GRID, GROUND, AND GLOBE: DISTANCES IN THE GPS ERA

THOMAS H. MEYER

1. Preamble

This section reviews several fundamental prerequisite concepts that will play a central role in all that is to follow.

1.1. **Distance.** The length l along a smooth curve C between distinct points **A** and **B** can be expressed as

$$l = \int_{\mathbf{A},\mathbf{B}} ds, \tag{1}$$

where ds is an infinitesimal chord. Obviously, there are an infinite number of curves between any two points. We define the **distance** between distinct points **A** and **B** to be the length of the shortest curve between **A** and **B**. Such a curve is called a **geodesic** and we will denote this curve by $\mathcal{G}$ and its length by d.

Eq. (1) often has a simple solution when $\mathcal{G}$ falls on a surface of constant curvature. If $\mathcal{G}$ is on a plane, then $\mathcal{G}$ is a straight line and (1) becomes

$$d = [\,(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2\,]^{1/2} \tag{2}$$

for the three-dimensional case. If $\mathcal{G}$ is on the intersection of a plane and a sphere, then $\mathcal{G}$ is a circular arc and (1) becomes

$$d = r \cdot \theta, \tag{3}$$

where r is the radius of the circle in the plane and θ is the angle subtended along the arc. When the plane includes the center of the sphere, the circle is called a **great circle**. It is easy to show that a great circle distance is minimal on a sphere so $\mathcal{G}$ is the shorter of the two great circle arcs connecting **A** and **B** on the sphere. A detailed discussion of great circle distances is given in section 4.5 on page 26.

If $\mathcal{G}$ falls on a surface that does not have constant curvature (e.g., an ellipsoid or a spheroid), then $\mathcal{G}$ is a complex curve and (1) does not have a simple form. When a map is compiled using an ellipsoidal Earth model, one must come to grips with solving (1) on a surface without constant curvature. This problem has no direction solution and is still an area of active research. A detailed discussion of geodesic distances on an ellipsoid is given in section 4.5 on page 26.

1.2. **Reference Ellipsoids.** The Earth is more nearly shaped like a spheroid than a sphere. A spheroid has a circular cross section and is elliptical in profile. The Earth is nearly circular in the equatorial plane but is foreshortened in the minor axis. Therefore, the Earth is modeled better by an oblate spheroid than a sphere. It is common in the United States to refer to this shape as an "ellipsoid," which is technically correct because a spheroid is an ellipsoid. Although spheroid is

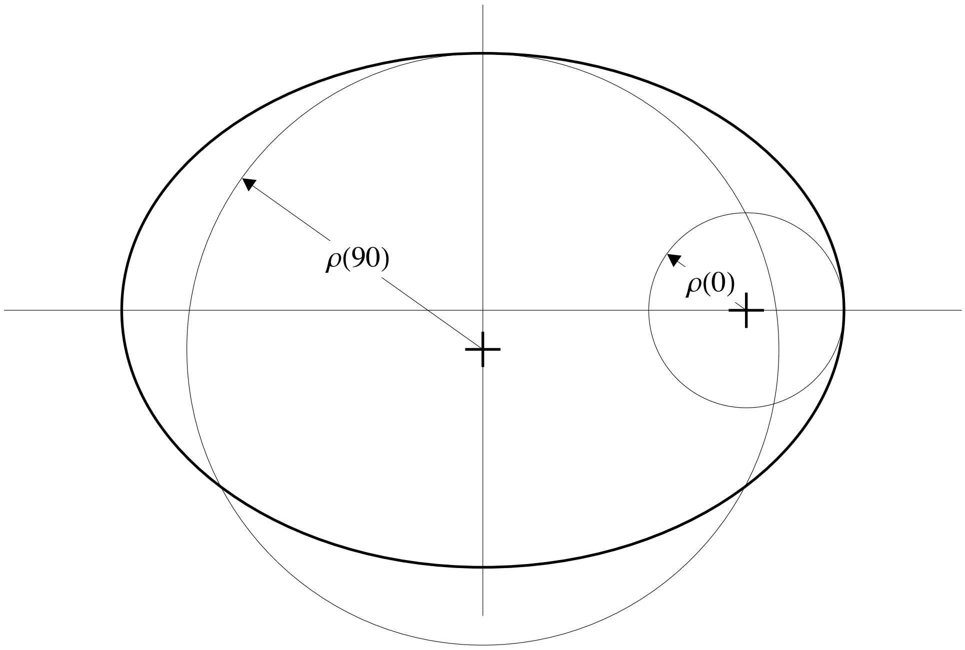

FIGURE 1. Radius of Curvature in the Prime Meridian.

the more correct of the two terms, ellipsoid will be used to conform to standard practice when the topic concerns geodetic reference ellipsoids, such as GRS 80 or WGS 84.

See Ewing and Mitchell (1970) for derivations of the following formulae. Four principle parameters are used to define the shape of spheroids: the length of its semimajor and semiminor axes, the foreshortening of the spheroid, and its eccentricity. The length of its semimajor and semiminor axes are denoted a and b respectively, and their values are taken from tables (see Table 1 on the facing page). The foreshortening of the spheroid is called its **flattening** and is defined as

$$f = (a - b)/a = 1 - b/a \tag{4}$$

Some reference ellipsoids are defined in terms of a and b. Others are defined in terms of a and $1/f$. The (first) eccentricity squared of an ellipse is defined in terms of the flattening as

$$e^2 = 2f - f^2 = \frac{a^2 - b^2}{a^2} = 1 - b^2/a^2 \tag{5}$$

The (second) eccentricity of an ellipse is defined in terms of the first eccentricy as

$$\epsilon = e^2/(1 - e^2) \tag{6}$$

1.2.1. *Radii of Curvature.* A radius of curvature is an instantaneous measure of how "tight" a curve is. For a curve possessing at least second-order derivatives and having the form $y = f(x)$, the radius of curvature R can be expressed formally as

$$R = \frac{[1 + (dy/dx)^2]^{3/2}}{d^2y/dx^2}$$

See (Casey 1996, McCleary 1994) for detailed discussions of curvature.

Semimajor(m)	1/f or Semiminor(m)	Name
6377563.396	b=6356256.910	Airy 1830
6377340.189	b=6356034.446	Modified Airy
6377104.43	rf=300.0	Andrae 1876 (Den., Iclnd.)
6378137.0	rf=298.25	Appl. Physics. 1965
6378160.0	rf=298.25	Australian Natl and S. Amer. 1969
6377397.155	rf=299.1528128	Bessel 1841
6377483.865	rf=299.1528128	Bessel 1841 (Namibia)
6378206.4	b=6356583.8	Clarke 1866
6378249.145	rf=293.4663	Clarke 1880 mod.
6375738.7	rf=334.29	Comm. des Poids et Mesures 1799
6376428.	rf=311.5	Delambre 1810 (Belgium)
6378136.05	rf=298.2566	Engelis 1985
6377298.556	rf=300.8017	Everest (Sabah and Sarawak)
6377276.345	rf=300.8017	Everest 1830
6377304.063	rf=300.8017	Everest 1948
6377301.243	rf=300.8017	Everest 1956
6377295.664	rf=300.8017	Everest 1969
6378166.	rf=298.3	Fischer (Mercury Datum) 1960
6378150.	rf=298.3	Fischer 1968
6378155.	rf=298.3	Modified Fischer 1960
6378137.0	rf=298.257222101	GRS 1980(IUGG, 1980)
6378160.0	rf=247.247167	GRS 67(IUGG 1967)
6378200.	rf=298.3	Helmert 1906
6378270.0	rf=297.	Hough
6378140.0	rf=298.257	IAU 1976
6378157.5	b=6356772.2	New International 1967
6378388.0	rf=297.	International 1909 (Hayford)
6378163.	rf=298.24	Kaula 1961
6378245.0	rf=298.3	Krassovsky, 1942
6378139.	rf=298.257	Lerch 1979
6397300.	rf=191.	Maupertius 1738
6378137.0	rf=298.257	MERIT 1983
6378145.0	rf=298.25	Naval Weapons Lab., 1965
6376523.	b=6355863.	Plessis 1817 (France)
6378136.0	rf=298.257	SGS 85
6378155.0	b=6356773.3205	Southeast Asia
6376896.0	b=6355834.8467	Walbeck
6378165.0	rf=298.3	WGS 60
6378145.0	rf=298.25	WGS 66
6378135.0	rf=298.26	WGS 72
6378137.0	rf=298.257223563	WGS 84

TABLE 1. Geodetic Reference Ellipsoid Parameters.

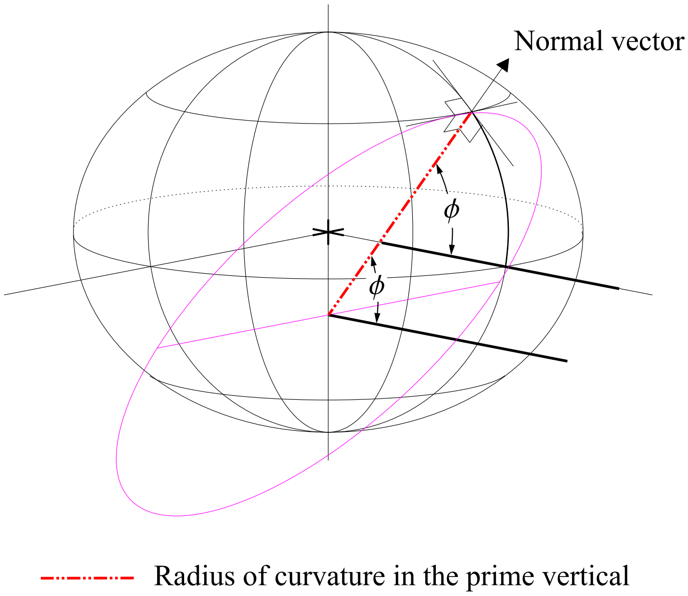

FIGURE 2. Geodetic latitude ϕ, the vector normal to the surface of the ellipsoid, and the radius of curvature in the prime vertical ν.

Radii of curvature play an important role in determining geodesic distances. Distance formulae on curved surfaces require a parameter indicating the radius of the surface upon which the computation is to be performed. When using a spherical model of the Earth, that number is simply the radius of the sphere. For a spheroid, however, things are more complicated. There are three principle radii of curvature for a spheroid, namely the radius of curvature in the meridian ρ, the radius of curvature in the prime vertical ν, and the radius of curvature in the normal section η.

1.2.2. *Radius of Curvature in the Meridian.* The radius of curvature in the meridian ρ is the radius of curvature of a spheroid in the plane of a meridian. It is the radius to use in the geodesic length computations in the special case that the geodesic falls exactly along a meridian. ρ varies with latitude, achieving a maximum at the equator and a minimum at the poles for an oblate spheroid. The equation for the radius of curvature in the meridian is

$$\rho = \frac{a(1-e^2)}{(1-e^2\sin^2\phi)^{3/2}}, \tag{7}$$

where the terms are as defined above. Note that $\rho(0^\circ) = b^2/a < b$ and $\rho(90^\circ) = a^2/b > b$; the inequalities hold when $e > 0$. See Figure 1 on page 2. The radius of curvature in the meridian is seldom (if ever) used directly. Instead, it is an input to a more general equation that is shown below.

1.2.3. *Radius of Curvature in the Prime Vertical.* The radius of curvature in the prime vertical ν is the length of a vector that begins at the z-axis, ends at the surface of the spheroid, and is normal to the spheroid. See Figure 2. The formula for the radius of curvature in the prime vertical is

$$\nu = \frac{a}{(1 - e^2 \sin^2 \phi)^{1/2}}, \tag{8}$$

where the terms are as defined above. Note that while $\nu(0°) = a,\ \nu(90°) = \rho(90°) = a^2/b > b$. It is also important to notice that neither ρ nor ν points in the same direction as the radius of the spheroid and they do not intersect the z-axis at the spheroid's center when $e^2 \neq 0$.

See Bowring (1987) for a simple proof by geometry that the center of curvature in the prime vertical section at a point on the reference ellipsoid lies on the minor axis.

1.2.4. *Radius of Curvature in the Normal Section.* It is evident by comparing equations (8) and (7) that the radius of curvature on a spheroid varies with azimuth. Let **A** be a point on a spheroid. The intersection of the spheroid and a plane containing the normal at **A** is called a **normal section** (Bowring 1971). See Figures 15, 17, and 18 on page 31. The radius of curvature in the normal section η is a combination of ρ and ν. Its formula, which is known as Euler's theorem (Ewing and Mitchell 1970, Casey 1996), is

$$\frac{1}{\eta_{\mathbf{AB}}} = \frac{\cos^2 \alpha_{\mathbf{AB}}}{\rho} + \frac{\sin^2 \alpha_{\mathbf{AB}}}{\nu}, \tag{9}$$

where $\alpha_{\mathbf{AB}}$ denotes the azimuth of the normal section at **A** containing **B**. By examining the formula, it is clear that $\eta_{\mathbf{AB}}$ equals ρ when $\alpha_{\mathbf{AB}}$ is zero and that $\eta_{\mathbf{AB}}$ equals ν when $\alpha_{\mathbf{AB}}$ is 90°, as expected.

1.2.5. *Examples.* The following examples pertain to the GRS 80 ellipsoid (Moritz 2000, p.131).

Example 1. The semimajor axis is defined as $a = 6,378,137$ m, exactly. □

Example 2. The flattening is derived to be $f = 0.003\,352\,810\,681\,18$, which gives $1/f = 298.257\,222\,101$. □

Example 3. The first eccentricity is derived to be

$$\begin{aligned} e^2 &= 2f - f^2 \\ &= 2 \cdot 0.003\,352\,810\,681\,18 - [0.003\,352\,810\,681\,18]^2 \\ &= 0.006\,694\,380\,022\,90 \end{aligned}$$

□

Example 4. The second eccentricity is derived to be

$$\begin{aligned} \epsilon &= e^2/(1 - e^2) \\ &= [0.006\,694\,380\,022\,90]^2/(1 - [0.006\,694\,380\,022\,90]^2) \\ &= 0.006\,739\,496\,775\,47 \end{aligned}$$

□

Example 5. At a geodetic latitude $\phi = 41.98097°$, the radius of curvature in the meridian is

$$\begin{aligned}\rho &= \frac{a(1-e^2)}{(1-e^2\sin^2\phi)^{3/2}}\\ &= \frac{6\,378\,137\,\mathrm{m}\,(1-0.006\,694\,380\,022\,90)}{(1-0.006\,694\,380\,022\,90\ \sin^2 41.98097°)^{3/2}}\\ &= 6\,364\,009.194\,79\,\mathrm{m}\end{aligned}$$

□

Example 6. At a geodetic latitude $\phi = 41.98097°$, the radius of curvature in the prime vertical is

$$\begin{aligned}\nu &= \frac{a}{(1-e^2\sin^2\phi)^{1/2}}\\ &= 6\,378\,137\,\mathrm{m}/\sqrt{(}1-0.006\,694\,380\,022\,90\ \sin^2 41.98097°),\\ &= 6\,387\,710.095\,74\,\mathrm{m}\end{aligned}$$

□

Example 7. At a geodetic latitude $\phi = 41.98097°$ and an azimuth of $\alpha = 45°$, the radius of curvature in the normal section is

$$\begin{aligned}\frac{1}{\eta} &= \frac{\cos^2\alpha}{\rho} + \frac{\sin^2\alpha}{\nu}\\ &= \frac{\cos^2 45°}{6\,364\,009.194\,79\,\mathrm{m}} + \frac{\sin^2 45°}{6\,387\,710.095\,74\,\mathrm{m}}\\ &= 1.56842137406544313398 \times 10^{-7}\,\mathrm{m}^{-1}\end{aligned}$$

so

$$\eta = 6\,375\,837.619\,50\,\mathrm{m}$$

□

2. The Flat-Earth Assumption

Before the Global Positioning System (GPS), it was usually possible to ignore the curvature of the Earth when making relatively short distance measurements and still maintain acceptable accuracy. This was possible in spite of the Earth being round because, of course, the Earth has a relatively large radius. The difference between the curved surface of the Earth and a straight line can safely be ignored over areas of the size of a typical subdivision and not incur unacceptable errors for, say, a typical surveying construction project. As will be shown, the difference between the length of the arc connecting two points relatively close to one another on the surface of the Earth and the length of the equivalent chord is very small, usually less than the precision of a typical electronic distance measuring device (EDM) found in a modern total station. However, GPS has changed things in such a way that a “flat Earth” assumption now causes everyday problems for surveyors.

GPS is becoming a surveying and mapping instrument whose importance rivals that of the total station. For example, for long baseline measurements, GPS has an advantage over opto-mechanical measuring devices in that the baseline stations do not need intervisibility, thus eliminating the need to traverse between them. Indeed, it is just as easy to process a GPS baseline connecting stations that are on opposite sides of the planet as a baseline between two stations 10 meters apart. Also,

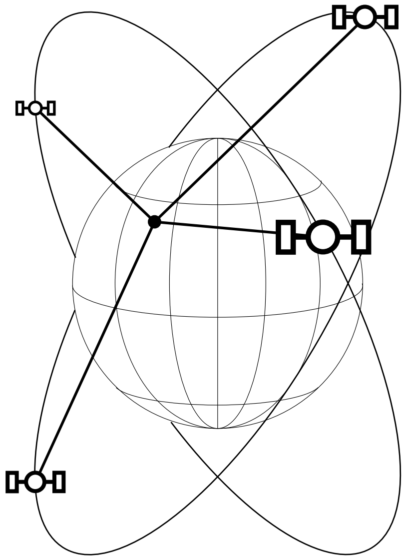

FIGURE 3. Distances from GPS NAVSTAR satellites to places on the Earth's surface are inherently three-dimensional.

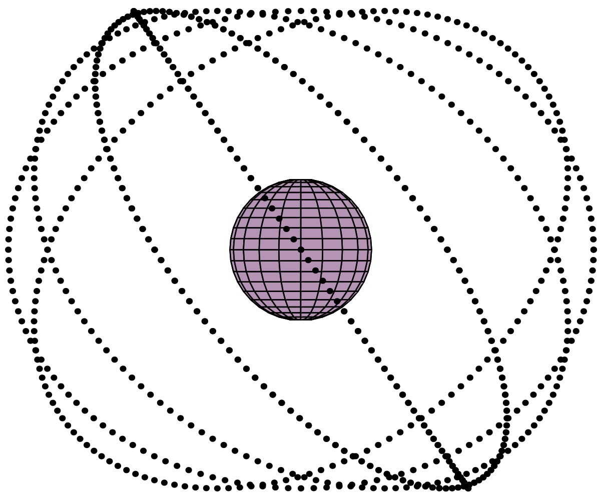

FIGURE 4. GPS satellite orbits viewed from the equatorial plane.

GPS baseline errors are far less sensitive to the length of the baseline than what is the case with a traditional instrument. Furthermore, using real-time kinematic (RTK) technologies, surveyors are able to take position measurements in an absolute coordinate system and with centimeter-precision while in the field (i.e., in "real time"). This is similar to total station data collector coordinate geometry (COGO) software that performs a traverse in the field. Its many advantages and high-accuracy insure that GPS is here to stay, at least for the foreseeable future.

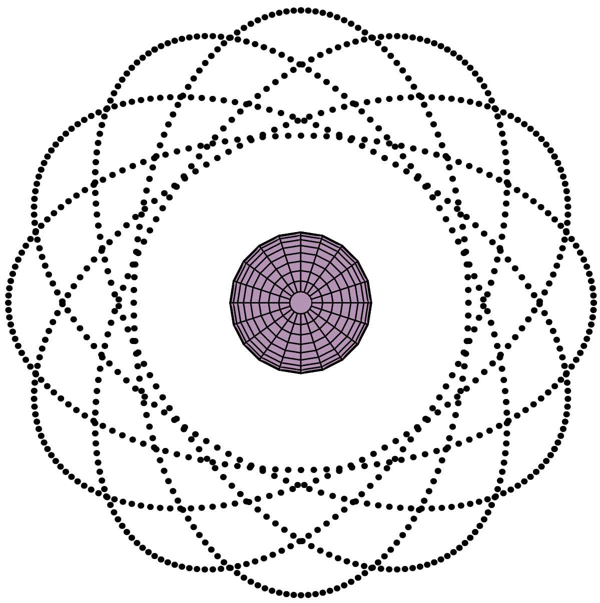

FIGURE 5. GPS satellite orbits viewed from the North Pole.

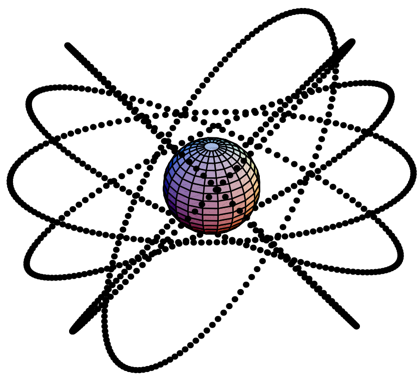

FIGURE 6. GPS satellite orbits viewed from space.

A "flat Earth" assumption is unnecessary and inappropriate for GPS measurements. Unlike measurements made with transits and levels, which are inherently two-dimensional and oriented according to gravity, GPS measurements are inherently three-dimensional and not oriented to the geoid in an obvious way (Figure 3). As shown in Figures 4, 5, and 6, there are six GPS satellite orbits. Each orbit is inclined 55° from the equatorial plane and is occupied by four satellites. For details about orbits and other GPS information see (Leick 1995, Hofmann-Wellenhof and others 1997).

The fundamental observable for GPS is the propagation time of a radio wave broadcast from the NAVSTAR satellites to the surveyor's GPS receiver (Leick 1995, Hofmann-Wellenhof et al. 1997, Van Sickle 1996). This propagation time is converted into a **pseudo-range**. This distance is called a "pseudo-range" because of an unknown clock bias between the receiver and the time standards in the satellites. This bias alters the derived distance from the true distance. Following the notation in Leick (1995), denote the pseudo-range from satellite s to receiver r as ρ_r^s. Then, using the well-known relationship, $\rho_r^s = c \cdot \Delta t_r^s$, where ρ_r^s is the pseudo-range just defined, c is the speed of light, and Δt_r^s is the propagation time of the radio wave from satellite s to receiver r. ρ_r^s is inherently three-dimensional because it is the distance from a satellite in space to a receiver on the Earth and the satellites are not all coplanar.

3. Coordinate Systems

There are four coordinate systems to consider for this discussion, namely, geocentric Cartesian, local geodetic Cartesian, geodetic, and planimetric Cartesian. The geocentric Cartesian coordinate system is often called "Earth-centered, Earth-fixed," "Conventional Terrestrial Reference System" (CTRS), and also "XYZ". We will use the latter in this discussion. Geodetic coordinates are geodetic latitude, longitude and ellipsoid height. Planimetric coordinates are usually called "eastings" and "northings." These will all be discussed in turn.

3.1. Geocentric Cartesian Coordinates.

The mathematics describing the spatial relationships of the GPS are defined in a geocentric Cartesian coordinate system (Leick 1995, Hofmann-Wellenhof et al. 1997). As an example of XYZ coordinates, Figure 7 on the following page depicts the XYZ coordinates of a place in Connecticut, USA. If the positions of a satellite and receiver[1] are expressed in geocentric Cartesian coordinates, then the actual (as opposed to a pseudo-range) distance between them is simply the Euclidean straight-line, three-dimensional distance given by (2) as

$$d_r^s = [\,(x^s - x_r)^2 + (y^s - y_r)^2 + (z^s - z_r)^2\,]^{1/2}, \tag{10}$$

where x, y, and z denote the geocentric Cartesian coordinates, and s and r denote the satellite and receiver, respectively. For mapping-grade GPS measurements, the pseudo-ranges to at least four satellites are mathematically manipulated using trilateration to determine the coordinates of the receiver. In (10), d_r^s is known, having been measured by the hardware. Also, x^s, y^s, and z^s are known because the coordinates of the satellites are derived from the part of the navigation message called the **broadcast ephemeris**. If one were to use (10) to compute distances, the result would be a **slant distance** and would be essentially the same thing that an EDM would have measured.

Mapping with slant distances is usually not done. Instead, slant distances are converted to horizontal and vertical distances. This transformation is simple when slant distances are measured

[1] To be specific, this is the distance to the phase center of the receiver's GPS antenna.

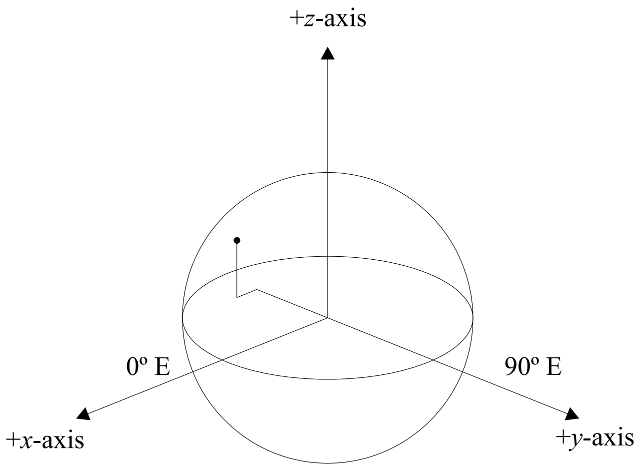

Figure 7. The WGS 84 XYZ coordinates of the south-east corner of the roof of the W.B. Young Building on the University of Connecticut's campus (x=1451658.2, y=-4534382.6, z=4230142.1). The coordinates indicate meters along the axes away for the Earth's center of mass.

with instruments like a total station because the instrument was presumably oriented according to the local vertical when the measurement was taken. Some trigonometry involving the measured angle of elevation (21) resolves a slant distance into its two parts. However, GPS measurements are not made with any notion of local vertical. The final output of a GPS post-processing computer program is a list of the coordinates of the occupied stations. If they are reported in their most natural form, these coordinates will be in the World Geodetic System of 1984 (WGS 84) datum and in its XYZ coordinate system. Finding the distance between such points is done using (2) and would result in a slant distance. However, the horizontal distance between the stations is required for map compilation, not the slant distance. To derive the horizontal distances from XYZ coordinates, the coordinates are first transformed to geodetic latitude, longitude, and ellipsoid height coordinates. Then, the horizontal distance is the length of the geodesic between these points.

3.2. **Local Geodetic Coordinate Systems: East-North-Up.** Another useful coordinate system is defined when local East/North/Up axes are centered at a place that is usually in the middle of a region of interest and on or near the Earth's surface, rather than at its center. Let **O** denote the origin of this coordinate system. The North axis is oriented along the meridian passing through **O**, positive to the North. The East axis oriented along the parallel passing through **O**, positive to the East. The Up axis passes through **O** and is parallel to the radius of curvature in the prime vertical at **O**. In Figure 8 on the next page, **O** defines the origin of a local ENU coordinate system. In this system, **P** has ENU coordinates of (P_e, P_n, P_u).

3.2.1. *ENU vs. XYZ for GPS.* The ENU coordinate system can be very helpful when examining GPS baseline statistics. The body of the Earth stops the radio waves transmitted from GPS satellites below the horizon. Consequently, GPS receivers get no observables from below them, only from above. Therefore, GPS measurements are less precise in the direction of the local vertical than horizontal and the precision estimates for GPS baselines are typically one and a half to twice as large vertically as horizontally. This effect is readily apparent when the baseline is presented

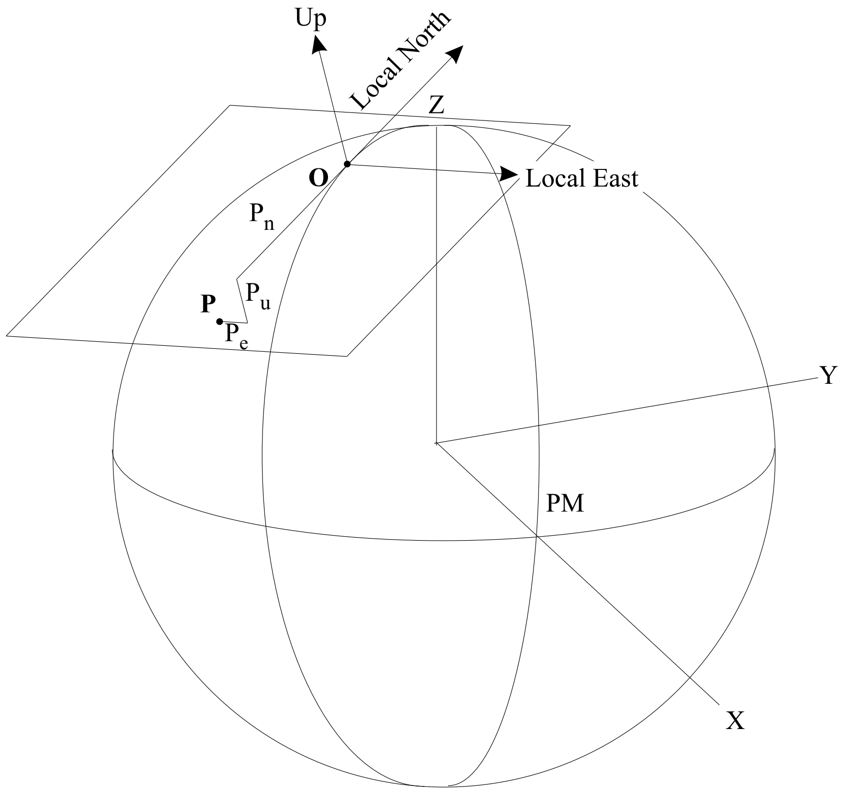

FIGURE 8. A Local Geodetic East-North-Up (ENU) Coordinate System.

in an ENU coordinate system. It is not apparent at all in an XYZ coordinate system because the three ENU directions get mixed together in non-obvious ways in the XYZ coordinate system. So, it is very helpful to see the precisions in the ENU coordinate system when checking that the vertical precision is twice the horizontal precisions.

3.2.2. *XYZ to ENU.* The ENU coordinate system is very similar to the geocentric Cartesian coordinate system presented in the previous section. Both are Cartesian, which means that they have three mutually perpendicular axes and the coordinates are distances from the origin. Indeed, any particular ENU coordinate system is simply the XYZ coordinate system translated from the center of the Earth to the local origin, and then rotated so the y-axis becomes the North axis, the x-axis becomes the East axis, and the z-axis becomes the Up axis. The details of this process follow. Converting XYZ to local geodetic ENU involves translating the coordinate system origin from the Earth's center to the local origin and then rotating the axes to align the y-axis with North, the x-axis with East, and the z-axis with Up (Wolf and Dewitt 2000, pp. 573–576). We need the coordinates of the ENU coordinate system local origin in both geodetic coordinate (ϕ_0, λ_0) and in XYZ coordinates (x_0, y_0, z_0). Let (x, y, z) be the XYZ coordinates of the point for which to find its

ENU coordinates, (e,n,u). Then,

$$\begin{bmatrix} e \\ n \\ u \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{bmatrix} = M \begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{bmatrix}, \tag{11}$$

where the matrix M is

$$\begin{aligned}
m_{11} &= -\sin\lambda_0 \\
m_{12} &= \cos\lambda_0 \\
m_{13} &= 0 \\
m_{21} &= -\sin\phi_0\cos\lambda_0 \\
m_{22} &= -\sin\phi_0\sin\lambda_0 \\
m_{23} &= \cos\phi_0 \\
m_{31} &= \cos\phi_0\cos\lambda_0 \\
m_{32} &= \cos\phi_0\sin\lambda_0 \\
m_{33} &= \sin\phi_0.
\end{aligned}$$

The inverse problem of finding the XYZ coordinates of a point whose coordinates are given in ENU is found from (11) by inverting M and solving for (x,y,z) as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = M^{-1} \begin{bmatrix} e \\ n \\ u \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \tag{12}$$

The ENU coordinate system might seem to not be a projected coordinate system. After all, it is merely a translation and rotation of the XYZ coordinate system, which is not projected. However, an ENU coordinate system defines a plane tangent to the reference ellipsoid at the local origin and the East/North/Up coordinates are projected onto this plane. Therefore, like most projected coordinate systems, the East and North directions in a ENU coordinate system increasingly deviate from geodetic east and north as distance to the origin increases. This phenomenon is called **convergence**. Also, the Up coordinate can be quite counter-intuitive. For example, with the exception of the local ENU origin, any point on the reference ellipsoid's surface will have a negative Up coordinate because the projection plane is tangent to the ellipsoid at the local origin. Any other point on the ellipsoid's surface is below this tangent plane.

3.3. **Geodetic Coordinates.** Geodetic coordinates are geodetic latitude, longitude and ellipsoid height. Geodetic latitude is the angle that the radius of curvature in the prime vertical makes with the equatorial plane (see Figure 2 on page 4). Ellipsoid height is the length of a vector that starts at the surface of and is normal to the reference ellipsoid and ends at the point of interest.

If one is using a spherical approximation for the Earth's figure, then converting from geodetic coordinates to XYZ coordinates is nothing more than the conversion from a spherical coordinate system to a Cartesian coordinate system that most people learned in calculus or in a course having analytical geometry. Converting to XYZ coordinates on an ellipsoid complicates the situation because the surface normal vector for an ellipsoid is the radius of curvature in the prime vertical, which does not intersect the z-axis at the center of the ellipsoid (Bowring 1976). Nevertheless, the formulae for spheres and ellipsoids are very similar in structure. The ellipsoidal formulae are

$$x = (\nu + h)\cos\phi\cos\lambda, \tag{13a}$$
$$y = (\nu + h)\cos\phi\sin\lambda, \tag{13b}$$
$$z = ((1 - e^2)\nu + h)\sin\phi \tag{13c}$$

where λ is longitude, h is the height above the ellipsoid and ν is the radius of curvature in the prime vertical at geodetic latitude ϕ.

The inverse problem of converting from XYZ coordinates back to geodetic coordinates on an ellipsoid has no direct solution (Heiskanen and Moritz 1967), so it has received a fair amount of attention (Heiskanen and Moritz 1967, Bowring 1976, Taff 1985, Borkowski 1989, Laskowski 1991, Lin and Wang 1995, Vaníček and Krakiwsky 1996, Fukushima 1999, Gerdan and Deakin 1999, Pollard 2002).

Bowring (1976) developed what is probably the most widely used procedure. Although it is iterative, it is always adequate when used without iteration for places on or near the Earth's surface, producing positions whose accuracy is about 1 μm. Bowring amended his original height formula by replacing it with one that is adequate for positions in outer space (Bowring 1985). Fukushima (1999) introduced a procedure that is accurate at all heights, thus improving a shortcoming of Bowring's procedure, which looses accuracy near the geocenter. Although there are many procedures to choose from , we present Bowring's procedure (Bowring 1976) for latitude and longitude due to its widespread use and his improved procedure for height (Bowring 1985). Define $p = \sqrt(x^2 + y^2)$. Then

$$\tan\lambda = y/x \tag{14a}$$
$$\tan\beta = \frac{z}{(1-f)p} \tag{14b}$$
$$\tan(90^\circ - \beta) = \frac{(1-f)p}{z} \tag{14c}$$
$$\tan\phi = \frac{z + \epsilon b \sin^3\beta}{p - ae^2\cos^3\beta} \tag{14d}$$
$$\tan(90^\circ - \phi) = \frac{p - ae^2\cos^3\beta}{z + \epsilon b \sin^3\beta} \tag{14e}$$
$$h = p\cos\phi + z\sin\phi - (a^2/\nu) \tag{14f}$$

where $\epsilon = e^2/(1-e^2)$ from (6). Use (14c) and (14e) when $\phi \approx 90^\circ$.

3.4. **Planimetric Cartesian Coordinates.** Planimetric Cartesian coordinates are often called "eastings" and "northings." They are the product of a cartographic projection from three-dimensional geodetic coordinates (ϕ, λ) into a two-dimensional Cartesian space (x, y), i.e., a map. In this work, eastings will be denoted by x and northings by y instead of e and n for two reasons. First, this emphasizes the Cartesian nature of projected coordinates. Second, using e for eastings could create confusion with e^2, the first eccentricity squared.

There are hundreds of unique cartographic projections (Maling 1973, Snyder 1987, Bugayevskiy and Snyder 1995, Iliffe 2000, Qihe and others 2000). All projections distort certain spatial characteristics or relationships of the features being projected. There are projections that do not distort one or more (but not all) of the following: great-circle distance (**gnomic**), direction (**azimuthal**),

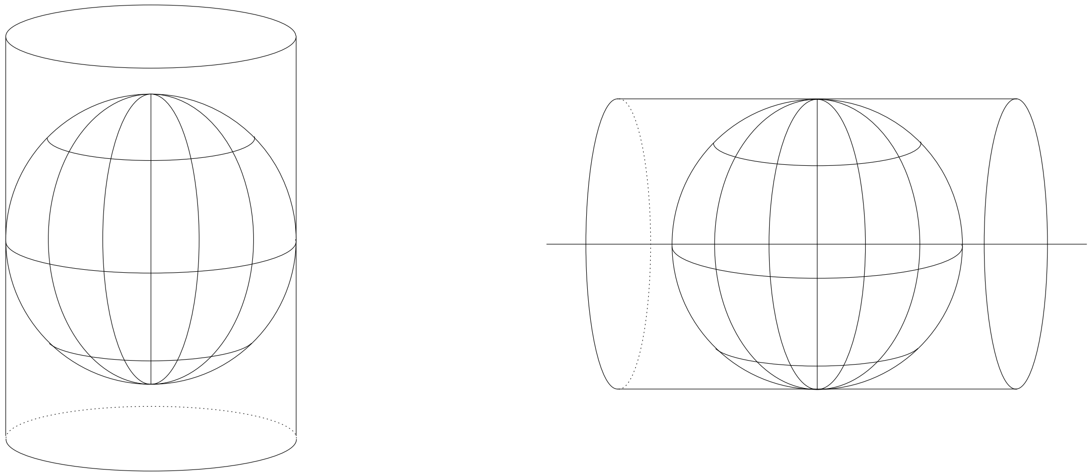

FIGURE 9. Two cylindrical projections. On the left is a standard cylindrical projection. The axis of the standard cylindrical projection is coincident with the minor axis. On the right is a transverse cylindrical projection. The axis of the transverse cylindrical projection is in the Equatorial plane.

shape (**conformal**), or area (**equal-area**). Maps are compiled with a particular projection that is usually chosen to display some characteristic of interest without distortion. For example, assessor's maps might be compiled using equal-area projections because taxation usually depends on the property's area.

Shape-preserving projections are said to be **conformal**. More specifically, a projection is conformal if and only if the projected angle between two lines equals the unprojected angle between those same lines. This property makes the mathematical analysis of a projection much easier so conformal projections are used and studied extensively.

3.4.1. *Scale Variation.* This section follows the material in Snyder (1987, pp. 20–28). No projection maintains constant scale throughout its range. Tissot (1881) developed **Tissot's Indicatrix**, which is a graphical representation of the scale and angular distortions produced by a projection. Tissot showed that an infinitely small circle on the Earth will be projected to a small ellipse. If the projection is conformal, the projected ellipse will have zero eccentricity, i.e., it will be a circle.

Scale distortion is most often calculated in the direction of the meridian and the parallel and are denoted h and k, respectively. It follows from Tissot's work that $h = k$ for conformal projections. h and k are scaled such that a value of identically one indicates no scale distortion in that direction. A curve of constant, unit scale distortion is called a **standard**.

For computing the scale factor for the transverse Mercator projection, see (Lee 1963, Stoughton 1982, Snyder 1987).

3.4.2. *Transverse Mercator.* There are far too many projections to even mention, much less present in the current work. Therefore, to illustrate grid coordinates, we will focus on two common conformal projections, the Transverse Mercator (TM) projection as employed in the Universal Transverse Mercator coordinate system (Snyder 1987, pp.48–65) and the Lambert Conformal Conic (LCC)

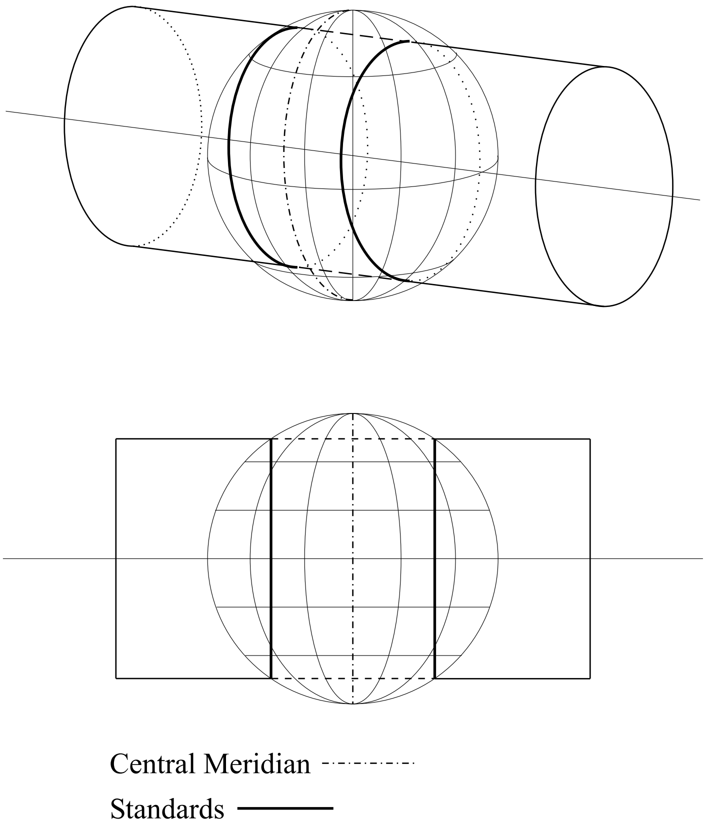

FIGURE 10. A transverse cylindrical projection with two standards. The standards are the heavier lines to the left and right of the central meridian. Note that the standards do not follow either meridians or parallels.

projection as employed in the State Plane Coordinate System for Connecticut (Stem 1995) and (Snyder 1987, pp.104–106).

As shown in Figure 9 on the preceding page, the TM is a cylindrical projection whose axis is in the Equatorial Plane. TM projections are defined by three parameters, λ_0, k_0, and ϕ_0. λ_0 defines the longitude of the central meridian. k_0 is the scale factor at λ_0. ϕ_0 is the latitude of the origin of the projection, which is at (ϕ_0, λ_0). Figure 10 shows a transverse Mercator projection with $k_0 < 1$. Thus, the region between the standards is projected smaller than reality and the regions outside the standards are projected larger than reality. TM projections are often used for regions whose extent is larger north-to-south than east-to-west because the standards more-or-less run north-to-south, at least near the Equator. Also, if the standards are fairly close, the projection naturally creates a narrow strip around the central meridian, a naturally north-south region.

The *forward* equations for a general Transverse Mercator projection: To project geodetic latitude and longitude (ϕ, λ) to eastings and northings (x, y):

$$x = k_0\,\nu[A + (1 - T + C)A^3/6 + (5 - 18T + T^2 + 72C - 58\epsilon)A^5/120] \tag{15a}$$

$$y = k_0\,\{M - M_0 + \nu \tan\phi[A^2/2 + (5 - T + 9C + 4C^2)A^4/24 + (61 - 58T + T^2 + 600C - 330\epsilon)A^6/720]\} \tag{15b}$$

$$k = k_0\,[1 + (1 + C)A^2/2 + (5 - 4T + 42C + 13C^2 - 28\epsilon)A^4/24 + (61 - 148T + 16T^2)A^6/720] \tag{15c}$$

$$T = \tan^2\phi \tag{15d}$$

$$C = \epsilon\cos^2\phi \tag{15e}$$

$$A = (\lambda - \lambda_0)\cos\phi, \tag{15f}$$

$$M = a\,[(1 - e^2/4 - 3e^2/64 - 5e^6/256 - \dots)\phi - (3e^2/8 + 3e^4/32 + 45e^6/1024 + \dots)\sin 2\phi + (15e^4/256 + 45e^6/1024 + \dots)\sin 4\phi - (35e^6/3072 + \dots)\sin 6\phi + \dots] \tag{15g}$$

with ϕ, λ, and λ_0 in radians. M is the true distance along the central meridian from the Equator to ϕ. $M_0 = M$ for ϕ_0. If $\phi = \pm\pi/2$, then only (15g) is used to compute M and M_0. Then $e = 0, n = k_0(M - M_0), k = k_0$.

For UTM, a is the semimajor axis in meters, $k_0 = 0.9996$, and $\phi_0 = 0$ so $M_0 = 0$, likewise. After computing x, add a false easting of 500,000 m. If $\phi < 0$, after computing y, add a false northing of 10,000,000 m.

The *inverse* equations for a general Transverse Mercator projection: To transform TM eastings and northings (x, y) to geodetic latitude and longitude (ϕ, λ):

$$\phi = \phi_1 - (\nu(\phi_1)\tan\phi_1/\rho(\phi_1))[D^2/2 - (5 + 3T_1 + 10C_1 - 4C_1^2 - 9\epsilon^2)D^4/24 + (61 + 90T_1 + 298C_1 + 45T_1^2 - 252\epsilon^2 - 3C_1^2)D^6/720] \tag{16a}$$

$$\lambda = \lambda_0 + [D - (1 + 2T_1 + C_1)D^3/6 + (5 - 2C_1 + 28T_1 - 3C_1^2 + 8\epsilon + 24T_1^2)D^5/120]/\cos\phi_1 \tag{16b}$$

ϕ_1 is the latitude at the central meridian which has the same northing coordinate as that of the point (ϕ, λ). It may be found by:

$$\phi_1 = \mu + (3e_1/2 - 27e_1^3/32 + \dots)\sin 2\mu + (21e_1^2/16 - 55e_1^4/32 + \dots)\sin 4\mu + (151e_1^3/96 + \dots)\sin 6\mu + (1097e_1^4/512 - \dots)\sin 8\mu + \dots \tag{16c}$$

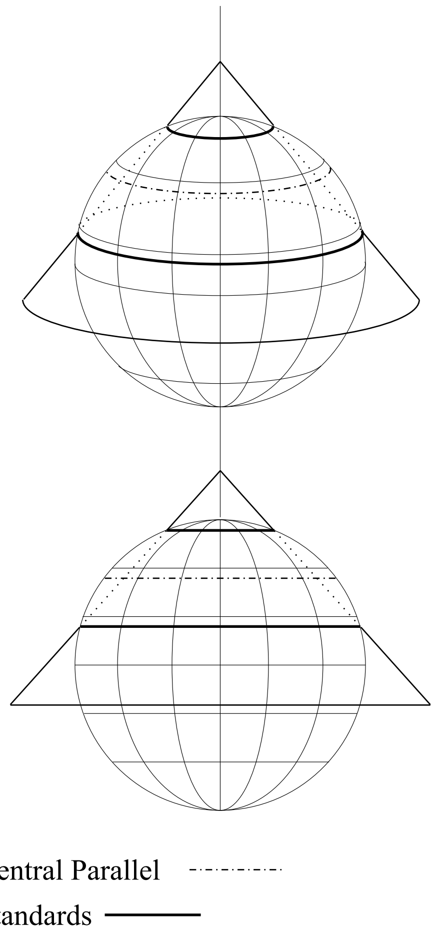

FIGURE 11. A conic projection with two standards. The standards are the heavier lines north and south of the central parallel. Note that the standards are also parallels.

where

$$e_1 = [1 - (1 - e^2)^{1/2}]/[1 + (1 - e^2)^{1/2}] \tag{16d}$$

$$\mu = M/[a(1 - e^2/4 - 3e^4/64 - 5e^6/256 - \dots)] \tag{16e}$$

$$M = M_0 + y/k_0 \tag{16f}$$

$$C_1 = \epsilon \cos^2 \phi_1 \tag{16g}$$

$$T_1 = \tan^2 \phi_1 \tag{16h}$$

$$D = x/(\nu(\phi_1)k_0) \tag{16i}$$

with M_0 calculated from (15g) for the given ϕ_0. If $\phi_1 = \pm\pi/2$, $k = k_0$, (16a), (16b), and (16d) - (16i) are omitted with $\phi = \pm 90°$, taking the sign of y, while λ is indeterminate and can be taken as λ_0 or any other convenient, valid value.

For UTM, a is the semimajor axis in meters, $k_0 = 0.9996$, and $\phi_0 = 0$ so $M_0 = 0$, likewise. Before using the eastings, subtract a false easting of 500,000 m. If y is in the Southern Hemisphere, subtract a false northing of 10,000,000 m before using y in the formulae.

3.4.3. *Lambert Conformal Conic.* The Lambert Conformal Conic projections are often used for regions whose extent is larger east-to-west than north-to-south because the standards run east-to-west (Figure 11 on the page before). Thus, the projection creates a strip around the central parallel, a naturally east-west region. LCC projections are parameterized by the semimajor axis of the reference ellipsoid a and the first eccentricity squared e^2, the latitudes of the standard parallels ϕ_0 and ϕ_1, and the latitude and longitude of the origin of the projection ϕ_0 and λ_0. The region between the standards is projected smaller than reality and the regions outside the standards are projected larger than reality.

The *forward* equations for a general Lambert Conformal Conic projection: To project geodetic latitude and longitude (ϕ, λ) to eastings and northings (x, y):

$$x = \rho \sin\theta \tag{17a}$$

$$y = \rho_0 - \rho\cos\theta \tag{17b}$$

$$k = \rho n/(am) \tag{17c}$$

where

$$\rho = aFt^n \tag{17d}$$

$$\theta = n(\lambda - \lambda_0) \tag{17e}$$

$$n = (\ln m_1 - \ln m_2)/(\ln t_1 - \ln t_2) \tag{17f}$$

$$m = \cos\phi(1 - e^2\sin^2\phi)^{1/2} \tag{17g}$$

$$t = \tan(\pi/4 - \phi/2)/[(1 - e\sin\phi)/(1 + e\sin\phi)]^{e/2} \tag{17h}$$

or

$$= \left[\left(\frac{1-\sin\phi}{1+\sin\phi}\right)\left(\frac{1+e\sin\phi}{1-e\sin\phi}\right)^e\right]^{1/2} \tag{17h$'$}$$

$$F = m_1/(nt_1^n) \tag{17i}$$

with $t_x = t(\phi_x)$, $m_x = m(\phi_x)$, and $\rho_0 = aFt_0^n$.

3.5. **Examples.**

Example 8. Problem: Convert NAD 83 $(41^\circ\ 21'\ 12.99487''\mathrm{N}, 72^\circ\ 01'\ 25.04041''\mathrm{W}, 635.478\,\mathrm{m})$ to XYZ coordinates.
Solution: In decimal degrees, the given coordinates are 41.3536096861°, -72.02362233611°. The formula for the radius of curvature in the prime vertical is (8), which is repeated here for convenience:

$$\nu = a/\sqrt{(1 - e^2\sin^2\phi)},$$

where a is the length of the semimajor axis of the ellipsoid, e^2 is the (first) eccentricity squared, and ϕ is geodetic latitude. So

$$\nu = 6\,378\,137\,\mathrm{m}/\sqrt{(}1 - 0.006\,694\,380\,022\,90\ \sin^2 41.3536096861°),$$
$$= 6\,387\,476.887\,01\,\mathrm{m}$$

From (13a),

$$x = (\nu + h)\cos\phi\cos\lambda,$$
$$= (6\,387\,476.887\,01\,\mathrm{m} + 635.478\,\mathrm{m})\cos(41.3536096861°)\cos(-72.02362233611°),$$
$$= 1\,479\,921.839\,\mathrm{m}$$

Similarly, from (13b),

$$y = (\nu + h)\cos\phi\sin\lambda,$$
$$= (6\,387\,476.887\,01\,\mathrm{m} + 635.478\,\mathrm{m})\cos(41.3536096861°)\sin(-72.02362233611°),$$
$$= -4\,561\,128.808\,\mathrm{m}$$

and from (13c),

$$z = ((1 - e^2)\nu + h)\sin\phi$$
$$= ((1 - 0.00669437999)\,6\,387\,476.887\,01\,\mathrm{m} + 635.478\,\mathrm{m})\sin(41.3536096861°),$$
$$= 4\,192\,401.531\,\mathrm{m}.$$

□

Example 9. Problem: Convert the XYZ coordinates from the previous example back into geodetic coordinates.
Solution: From (14a),

$$\lambda = \tan^{-1}(y/x)$$
$$= \tan^{-1}(-4\,561\,128.808\,\mathrm{m}/1\,479\,921.839\,\mathrm{m})$$
$$= \tan^{-1}(-0.324463943364)$$
$$= -72.0236223361°.$$

Units will be omitted from the following formulae when necessary for clarity. From (14b),

$$\beta = \tan^{-1}\left(\frac{z}{(1-f)\sqrt{(}x^2 + y^2)}\right)$$
$$= \tan^{-1}\left(\frac{4\,192\,401.531}{(1 - 1/298.257\,222\,101)\sqrt{(}1\,479\,921.839^2 + (-4\,561\,128.808)^2)}\right)$$
$$= \tan^{-1}(0.877\,230\,152)$$
$$= 41.258\,215\,228°$$

Notice that β is actually just an intermediate result. It's only purpose is its use as an input to the next equation. From $\epsilon = e^2/(1-e^2)$, $\epsilon = 0.00669437999/(1-0.00669437999) = 0.006\,739\,496\,775\,47$.

Since a is the length of the semimajor axis and b is the length of the minor axis, from (14d)

$$\begin{aligned}
\phi &= \tan^{-1}\left(\frac{z+\epsilon b\sin^3\beta}{\sqrt{(x^2+y^2)}-ae^2\cos^3\beta}\right)\\
&= \tan^{-1}\left(\frac{4\,192\,401.531+0.0067395^2\times 6\,356\,752.314\sin^3 41.3536096861^\circ}{\sqrt{(1\,479\,921.839^2+(-4\,561\,128.808)^2)}-6\,378\,137\times 0.00669437999\cos^3 41.3536096861^\circ}\right)\\
&= \tan^{-1}(0.8801806425481159)\\
&= 41.353\,609\,686^\circ
\end{aligned}$$

And, finally, from (14f)

$$\begin{aligned}
h &= \frac{\sqrt{(x^2+y^2)}}{\cos\phi}-\nu(\phi)\\
&= \frac{\sqrt{(1\,479\,921.839^2+(-4\,561\,128.808)^2)}}{\cos 41.3536096861^\circ}-6\,387\,476.887\,01\\
&= 635.478\,\mathrm{m}
\end{aligned}$$

□

4. Distances

This section defines and compares the distances that are computed in the various coordinate systems presented in section 3. Ground distances, geodesic distances and grid distances are all different from one another. This section will present and explain the differences and give procedures to convert between them. Ground distances are presented first.

4.1. Ground Distances.

Distances between objects on the topographic surface of the Earth are called **ground distances.** Ground distances can be measured with chains or calibrated wheels with odometers and are needed whenever an analysis depends upon actual distances on the Earth's surface. This can be the case for highway construction, hydrology, and meteorology, for example. Mappers and surveyors seldom measure ground distances because most maps are compiled with horizontal or reduced distances. Therefore, ground distances will not be considered further in this material.

4.2. Spatial Distances.

The distance measured by an EDM will be called an **EDM distance**. EDM distances are generally not straight-line distances because the light emitted by the instrument is refracted by the Earth's atmosphere. A Euclidean straight-line distance is called a **spatial distance**. One can attempt to remove the atmosphere's effect on EDM distances by measuring atmospheric temperature, humidity, and pressure. However, this cannot be done accurately: see (Bomford 1980, p. 31) and (Torge 1997, pp. 89-93, 96-107). C.F. Gauss determined an average value for the coefficient of refraction to be 0.13 and this value is widely used and adjusted to account for local conditions.

Spatial distances and zenith angle observations are used for **trigonometric heighting**. The change in ellipsoid height between two stations is given by (Strang and Borre 1997, p. 360)

$$h_{target}-h_{station}=s\cos z+\frac{1-k}{2\eta}s^2\sin^2 z+HI-HR \tag{18}$$

where $h_{station}$ is the ellipsoid height of the station, s is the spatial distance, z is the observed zenith angle, k is the coefficient of refraction, η is the radius of curvature in the normal section, HI is the height of the instrument, and HR is the height of the target (reflector). In most cases, η will not be known for use in this equation – it is often what is ultimately trying to be determined! However, this equation is not very sensitive to the value of η and the value of the semimajor axis is often an acceptable substitute. If the zenith angle is the appropriate combination of left and right face observations, then one need not explicitly account for the deflection of the vertical, unless the stations cannot be thought to have nearly deflections of the vertical (section 5.1 on page 29). This might be the case for long lines in mountainous regions.

The height of the target station is found using (18) and reduced to the length of a normal section chord by (Torge 1997, p. 103)

$$s_0 = \sqrt{\frac{s^2 - (h_t - h_s)^2}{(1 + h_s/\eta)(1 + h_t/\eta)}} \tag{19}$$

where h_t and h_s are the ellipsoid heights of the target and station, respectively, and the other quantities are as in (18). s_0 can be reduced to the length of the normal section by

$$S = 2\eta \arcsin \frac{s_0}{2\eta} \tag{20}$$

One can convert spatial distances to horizontal distances with the relationship

$$d_{hor} = d_{spatial} \cdot \sin z, \tag{21}$$

where z is the zenith angle.

4.2.1. *Chord Length vs. Arc Length.* It is interesting to consider the type of distance measured with chains compared to those measured with an EDM. In the past, there have been long base lines measured with chains that provided the linear basis for large-extent geodetic surveys. In the survey of India (Keay 2000), these base lines were mechanically flattened to remove topographic variation and the chains were placed on leveled tables to remove the catenary shape from the chain. Some of these base lines were nearly ten miles long and took many months to measure. Today, one could simply set up an EDM and prism at the endpoints of the base line and make what might seem to be the same measurement in a matter of seconds, assuming intervisibility of the stations. However, speaking conceptually, these are really not the same measurement at all. Let us consider the difference between straight-line distances and distances measured along the surface of a reference ellipsoid (i.e., on an arc). To motivate and illustrate this idea, we will consider the Verrazano Narrows Bridge located at the mouth of upper New York Bay. As sketched in Figure 12 on the next page, the deck is 4,260 ft long and is 228 ft (69.5 m) above mean high water, and the towers are 693 ft tall. The deck of the bridge is shown as an arc as is the surface of the Bay. Although exaggerated, these are shown as arcs to emphasize the curvature of the Earth. If the bridge's deck were level, then it would have to follow the Earth's curvature. Thus, *level* objects on the Earth's surface are not *straight.* Suppose there is a total station and a prism setup and leveled as shown. The straight line between the total station and the prism depicts the path that the light from an EDM would take, ignoring refraction effects in the atmosphere and the extremely small deflections caused by the Earth's gravitational field (Höpcke 1966). Let us undertake to determine how much different the length of the deck is compared to the distance measured by the EDM.

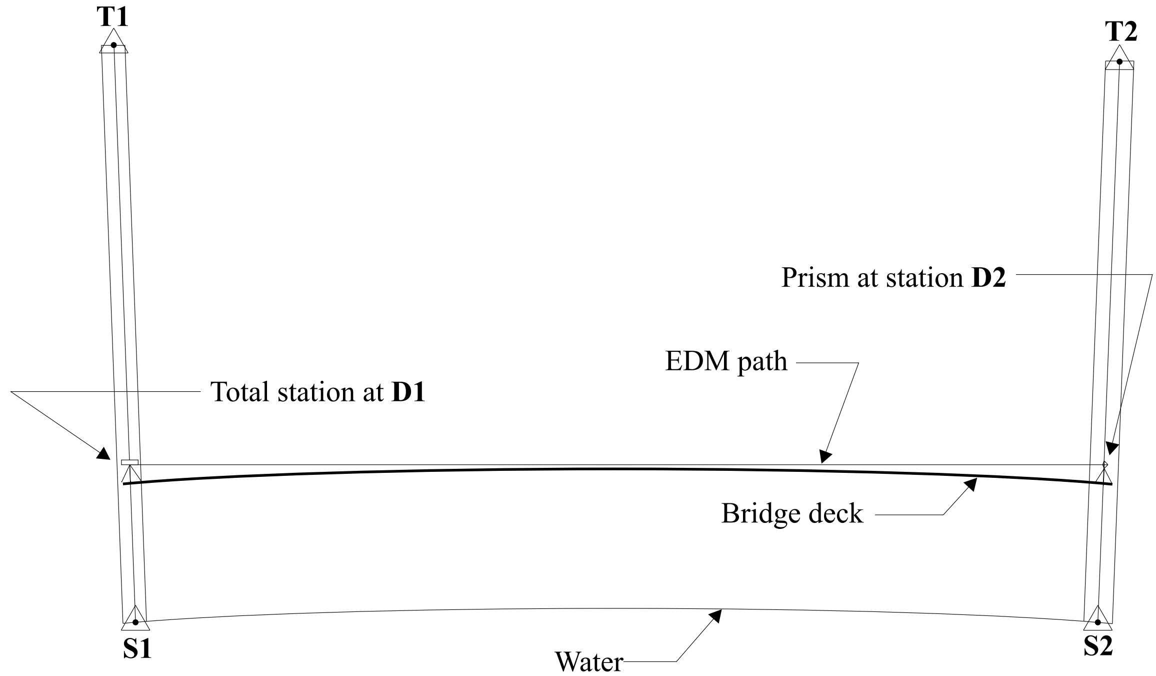

FIGURE 12. A simplistic rendering of the Verrazano Narrows Bridge showing the two towers, the bridge deck, and the water. This illustrates straight line EDM distances (chords) vs. arcs on the ellipsoid.

Consider a circle whose radius equals that of the semimajor axis of the GRS 80 ellipsoid, 6,378,137 m plus 69.5 m for the height of the deck. Denote this distance as $r = 6,378,206.5$ m. Refer to Figure 13 on the facing page. Let **O** denote the point at the center of the circle and define two points **A** and **B** on the circle. Denote the arc distance between **A** and **B** as d_{arc}. From (3), the arc length between **A** and **B** is $d_{arc} = r \cdot \theta$. The length of the chord $\overline{\mathbf{AB}} = c$ is given by the Law of Cosines as $c = \sqrt{2r^2 - 2r^2 \cos\theta}$. For the example of the Verrazano Narrows Bridge, d_{arc} was given as 4,260 ft so

$$\begin{aligned}\theta &= d_{arc}/r \\ &= (4260\,\mathrm{ft} \times 0.3048\,\mathrm{ft/m})/6\,378\,206.5\,\mathrm{m} \\ &= 0.000\,203\,576 \end{aligned} \tag{22}$$

Then

$$\begin{aligned} c &= \sqrt{2r^2 - 2r^2 \cos\theta} \\ &= \sqrt{2 \times 6\,378\,206.5^2 - 2 \times 6\,378\,206.5^2 \cos(0.000\,203\,576)} \\ &= 1298.447997851666\,\mathrm{m}/(0.3048\,\mathrm{ft/m}) \\ &= 4259.999992951660\,\mathrm{ft}, \end{aligned}$$

which is a minute difference. The chord vs. arc distance correction is very small for short baselines and often can safely be ignored. However, Table 2 on the next page lists increasing values of d_{arc} and

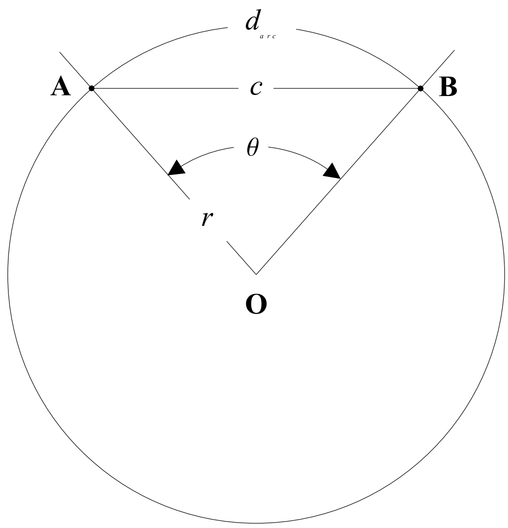

FIGURE 13. Chord distances vs. arc distances.

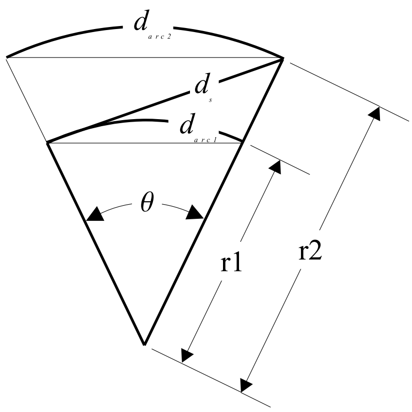

FIGURE 14. Chord distances vs. arc distances.

the corresponding value of c at sea level for comparison. As would be expected, the chord distance is marginally shorter than the arc distance. At 10 km the difference is roughly one millimeter although at 100 km the difference is roughly one meter.

d_{arc}	1 500	10 000	100 000	1 000 000
c	1499.9999965	9 999.998975	99 998.976	998,976.098
c/d_{arc}	0.999999998	0.999999898	0.99998975	0.998976

TABLE 2. Arc distances d, equivalent chord distances c, and their ratio. Distances are given in meters.

We now generalize this discussion and derive a procedure to convert slant distances into the corresponding arc distances. As shown in Figure 14 on the preceding page, an infinite number of horizontal arc distances can be derived from the single slant distance d_s, one for each $r1 \leq r \leq r2$. The chord distance for any $r1 \leq r \leq r2$ is

$$\theta = \arccos\left(\frac{r1^2 + r2^2 - d_s^2}{2\, r1\, r2}\right) \qquad \text{Law of Cosines}$$

so, since the horizontal arc distance $d_{arc} = r \cdot \theta$,

$$d_{arc} = r \cdot \arccos\left(\frac{r1^2 + r2^2 - d_s^2}{2\, r1\, r2}\right) \tag{23}$$

In (23), $r1 = \eta + h$ and $r2 = \eta + h2$, where η is the radius of curvature in the normal section and h is ellipsoid height.

4.3. **Reduced Distances on the Sphere.** This section presents the concept of **reduced distances**, which pertains to the seemingly contradictory idea that two pairs of stations separated only in height are not separated by the same horizontal distance (Clark 1963, Heiskanen and Moritz 1967, Bomford 1980). To elaborate, suppose there are two stations **S1** and **S2** at sea level as shown in Figure 12 on page 22. Then imagine that there are two other stations **D1** and **D2** exactly above **S1** and **S2** at the height of the bridge deck. That is, **S1** has the same latitude and longitude as **D1** and **S2** has the same latitude and longitude as **D2**, but **D1** and **D2** are at a height of 228 ft rather than at sea level. It happens that the arc distance between **S1** and **S2** is less than the arc distance between **D1** and **D2**. As shown in the figure, the essence of the matter is that both pairs of stations are on arcs but the radius for **D1** and **D2** is greater than that for **S1** and **S2**.

The arc distance between **D1** and **D2** was given as 4260 ft or 1,298.448 m. To find the arc distance between **S1** and **S2**, simply apply (3) with a radius that does not include the height of the deck. θ is supplied from (22), so

$$\begin{aligned} d_{\mathbf{S1},\mathbf{S2}} &= a\,\theta \\ &= 6378137\,\text{m} \times 0.000\,203\,576 \\ &= 1298.44\,\text{m} \times 3.28084\,\text{ft/m} \\ &= 4259.96\,\text{ft} \end{aligned}$$

Thus, the distance along the deck is 0.04 ft or 0.49 in longer than the distance across the surface of the bay. The towers stand 693 ft high so their tops are separated by 4260.12 ft, which is greater than the distance on the water by 1.4525 in. These differences are enough to be of concern and they are taken into consideration in the design of such large structures.

For mapping, the situation is usually the opposite. Consider a land survey done in Bogota, Columbia at an elevation of 8,500 ft. Suppose there are two stations connected by a baseline measured on the ground having a length of 10 km. This same baseline measured on the ellipsoid would measure only 9995.94 m. Table 3 on the facing page shows separations of stations on a 10 km baseline on the surface of the sphere and increasing in height to 5000 m. Unlike the distances given in Table 2, these distances increase with height. Distances on a reference ellipsoid are said to be the **reduced** equivalent of the ground distances they correspond with. Reduced distances are often smaller than ground distances, but not necessarily.

Height (m)	Baseline Length (m)
0	10 000.000
100	10 000.157
500	10 000.784
1000	10 001.568
1500	10 002.352
2000	10 003.136
3000	10 004.704
4000	10 006.271
5000	10 007.839

TABLE 3. The length of baselines separating stations at identical latitude and longitude, differing only in height.

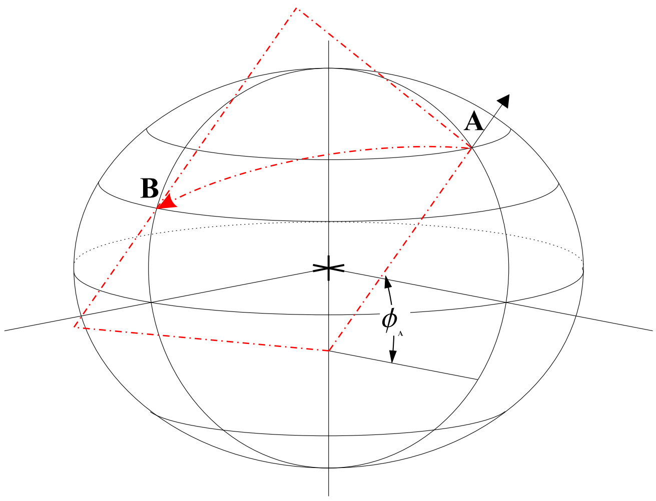

FIGURE 15. The normal section from **A** to **B**. Notice that the section plane includes the normal vector at **A** and contains **B**.

Unlike the discrepancy between arc length and chord length, this is a matter of concern for at least two reasons. First, distances reported by GPS post-processing software are given *on the surface of the reference ellipsoid, not on the ground.* Second, maps compiled in an absolute coordinate system such as the State Plane Coordinate System typically employ reduced distances, as well. Therefore, significant inconsistencies arise when a surveyor uses ground distances to compile such a map or compares ground distances with GPS-derived distances. See Stem (1995) for a more complete treatment of how to use reduced distances (including traverses), particularly when mapping in the State Plane Coordinate System.

4.4. **Normal Section Distances.** Suppose there is an observation made from a point **A** to a target **B**; see Figure 15. Let $\nu_{\mathbf{A}}$ be the normal at **A**. The **normal section** from **A** to **B** is the curve on a spheroid made by the intersection of a plane containing $\nu_{\mathbf{A}}$ and **B**.

Consider how an EDM measures a distance. Suppose that an EDM and a target reflector have been leveled perfectly with respect to some reference ellipsoid. Then, the EDM will measure a straight-line to the reflector, forming a plane containing three points: the intersection of the radius of curvature in the prime vertical with the minor axis, the optical center of the EDM, and the reflector. The intersection of this plane with the surface of the ellipsoid forms a normal section. It might seem obvious that the length of a normal section should be a geodesic since it seems to somehow be a straight-line path between the station and the target. Were the reference surface a sphere, this would be so; it is the definition of a great circle. However, this is generally not the case on a spheroid. This section presents formulae to compute the length of the normal section.

Clarke (1880) first presented the original formulae. Several methods were presented and compared in (A.C.I.C. 1959). Robbins (1962) developed these further, giving formulae with a maximum error of 1/180 M over 1600 km. Bowring (1971) extended Robbins (1962) so that distances up to 6000 miles can be computed. Bomford (1980, p. 102) notes that, "The length of the normal section differs from that of the geodesic by under 1 in 150 000 000 in a side of 3 000 km, and the difference can always be ignored." Even so, Bomford gives the corrections on page 95 and 499. Thus, normal section distance is practically equivalent to geodesic difference, the distinction being purely mathematical in nature. Therefore, separate formulae for normal section distances will be presented and the interested reader is referred to the citations for details.

4.5. **Geodesic Distances.** Geodesic distances are computed using flat-Earth models, spherical-Earth models, and ellipsoidal-Earth models.

4.5.1. *Flat-Earth Geodesic Distances.* For flat-Earth models, distances are given by the Euclidean distance formula (2) in either two or three dimensions, whichever is appropriate.

4.5.2. *Spherical-Earth Geodesic Distances.* On the sphere, the shortest distance between two points is the length of the (shorter) arc of the great circle common to the points. The length of this arc is given by (3), where r is the radius of the circle, which in this case equals the radius of the sphere and θ is the angle subtended by the arc in the plane of the great circle. θ can be found using the Law of Cosines for spherical triangles,

$$\theta = \arccos[\sin\phi_1 \sin\phi_2 + \cos\phi_1 \cos\phi_2 \cos(\lambda_1 - \lambda_2)], \tag{24}$$

where ϕ is latitude, λ is longitude, and the subscripts denote points 1 and 2, respectively. Substituting (24) into (3) gives

$$d = r \cdot \arccos[\sin\phi_1 \sin\phi_2 + \cos\phi_1 \cos\phi_2 \cos(\lambda_1 - \lambda_2)], \tag{25}$$

where r is the radius of the sphere. The computation of (25) looses a great deal of numerical accuracy for very short distances. The following formula is suitable for all distances.

$$\sin^2(d/2r) = \sin^2([\phi_2 - \phi_1]/2) + \cos\phi_1 \cdot \cos\phi_2 \cdot \sin^2([\lambda_2 - \lambda_1]/2) \tag{26}$$

4.5.3. *Spheroidal-Earth Geodesic Distances.* Even though (26) is computationally stable at all distances, it is nonetheless inappropriate for use with spheroids. A great deal of attention has been given to finding the length of geodesics on spheroids. Bomford (1980, p. 108-110) provides a review and (McCleary 1994, Casey 1996) provide a theoretical treatment of the general subject. Geodesy-specific treatments include (Rainsford 1949, Rainsford 1955, A.C.I.C. 1959, Sodano 1965, Bowring 1969, Bowring 1972, Vincenty 1971, Vincenty 1975, Murphy 1981, Bowring 1985, Day

1987, Danielsen 1994, Vassallo and Secci 1995). Bowring (1996) presented what seems to be a definitive work that claims, "The inverse problem for all possible geodesics on the spheroid is solved in ways that are selected by the programme in a manner appropriate to any two given end positions. The comparatively simple total inverse solution for the great elliptic is also given."

The problem of finding the geodesic distance between two geodetic points on a spheroid is related to finding the forward and backward azimuth between those points. We present Robbins's formula (Robbins 1962) as given in (Bomford 1980, p. 108-110). This formula is correct to 1 in 10^8 at 1600 km.

$$\tan\psi_2 = \left(1 - e^2 + e^2\frac{\nu_1\sin\phi_1}{\nu_2\sin\phi_2}\right)\tan\phi_2,\ \delta\phi_2 = (\phi_2 - \psi_2), \tag{27a}$$

$$\cot\alpha_{12} = (\cos\phi_1\tan\psi_2 - \sin\phi_1\cos\Delta\lambda)\csc\Delta\lambda, \tag{27b}$$

$$\cot\alpha'_{21} = (\sin\phi_2\cos\Delta\lambda - \cos\phi_2\tan\psi_1)\csc\Delta\lambda, \tag{27c}$$

$$\sin\sigma = \sin\Delta\lambda\cos\psi_2\csc\alpha_{12}, \tag{27d}$$

$$= -\sin\Delta\lambda\cos\phi_1\csc\alpha'_{21}, \qquad \text{(for check)}$$

$$\alpha_{21} = \alpha'_{21} - \delta\phi_2\sin\alpha'_{21}\tan\frac{1}{2}\sigma \tag{27e}$$

Eq. 27b provides the forward azimuth and (27e) provides the back azimuth. The geodesic distance is given by

$$d = \nu_1\sigma\left[1 - \frac{\sigma^2}{6}h^2(1-h^2) + \frac{\sigma^3}{8}gh(1-2h^2) + \frac{\sigma^4}{120}\left\{h^2(4-7h^2) - 3g^2(1-7h^2)\right\} - \frac{\sigma^5}{48}gh\right], \tag{28}$$

where

$$g^2 = \epsilon\sin^2\phi_1 \text{ and}$$

$$h^2 = \epsilon\cos^2\phi_1\cos^2\alpha_{12}$$

and ϵ is from (6).

When α_{12} is near 0° or 180°, get $(\sin\Delta\lambda\csc\alpha_{12})$ for (27d) from (27b) written as $\sin\Delta\lambda\csc\alpha_{12} = (\cos\phi_1\tan\psi_2 - \sin\phi_1\cos\Delta\lambda)\sec\alpha_{12}$ and similary from (27c) when α'_{21} is near 0° or 180°. α_{21} may be got from (27a) and (27b) with suffixes 1 and 2 interchanged, instead of from (27d) and (27e). The second expression in (27d) cannot then be used.

The reader is encouraged to compare (28) with (25). The former accounts for the ellipsoid whereas the latter is a great circle distance.

4.6. **Grid Distances.** Grid distances are computed with the two-dimensional version of (2). So, for two points **A** and **B** in a projected cartographic grid,

$$d_g = [(\Delta x)^2 + (\Delta y)^2]^{1/2}, \tag{29}$$

where $\Delta x = x_{\mathbf{A}} - x_{\mathbf{B}}$ and $\Delta y = y_{\mathbf{A}} - y_{\mathbf{B}}$. Grid distances are akin to pseudo-ranges in that both contain intrinsic errors. Pseudo-ranges are in error due to the unmeasurable time bias in the GPS clocks, whereas grid distances as given in (29) are biased by the projection scale factor discussed in section 3.4.1 on page 14. In general, scale factors are functions of latitude, longitude and azimuth. For conformal projections, they are not functions of azimuth but will be functions of either latitude or longitude, or both.

The projection scale factor k is the ratio of the grid distance given by (29) to geodetic (ellipsoidal) distance d_e. k is defined on infinitely small curves ds. So, $ds_e = ds_g/k$. For a curve on the ellipsoid $\mathcal{C}_e$ and its projection on the grid $\mathcal{C}_g$,

$$d_e = \int_{\mathcal{C}_e} ds_e \qquad \text{from (1)}$$

and by $ds_e = ds_g/k$,

$$= \int_{\mathcal{C}_g} 1/k \cdot ds_g. \tag{30}$$

If k is essentially a constant over $\mathcal{C}_g$, then

$$\approx 1/k \int_{\mathcal{C}_g} ds_g.$$

If $\mathcal{C}_g$ is a straight line, then

$$\approx 1/k \, [(\Delta x)^2 + (\Delta y)^2]^{1/2}. \qquad \text{from (29)}$$

And finally, if $\mathcal{C}_e \approx \mathcal{G}$, then

$$d \approx 1/k \, [(\Delta x)^2 + (\Delta y)^2]^{1/2} \tag{31}$$

where d is the geodesic distance. It is the goal of the State Plane Coordinate System to ensure that (31) holds over the extent of the map. This is done by ensuring that the standards are sufficiently close together and that the map extent is not so great that k exceeds one part in 10,000. Under these circumstances, (31) can be used in lieu of (28) while maintaining survey-grade accuracy. The simplicity of (31) compared with (28) justifies the effort needed to select the projection correctly.

It is appropriate to quantify the effects of the assumptions going into (31). The following sections will describe and analyze the errors these assumptions create.

4.6.1. *Constant Scale Factor.* The projection scale factor is, in general, a function of azimuth, latitude and longitude. When computing distances we are confronted with general integrals such as (30), which could be approached with a numerical integration technique, such as Gaussian Quadrature. However, if the base line is so long as to warrant such a technique, it is probably best to convert the grid coordinates to geodetic coordinates and find the distance using (28).

It is often acceptable to use

- a single k for an entire project,
- a single k for each base line, or
- an average value for k taken from various points on the base line.

If the project's extent is fairly small, it is possible that a single value for k will suffice for all base lines. This can be checked by computing k at the extremes of the project and deciding if the average of these values differs from the extremes too much. If it does, then scale factors can be assigned to each base line individually. The simplest approach is to use the scale factor at the midpoint of the line. If this is unacceptable, the average of the scale factors at the endpoints is better. However, if this is still unacceptable, use

$$k_{\mathbf{A},\mathbf{B}} = \frac{k_{\mathbf{A}} + 4\,k_{midpoint} + k_{\mathbf{B}}}{6}$$

where **A,B** are the endpoints of the base line (Stem 1995, p. 50).

4.6.2. *Geodesics vs. Grid Straight lines.* The last assumption for (31) is that a straight line on the grid is an acceptable approximation for the corresponding geodesic. This bares a similarity to the situation regarding normal sections (see section 4.4 on page 25), but it is different. Only gnomic projections project geodesics into straight lines. All other projections project geodesics into complex curves, in general.

5. Observation Corrections

Field measurements differ from geodetic measurements in at least three ways. First, field measurements made with instruments that are leveled to the local vertical (i.e., with respect to gravity) use a different horizontal plane than the corresponding plane on the ellipsoid. Second, observations over relatively long base lines whose end points are greatly different in height incur a bias in the observed versus the geodetic azimuth. Third, field measurements made with opto-mechanical instruments occur along normal sections (ignoring atmospheric refraction and other small factors that bend light). The length of a normal section is longer than the geodesic between the observation station and the target. The third point was discussed in section 4.4 on page 25. This section discusses the first two issues and present corrections.

5.1. **Gravity: Deflection of the Vertical.** The **geoid** is the gravitational equipotential surface that nominally defines mean sea level (Kellogg 1953, Ramsey 1981, Torge 1997, Blakely 1995). If the Earth were homogeneous, spherical, and not rotating, the geoid would be spherical. However, the Earth's mass is very inhomogeneously distributed, its shape is more nearly a spheroid than a sphere, and it is rotating. These factors plus others cause the geoid to not be spherical and, instead, adopt a shape that roughly reflects that of the topographic surface of the Earth. One consequence of this is that local vertical ("up") is generally not parallel to the normal to the reference ellipsoid's surface. The difference between these two is called the **deflection of the vertical**.

The deflection of the vertical causes angular measurements that are horizontal with respect to vertical to differ from angular measurements that are horizontal with respect to normal. See Figure 16 on the next page, which is based on Bomford (1980, p. 94). Observations are taken at station **A** of a target at **B**. $\overline{\mathbf{Aa'}}$ is the vertical and $\overline{\mathbf{Aa}}$ is the normal. The deflection of the vertical is usually broken into two parts, deflection in the plane of the meridian ξ and deflection in the plane of the prime vertical η. Define $\zeta = \xi \sin\alpha - \eta\cos\alpha$, where α is the measured azimuth of the normal section from the transit to the target reckoned clockwise from north. ζ is the angular difference of what was measured versus what would have been measured had the instrument been level with respect to the ellipsoid instead of the geoid.

Call the angle of elevation θ, the slant distance to the target d_s, and the vertical distance to the target d_v. Then, $\tan\theta = d_v/d_h$ so $d_v = d_h \cdot \tan\theta$. Therefore, $\mathbf{t'}\mathbf{t''} = \zeta \cdot d_h\ \tan\theta$. Then the correction

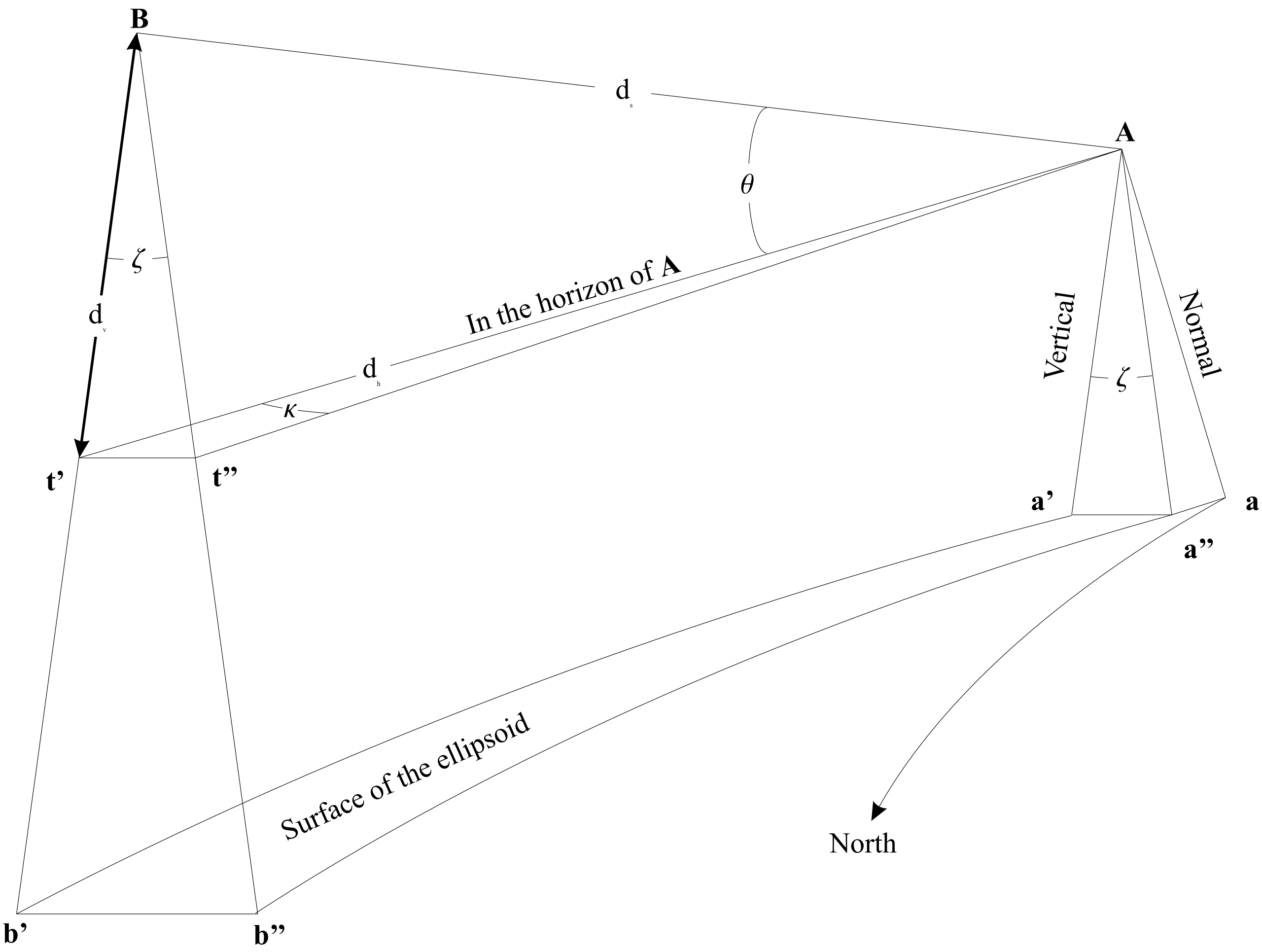

FIGURE 16. The deflection of the vertical.

factor κ to be added to the measured azimuth is

$$t't'' = d_h \cdot \kappa \qquad \text{from (3)}$$
$$\kappa = t't''/d_h$$
$$= (\zeta \cdot d_h \tan\theta)/d_h \qquad \text{sub for } t't''$$
$$= -\zeta \tan\theta \qquad (32)$$

The minus sign in (32) appears so that $\alpha = A + \kappa$ produces the reduced geodetic azimuth, where A is the observed azimuth.

Clearly, this is a factor that is more relevant to baselines that have significant angles of elevation, such as ones that might be observed from the base of mountains to their peaks. In practice this correction is usually less than five arc-seconds but can approach 30 arc-seconds in mountainous areas.

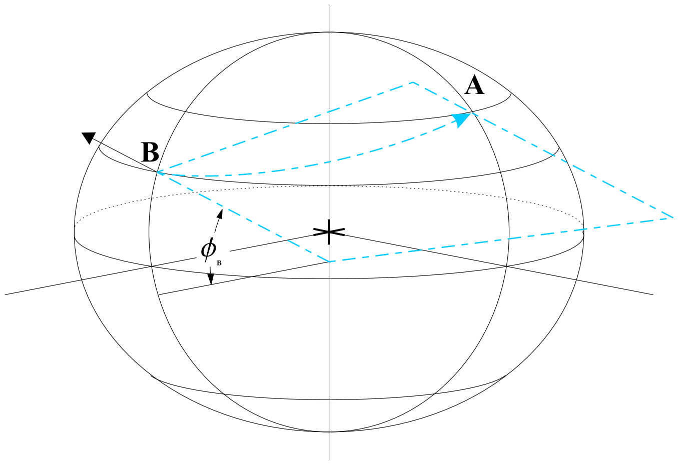

FIGURE 17. The normal section from **B** to **A**. Notice that the section plane includes the normal vector at **B** and contains **A**.

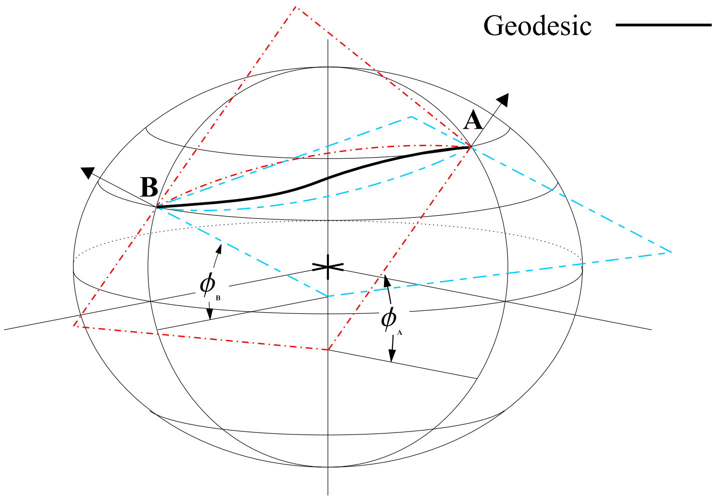

FIGURE 18. A composite picture putting Figure 15 with Figure 17. Note how the normal vectors intersect the z-axis in different locations, and that they are not coplanar. The geodesic from **A** to **B** is shown between the normal sections as a bold solid curve.

The National Geodetic Survey freely distributes a software package that estimates the deflection of the vertical anywhere in the United States. This program can be used to produce a value for ζ, which can then be used in (32).

Thomson and Vaníček (1974) discuss two sources of error in this derivation: it assumes the normals intersect and that the mean Eulerian radii is an adequate surrogate for the truth. The worst case is analyzed and found to have negligible error magnitudes. They conclude the standard formulae as presented are adequate.

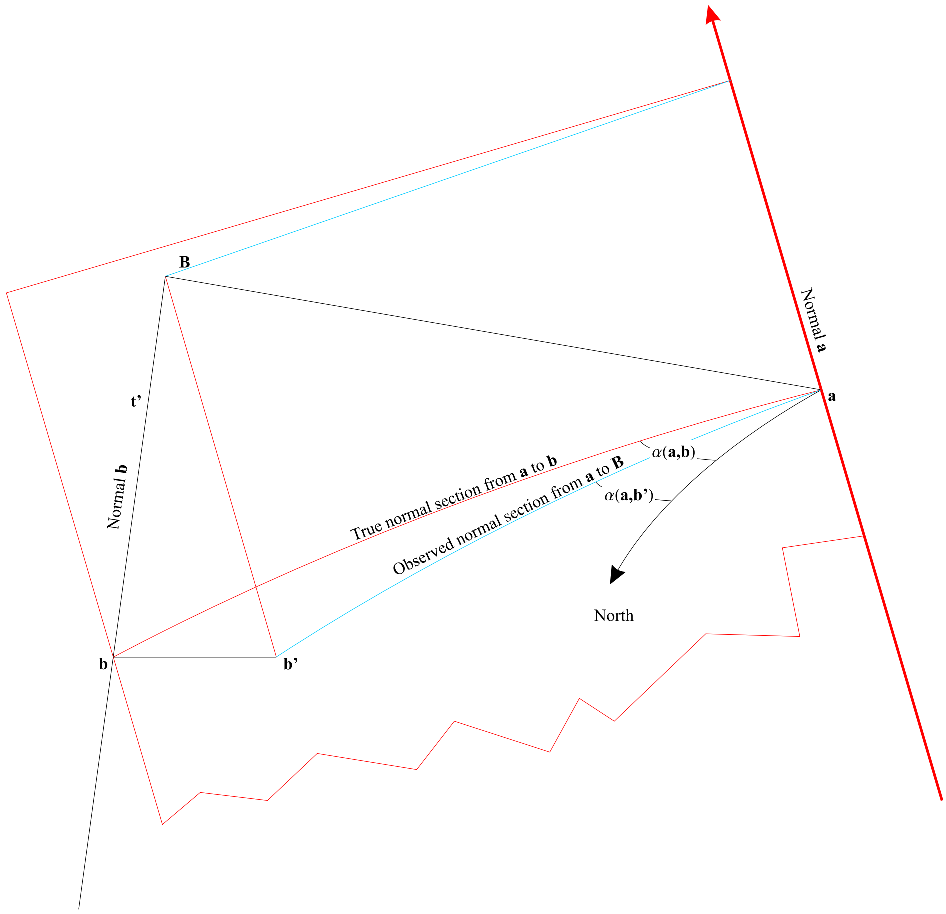

FIGURE 19. The observing station is on the normal at **a** and the target station is at **B**. When **B** is not on the ellipsoid, the observed azimuth from **a** to **B** varies from the true value to do the normal at **b** being skew to the normal at **a**.

5.2. **Surveys at Altitude: Skew of the Normals.** As shown in Figure 17 on the preceding page, the normal section from **B** to **A** is the curve on a spheroid made by the intersection of a plane containing $\nu_{\mathbf{B}}$ and **A**. Figure 17 is the counterpart to Figure 15 on page 25, which shows the normal section from **A** to **B**. These two figures are drawn together in Figure 18 for comparison. As shown in the figure, if $\phi_{\mathbf{A}} \neq \phi_{\mathbf{B}}$, then $\nu_{\mathbf{A}}$ and $\nu_{\mathbf{A}}$ do not intersect the minor axis at the same place (Bowring 1971, Bowring 1987). Consequently, the two normal sections are different curves.

When the target is not on the ellipsoid, there is a minor offset in space due to the skew of the normals (Figure 19). The normal plane at **a** is shown in red and the true normal section from **a** to **b** is also drawn in red. However, the line $\overline{\mathbf{Bb'}}$ is parallel to the normal at **a** and, when intersected with the ellipsoid, produces point **b′** rather than **b**. As indicated in the figure, this is caused by the normal at **b** being skew to the normal at **a**. This results in a small pointing error, called the

FIGURE 20. Process of converting between grid, ground, and geodesic distances.

skew of the normals (Torge 1997, p. 141,215), (Bomford 1980, p. 94). The correction is

$$\psi = -\frac{h_2 \epsilon}{2\nu(\phi_{\mathbf{B}})} \sin 2\alpha_{\mathbf{A},\mathbf{B}} \cos^2 \phi_{\mathbf{B}} \tag{33}$$

where h_2 is the ellipsoid height of **B**, ϵ is the second eccentricity, $\alpha_{\mathbf{A},\mathbf{B}}$ is the geodetic azimuth from **A** to **B**, and $\phi_{\mathbf{B}}$ is the geodetic latitude of **B**. As with (32), the minus sign in (33) appears so that this correction is always added to the observed azimuth: $\alpha = A + \psi$. According to Bomford (1980, p. 95), "This is a small correction, not more than $0.3''$ if $h_2 = 3000$ meters. Except in mountainous country it can reasonably be ignored..."

6. Making it Work

Having laid out all the theoretical considerations, this section presents how to start with a grid, ground, or geodesic distance and convert it to one of the others. The process is conceptually straightforward and is shown in Figure 20. Always convert ground to geodesic, geodesic to either, and grid to geodesic. For example, to convert grid to ground, first convert grid to geodesic and then convert the geodesic to ground.

Example 10. given a slant distance d_s measured with an EDM at an observed azimuth α_s and an observed zenith angle of z_s at a geodetic position of ϕ, λ, h.
Required: the geodesic distance and azimuth.
Solution

1. Correct the observed EDM distance to a spatial distance for atmospheric refraction.
2. Correct the observed azimuth for the deflection of the vertical.
3. Correct the observed azimuth for skew of the normals.
4. Reduce the spatial distance to the ellipsoid forming the length of a normal section.

In theory, this process is straightforward. However, in practice, there is an issue involving trigonometric leveling. Not surprisingly, the process of reducing an EDM measurement to the ellipsoid requires knowledge of the horizontal distance to the target. Of course, computing a horizontal distance is equivalent to trigonometric leveling. According to Bomford (1980, p. 31),"Less accuracy is required (for vertical angles) than for horizontal angles, since atmospheric refraction makes high accuracy impossible in any case, and since the final height control will be by spirit leveling." If it is impossible to make high accuracy vertical angle measurements, the question becomes how sensitive is the answer to errors in vertical angle measurements. To provide a real-world example, three temporary survey markers were constructed and surveyed on the campus of the University of Connecticut in the Fall of 2002. The markers were surveyed in three sessions having simultaneous occupation of the three markers using two Javad Positioning System and one Trimble Navigation L1/L2, C/A-P code, phase receivers using a nearby cooperative CORS station (NRME) for control. Table 4 on the following page shows the results of the survey.

These coordinates were used to inverse the geodesic distance (28) and spatial distance (2), as shown in Table 5 on the next page.

	Coordinates NAD83 CORS			Sigmas (mm)			Corr. (%)		
Point Name	Latitude	Longitude	Height (m)	s(N)	s(E)	s(U)	N-E	N-U	E-U
HBH1	$41°49'08.49900''N$	$72°15'10.88705''W$	187.3853	1.3	1.2	2.8	-25	-16	25
HBH2	$41°48'59.20173''N$	$72°14'49.14831''W$	184.5880	1.2	1.1	2.6	-23	-9	19
HBH3	$41°48'53.30021''N$	$72°14'50.94347''W$	178.0003	1.3	1.2	2.8	-30	-18	26

TABLE 4. Geodetic coordinates of the three survey markers.

	HBH1	HBH2	HBH3
HBH1	0	577.933 (577.956)	657.087 (657.172)
HBH2		0	186.732 (186.854)

TABLE 5. Geodesic and (spatial distances) between stations in meters.

Average EDM Distance	Spatial (computed)	EDM - Spatial
578.028	577.956	0.072
657.246	657.172	0.074
186.926	186.854	0.072

TABLE 6. Geodesic and (spatial distances) between stations in meters.

These stations were occupied with a Leica TC307 total station. Distances and bearing to all combinations of markers were measured with at least two sets of left and right face angular measurements plus all reciprocal vertical angles. The distances were measured using the total station's EDM with atmospheric temperature, pressure and humidity being input from a meteorological station on site to be used for refraction corrections. Table 6 shows that the corrected EDM measurements (refraction and prism constant) agree well with the computed spatial distances but that they are longer, as would be expected.

Table 7 on the next page shows the measured vertical angles, truth (as computed from the coordinates), and the residual. Notice that the errors are huge compared to the precision of the instrument, which is seven arc-seconds. This is due to the refraction of the atmosphere.

Without reciprocal zenith angle measurements, the measurements as given will be used to find the horizontal distance to be reduced to the ellipsoid. Table 8 on the facing page shows the results. The EDM distance is shown for reference. The geodesic distance was computed by inversing between the station coordinates. The reduced distance is derived by applying equations to the observed slant distance and zenith angle. The reduced error shows the accuracy of the results, which generally agree with the GPS-derived values at the decimeter level.

This example illustrates two points. First: EDM measured distances and GPS derived spatial distances can agree closely. Second, EDM measured distance, if used with trigonometric leveling, can produce reduced horizontal distances or height differences that agree at the decimeter level with GPS-derived results.

The situation is improved somewhat with reciprocal zenith angle measurements. Compare the average of the computed change in height for station HBH2 to HBH3 (-6.076) and the computed change in height for station HBH3 to HBH2 (6.941) = -6.509 against the true value, -6.588. This

Station	Target	Computed Zenith	Observed Zenith	Zenith Error (DMS)		
HBH1	HBH2	90.2799	90.3265	0°	−2′	−47.6479
HBH1	HBH3	90.8212	90.6552	0°	9′	57.6165
HBH2	HBH1	89.7253	89.6051	0°	7′	12.8372
HBH2	HBH1	89.7253	89.6044	0°	7′	15.1412
HBH2	HBH3	92.0213	91.7150	0°	18′	22.6509
HBH3	HBH1	89.1847	89.2851	0°	−6′	−1.5099
HBH3	HBH1	89.1847	89.2750	0°	−5′	−24.9699
HBH3	HBH1	89.1847	89.2855	0°	−6′	−2.9499
HBH3	HBH2	87.9804	87.6597	0°	19′	14.6423
HBH3	HBH2	87.9804	87.6572	0°	19′	23.4623

TABLE 7. Measured zenith angle, computed zenith angles, and the error.

Station	Target	EDM Distance	Geodesic Distance	Reduced EDM Distance	Error (cm)
HBH1	HBH2	578.021	577.933	577.990	5.75
HBH1	HBH3	328.612	657.087	657.155	6.86
HBH2	HBH1	578.031	577.933	578.003	7.06
HBH2	HBH1	578.031	577.933	578.003	7.05
HBH2	HBH3	186.903	186.732	186.799	6.66
HBH3	HBH1	657.244	657.087	657.180	9.34
HBH3	HBH1	657.243	657.087	657.180	9.31
HBH3	HBH1	657.249	657.087	657.186	9.92
HBH3	HBH2	186.949	186.732	186.815	8.22
HBH3	HBH2	186.948	186.732	186.814	8.14

TABLE 8. Slant, geodesic, and reduced distances (m).

is an error of eight centimeters versus one-half meter for the unaveraged values. The computed reduced distance is 186.785 versus the geodesic distance of 186.732, an error of 53 millimeters. This is an improvement of about 15 millimeters compared to the non-reciprocal scenario.

However, in spite of all the redundant measurements and attention to detail, at this height the EDM-measured reduced distances exceed the GPS values by roughly a factor of three or worse. The problem is evidently the inability to account for the effective refractive index, which can be seen by the correlation between the EDM slant distances compared to the spatial distances (Table 6). This 300% error level cannot be considered acceptable, which reconfirms what was already known: trigonometric leveling is unacceptable for geodetic-level work.

□

Example 11. : The geodetic coordinates for HBH1, HBH2, and HBH3 have been converted into State Plane Coordinate System, CT0600, NAD83 meters and into UTM 18N meters (Table 9 on the next page).

	State Plane NAD83 CT0600			UTM NAD83 18N		
Point Name	Easting	Northing	Height	Easting	Northing	Height
HBH1	346091.482	261990.665	187.385	728151.206	4633332.566	186.146
HBH2	346594.854	261706.728	184.588	728661.977	4633061.862	183.348
HBH3	346554.481	261524.413	178.000	728626.391	4632878.506	176.760

Table 9. Cartographic coordinates of the three survey markers. All coordinates are in meters.

	HBH1-HBH2	HBH1-HBH3	HBH2-HBH3
Geodesic	577.933	657.087	186.732
SPCS	577.930	657.084	186.732
UTM	578.073	657.246	186.778

Table 10. Geodesic and uncorrected grid distances (m) between all stations.

	HBH1-HBH2	HBH1-HBH3	HBH2-HBH3
Geodesic	577.933	657.087	186.732
SPCS	577.933	657.087	186.733
UTM	578.932	657.086	186.733

Table 11. Geodesic and corrected grid distances (m) between all stations.

Table 10 shows the distances between the stations in the various coordinate systems. As expected, the SPCS distances match the geodesic distances fairly closely, in this example not being different by more than three millimeters. The UTM distances, however, are different by 140, 159, and 56 millimeters for the three distanes. The UTM scale factors for the three stations are 1.000240581, 1.000243453, and 1.000243253, with an average of 1.000242429. The SPCS scale factors for the three stations are 0.999995519, 0.999995295, and 0.999995155, with an average of 0.999995323. Correcting the grid distances with the average scale factor produces good agreement with the geodesic distances, with the maximum difference being only one millimeter (Table 11).

References

A.C.I.C., 1959, Geodetic distance and azimuth computations for lines over 500 miles: technical report 80, United States Air Force Aeronautical Chart and Information Center.

Blakely, Richard J., 1995, Potential Theory in Gravity and Magnetic Applications: Cambridge University Press, Cambridge.

Bomford, Guy, 1980, Geodesy: Clarendon Press, Oxford, 4th ed.

Borkowski, Kazimierz M., 1989, Accurate algorithms to transform geocentric to geodetic coordinates: Bulletin Géodesique, vol. 63, no. 1, pp. 50–56.

Bowring, B. R., 1969, The further extension of the Gauss inverse problem: Survey Review, vol. XX, no. 151, pp. 40–43.

Bowring, B. R., 1971, The normal section – forward and inverse formulae at any distance: Survey Review, vol. XX, no. 161, pp. 131–135.

Bowring, B. R., 1972, Distance and the spheroid (correspondence): Survey Review, vol. XXI, no. 164, pp. 281–284.

Bowring, B. R., 1976, Transformation from spatial to geographical coordinates: Survey Review, vol. XXIII, no. 181, pp. 323–327.

Bowring, B. R., 1985, The accuracy of geodetic latitude and height equations: Survey Review, vol. 28, no. 218, pp. 202–206.

Bowring, B. R., 1987, Notes on the curvature in the prime vertical section: Survey Review, vol. 29, no. 226, pp. 195–196.

Bowring, B. R., 1996, Total inverse solutions for the geodesic and great elliptic: Survey Review, vol. 33, no. 261, pp. 461–476.

Bugayevskiy, Lev M. and Snyder, John P., 1995, Map projections: A reference manual: Taylor & Francis, Philadelphia, PA.

Casey, James, 1996, Exploring curvature: Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig/Wiesbaden, Germany.

Clark, D., 1963, Plane and Geodetic Surveying *II*: Constable and Co. Ltd., 5th ed.

Clarke, A.R., 1880, Geodesy: Clarendon Press, Oxford.

Danielsen, J.S., 1994, The use of Bessel-spheres for solution of problems related to geodesics on the ellipsoid: Survey Review, vol. 32, no. 253, pp. 445–449.

Day, J.W.R., 1987, A refined chord-arc method of calculating geodesics: Survey Review, vol. 29, no. 226, pp. 191–194.

Ewing, Clair E. and Mitchell, Michael., 1970, Introduction to Geodesy: American Elsevier Publishing Company, Inc., New York.

Fukushima, T., 1999, Fast transform from geocentric to geodetic coordinates: Journal of Geodesy, vol. 73, pp. 603–610.

Gerdan, G.P. and Deakin, R.E., 1999, Transforming Cartesian coordinates X,Y,Z to geographical coordinates ϕ, λ, h: Australian Surveying, vol. 44, no. 1, pp. 55–63.

Heiskanen, Weikko Aleksanteri and Moritz, H., 1967, Physical Geodesy: W. H. Freeman and Company, San Fransisco.

Hofmann-Wellenhof, B.; Lichtenegger, H.; and Collins, J., 1997, Global Position System: Theory and Practice: Springer–Verlag Wien New York, New York, 4th ed.

Höpcke, W., 1966, On the curvature of electromagnetic waves and its effect on measurement of distance: Survey Review, vol. 141, pp. 298–312.

Iliffe, Jonathan, 2000, Datums and map projections for remote sensing, GIS, and surveying: Whittles Publishing, New York.

Keay, John, 2000, The Great Arc: the dramatic tale of how India was mapped and Everest was named: Harper Collins College Publishers, New York.

Kellogg, Oliver Dimon, 1953, Foundations of Potential Theory: Dover Publications, Inc., New York.

Laskowski, P., 1991, Is Newton's iteration faster than simple iteration for transformation between geocentric and geodetic coordinates?: Bulletin Géodesique, vol. 65, pp. 14–17.

Lee, L. P., 1963, Scale and convergence in the Transverse Mercator projection of the entire spheroid: Survey Review, vol. XVII, no. 127, pp. 49–50.

Leick, Alfred, 1995, GPS Satellite Surveying: John Wiley & Sons, New York, 2nd ed.

Lin, K.C. and Wang, J., 1995, Transformation from geocentric to geodetic coordinates using Newton's iteration: Bulletin Géodesique, vol. 69, no. 4, pp. 14–17.

Maling, D.H., 1973, Coordinate systems and map projections: George Philip & Son, Ltd., London.

McCleary, John, 1994, Geometry from a differentiable viewpoint: Cambridge University Press, Cambridge.

Moritz, H., 2000, Geodetic Reference System 1980: Journal of Geodesy, vol. 74, no. 1, pp. 128–162.

Murphy, D. W., 1981, Direct problem geodetic computation using a programmable pocket calculator: Survey Review, vol. 26, no. 199, pp. 11–16.

Pollard, J., 2002, Iterative vector methods for computing geodetic latitude and height from rectangular coordinates: Journal of Geodesy, vol. 76, no. 1.

Qihe, Yang; Snyder, John P.; and Tobler, Waldo R., 2000, Map projection transformation principles and applications: Taylor & Francis, Philadelphia, PA.

Rainsford, Hume F., 1949, Long lines on the Earth: Various formulae: Empire Survey Review, vol. 71,72.

Rainsford, Hume F., 1955, Long geodesics on the ellipsoid: Bulletin Géodesique, vol. 37.

Ramsey, A.S., 1981, Newtonian Attraction: Cambridge University Press, Cambridge.

Robbins, A. R., 1962, Long Lines on the Spheroid: Emperial Survey Review, vol. xvi, no. 125, pp. 301–309.

Snyder, John P., 1987, Map Projections – A Working Manual: Professional Paper 1395, U.S. Geological Survey, Washington, DC.

Sodano, E.M., 1965, General non-iterative solution of the inverse and direct geodetic problems: Bulletin Géodesique, vol. 75.

Stem, James E., 1995, State Plane Coordinate System of 1983: NOAA Manual NOS NGS 5, National Oceanic and Atmospheric Administration, Rockville, MD.

Stoughton, Herbert W., 1982, Simple algorithms for calculation of scale factors for Transverse Mercator systems: Survey Review, vol. 26, no. 206, pp. 386–391.

Strang, Gilbert and Borre, Kai, 1997, Linear algebra, geodesy, and GPS: Wellesley-Cambridge Press, Wellesley MA.

Taff, L.G., 1985, Celestial mechanics. A computational guide for the practitioner: North–Holland, New York.

Thomson, D. B. and Vaníček, P., 1974, Note on the reduction of spatial distances to a reference ellipsoid: Survey Review, vol. XXII, no. 173, pp. 309–312.

Tissot, Nicholas Auguste, 1881, Mémoire sur la représentation des surfaces et les projections des cartes géographiques: Gauthier Villars, Paris.

Torge, Wolfgang, 1997, Geodesy, 2nd ed.: Walter de Gruyter, New York.

Van Sickle, Jan, 1996, GPS for Land Surveyors: Ann Arbor Press, Inc., Chelsea, MI.

Vaníček, Petr and Krakiwsky, E.J., 1996, Geodesy: The Concepts: Elsevier Scientific Publishing Company, Amsterdam, 3rd ed.

Vassallo, A. and Secci, A., 1995, Unique algorithm for the calculation of geodetic distances with the solution of the first fundamental geodetic problem: Survey Review, vol. 33, no. 255, pp. 50–58.

Vincenty, T., 1971, The meridional distance problem for desk computers: Survey Review, vol. XX, no. 161.

Vincenty, T., 1975, Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations: Survey Review, vol. XXIII, no. 176.

Wolf, Paul R. and Dewitt, Bon A., 2000, Elements of Photogrammetry with Applications in GIS: McGraw Hill, Boston, 3rd ed.

Department of Natural Resources Management and Engineering, University of Connecticut, Storrs, CT 06269-4087

E-mail address: `tmeyer@canr.uconn.edu`