Chapter 1

The Earth and the celestial sphere

Step outside on a clear dark night and be welcomed by a lavish celestial display. Quite possibly the Moon will be visible, and it will surely be the first thing you notice. If it is not too bright, you will see a profusion of stars arranged in distinctive patterns or *constellations*. At first the number of stars looks countless; some few are quite bright, but the dimmer ones are far more numerous. If you had the patience, though, and if the Moon were not present, you would find you could count about 2000. There might also be one or more very bright bodies that you would recognize as planets. Finally, if the night were very dark, you might see a wide luminous encircling band, the Milky Way, the combined light of millions of faint stars.

If you watch for perhaps 15 minutes, you will see that the sky is not stationary. Stars set below the western horizon and new ones take their places in the east. Over the course of the night, the heavens change considerably in appearance. Then the eastern horizon begins to glow and the stars start to fade away. As sunrise approaches, they disappear altogether and the sky becomes dominated by the westward-moving Sun. You can now see that the whole celestial vault – stars, Moon and Sun – appears to turn about points that are fixed in the north and south.

If you were to watch the sky for many days, you would see more subtle effects and slower changes. The most obvious would be that the Moon appears to move to the east against the stellar background. On successive nights it rises and sets steadily later, changing its shape as it goes. You would also note that the stars are never found in the same place at any given time from night to night. Each evening they are displaced a little to the west, setting progressively earlier until they are lost in twilight. Simultaneously, new ones arrive in the east, until after several months of watching, the entire aspect of the sky has changed. Observations made over several days or weeks also show that the planets are discriminated from stars not by their brightness but by their seemingly erratic movements, shifting east then west but on the whole drifting eastward against the starry display, very roughly following the path of the Moon.

In the daytime the position of the Sun also changes. Your first viewing might be on a warm summer day somewhere in northern climes. The Sun passes nearly overhead at noon and sets far to the northwest, providing long daylight hours. But as you watch, it appears to slip south day by day. Sunrise and sunset move to the southeast and southwest, and the days become colder and shorter until you might think that the Sun will disappear forever from view. But to your relief, the southerly movement eventually stops and the Sun starts northward again, the whole performance to be repeated year after year.

If you could now watch not just for days or even a few years but instead for centuries, you could detect subtler movements in the heavens. The whole sky would appear to be shifting slowly in an independent manner as stars creep past the fixed points of daily rotation. Stars and constellations never before seen would appear above the southern horizon while some others would disappear altogether from view. The familiar constellations of each season would slowly change. The sight of Leo in the evening announces Spring for northerners. Many thousands of years from now it will proclaim Autumn. We cannot see these changes in a mere lifetime, but they are obvious over many generations, and we have our written records, which can give us humans the effect of having lived for the required millennia.

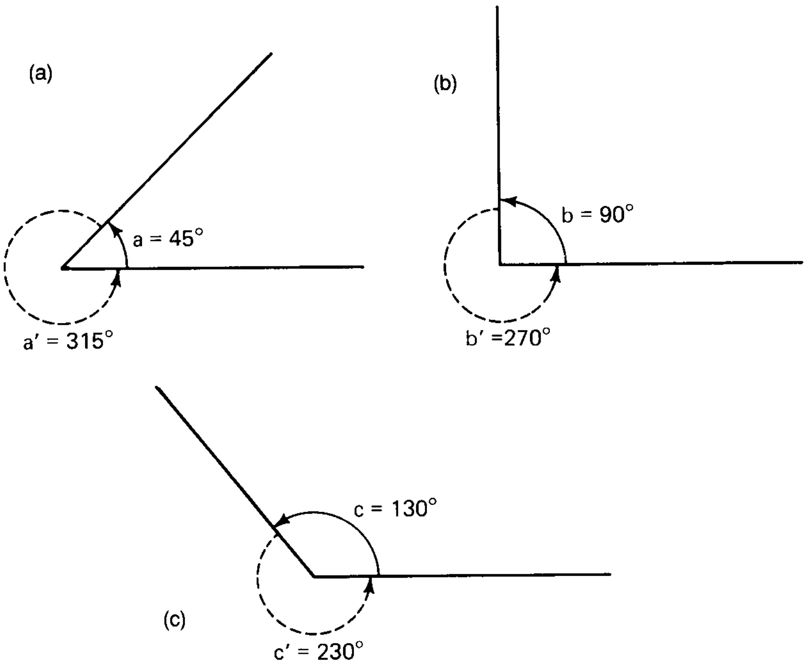
On a still longer time scale of thousands of centuries even the constellations would start to dissolve, the result of the motions of the stars themselves and our changing perspective, as the Sun, with its family of planets, pursues a relentless path through space. Thousands of millennia hence, the sky will still be filled with twinkling stars, but almost none you see today will be visible. We live in a world of constant and complex motion, a truly dynamic universe. That is the point and message of the tale that follows.

In the succeeding sections and chapters we will examine the appearance and motions of the sky in detail, approaching the subject from two perspectives. First, the sky will be described as it appears to us as observers on the Earth. As we watch, it does not seem that our planet is moving at all, but that it is the *sky* that is in motion. Second, the reasons behind the actions will be explained. That is, we will deal with a given problem not just from an observer's standpoint, but from a physical one as well. In this way the reader will be able to make the generalizations necessary for the understanding of a given phenomenon as seen from any point in time or any place on Earth.

1.1 Angles on a plane

To understand the mechanism of the heavens, we must deal with spheres. Indeed, the term *spherical astronomy* is commonly used to describe much of the subject matter to be covered. The Earth is a very good approximation to a sphere, and the sky can be thought of as an exact sphere. We must be able to define *position* on a sphere so that we can locate London with respect to Buenos Aires, the constellation Orion (the famed celestial hunter: see Chapter 4) relative to the Big Dipper (or the Plough, as it is known in Britain), and Orion with regard to the terrestrial observer.

To measure position, we use the *angle*, which is gauged at the intersections of the line pairs in Figure 1.1. We can measure the angle either interior to the two lines (a, b, c), or exterior to them (a', b', c'). The sum ($a + a'$ etc.) is always 360°, the maximum

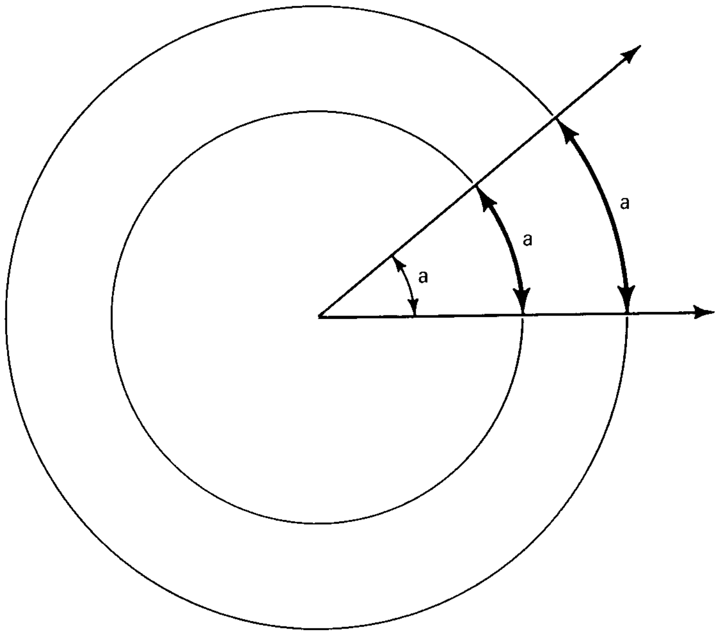


1.1. Various angles, measured at the intersections of two lines. The solid and dashed arrows respectively denote interior and exterior angles, which must sum to 360°. Angles *a* and *c* are respectively acute and obtuse. The 90° angle (*b*) formed by the perpendicular lines is a right angle.

angle that can be measured through or around a full circle. Why 360°? Whence comes this odd number? Why not divide the circle by 100 or 1000 units (as is actually done in some instances)? The reason immediately illustrates the ancient and abiding influence of the sky on daily life. There are 365 days to the year, and 360 is the closest number that is easily divisible into parts. With this definition, the Earth then orbits simply and steadily through an angle centered on the Sun of about 1° every day. Or, more significantly, the Sun, reflecting the Earth's motion, *appears* to move about 1° per day against the backdrop of the constellations.

Certain angles, especially those that quarter the circle, are of special significance. In particular, the 90° angle, formed by two perpendicular lines (Figure 1.1b), is called a *right angle*. Its exterior complement is obviously 270°. Angles less than 90° (Figure 1.1a) are called *acute*, those greater (Figure 1.1c) *obtuse*. The 180° angle is sometimes used synonymously for "opposite." Others – 45°, 60°, 30° – do not carry specific names but are commonly used in setting examples and are important in astronomy's parallel in the pseudoscientific world, astrology.

Angles are also measured around the circumference of a circle. Figure 1.2 shows two lines intersecting to form an angle called *a*. Draw a circle centered on the intersection

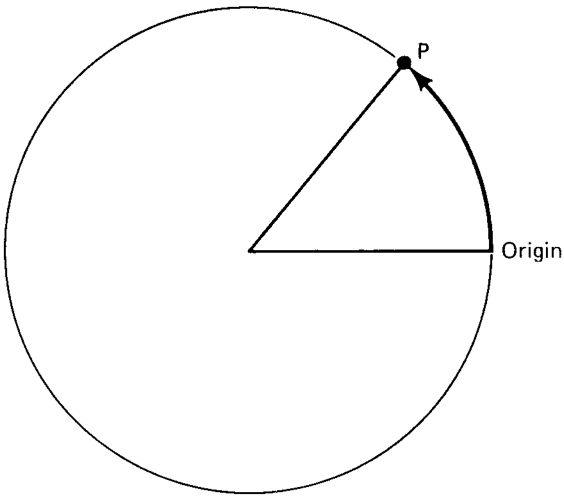


1.2. Arc measurement on a circle. The angle *a* is 40° and the sections of the circles between the lines have arcs of 40° as well. The length of an arc in degrees does not depend upon the physical radius of the circle.

and extend the two lines outward until they intersect it. The lines that form angle *a* define a part of the circle's circumference called the *arc* of angle *a*. We can then measure that section in *degrees of arc*, there being 360° of arc on the entire circle. We say that the arc *a subtends* (is opposite to) the angle *a*, or vice versa. Arc measurement does not depend on the physical dimension, or radius, of the circle, only on the fraction of its circumference defined by the lines. Two circles of different radius are drawn in Figure 1.2 and it is obvious that angle *a* subtends arcs *a* in both cases.

We can now define positions along the circle not by inches or centimeters, which would be dependent on the radius, but through angular measurement. To do so we must define an *origin* or starting point. The origin may be picked with some basis in mind, or it may be selected arbitrarily. Point *P* on the circle (Figure 1.3) can be located precisely and uniquely by specifying its arc around the circle from the origin and the direction of measurement.

The human eye can easily notice an angle of 1°, and specialized instruments can do much better. The degree is subdivided into 60 divisions called *minutes* (written 60′) and the minute into 60 *seconds* (60″). Do not confuse these terms with the more common time units of minutes and seconds. One word is used to express two concepts and the meaning must be taken from context. We might specify, for example, that in



1.3. Position by angle or arc on a circle. P is about 50° from the origin, measured counterclockwise.

Figure 1.3 point P is 50° 16′ 32″.2 from the origin, as measured counterclockwise, or that it is 360° – 50° 16′ 32″.2 from the origin as measured clockwise.

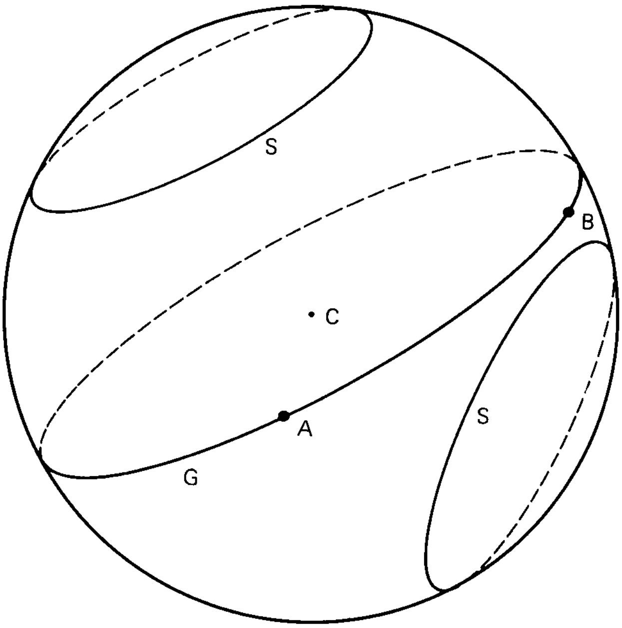
In the subtraction one must always subtract seconds from seconds, minutes from minutes, and degrees from degrees. The larger unit must first be converted into a form to make the subtraction possible. Since 1° equals 60′, 360° can be written as 359° 60′, and similarly this number can be expressed as 359° 59′ 60″. Thus

$$\begin{array}{r} 359^\circ\ 59'\ 60''.0 \\ \underline{-50^\circ\ 16'\ 32''.2} \\ 309^\circ\ 43'\ 27''.8, \end{array}$$

the remainder being the obtuse angle in Figure 1.3.

1.2 Angles on a sphere

These concepts can be transferred, with some modification, from the plane to the sphere. The sphere must be drawn on the flat page: with practice, the perspective will become familiar. Examine the sphere drawn in Figure 1.4. The small dot (labeled C) represents its center, and A and B are two points arbitrarily located on its surface with the singular qualification that they are not on a line with the center. Next draw a circle on the sphere around the center through the points. The circle defines a plane that is the same as those in Figures 1.2 and 1.3. The only difference is that in Figure 1.4 we are looking at it not from overhead but at an angle, so that the circle appears distorted. You can get the same effect in either of the two previous figures if you place your eye



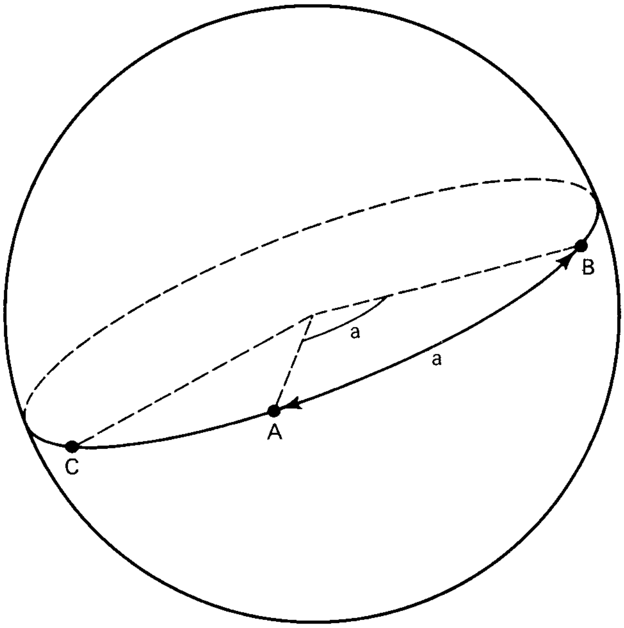
1.4. Great and small circles on a sphere. *C* is the center of the sphere. A great circle (G) is defined as a circle through arbitrary points *A* and *B* that also has *C* as its center. The small circles (S) are not centered on *C*.

near the edge of the page. The sphere is a solid object that may or may not be transparent. The half of the circle through *A* and *B* that passes around the back is shown as a dashed line that may in practice be faint or actually invisible. The result is a three-dimensional visualization of a sphere.

It is obvious that we can draw one *and only one* circle on the sphere through *A* and *B* that has as its center the center of the sphere. If you have any doubts about this restriction try it on a real globe, one of the Earth, or perhaps a basketball. This kind of circle is special. It is called a *great circle* and is marked G in Figure 1.4. By definition a great circle is a circle on a sphere whose center (*C*) is coincident with the center of the sphere. A great circle is a circle of largest radius that can be drawn on a sphere. It is a diametric slice that always divides the sphere into two equal hemispheres. The shortest distance on the surface of the sphere between any two points *A* and *B* is along the great circle connecting them.

Any two points on a sphere, as long as they are not opposite each other, define a unique great circle. There is an infinite number of point pairs, and consequently an infinite number of great circles. Any circle that is not a great circle (one whose center is not coincident with that of the sphere) is called a *small circle*; two of these are labeled S in Figure 1.4.

Measurement of angle on the sphere is the same as it is for the circle. The sphere with the great circle through *A* and *B* reappears in Figure 1.5 with radii

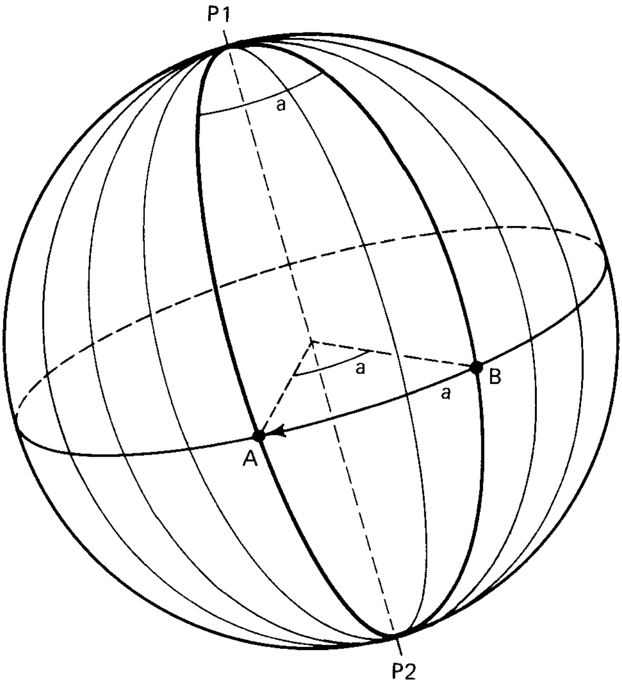


1.5. Angular measurement on a sphere. The angle *a* at the center, defined by the radii to *A* and *B*, is the same as the arc between *A* and *B* measured along the great circle. The great circle is identical to the circle in Figure 1.3 except that it is viewed at an angle. *AB* is about 90°, *AC* about 60°.

extended from the center to each of the points. The plane defined by the circle is identical to that in Figure 1.3 except that we see it from a tilted perspective. The radii define angle *a* at the center, which is numerically equivalent to the arc along the great circle from *A* to *B*.

We might then express the distance between two cities on the Earth in degrees along the great circle that connects them. Try it on a globe. Select London, England, and Montreal, Quebec. Fix a piece of string tightly end-to-end around the globe to construct the great circle through them. Estimate (or measure) the fraction of the string that lies between the cities, multiply by 360° (the full circle) and you find an arc of about 45°.

A globe is rarely readily at hand and it would be cumbersome to use one constantly as you read. In this book we must consequently rely on eye estimates of arc and angle from the perspective representations. In Figure 1.5 the distance between *A* and *B* is about one-half of the solid-line semicircle on the sphere's front half. The angular length of the semicircle is 180°, so the arc between *A* and *B* (and the angle *a*) must be about 90°; that between *A* and *C* is roughly 60°. This kind of estimation is not meant to be precise but only approximate for purposes of learning.



1.6. The poles and secondaries of a great circle. The poles (named *P*1 and *P*2) of the great circle defined by *A* and *B* are the points of intersection between the sphere and the line through its center perpendicular to the circle's plane. Several secondary great semi-circles are drawn through the poles perpendicular to the primary great circle defined by *A* and *B*. The specific secondaries through *A* and *B* meet at the poles with the polar angle *a*, which is numerically equal to the angle *a* at the center and the arc *a* along the primary.

1.3 Poles and secondaries

The practice sphere in Figure 1.5 is repeated in Figure 1.6 with point *B* somewhat shifted. Run a line through the sphere's center perpendicular to the circle's plane. It will exit the sphere at two opposite points that are called the *poles* of the great circle, each of which is in every direction 90° of arc from its parent great circle, that is, the poles are centered on the hemispherical surface created by the great circle. We distinguish between the poles by assigning names, here just *P*1 and *P*2. The most familiar example of this concept is the equator of the Earth. It is clearly a great circle as is evident from a globe. The poles of the equator are simply the north and south poles of the Earth, which define our planet's rotation axis. Every great circle has two opposite poles associated with it, and we will look at many additional examples in the pages to come.

Note that Figure 1.6 is not a true perspective construction. The great circle is seen tilted, yet the poles are drawn *on* the circle that represents the sphere, and are not tilted forward and backward as they would be in real perspective. Such a stylized, schematic rendering of a sphere is traditional in classical astronomy, the perspective being distorted for ease in drawing and learning.

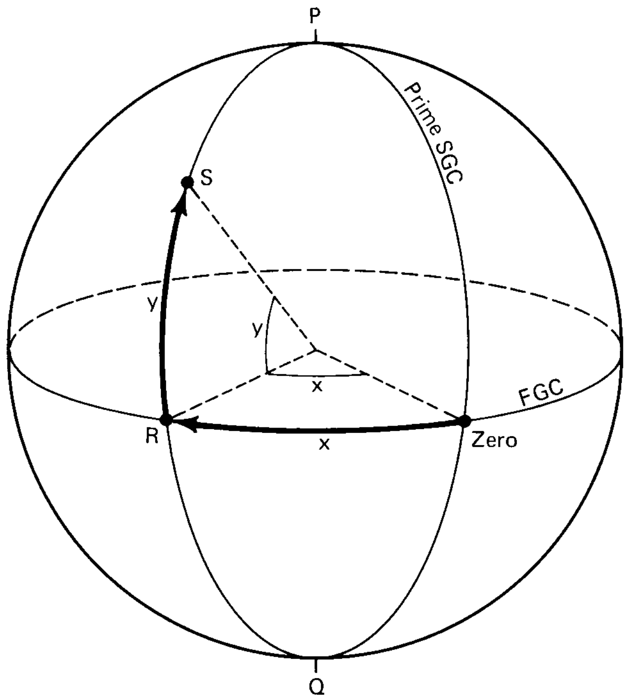
Pick pole *P*1 and point *A* in Figure 1.6. A great circle passed through these two points must be perpendicular to the *primary great circle* (that through *A* and *B*) and must pass through the opposite pole as well. This second circle is called, appropriately, a *secondary great circle*. Every point on the primary defines another secondary, and consequently there must be a whole family – an infinite number – of them. Several are lightly drawn in Figure 1.6, with those through points *A* and *B* singled out as special. Only the front halves are indicated to avoid confusion. Every great circle on the sphere must have its own poles and family of secondaries.

We now have a third way of defining the angle between *A* and *B*. The two secondaries through these points converge at either pole, and make an angle (the *polar angle*) that is also equal to *a*, the angle at the center, and the arc *a* between *A* and *B*.

1.4 Coordinates

We may now use these great and small circles, and the angles that have been defined, for the measurement of place or position on a sphere. On the sphere in Figure 1.7 we draw a great circle that we will designate the *fundamental great circle* (or FGC). This circle may have some physical basis or it may be chosen arbitrarily. The poles are again defined by the perpendicular to the circle's plane. Since we are considering only a generic sphere, they are distinguished by calling them *P* and *Q*; north and south or plus and minus would be equally valid. As before, a set of secondaries must run from pole to pole perpendicular to the FGC.

We are now prepared to construct a *coordinate system* that will generate a pair of numbers and will define any point on the sphere. As an example, we wish to define the coordinates of a point *S* in Figure 1.7. We first establish a *zero point*, or origin, on the FGC from which to measure, as required in Figure 1.3. Again the point may have some significance, or it may be arbitrary. The secondary great circle (or SGC) through the zero point is now a special case and is called the *prime secondary*. From here on, we deal with secondaries only as semicircles. That is, only the semicircle through the origin is called the prime secondary; the part of the circle opposite the origin is not included. We also draw another secondary (semi)circle through point *S*, which intersects the FGC at point *R*. We have now defined two arcs suitable for angular measurement. One is on the FGC between zero and *R*, and is analogous to the arc between *A* and *B* in Figures 1.5 and 1.6. We can obviously go between zero and *R* in two directions, to the right or left of the zero point, so we must specify which direction we want. The other arc is between the FGC at *R* and point *S* along the SGC. Here we have to specify on which side of the FGC we find *S* – toward pole *P* or pole *Q*. We refer to the two arcs as *x* and *y* in Figure 1.7. Every point *S* on the sphere now has a unique set of *x* and *y*, specified according to direction. In the diagram, *S* has coordinates $x = 80°$ (measured clockwise from zero, looking down from pole *P*), and $y = 45°$, measured toward *P*. Note again that direction *must* be specified.



1.7. A general system of spherical coordinates. The FGC is the fundamental great circle, which then defines poles *P* and *Q*. The zero point is arbitrarily chosen. Secondary great circles run from the poles through zero and point *S*. The one through zero is the prime secondary. Arc *x* is measured along the FGC between zero and point *R*, the intersection of the secondary through *S* and the FGC. Arc *y* is measured along the secondary between *R* and *S*. Any point on the sphere can now be located by the unique pair of numbers *x* and *y* plus their directions of measurement.

All points on the FGC have coordinate $y = 0°$. The poles *P* and *Q* do not have a coordinate *x*, since all the secondaries converge there. That is, *x* does not exist at the poles. For the poles, specifying that $y = 90°$ (in the appropriate direction) is sufficient.

1.5 The Earth

The most important example of a coordinate system is the one used to specify position on Earth. The Earth is not a perfect sphere but an *oblate spheroid* with a diameter through the equator (12 756 km) that is 41 kilometers greater than that through the poles. For our purposes here, however, we will assume that the Earth is indeed perfectly spherical, the deviation from the sphere to be addressed in Section 5.2.

How do we indeed know that the Earth is a sphere? At first glance it looks to be flat, and for centuries was considered so. The most obvious evidence is that if we look at an object that is sufficiently distant – the classic example is a ship out at sea – we only