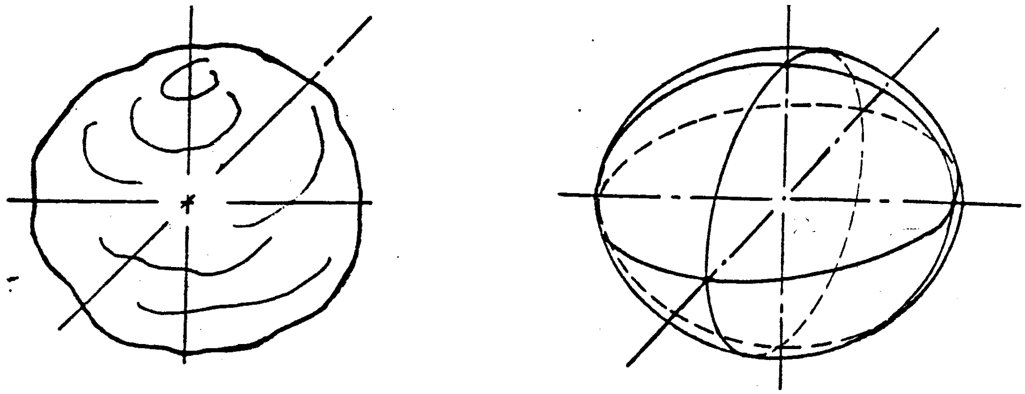
Most surveyors are familiar with the methods of so-called plane surveying. This is surveying which is done on a flat surface or plane. For many types of surveys of limited extent, this system is quite adequate. For example, surveys of small tracts of land and surveys of construction projects that are local in nature can be handled by this method of plane surveying since the effects of the earth's curvature are negligible. However, when we become involved in surveys extending over a large portion of the surface of the earth, we find that we are confronted with an entirely different problem.

Reduced to its simplest terms, we are confronted with the problem of portraying a curved surface on a flat piece of paper. If the surface were curved only in one direction, such as a cylindrical surface, it would not be too difficult to map this on a flat piece of paper. However, the surface with which we are working is curving in all directions at once. If this surface were a spherical one the problem would be bad enough, but the surface with which we are concerned is actually much more complicated than a sphere.

The name given to the shape of the earth is a geoid, which is a highly irregular shape because of the high and low portions of the earth's crust created by ocean basins and mountain ranges. This surface is geometrically so complicated that it would be virtually impossible to map it in a very accurate manner.

As a result, many geodetic surveys are made on the assumption that the shape of the earth is an oblate spheroid. This is the type of solid figure generated by revolving an ellipse around its shorter axis. All of the diameters through the equatorial plane are equal, but the diameter along the polar axis is less than the equatorial diameter.

Occasionally, the shape of the earth is referred to as an ellipsoid. Though this is technically true, the term ellipsoid is better used to describe a solid in which the X-axis, Y-axis, and Z-axis are all unequal. The spheroid is a special case of the ellipsoid in which two of the major axes are equal. See Figure 1.



Geoid Ellipsoid

Figure 1

The spheroid that has been chosen as a basis for state plane coordinate surveys approximates very closely the shape of the geoid itself, so that the minor variations can be ignored. Even so, trying to represent the position of points on this spheroid on a flat piece of paper becomes quite a chore. Cartographers and navigators have been wrestling with this problem ever since men realized that the earth wasn't flat. When we review the work done by early mapmakers we must marvel at the amount of information they were able to develop with the limited tools at their disposal. Many different systems were developed in an attempt to chart this spheroidal surface on paper, some of which were more practical than others. Some of the earliest methods have survived to this day, in essentially their original form, which is certainly a tribute to the genius of these pioneer cartographers.

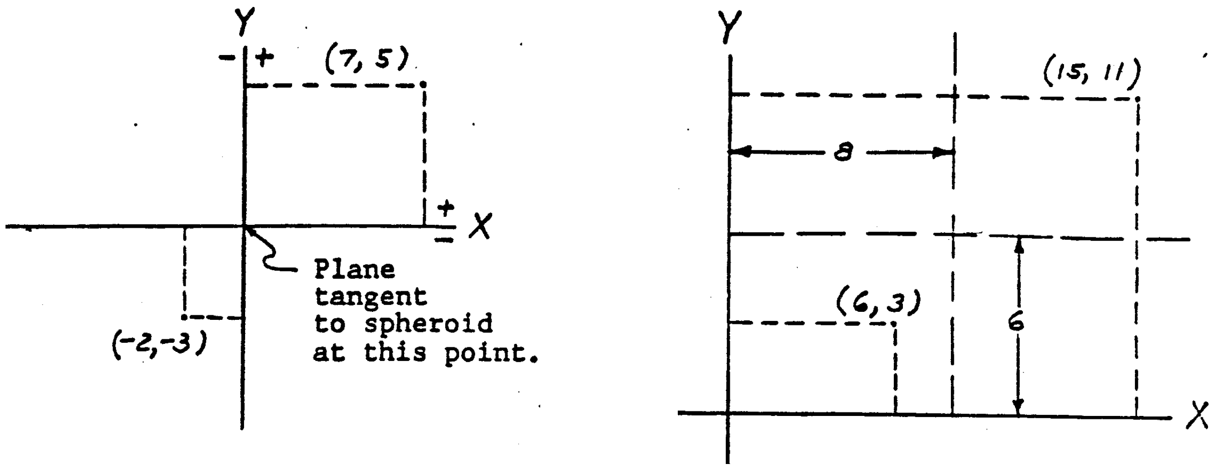
Another unfortunate use of terms should be pointed out at this time. Since we have replaced the geoid with a spheroid, it would be more logical to use a term such as "spheroidal" or "spheroidetic" rather than "geodetic" to described operations carried out on the spheroid. In advanced geodetic work, the actual shape of the geoid is considered together with the magnitude and direction of gravitational forces. For most work with the state plane coordinate systems, these refinements are not necessary. However, the student should be aware that they exist.

We return now to a consideration of mapping procedures as carried out on a plane. Two of the most common systems are the rectangular coordinate system and the polar coordinate system. Most people are familiar with the rectangular coordinate system in one form or another because it has so many applications in everyday life. The development of this system is credited to René Descartes, a French mathematician, and the system sometimes is known by his name, that is, the "Cartesian Coordinate System." Many graphs, progress charts, and schedules are of this form wherein one piece of information is plotted along the base of the chart, and related information is plotted vertically along the side of the chart. For example, a graph of stock market prices can be plotted using the horizontal axis for dates, increasing in number from left to right, and the vertical axis for the corresponding market value or other price information.

In the case of mapping activities, we establish the axes at right angles to each other in the mapping plane. Then the location of any point on this map may be described in terms of its distances from each of these two rectangular axes. Conversely, if we wish to plot a point on this map, knowing its coordinates, we simply measure the required distances from each of the control axes and locate the point accordingly. That is, any point on the map can be represented by a set of coordinates and any set of coordinates represents one and only one point on the map. A given point may be on either side of the axes, and we normally use plus and minus signs to indicate the side intended. Plus values are conventionally above the horizontal axis and to the right of the vertical axis.

When this system is applied to surveying and mapping, we assume the plane to be tangent to the spheroid as some point near the center of the working area. Occasionally, this point is at the intersection of the two coordinate axes, although the major axes may be renumbered to make all of the coordinate values positive. See Figure 2. Sometimes these axes are oriented in a north-south and east-west arrangement, but in other situations they may be rotated to make a more workable system for the project involved.

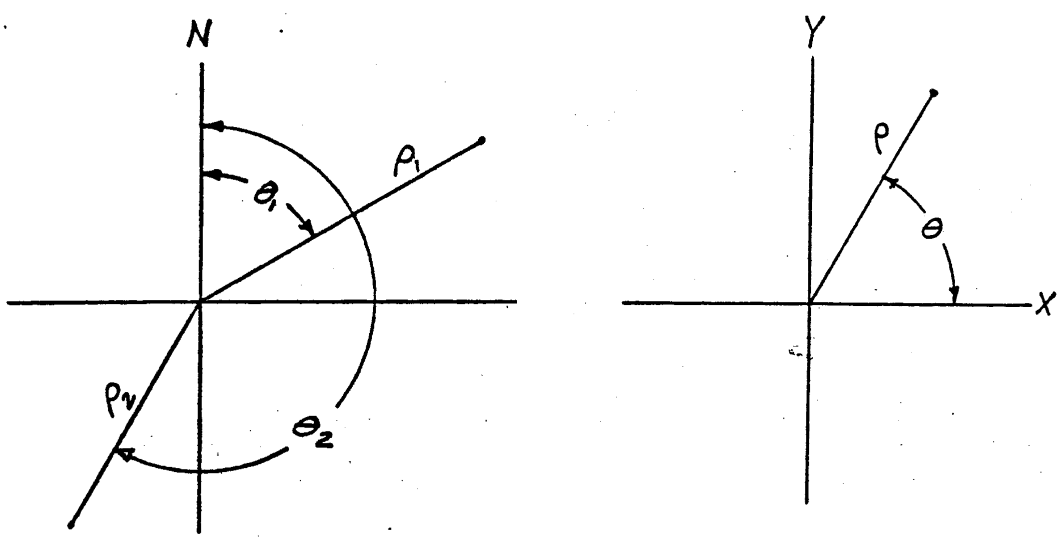
Polar coordinates offer another method of locating points on a plane surface. In this system, a single point is chosen at a convenient location near the center of the work area. A directional reference line is established from this point that may extend northward or in some other convenient direction. Points in the plane are then described in terms or their distance from this origin and the angle measured <u>clockwise</u> from the reference line. It is common in this system to designate the distance from the origin to the point by the Greek letter rho (ρ) and the reference angle by the Greek letter theta (θ). This system also gives us a one-to-one relationship between a location and its coordinates as long as the angle is positive and less than 360°. See Figure 3. It should be noted that in trigonometry the reference angle if usually measured <u>counterclockwise</u> from the <u>X-axis</u>.



Origin is in center of work area.

Axes translated to keep all coordinates positive.

Figure 2



North Azimuth Method

Trigonometric Method

Figure 3

We now turn our attention to the problem of locating and describing points in three-dimensional space. If we expand the rectangular coordinate system by establishing a third major axis, perpendicular to the two we already have, we then find that we can describe the position of a point in space in terms of its distance from each of the three coordinate planes. In this system, as in the plane rectangular system, we have a method for designating whether the point is on one side or the other side of the major axis or coordinate plane. Adopting simple algebraic notation as before, we may then designate the location of any point in terms of x, y, and z coordinates, the necessary additional information being furnished by a plus or minus sign in front of the coordinate. The coordinates (x, y, and z) are the distances from the YZ, XZ, and XY planes respectively. This situation is shown in a perspective view in Figure 4.

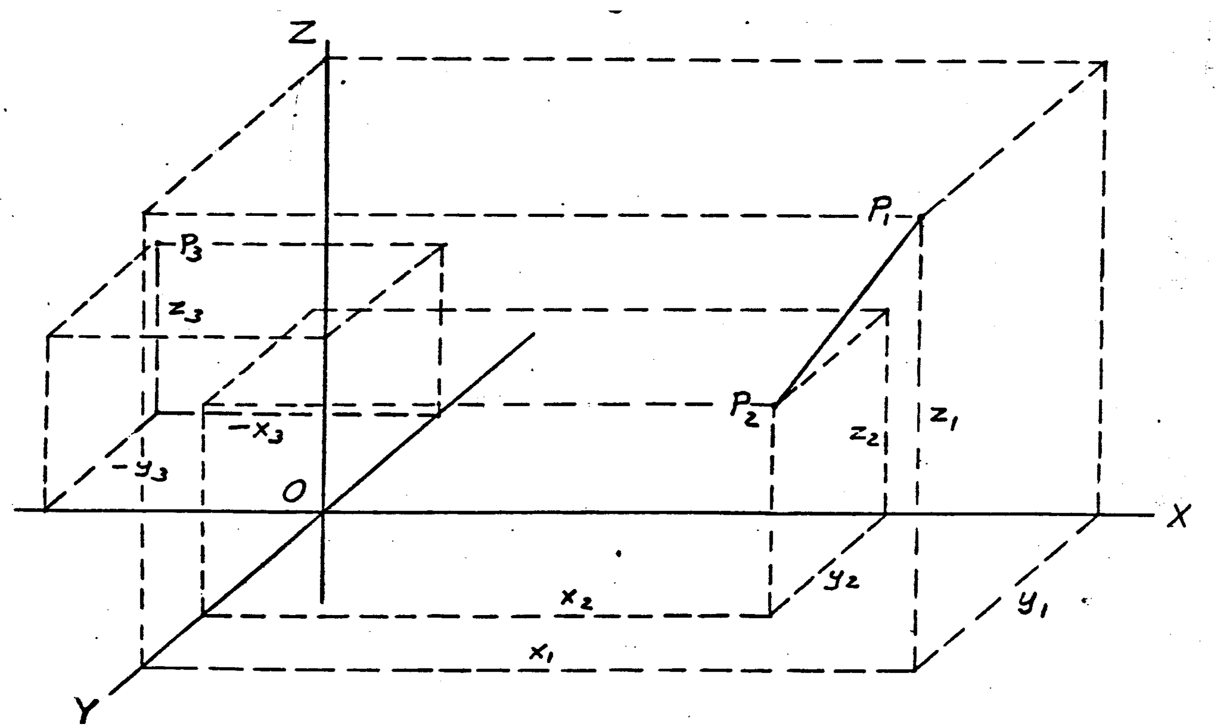


Figure 4

If we now return to the polar coordinate system and expand it into three-dimensional space, we can follow a similar procedure. In this case, we state the distance (rho) from the origin to the point and the angle (theta) between the YZ plane and the plane that passes through the point and the Z-axis. We then add a second angular value that indicates the angle of elevation or depression of our point from the XY plane, as measured in the previously mentioned plane through the point and the Z-axis. This second reference angle is designated by the Greek letter phi (φ). In order to avoid redundant values and maintain the one-to-one relationship, phi should be restricted to values between -90° and +90°. In mapping situations the angle corresponding to theta does not span from 0° to 360°. Instead, it usually goes from 0° to +180° and 0° to -180° (or in more conventional terms: 180° west and 180° east). This angle corresponds to longitude, designated by the Greek letter lambda (λ), while the phi angle corresponds to latitude. In this case the YZ plane represents a plane through the earth's polar axis and Greenwich, England. See Figure 5.

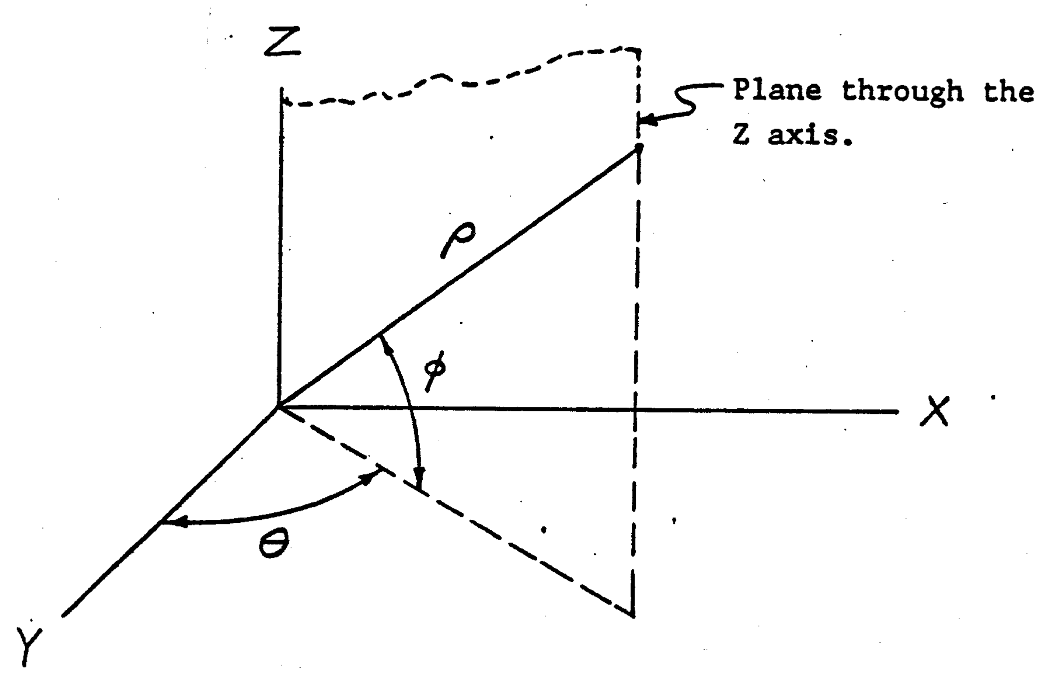
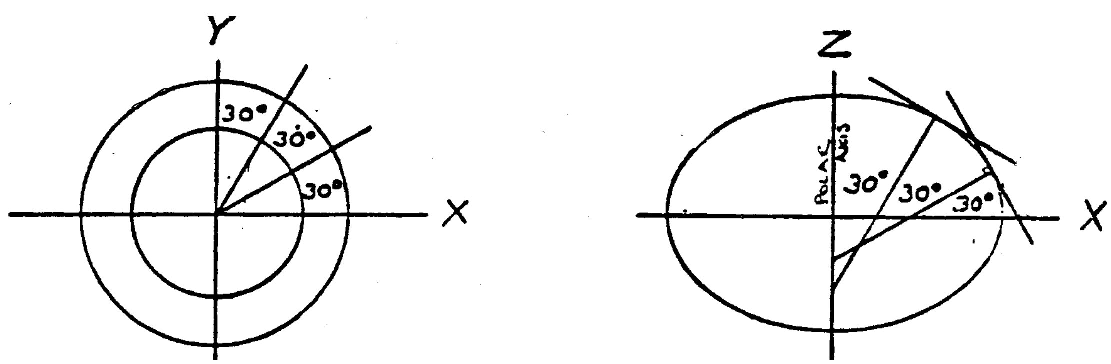


Figure 5

If we are working on a sphere so that the radius (rho) is the same at all points, we can locate any point on that sphere by only two coordinates, namely the longitude (λ) and latitude (φ). Again, we have a one-to-one relationship, namely, every point can be identified by its latitude and longitude, and a stated latitude and longitude identify one and only one point on the sphere. At this point, the student should realize that if the earth were a perfect sphere, instead of a spheroid, mapping procedures would be much easier. In terms of longitude, no problem is encountered with a spheroid, since the spheroid is symmetrical about the polar axis and rho is a constant for a given latitude. However, in terms of latitude, we run into a problem, because a difference in latitude does not correspond to a given distance on the surface of the spheroid, and the radius varies with the latitude for any given longitude. This situation is shown in Figure 6.



For a constant difference in <u>longitude</u>, the distance along any <u>latitude</u> circle is constant.

For a constant difference in <u>latitude</u>, the distance along a <u>longitude</u> ellipse is <u>not</u> constant.

Figure 6

If maps were made on spheroidal surfaces, such as on the surface of large globes, many of the problems of mapping would be taken care of automatically. In this case, we would simply be making a scale model of a portion of the surface of the earth in which the dimensions of our model would be proportional to the actual earth in all directions. In most cases, practical considerations do not permit this to be done and make it more desirable to put the map on a flat surface. Hence, we have to have some other way of transferring or projecting the information from the spheroidal surface to a plane surface. The methods by which this is done are known as map projections.

The physical problem involved is often illustrated by the example of trying to flatten out a piece of orange peel. If a small piece of the orange peel is used, such as half an inch square, it can be flattened out without a great deal of distortion. However, if a larger piece is used, such as a third or half of the orange, and an attempt is made to flatten it out, it obviously cannot be done without tearing and distorting the orange peel in many places. If we have a map drawn on our orange peel and then flatten it out, introducing many gaps and tears, we will find that our map would be of vary little value, since it will be very difficult, if not impossible, to relate the original position of a point on one fragment of the torn peel to the original position of a point on another fragment. In spite of this fact, many map projections look pretty much as if someone had indeed stepped on an orange peel and flattened it out. Several typical examples are shown in Figure 7.

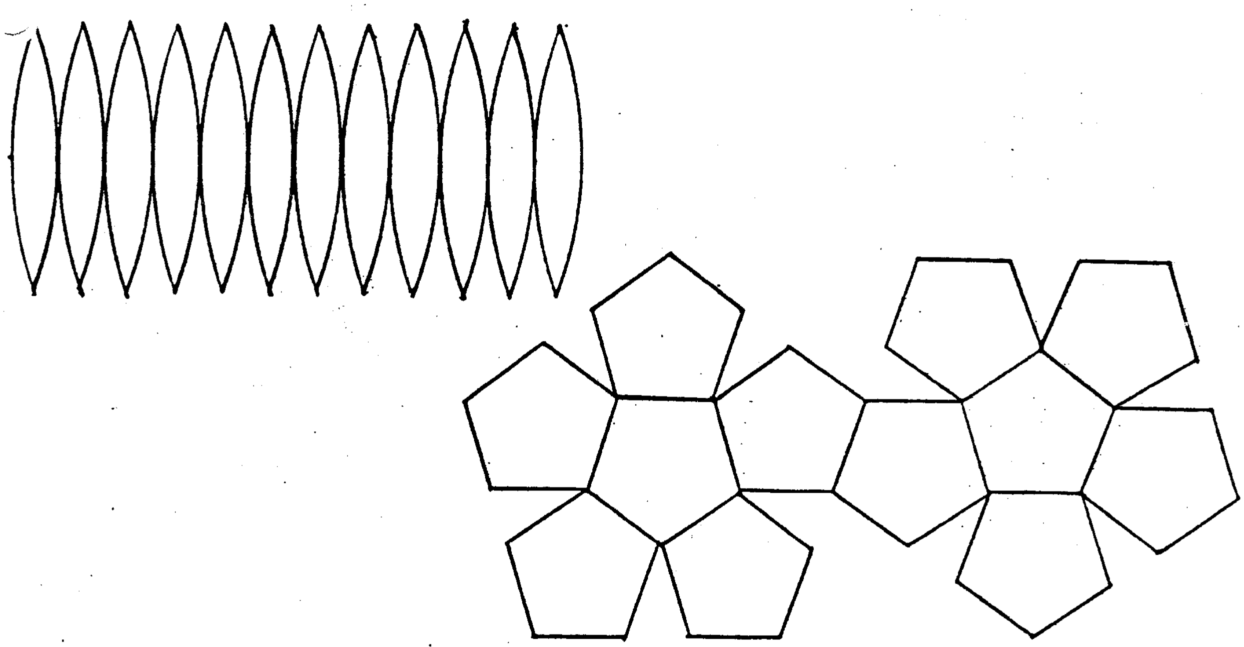
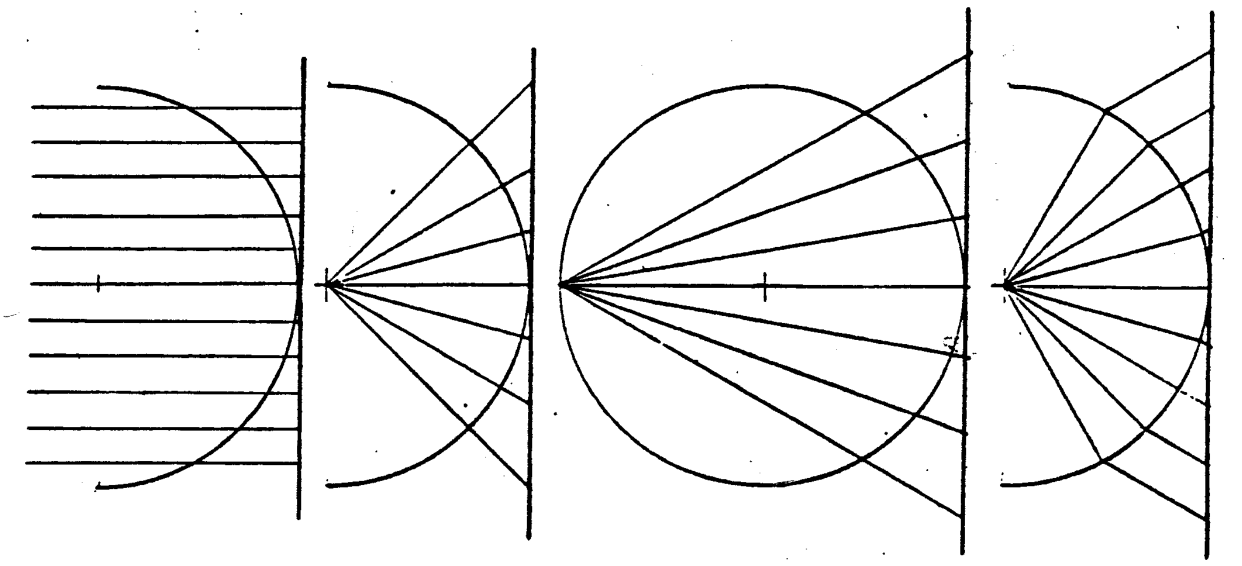


Figure 7

The more useful map projections are those in which there is some convenient graphical or mathematical relationship between the points on the map and the points on the earth. In mapping techniques, there are several characteristics that are considered desirable. For example, directions and distances on the earth should correspond to the directions and distances shown on the map, and vice versa; areas on the earth should be in constant ratio to areas on the map and vice versa; angles between intersecting lines on the sphere should be the same as the angles between the corresponding lines on the map; and if possible, great circles on the earth should appear as straight lines on the map. Unfortunately, cartographers have had to admit that all of these desirable features cannot be attained in one map projection. Hence, the various map projections which have been developed concentrate on meeting one of the basic criteria and ignore the others, or attempt to reach a compromise by retaining various amounts of several desirable characteristics.

The type of map projection that is used will be determined a great deal by the use to which the map is to be put. Most of the map projection methods consist of transferring the information on the spheroid to a plane or to some sort of solid figure that can be “developed” into a flat figure or flat surface. The most common geometric figures used for this purpose are the cone and the cylinder, and in a sense, the cylinder may be considered as a special case of the cone in which the vertex is at infinity. Either the cone or the cylinder may be cut along one of its elements and then unrolled or “developed” to form a plane surface. If a large portion of a spheroid is to be mapped, a series of cones or cylinders may be employed in order to keep the amount of distortion in each area to a minimum.

We will first consider projecting the curved surface directly onto a plane, and for the time being we will assume that we are working with a spherical rather than a spheroidal surface. This method consists of constructing a plane surface tangent to the sphere at some convenient point and then projecting image points from the curved surface onto the plane. The projection lines in this case could be parallel lines, radial lines, or lines that meet some other geometric criterion. However, there may be no actual projection lines as such; in which case, the transfer of points from one surface to the other is accomplished by having the corresponding points meet certain mathematical requirements on their respective surfaces. Some of the geometric possibilities are shown in Figure 8, and the mathematical possibilities will be covered in more detail under "Transformations."

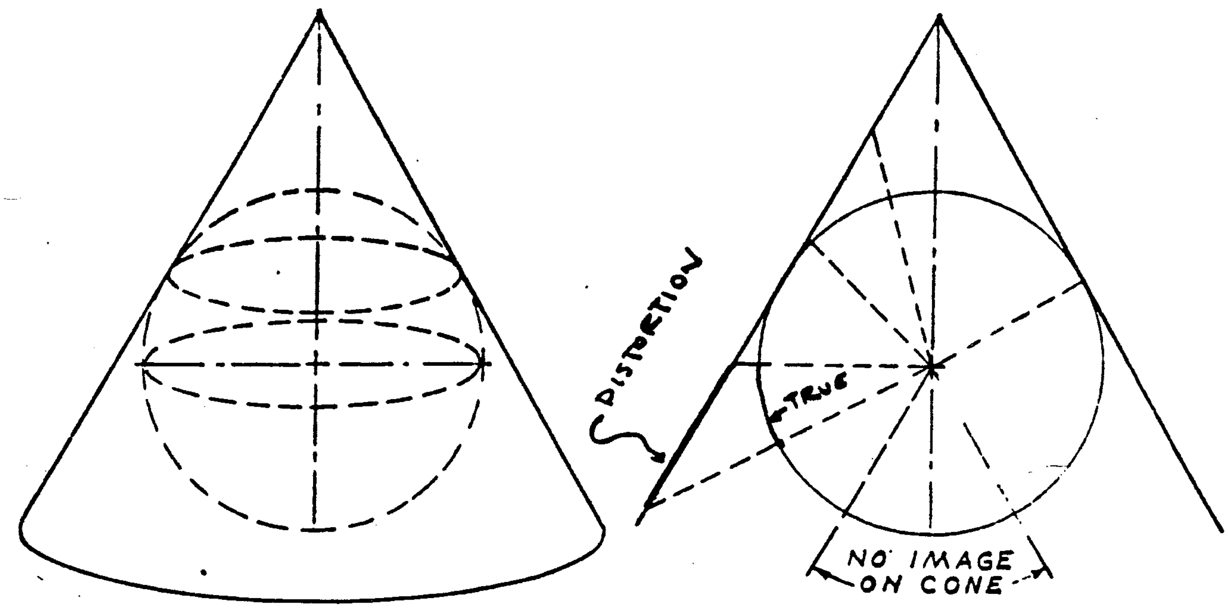


Stereographic Projection

Gnomic Projection

Figure 8

The next projection we will consider is known as the tangent cone method, in which an imaginary cone is constructed with its vertex on the prolongation of the earth's polar axis and with the cone tangent to the earth at some circle of latitude. This is shown in Figure 9A.

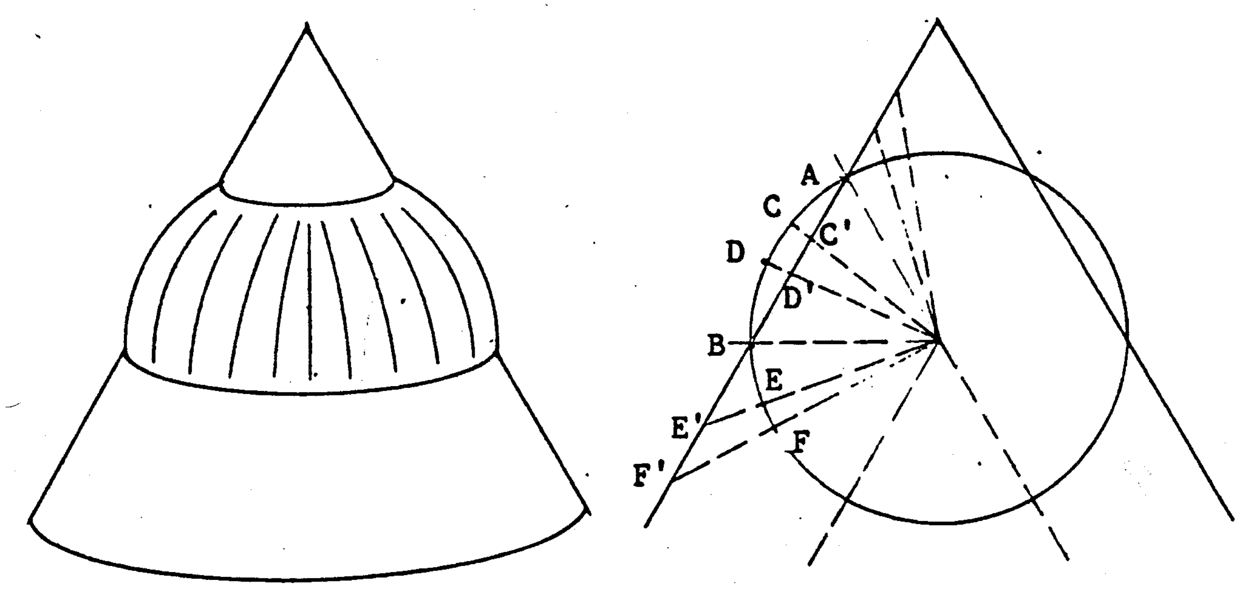


A B

Figure 9

If you now imagine that the earth is a large glass ball with the map features drawn on the surface of it and imagine a very bright light located at the center of the earth, you can see that the image of these map features would be projected onto the surface of the cone. In the vicinity of the line of tangency between the sphere and the cone there would be a fairly accurate relationship between the points on the sphere and the points on the cone. But as we get farther above or below this line of tangency, we see that there will be more and more distortion as the image from the glass ball is projected onto the surface of the cone. In fact, as we get farther toward the South Pole, we reach a condition where the light rays from the center of the earth would not strike the cone at all. Hence, the vicinity of the South Pole could not be mapped on this cone with this type of projection. This is shown in cross section in Figure 9B.

The next refinement is to consider the so-called “secant cone” or cutting cone. This is an imaginary cone that cuts through the surface of the sphere along two circles of latitude as seen in Figure 10A. In this case, the analogy to the glass ball and the projection light breaks down somewhat, because, as is shown in Figure 10B, the portion of the sphere between A and B is outside of the cone. However, we could replace the light rays with straight lines and then consider the intersection of these straight lines with the cone as they are drawn from the center of the earth to a particular point on the sphere. In this case it would be seen that distances between A and B which are “projected” inward to the cone will be shorter than the original distances on the sphere, whereas, distances above A and below B when projected from the sphere to the cone will be longer. Hence, C’D’ is shorter than CD and E'F' is longer than EF.



A B

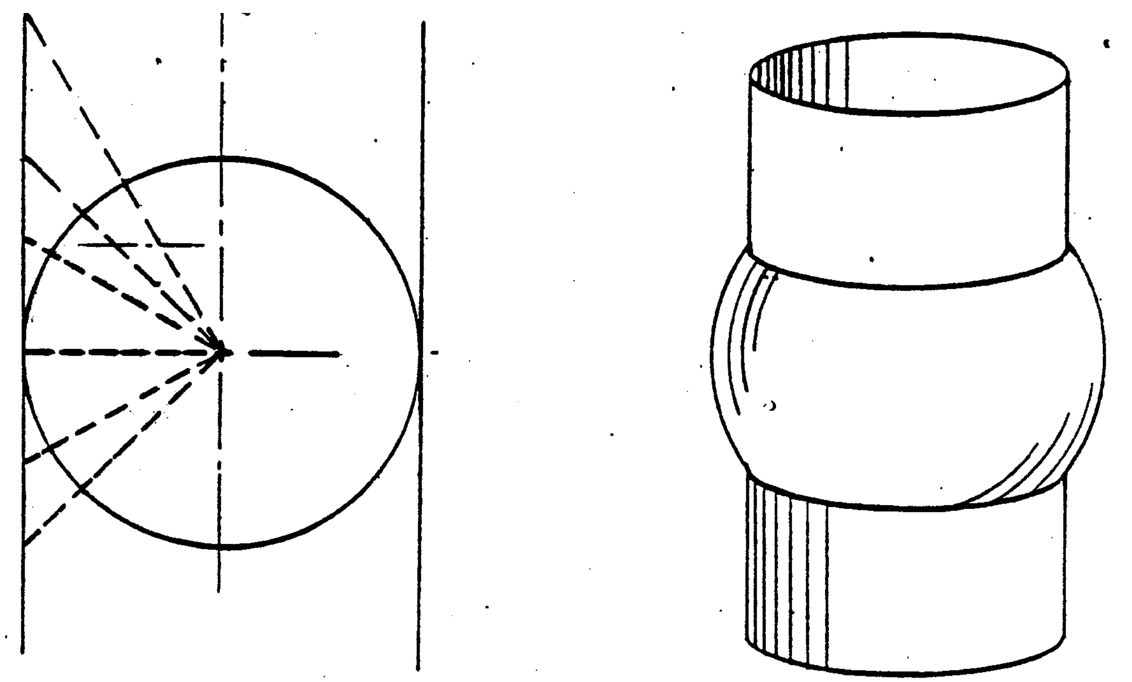
Figure 10

The advantage of the secant cone method over the tangent cone method is that it is possible to map or "project" a wider band without encountering excessive distortion than in the case of the tangent cone. The limitations as far as mapping the southern part of the earth would be the same as in the previous case.

While it is convenient in the case of the plane or conic projection to have the projection surface symmetrical about the polar axis, this is not essential. Projections that do not meet this criterion are called “oblique” projections.

As we mentioned earlier, the cylinder is a special case of the cone that has its vertex at infinity. (It is interesting to note also, that the plane represents the opposite extreme, wherein, the vertex coincides with the point of tangency.)

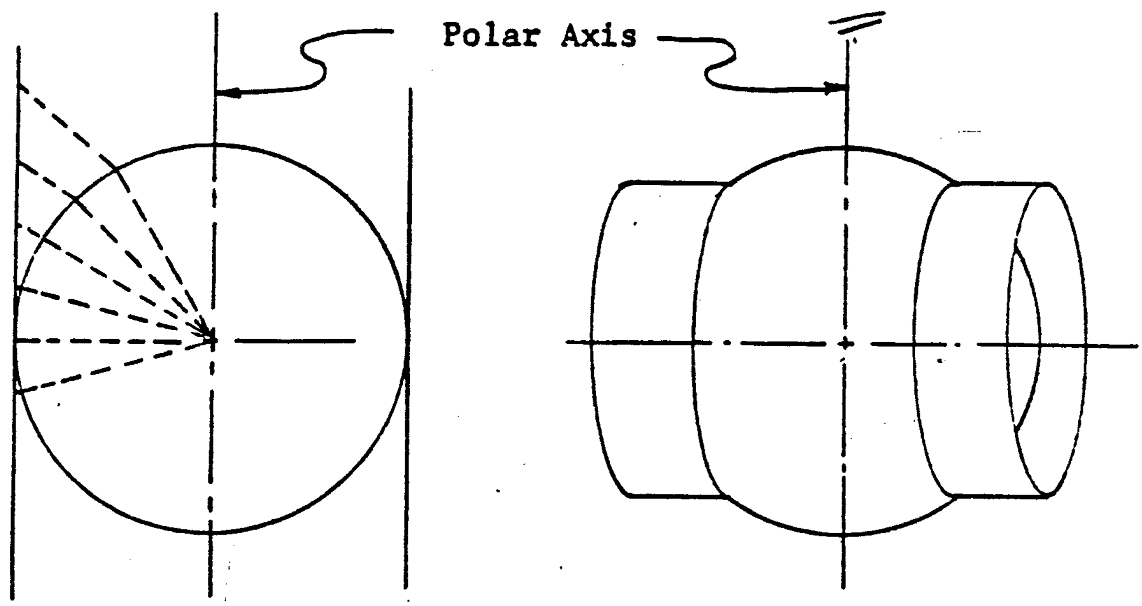
As in the case of the usual conic projections, cylindrical projections may be either on a "tangent" cylinder or a "secant" cylinder. These are shown diagrammatically in Figure 11A and 11B. The same "projection light" analogy used for cones can also be used for cylinders. In this case, we see that there are only two points that will not be projected onto the cylinder. These are the two points at which the axis of the cylinder intersects the sphere.



A B

Figure 11

It will be noted with this system also, that as we progress away from the line or lines of tangency, a great deal of distortion is encountered. As in the case of the plane projections, “bending” the projection lines according to some geometric or mathematical criterion can reduce this. One geometric solution, in which the amount of bending is a function of the latitude, is shown in Figure 12A.



A B

Figure 12

As in the case of the plane and conic projections, the axis of the cylinder need not coincide with the polar axis of the sphere. If the angle between these two axes is between 0° and 90°, the projection is called an oblique projection. If the angle is 90°, the projection is called a <u>transverse projection</u>. See Figure 12B.

Returning now to a consideration of the earth as a spheroid, we see some interesting refinements. If the axis of the cylinder coincides with the polar axis of the spheroid, and if we use a tangent cylinder, the line of tangency will be a circle, namely the equator. If we use a secant cylinder with a similar orientation, the lines of intersection will be two circles of latitude.