapezoid.

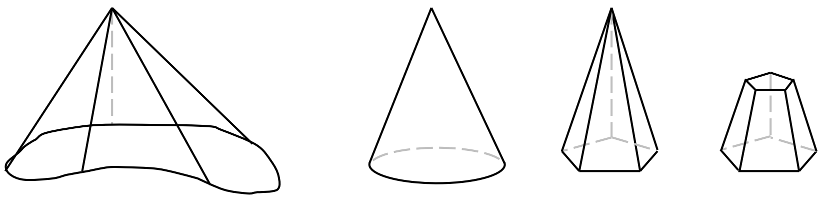

Figure 2.10: Cone, right circular cone, polygonal cone, and polygonal frustum.

Proposition 7.1 *Suppose that a frustum of a right polygonal cone has bases of n sides each. Suppose that the perimeter of the bases are p_1 and p_2, and the frustum has slant height ℓ. Then the lateral area (i.e., the area without the bases) of the frustum is $\frac{\ell}{2}(p_1 + p_2)$.*

Proof. The frustum is the union of n trapezoids. Each has height ℓ and bases of length p_1/n and p_2/n, since the bases have n sides each of equal length. By §1, Exercise 4, the area of each trapezoid is $\frac{\ell}{2}(\frac{p_1}{n} + \frac{p_2}{n})$. Since there are n of them the total area is $n\frac{\ell}{2}(\frac{p_1}{n} + \frac{p_2}{n})$, or $\frac{\ell}{2}(p_1 + p_2)$. ♢

Proposition 7.2 *Suppose that a frustum of a right circular cone has two bases with perimeter s_1 and s_2 and slant height ℓ. Then the (lateral) surface area of the cone is $\frac{\ell}{2}(s_1 + s_2)$.*

Proof. We approximate the circular base of the one with a regular polygon of n sides formed by spacing n points equally around the perimeter of the base, and obtain the formula for the area of the polygonal frustum from Proposition 7.1. For large values of n, the slant height ℓ_n of the polygonal frustum is close to the slant height ℓ of the circular frustum, and the perimeters p_1^n, p_2^n of the bases of the polygonal frustum are close to the perimeters s_1, s_2 of the bases of the circular frustum, so the formula $\frac{\ell_n}{2}(p_1^n + p_2^n)$ obtained from Proposition 7.1 approaches the value $\frac{\ell}{2}(s_1 + s_2)$, as desired. ♢

We may see this another way as follows. We unroll the frustum on a plane; the lateral surface lays out onto the sector of an annulus. The lengths of the circular arcs are s_1 and s_2. Because the slant height of the frustum is ℓ, the difference of the two radii of the annulus is ℓ. By §1, Exercise 6, the area rolled out is $\frac{\ell}{2}(s_1 + s_2)$.

Note that if the radii of the bases are r_1 and r_2, respectively, then $s_1 = 2\pi r_1$ and $s_2 = 2\pi r_2$, so the formula from Proposition 7.2 gives a surface area of

$$\pi\ell(r_1 + r_2) \tag{2.1}$$

We now proceed to derive the formula for the surface area of a sphere of radius r. We view the sphere as being obtained by revolving a semicircle of radius r about its diameter. If we view the semicircle as being approximated by a union of small chords of the semicircle, then the whole sphere is approximated by revolving the chords around the axis of the diameter of the semicircle.

Suppose that an arc of a semicircle is revolved around a diameter of that semicircle. The result is a surface called a *zone* of the sphere.

Theorem 7.3 *Suppose that an arc of a semicircle of radius r is revolved around the diameter of the semicircle. Suppose that the projection of the arc on the diameter has length d. Then the zone generated has surface area $2\pi rd$.*

Proof. (See Figure 2.11.) Let us suppose that $n-1$ distinct points are chosen

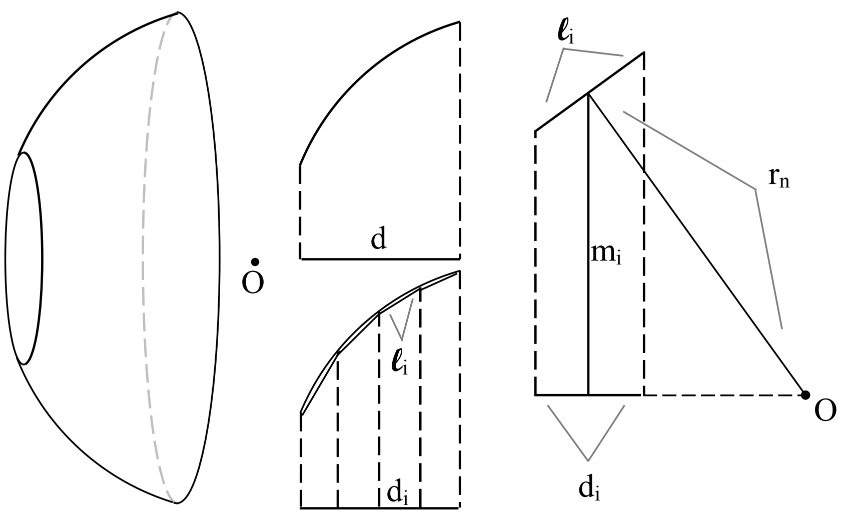

Figure 2.11: A zone, the arc being revolved, its approximation by segments, and the trapezoid of Theorem 7.3.

equally spaced on the arc being revolved. We connect the points in order — along with the endpoints of the arc — to form n chords of the arc of equal length. If we view the arc as being approximated by the n chords, then the zone is approximated by revolving the n chords around the axis of the diameter of the semicircle. Each chord along with its projection to the diameter form a trapezoid. For $i = 1, 2, \ldots, n$, let ℓ_i denote the length of the i^{th} chord, let m_i be the length of the midline of the i^{th} trapezoid, and let d_i be the length of the projection of the i^{th} chord on the diameter. Since every chord has the same length, the midpoint of each is at the same distance r_n from the center of the semicircle. Then the i^{th} chord generates a surface area of $2\pi\ell_i m_i$, according to (2.1) and the formula for the length of the midline of a trapezoid (the length of the midline is the average of the lengths of the bases). By Lemma 1.4, $2\pi\ell_i m_i = 2\pi d_i r_n$. Then the union of all the chords generates a surface of area

$$\sum_{i=1}^{n} 2\pi\ell_i m_i = \sum_{i=1}^{n} 2\pi d_i r_n = 2\pi d r_n, \tag{2.2}$$

where $\sum_{i=1}^{n} d_i = d$ because the union of the projections of the chords is the same as the projection of the arc, which has length d. Now as the number of arcs n is allowed to become arbitrarily large, r_n approaches r, so the surface area generated by the union of the arcs approaches $2\pi r d$; this is the surface area of the zone. $\diamondsuit$

Corollary 7.4 *The surface area of a sphere of radius r is $4\pi r^2$.*

Proof. Using Theorem 7.3, we let the arc be the whole semicircle. Since the projection of the whole semicircle has length $2r$, the surface area of the whole sphere is $2\pi r d = 2\pi 2r^2 = 4\pi r^2$. $\diamondsuit$

We conclude by discussing the area enclosed by a lune. Since a lune itself is merely the union of two semicircles, it technically has area zero. But it is common to make a small abuse of language and write "area of a lune" when we mean "area enclosed by a lune." We will also do so as long as there is no reason for confusion.

Proposition 7.5 *In a sphere of radius r, the area of a lune whose angle has radian measure θ is $2\theta r^2$.*

The simplest justification for this involves use of the rotational symmetry of the sphere. If we rotate any lune on its axis, we obtain a lune of the same size and shape and hence the same area. Then the proportion of the area of the sphere that lies in the lune is the same as the proportion of θ to 2π. Thus the area of the lune is $\frac{\theta}{2\pi} 4\pi r^2 = 2\theta r^2$, as desired.

Exercises §7.

1. Suppose that an arc of a semicircle is revolved about the diameter of the semicircle through an angle of θ (measured in radians). If the semicircle

has radius r and the projection of the arc on the diameter has length d, prove that the surface area of the region generated (called a *zone of a lune*) is $rd\theta$.

2. Prove that on a sphere of radius r, the area of the interior of a small circle of spherical radius θ is $4\pi r^2 \sin^2(\frac{\theta}{2})$.

3. Let r be the radius of the earth. What proportion of the earth's surface is visible at an altitude of a above the surface of the earth?

4. Two spheres of radius 3 and 4 have centers 5 units apart. Determine the area of each sphere which is contained inside the other.

5. What proportion of the earth's surface lies between the latitudes of $45°N$ and $45°S$?

8 Spherical coordinates

We now introduce two closely related kinds of coordinates on a sphere: polar coordinates and latitude-longitude coordinates. Suppose that s is a sphere of radius r. We shall specify coordinates (ϕ, θ) for points on s. The coordinates shall be determined as follows, referring to Figure 2.12. Choose a point P of s, which we will call the (north) pole of our coordinate system, and a great semicircle PQP^a (called the *zero meridian* or *prime meridian*,) where P^a is the antipode of P and Q is any point of s other than P and P^a.

Here is how we specify polar coordinates (ϕ, θ) for a point X on the sphere. We let ϕ be the measure of the great circle arc from X to P. (If $X = P$ or $X = P^a$ then $\phi = 0$ or $\phi = \pi$, respectively.) The angle θ is the measure through which we rotate greate semicircle PQP^a to obtain great semicircle PXP^a. (We must specify whether a positive θ is a rotation in the clockwise or counterclockwise direction when viewed from above P. If $\theta > 0$ is a clockwise rotation by θ, then a negative value of θ gives a counterclockwise rotation by angle $-\theta$.) A full rotation by measure 2π brings PQP^a back to itself. Thus θ is determined only up to integer multiples of 2π, so as in the plane θ is often restricted to intervals of length 2π, e.g., $-\pi < \theta \leq \pi$ or $0 \leq \theta < 2\pi$. We could also use degree measure for ϕ and θ.

If $X = P$ or $X = P^a$ then θ could have any value.

A great semicircle where θ is constant is called a *meridian*. The great circle where ϕ equals $\frac{\pi}{2}$ is called the *equator*. From this we also obtain *latitude-longitude coordinates*, closely related to the polar coordinates above. Given polar coordinates (ϕ, θ), the associated latitude-longitude coordinates for s are $(\ell_1, \ell_2) \equiv (\frac{\pi}{2} - \phi, \theta)$. Here the *latitude* variable ℓ_1 is zero on the equator E. Then ℓ_1 is positive in the hemisphere of s containing P and negative in the hemisphere of s containing P^a, so $-\frac{\pi}{2} \leq \ell_1 \leq \frac{\pi}{2}$. Furthermore, absolute value of ℓ_1 turns out to be the measure of an arc from X perpendicular to the equator. The *longitude* variable ℓ_2 is zero on the *zero meridian*. As with

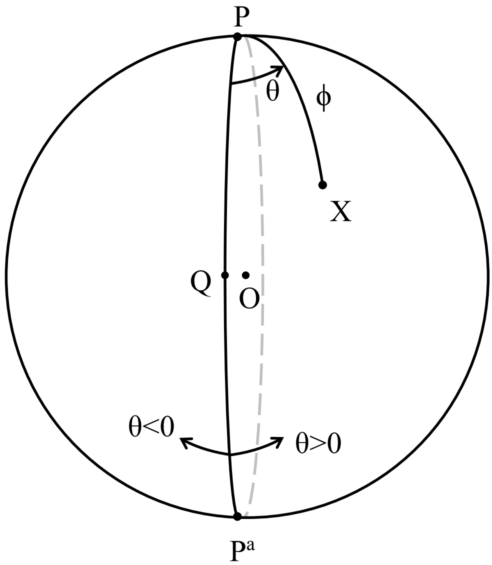

Figure 2.12: Spherical polar coordinates.

polar coordinates, we continue to say that a *meridian* is any great semicircle where longitude is constant.

The natural example of latitude-longitude coordinates are the latitude-longitude coordinates for the surface of the earth. Here P is the geographic north pole of the earth. The zero meridian (usually called the *prime meridian*) is the great semicircle between the geographic north and south poles which passes through Greenwich, England. Longitude is measured positively (usually in degrees) in the direction east of the zero meridian.

Since any point in space must lie on a sphere of some (possibly zero) radius r about the origin, the ordered triple (r, ϕ, θ) serves as what is known as a triple of spherical coordinates for a point in space.

Note that different ordered triples may determine the same point X. If $r = 0$, the point X must be O regardless of the values of ϕ or θ since only O is at distance zero from O. Also, given $r > 0$, any point where $\phi = 0$ must be P, regardless of the value of θ. Also, if $\phi = \pi$, X must be P^a regardless of the value of θ. Lastly, as noted above, given a triple (r, ϕ, θ), if we add an integer multiple of 2π to θ, we arrive at the same point. Our choice above that $-\pi < \theta \leq \pi$ is a matter of convenience. So a point in space does not have a unique set of spherical coordinates. However, we have the following theorem.

Theorem 8.1 *Suppose that (ϕ, θ) are polar coordinates for a sphere of radius r. Suppose that x, y, and z axes are chosen for space in the following manner. Let the center of the sphere be the origin. Let the positive z axis point in the* *direction of the pole of the coordinate system. Let the positive x axis point in the direction of the point $(\frac{\pi}{2}, 0)$ and let the positive y axis point in the direction of the point $(\frac{\pi}{2}, \frac{\pi}{2})$. Then a point A in space on the sphere of radius r with xyz coordinates (x, y, z) and polar coordinates (ϕ, θ) satisfies*

$$(x, y, z) = (r \sin\phi \cos\theta, r \sin\phi \sin\theta, r \cos\phi). \tag{2.3}$$

Proof.

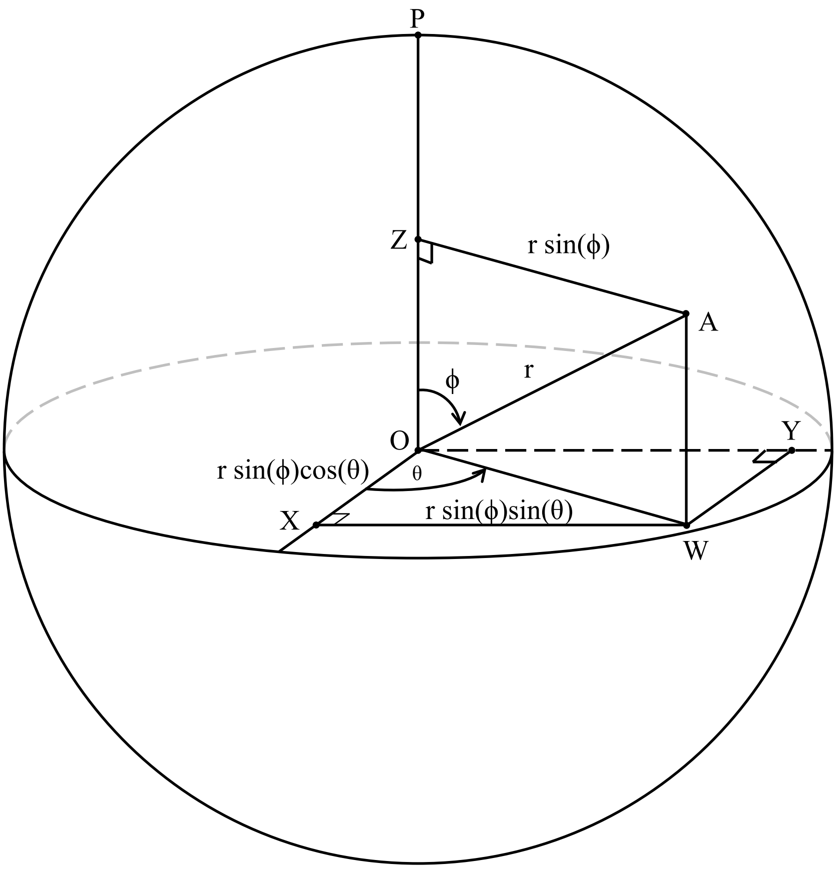

Figure 2.13: Theorem 8.1.

The discussion above shows how every (ϕ, θ) determines a point A on the sphere. Let X, Y, and Z be the projections of the point A to the corresponding axes. Then $\triangle AZO$ is a right triangle with a right angle at Z. By definition of arc measure, $m\angle AOZ = m\ \widehat{AP}$, which is ϕ by definition. Since $\triangle AZO$ is a right triangle and the length of the segment $\overline{OA}$ is r, the coordinate of Z is $r\cos\phi$ and the length of $\overline{AZ}$ is $r\sin\phi$. Thus $z = r\cos\phi$, as desired. Next, let W be the projection of A to the XY plane. Then the length of $\overline{OW}$ is the same as the length of $\overline{AZ}$, which is $r \sin \phi$. The projections of W to the x and y axes are X and Y, respectively. Next, the absolute value of θ is the measure of planar angle $\angle XOW$. Then since W lies at radius $r \sin \phi$ from O and $\overline{OW}$ forms a signed angle of θ with the positive x axis, its xy coordinates must be $x = r \sin \phi \cos \theta$ and $y = r \sin \phi \sin \theta$. $\diamondsuit$

Exercises §8

1. Extend §6, Exercise 1c as follows: if a point X has latitude-longitude coordinates (a, b) with respect to some system of cooordinates on the sphere, then the distance c from X to the point with coordinates $(0, 0)$ satisfies $\cos(c) = \cos(a) \cos(b)$.