APPENDIX H

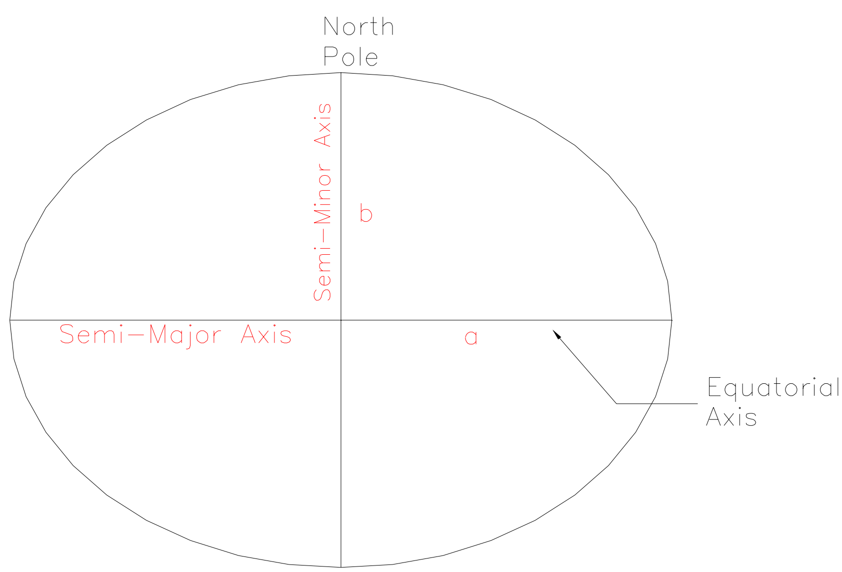
INTEGRATING GPS OBSERVATIONS WITH CLASSICAL SURVEYING MEASUREMENTS

H.1 The Model

The latter half of the twentieth century has brought about rapid and significant technological change to surveying and mapping. Geodesy matured from precision taping and triangulation to traversing and trilateration with the advent of electronic distance measuring (EDM) equipment to very accurate positioning with the global positioning system (GPS).

With these technological achievements came increased accuracy in our determination of the shape and size of the earth. While it is well known that the earth can be mathematically approximated by an ellipsoid of revolution, the increased precision of our measurement tools also began to highlight the discrepancies between the real earth and its math model.

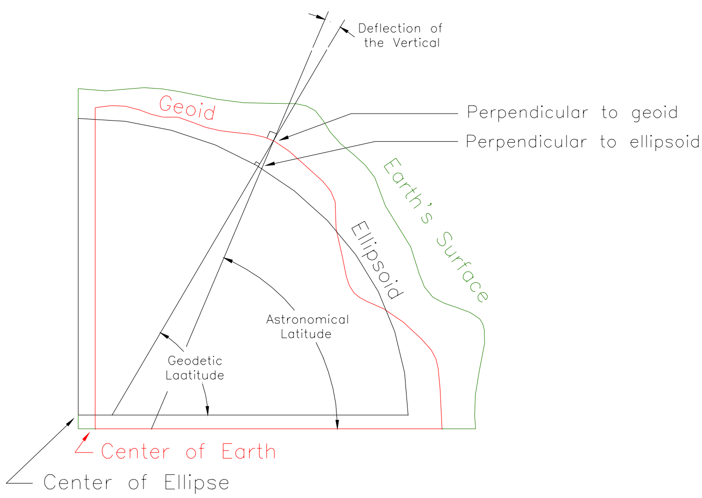
Figure H.1
Geometry of the ellipse



The ellipsoid is shown in figure H.1. The distance from the center of the ellipse to the surface along the equatorial plane is called the semi-major axis (a) while the distance to the poles is referred to as the semi-minor axis (b). It is also well established that a > b making the earth flattened at the poles and elongated at the equator.

While the ellipse is very good at defining location, a different model is used to measure elevations. Here the geoid is used (figure H.2). The geoid is an equipotential surface where, at every point, the surface is perpendicular to the direction of gravity. The difference between the normal to the ellipsoid and the normal to the geoid (this is the vertical line or the direction of gravity) is called the deflection of the vertical (δ).

Figure H.2
Relationship between the geoid, ellipsoid, and the earth's surface.

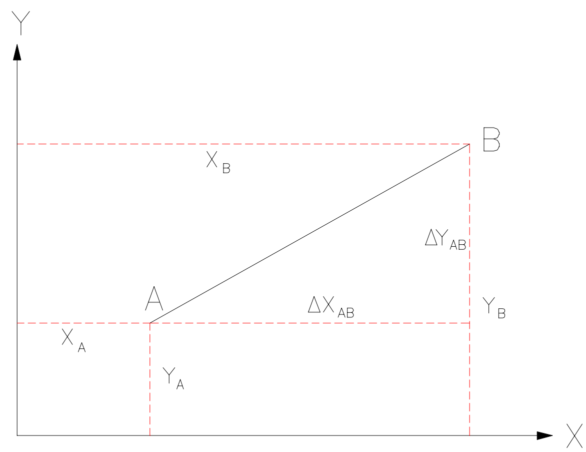


H.2 Coordinate Systems

To locate features or points in space, some type of coordinate system is necessary. Thinking back to the Cartesian coordinate system that is often employed in surveying, one can see that locating the X and Y coordinates of any two points gives the user an excellent ability to obtain other types of information (figure H.3.). For example, if the X-axis is oriented to the east and the Y-axis to the north then, using the coordinates of points A and B, the direction of the line from A to B and its distance can easily be computed. The distance is

$$D_{AB} = \sqrt{(X_B - X_A)^2 + (Y_B - Y_A)^2}$$

Figure H.3
Cartesian coordinate system.



and the azimuth becomes

$$Az_{AB} = \tan^{-1}\left[\frac{X_B - X_A}{Y_B - Y_A}\right]$$

Thus, a coordinate system is essential in establishing the relationship that exists between points. Since we know that the earth is an ellipsoid, the establishment of a different coordinate system is necessary.

H.2.1 Geodetic Coordinate System

The geodetic coordinate system is broken down into two components: horizontal and vertical coordinate systems. The location of a point on the ellipsoid is found using latitude and longitude (figure H.4). A meridian can be defined as the cross-section of an ellipse that contains both the poles. For convenience, one meridian is designated as the prime meridian (Greenwich meridian) from which all other meridians are referenced. The longitude is the angle measured in the equatorial plane from the prime meridian to the meridian containing the point of interest (called the local meridian).

Geodetic longitude is designated as λ. Geodetic latitude (φ) is defined as the angle measured in the meridian from the intersection of the line normal to the ellipsoid at the point with the equatorial plane.

Height is a little different in that there are two possible reference surfaces from which it can be measured. Elevation is defined as the distance along the vertical line from the geoid to the point in questions. Figure H.5 shows that a point on the surface can be expressed in terms of an ellipsoid height or an orthometric height (used for elevations). The ellipsoid height (h) is the distance

Figure H.4
Geodetic latitude and longitude

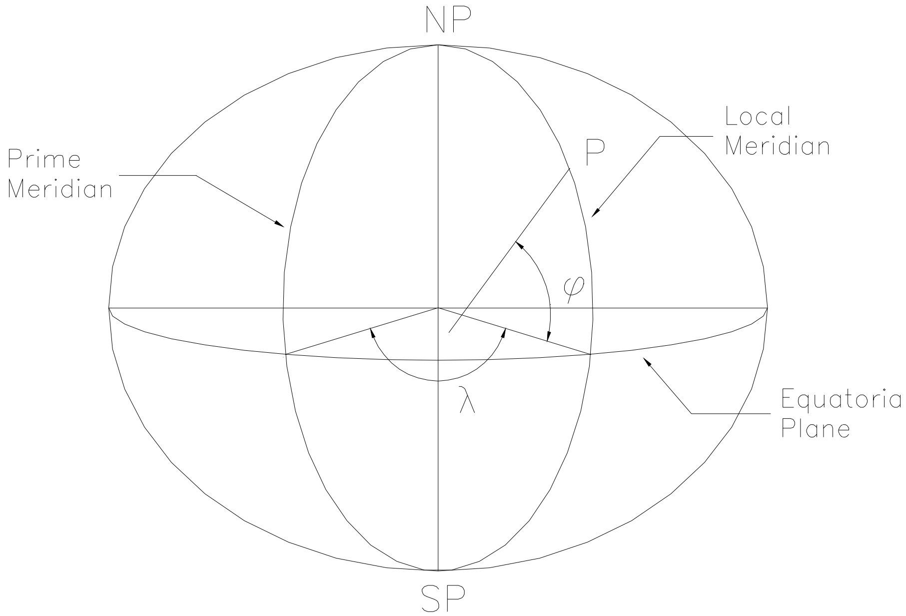
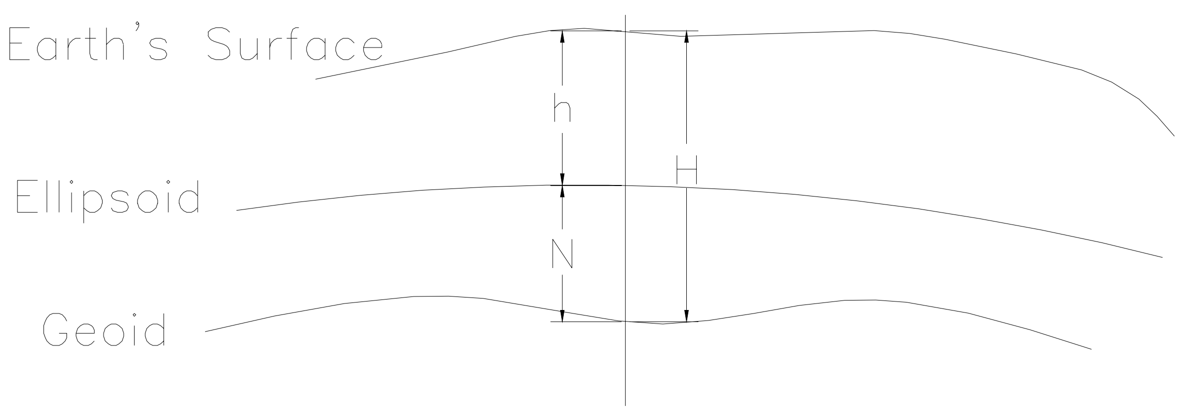


Figure H.5
Relationship between ellipsoid height, orthometric height, and the geoid separation.



measured from the normal to the ellipsoid to the point whereas the orthometric height (H) is the distance along the vertical from the geoid to the point.

The distance between the geoid and the ellipsoid is generally referred to as the geoid separation or geoid undulation. From figure H.5, it is very evident that the geoid separation can be easily computed using the relationship:

$$N = h - H.$$

The relationship between the orthometric and geoid height is dependent upon the geoid model being used. NGS has developed a series of refinements to the geoid model. The last two models were called GEOID96 and GEOID99. This new geoid model is based on over 2 million points and has an estimated accuracy of about 0.15 ft. This compares to the older GEOID96 model with an spproximate accuracy of 0.18 ft. It is recommended that the new GEOID99 geoid model be used for MDOT surveys.

H.2.2 Geocentric Coordinate System

Another method of describing the location of a point is to use a three-dimensional Cartesian coordinate system from the origin of the ellipsoid. This is referred to as the geocentric coordinate and it is depicted in figure H.6. The coordinates can be computed using the following relationships:

$$X = (N + h)\cos\varphi\cos\lambda$$

$$Y = (N + h)\cos\varphi\sin\lambda$$

$$Z = \left[\left(1 - e^2\right) N + h\right]\sin\varphi$$

Given the X, Y, and Z coordinates of a point, computing the latitude, longitude, and height is performed in an iterative fashion using the following formula:

$$\varphi = \tan^{-1}\left[\frac{Z}{\sqrt{X^2 + Y^2}}\left(1 - e^2\frac{N}{N + h}\right)^{-1}\right]$$

$$h = \frac{\sqrt{X^2 + Y^2}}{\cos\varphi}$$

$$\lambda = \tan^{-1}\frac{Y}{X}$$

H.2.3 Topocentric Coordinate System

A local Cartesian coordinate system is referred to as a topocentric coordinate system. Here the reference system is taken to be tangent to the site area. For convenience, the coordinate system is oriented such that the X-direction is to the east and the Y-direction is north. The coordinates are referred to as northings and eastings. Height is the up direction.

H.3 Conversion Strategies

Plane surveying offers the advantage of simplifying calculations and is extremely efficient on sites of limited extent. Unfortunately, in large area project, such as highways, this simplification leads to problems. Because of the distortions involved with working on a plane, highway surveying and mapping projects need to take into account the curvature of the earth. This is explained in more depth in **Part IV, Coordinates and Map Projections**. When the United States Coast and Geodetic Survey performed their adjustment of the North American datum in 1927, they selected a particular ellipsoid, Clarke's 1866 Ellipsoid, that fit the conterminous U.S. very well. Moreover, a single point, Meades Ranch, was picked as the origin of this system.

Conversion of coordinates with the N-E-U system to another rectangular coordinate system is pretty straight forward and can be easily be done by hand. The easiest process is to perform an affine transformation to the data. This transformation is ideal in that the transformation can include a number of different parameters that can account for non-orthogonality of the axes and differential scale factor along different axes. Here data is required in both systems from which the transformation parameters can be computed and the new coordinates calculated.

In 1983 the National Geodetic Survey (NGS) re-adjusted the North American datum. This was a long project for the NGS. The actual completion date was 1986. In the planning process, NGS recognized the importance of satellite geodetic observations on the national network. This realization came because of the Survey's activities in very long baseline interferometry and Doppler satellite observations. GPS was not operational at this time. Since these space borne systems utilize an orbit that is the center of mass of the earth, it was decided to base this new datum on the World Geodetic System of 1984. This datum was earth-centered and earth-fixed.

There is no direct correspondence between the NAD 27 and NAD 83 datums. Therefore, any transformation to convert coordinates in one system to the other is only an approximation. The NGS has analyzed the shift between NAD 27 and NAD 83. From this analysis of control points with values in both systems, transformation parameters have been developed. In addition, the program NADCON is available through the NGS that will convert coordinates from NAD 27 to NAD 83.

Another program developed by the U.S. Army Corps of Engineers is also available. It is a little easier to use because it is Windows-based. This program, called CORPSCON, is based on the NGS NADCON program. There is an inherent weakness in using these programs. The model is based on a global model. In other words, it is a best fit of the NAD 83 datum to the NAD 27 datum. It does not take into account a denser control network that may be in place in a local network. Thus, alternative conversion routines should be used for local transformations.

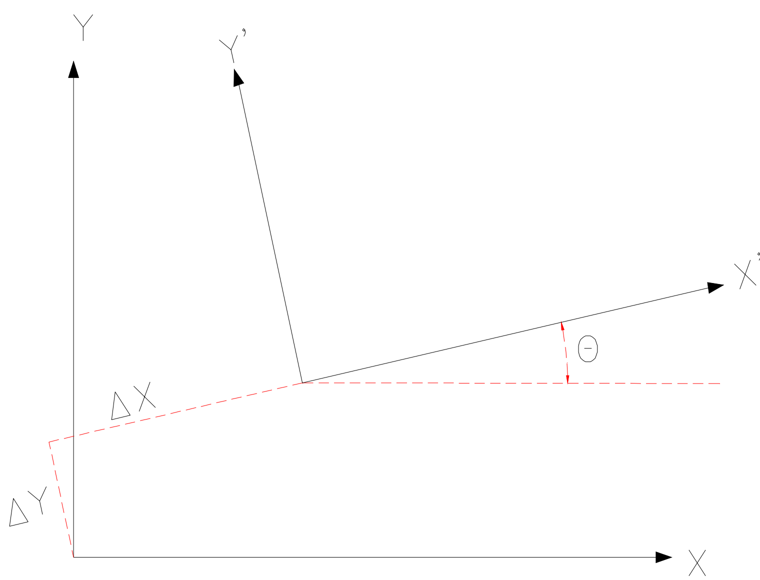
A very simple and common two-dimensional transformation is the affine transformation. This is usually given as:

$$X' = SX\cos\theta + SY\sin\theta + \Delta X$$

$$Y' = -SX\sin\theta + SY\cos\theta + \Delta Y$$

where θ is the rotation angle between the two coordinate systems (figure H.6), X and Y are the original coordinates, X' and Y' are the transformed coordinates, S is the scale factor, and ΔX and ΔY are the translations between the centers of the two coordinate systems. If we designate a = S cos θ, b = S sin θ, c = ΔX, and d = ΔY, then the transformation is shown to be

Figure H.6
Relationship between two coordinate systems for the affine transformation



$$aX + bY + c = 0$$

$$-bX + aY + d = 0$$

The weakness of the affine transformation is that it is a simplification of the model. This transformation works for areas of limited extent but fails as the area grows in size. A better approach, particularly when using the 3-D GPS mode, is to perform the transformation in a series of step. These transformations involve going from geodetic to geocentric to local rectangular space and the inverse, depending upon the projects needs.

The transformation from geodetic (φ, λ, h) to geocentric (X, Y, Z) follows the well-known formulas:

$$X = (N + h) \cos\varphi \cos\lambda$$

$$Y = (N + h) \cos\varphi \sin\lambda$$

$$Z = \left[N(1 - e^2) + h\right] \sin\varphi$$

where: N is the radius of curvature in the prime vertical and e is the eccentricity of the ellipse. The radius of curvature is computed as:

$$N = \frac{a}{(1 - e^2 \sin^2\varphi)^2}$$

The transformation from geocentric to geodetic involves an iterative approach for the computation of the latitude. The longitude is computed directly as:

$$\lambda = \tan^{-1}\frac{Y}{X}$$

Iterate to compute the latitude. The initial estimate of the latitude is

$$\varphi_o = \tan^{-1}\left(\frac{Z}{\sqrt{X^2 + Y^2}}\right)$$

From this, compute an adjusted value for the final values of the latitude and height. With the new value for the latitude, check with the current estimate (φ_o). If the difference is larger than the desired difference, then replace the estimated value with the new value and recompute. These formulas are given as:

$$h = \frac{\sqrt{X^2 + Y^2}}{\cos\varphi} - N$$

$$\varphi = \tan^{-1}\left[\frac{Z}{\sqrt{X^2 + Y^2}}\left(1 - e^2\frac{N}{N + h}\right)^{-1}\right]$$

There are numerous solutions to this problem. One non-iterative approach is based on the theory developed by Bowring and it is shown as follows:

$$\tan\varphi = \frac{Z + e^2 a\sin^3\mu}{\rho - e^2 a\cos^3\mu}$$

$$h = \rho\cos\varphi + Z\sin\varphi - \left(\frac{a^2}{N}\right)$$

where a is the semi-major axis of the ellipse and ρ and μ are computed from

$$\rho = \sqrt{X^2 + Y^2}$$

$$r = \sqrt{\rho^2 + Z^2}$$

$$\tan\mu = \frac{Z(1 - f)}{\rho}\left(1 + \frac{e^2 a}{r}\right)$$

The transformation from geocentric (X, Y, Z) to local rectangular space (x, y, z) involves both a rotation and translation to the origin. It is given as:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} X - X_o \\ Y - Y_o \\ Z - Z_o \end{bmatrix}$$

where R is the rotation matrix defined as:

$$R = \begin{bmatrix} -\sin\lambda & \cos\lambda & 0 \\ -\sin\varphi\cos\lambda & -\sin\varphi\sin\lambda & \cos\varphi \\ \cos\varphi\cos\lambda & \cos\varphi\sin\lambda & \sin\varphi \end{bmatrix}$$

The transformation from the local rectangular coordinate system to the geocentric coordinate system is the inverse of the relationship above. Since R is an orthogonal matrix, the transpose can be used instead of the inverse. Thus, it is presented as:

$$\begin{bmatrix} X - X_o \\ Y - Y_o \\ Z - Z_o \end{bmatrix} = R^T \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

The NAD 83 adjustment utilized historical survey data in the adjustment of the North American datum. Since that time, GPS has become the instrument of choice for geodetic surveys. The results of these GPS surveys are far superior than the expected results from the initial NAD 83 adjustment. To aid the surveyors in achieving higher accuracies with GPS, all of the states have initiated a High Accuracy Reference Network (HARN). NGS has been forced to adjust these networks for each of the participating states. The HARN system is still based on the NAD 83 datum. Therefore, to assist the user in defining what set of coordinates are being reported, the adjustment date is also included in the definition of the reference framework. For example, NAD 83 (86) coordinates refer to the initial adjustment of the 1983 datum. NAD 83 (94) coordinates refer to the adjustment based on the HARN system in Michigan. As our ability to define the datum increases in accuracy, so will the need to readjust these networks in the future. There is also a push to make the NAD 83 system agree to the International Terrestrial Reference Framework (ITRF).