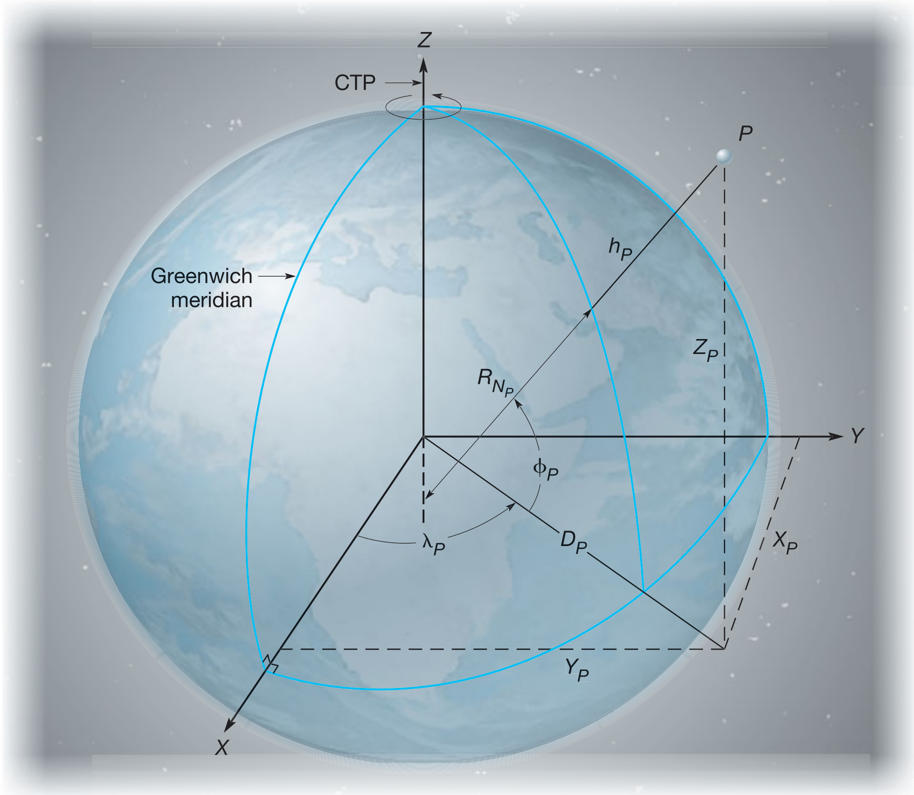


Figure 13.6 The geodetic and geocentric coordinate systems.

where

$$R_{N_P} = \frac{a}{\sqrt{1 - e^2 \sin^2 \phi_P}} \qquad \textbf{(13.2)}$$

In Equations (13.1), X_P, Y_P, and Z_P are the geocentric coordinates of any point P, and the term e, which appears in both Equations (13.1) and (13.2), is the *eccentricity* of the WGS84 reference ellipsoid. Its value is 0.08181919084. In Equation (13.2), R_{N_P} is the *radius in the prime vertical*[4] of the ellipsoid at point P, and a, as noted earlier, is the semimajor axis of the ellipsoid. In Equations (13.1) and (13.2), north latitudes are considered positive and south latitudes negative. Similarly, east longitudes are considered positive and west longitudes negative. Additionally, the programming for the conversion of geodetic coordinates to geocentric coordinates and vice versa is demonstrated in Mathcad® worksheet *C13.xcmd*, which is on the Prentice Hall companion website for this book.

[4]The eccentricity and radius in the prime vertical are both described in Chapter 20.

Example 13.1

The geodetic latitude, longitude, and height of a point A are 41°15′18.2106″N, 75°00′58.6127″ W, and 312.391 m, respectively. Using WGS84 values, what are the geocentric coordinates of the point?

Solution

Substituting the appropriate values into Equations (13.1) and (13.2) yields

$$R_{N_A} = \frac{6{,}378{,}137}{\sqrt{1 - 0.0066943799 \sin^2(41°15'18.2106'')}} = 6{,}387{,}440.3113 \text{ m}$$

$$\begin{aligned} X_A &= (6{,}387{,}440.3113 + 312.391) \cos 41°15'18.2106'' \cos(-75°00'58.6127'') \\ &= 1{,}241{,}581.343 \text{ m} \\ Y_A &= (6{,}387{,}440.3113 + 312.391) \cos 41°15'18.2106'' \sin(-75°00'58.6127'') \\ &= -4{,}638{,}917.074 \text{ m} \\ Z_A &= [6{,}387{,}440.3113(1 - 0.00669437999) + 312.391)] \sin(41°15'18.2106'') \\ &= 4{,}183{,}965.568 \text{ m} \end{aligned}$$

Conversion of geocentric coordinates of any point P to its geodetic values is accomplished using the following steps (refer again to Figure 13.6).

Step 1: Compute D_P as

$$D_P = \sqrt{X_P^2 + Y_P^2} \tag{13.3}$$

Step 2: Compute the longitude as[5]

$$\lambda_P = 2 \tan^{-1}\left(\frac{D_P - X_P}{Y_P}\right) \tag{13.4}$$

Step 3: Calculate approximate latitude, ϕ_0[6]

$$\phi_0 = \tan^{-1}\left[\frac{Z_P}{D_P(1 - e^2)}\right] \tag{13.5}$$

Step 4: Calculate the approximate radius of the prime vertical, R_N, using ϕ_0 from step 3, and Equation (13.2).

Step 5: Calculate an improved value for the latitude from

$$\phi = \tan^{-1}\left(\frac{Z_P + e^2 R_{N_P} \sin(\phi_0)}{D_P}\right) \tag{13.6}$$

[5]This formula can conveniently be implemented in software with the function atan2(X_P, Y_P).

[6]A closed-form set of formulas for computing latitude of the station is demonstrated in the Mathcad electronic book on the companion website for this book.

Step 6: Repeat the computations of steps 4 and 5 until the change in ϕ between iterations becomes negligible. This final value, ϕ_P, is the latitude of the station.

Step 7: Use the following formulas to compute the geodetic height of the station. For latitudes less than 45°, use

$$h_P = \frac{D_P}{\cos(\phi_P)} - R_{N_P} \tag{13.7a}$$

For latitudes greater than 45° use the formula

$$h_P = \left[\frac{Z_P}{\sin(\phi_P)}\right] - R_{N_P}(1 - e^2) \tag{13.7b}$$

Example 13.2

What are the geodetic coordinates of a point that has X,Y,Z geocentric coordinates of 1,241,581.343, −4,638,917.074, and 4,183,965.568, respectively? (Note: Units are meters.)

Solution

To visualize the solution, refer to Figure 13.6. Since the X coordinate value is positive, the longitude of the point is between 0° and 90°. Also, since the Y coordinate value is negative, the point is in the western hemisphere. Similarly since the Z coordinate value is positive, the point is in the northern hemisphere.

Substituting the appropriate values into Equations (13.3) through (13.7) yields

Step 1:

$$D = \sqrt{(1{,}241{,}581.343)^2 + (-4{,}638{,}917.074)^2} = 4{,}802{,}194.8993$$

Step 2:

$$\lambda = 2\tan^{-1}\left(\frac{4{,}802{,}194.8993 - 1{,}241{,}581.343}{-4{,}638{,}917.074}\right) = -75°00'58.6127'' \text{ (West)}$$

Step 3:

$$\phi_0 = \tan^{-1}\left[\frac{4{,}183{,}965.568}{4{,}802{,}194.8993(1 - 0.00669437999)}\right] = 41°15'18.2443''$$

Step 4:

$$R_N = \frac{6{,}378{,}137}{\sqrt{1 - 0.00669437999\ \sin^2(41°15'18.2443'')}} = 6{,}387{,}440.3148$$

Step 5:

$$\phi_0 = \tan^{-1}\left[\frac{4{,}183{,}965.568 + e^2 6{,}387{,}440.3148 \sin 41°15'18.2443''}{4{,}802{,}194.8993}\right]$$

$$= 41°15'18.2107''$$

Step 6: Repeat steps 4 and 5 until the latitude converges. The values for the next iteration are

$$R_N = 6{,}387{,}440.3113$$
$$\phi_0 = 41°15'18.2106''$$

Repeating with the above values results in the same value for latitude to four decimal places, so the latitude of the station is $41°15'18.2106''$ N.

Step 7: Compute the geodetic height using Equation (13.7a) as

$$h = \frac{4{,}802{,}194.8993}{\cos 41°15'18.2106''} - 6{,}387{,}440.3113 = 312.391$$

The geodetic coordinates of the station are latitude $= 41°15'18.2106''$ N, longitude $= 75°00'58.6127''$ W, and height $= 312.391$ m. Note that this example was the reverse computations of Example 13.1, and it reproduced the starting geodetic coordinate values for that example.

It is important to note that geodetic heights obtained with satellite surveys are measured with respect to the ellipsoid. That is, the geodetic height of a point is the vertical distance between the ellipsoid and the point as illustrated in Figure 13.7. As shown, these are not equivalent to *elevations* (also called *orthometric heights*) given with respect to the geoid. Recall from Chapter 4 that the geoid is an equipotential gravitational reference surface that is used as a datum for elevations. To convert geodetic heights to elevations, the *geoid height* (vertical distance between ellipsoid and geoid) must be known. Then elevations can be expressed as

$$H = h - N \tag{13.8}$$

where H is elevation above the geoid (orthometric height), h the geodetic height (determined from satellite surveys), and N the geoidal height. Figure 13.7 shows the correct relationship of the geoid and the WGS84 ellipsoid in the continental United States. Here the ellipsoid is above the geoid, and geoid height (*measured from the ellipsoid*) is negative. The geoid height at any point can be estimated

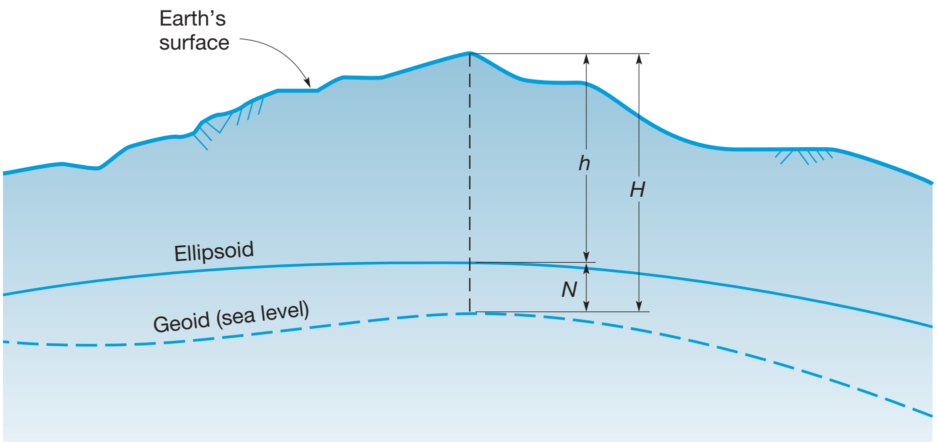


Figure 13.7 Relationships between elevation H, geodetic height h, and geoid undulation N.

with mathematical models developed by combining gravimetric data with distributed networks of points where geoidal height has been observed. One such model, *GEOID09*, is a high-resolution model for the United States available from the National Geodetic Survey.[7] It uses latitude and longitude as arguments for determining geoid heights at any location in the conterminous United States (CONUS), Hawaii, Puerto Rico, and the Virgin Islands.

Example 13.3

Compute the elevation (orthometric height) for a station whose geodetic height is 59.1 m, if the geoid undulation in the area is −21.3 m.

Solution

By Equation (13.8):

$$H = 59.1 - (-21.3) = 80.4 \text{ m}$$

Since the geoid height generally changes gradually, a value that can be applied for it over a limited area can be determined. Including NAVD 88 benchmarks in the area in a GNSS survey can do this. Then with the ellipsoid heights and elevations known for these benchmarks, the following rearranged form of Equation (13.8) is used to determine GPS observed geoidal heights:

$$N_{GPS} = h - H \tag{13.9}$$

The value for N_{GPS} obtained in this manner should be compared with that derived from the model supplied by the NGS, and the difference should be computed as $\Delta N = N_{GPS} - N_{\text{model}}$. It is best to perform this procedure on several well-dispersed benchmarks in an area whenever possible. Then using an average ΔN for the survey area, the corrected orthometric height is

$$H = h - (N_{\text{model}} + \Delta N_{\text{avg}}) \tag{13.10}$$

Example 13.4

The GNSS observed geodetic heights of benchmark stations *Red, White*, and *Blue* are 412.345, 408.617, and 386.945 m, respectively. The model geoidal heights for the stations are −29.894, − 29.902, and −29.901 m, respectively, and their published elevations are 442.214, 438.490, and 416.822 m, respectively. What is the elevation of station *Brown*, which has an observed GNSS height of 397.519 m, if the model geoid height is determined to be −29.898 m?

[7]A disk containing GEOID09 can be obtained by writing to the National Geodetic Information Center, NOAA, National Geodetic Survey, N/CG17, SSMC3 Station 09535, 1315 East West Highway, Silver Spring, MD 20910, telephone (301) 713-3242, or it can be downloaded over the Internet at http://www.ngs.noaa.gov/PC_PROD/pc_prod.shtml.

Solution

By Equation (13.9), the observed geoid heights and ΔN's are

Station	N	ΔN
Red	$412.345 - 442.214 = -29.869$	$-29.869 - (-29.894) = 0.025$
White	$408.617 - 438.490 = -29.873$	$-29.873 - (-29.902) = 0.029$
Blue	$386.945 - 416.822 = -29.877$	$-29.877 - (-29.901) = 0.024$
		$\Delta N_{avg} = 0.026$

By Equation (13.10), the elevation of *Brown* is

$$Elev_{Brown} = 397.519 - (-29.898 + 0.026) = 427.391 \text{ m}$$

A word of caution should be added. Because the exact nature of the geoid is unknown, interpolated or extrapolated values of geoidal heights from an observed network of points, or those obtained from mathematical models, are not exact. Thus orthometric heights obtained from ellipsoid heights will be close to their true values, but they may not be accurate enough to meet project requirements. Thus, for work that requires extremely accurate elevation differences, it is best to obtain them by differential leveling from nearby benchmarks.

■ 13.5 FUNDAMENTALS OF SATELLITE POSITIONING

As discussed in Section 13.3, the precise travel time of the signal is necessary to determine the distance, or so-called *range*, to the satellite. Since the satellite is in an orbit approximately 20,200 km above the Earth, the travel time of the signal will be roughly 0.07 sec after the receiver generates the same signal. If this time delay between the two signals is multiplied by the signal velocity (speed of light in a vacuum) c, the range to the satellite can be determined from

$$r = c \times t \tag{13.11}$$

where r is the range to the satellite and t the elapsed time for the wave to travel from the satellite to the receiver.

Satellite receivers in determining distances to satellites employ two fundamental methods: *code ranging* and *carrier phase-shift measurements*. Those that employ the former method are often called *mapping grade* receivers; those using the latter procedure are called *survey-grade* receivers. From distance observations made to multiple satellites, receiver positions can be calculated. Descriptions of the two methods and their mathematical models are presented in the subsections that follow. These mathematical models are presented to help students better understand the underlying principles of GPS operation. Computers that employ software provided by manufacturers of the equipment perform solutions of the equations.

13.5.1 Code Ranging

The code ranging (also called *code matching*) method of determining the time it takes the signals to travel from satellites to receivers was the procedure briefly described in Section 13.3. With the travel times known, the corresponding distances to the satellites can then be calculated by applying Equation (13.11). With one range known, the receiver would lie on a sphere. If the range were determined from two satellites, the results would be two intersecting spheres. As shown in Figure 13.8(a), the intersection of two spheres is a circle. Thus, two ranges from two satellites would place the receiver somewhere on this circle. Now if the range for a third satellite is added, this range would add an additional sphere, which when intersected with one of the other two spheres would produce another circle of intersection. As shown in Figure 13.8(b), the intersection of two circles would leave only two possible locations for the position of the receiver. A "seed position" is used to quickly eliminate one of these two intersections.

For observations taken on three satellites, the system of equations that could be used to determine the position of a receiver at station A is

$$\begin{aligned}\rho_A^1 &= \sqrt{(X^1 - X_A)^2 + (Y^1 - Y_A)^2 + (Z^1 - Z_A)^2}\\ \rho_A^2 &= \sqrt{(X^2 - X_A)^2 + (Y^2 - Y_A)^2 + (Z^2 - Z_A)^2}\\ \rho_A^3 &= \sqrt{(X^3 - X_A)^2 + (Y^3 - Y_A)^2 + (Z^3 - Z_A)^2}\end{aligned} \tag{13.12}$$

where ρ_A^n are the *geometric ranges* for the three satellites to the receiver at station A, (X^n, Y^n, Z^n) are the geocentric coordinates of the satellites at the time of the signal transmission, and (X_A, Y_A, Z_A) are the geocentric coordinates of the receiver at transmission time. Note that the variable n pertains to superscripts and takes on values of 1, 2, or 3.

However, in order to obtain a valid time observation, the systematic error (known as *bias*) in the clocks, and the refraction of the wave as it passes through the Earth's atmosphere, must also be considered. In this example, the receiver clock bias is the same for all three ranges since the same receiver is observing

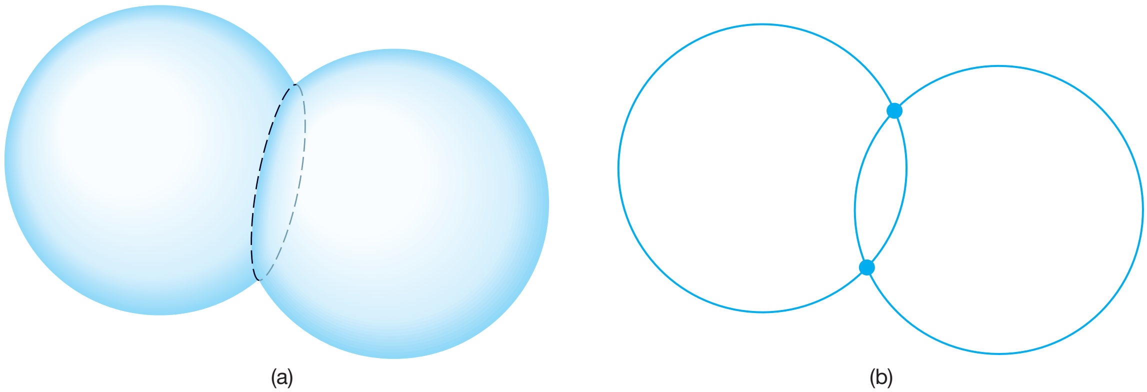


Figure 13.8 (a) The intersection of two spheres and (b) the intersection of two circles.

each range. With the introduction of a fourth satellite range, the receiver clock bias can be mathematically determined. This solution procedure allows the receiver to have a less accurate (and less expensive) clock. Algebraically, the system of equations used to solve for the position of the receiver and clock bias are:

$$\begin{aligned} R_A^1(t) &= \rho_A^1(t) + c(\delta^1(t) - \delta_A(t)) \\ R_A^2(t) &= \rho_A^2(t) + c(\delta^2(t) - \delta_A(t)) \\ R_A^3(t) &= \rho_A^3(t) + c(\delta^3(t) - \delta_A(t)) \\ R_A^4(t) &= \rho_A^4(t) + c(\delta^4(t) - \delta_A(t)) \end{aligned} \tag{13.13}$$

where $R_A^n(t)$ is the observed *range* (also called *pseudorange*) from receiver A to satellites 1 through 4 at epoch (time) t, $\rho_A^n(t)$ the geometric range as defined in Equation (13.12), c the speed of light in a vacuum, $\delta_A(t)$ the receiver clock bias, and $\delta^n(t)$ the satellite clock bias, which can be modeled using the coefficients supplied in the broadcast message. These four equations can be simultaneously solved yielding the position of the receiver (X_A, Y_A, Z_A), and the receiver clock bias $\delta_A(t)$. Equations (13.13) are known as the *point positioning equations* and as noted earlier they apply to code-based receivers.

As will be shown in Section 13.6, in addition to timing there are several additional sources of error that affect the satellite's signals. Because of the clock biases and other sources of error, the observed range from the satellite to receiver is not the true range, and thus it is called a *pseudorange*. Equations (13.13) are commonly called the *code pseudorange model*.

13.5.2 Carrier Phase-Shift Measurements

Better accuracy in measuring ranges to satellites can be obtained by observing *phase-shifts* of the satellite signals. In this approach, the phase-shift in the signal that occurs from the instant it is transmitted by the satellite until it is received at the ground station, is observed. This procedure, which is similar to that used by EDM instruments (see Section 6.19), yields the fractional cycle of the signal from satellite to receiver.[8] However, it does not account for the number of full wavelengths or *cycles* that occurred as the signal traveled between the satellite and receiver. This number is called the *integer ambiguity* or simply *ambiguity*. Unlike EDM instruments, the satellites utilize one-way communication, but because the satellites are moving and thus their ranges are constantly changing, the ambiguity cannot be determined by simply transmitting additional frequencies. There are different techniques used to determine the ambiguity. All of these techniques require that additional observations be obtained. One such technique is discussed in Section 13.6. Once the ambiguity is determined, the mathematical model for carrier phase-shift, corrected for clock biases, is

$$\Phi_i^j(t) = \frac{1}{\lambda}\rho_i^j(t) + N_i^j + f^j[\delta^j(t) - \delta_i(t)] \tag{13.14}$$

[8]The phase-shift can be measured to approximately 1/100 of a cycle.

where for any particular epoch in time, t, $\Phi_i^j(t)$ is the carrier phase-shift measurement between satellite j and receiver i, f^j the frequency of the broadcast signal generated by satellite j, $\delta^j(t)$ the clock bias for satellite j, λ the wavelength of the signal, $\rho_i^j(t)$ the range as defined in Equations (13.12) between receiver i and satellite j, N_i^j the integer ambiguity of the signal from satellite j to receiver i, and $\delta_i(t)$ the receiver clock bias.

13.6 ERRORS IN OBSERVATIONS

Electromagnetic waves can be affected by several sources of error during their transmission. Some of the larger errors include (1) satellite and receiver clock biases and (2) ionospheric and tropospheric refraction. Other errors in satellite surveying work stem from (a) satellite ephemeris errors, (b) multipathing, (c) instrument miscentering, (d) antenna height measurements, (e) satellite geometry, and (f) before May 1, 2000, selective availability. All of these errors contribute to the total error of satellite-derived coordinates in the ground stations. These errors are discussed in the subsections that follow.

13.6.1 Clock Bias

Two errors already discussed in Section 13.5 were the satellite and receiver clock biases. The satellite clock bias can be modeled by applying coefficients that are part of the broadcast message using the polynomial

$$\delta^j(t) = a_0 + a_1(t - t_0) + a_2(t - t_0)^2 \tag{13.15}$$

where $\delta^j(t)$ is the satellite clock bias for epoch t, t_0 the satellite clock reference epoch, and a_0, a_1, and a_2 the satellite clock offset, drift, and frequency drift, respectively. The parameters a_0, a_1, a_2 and t_0 are part of the broadcast message. When using *relative-positioning* techniques, and specifically *single differencing* (see Section 13.8), the satellite clock bias can be mathematically removed during post processing.

As was shown in Section 13.5, the receiver clock bias can be treated as an unknown and computed using Equations (13.13) or (13.14). When using relative positioning techniques, however, it can be eliminated through *double differencing* during post processing of the survey data. This method is discussed in Section 13.8.

13.6.2 Refraction

As discussed in Section 6.18, the velocities of electromagnetic waves change as they pass through media with different refractive indexes. The atmosphere is generally subdivided into regions. The subregions of the atmosphere that have similar composition and properties are known as *spheres*. The boundary layers between the spheres are called *pauses*. The two spheres that have the greatest effect on satellite signals are the *troposphere* and *ionosphere*. The troposphere is the lowest part of the atmosphere, and is generally considered to exist up to 10–12 km in altitude. The *tropopause* separates the troposphere from the *stratosphere*. The stratosphere goes up to about 50 km. The combined refraction in the stratosphere, tropopause, and troposphere is known as *tropospheric refraction*.

There are several other layers of atmosphere above 50 km, but the one of most interest in satellite surveying is the *ionosphere* that extends from 50 to 1500 km above the Earth. As the satellite signals pass through the ionosphere and troposphere, they are refracted. This produces range errors similar to timing errors and is one of the reasons why observed ranges are referred to as pseudoranges.

The ionosphere is primarily composed of ions—positively charged atoms and molecules, and free negatively charged electrons. The free electrons affect the propagation of electromagnetic waves. The number of ions at any given time in the ionosphere is dependent on the sun's ultraviolet radiation. Solar flare activity known as space weather can dramatically increase the number of ions in the ionosphere, and thus can be reason for concern when working with satellite surveying during periods of high sunspot activity, which follows a periodic peak variation of 11 years.[9] Since ionospheric refraction is the single largest error in satellite positioning, it is important to explore the space weather when performing surveys. This topic is further discussed in Section 15.2.

A term for both the ionospheric and tropospheric refraction can be incorporated into Equations (13.13) and (13.14) to account for those errors in the signal. Letting $\Delta\delta^j$ equal the difference between the clock bias for satellite j and the receiver at A for epoch t [i.e., $\Delta\delta^j = \delta^j(t) - \delta_A(t)$], then for any particular range listed in Equation (13.13) the incorporation of tropospheric and ionospheric refraction on the code pseudorange model yields

$$\begin{aligned} R^j_{L1}(t) &= \rho^j(t) + c\Delta\delta^j + c[\delta^{iono}_{f_{L1}} + \delta^{trop}(t)] \\ R^j_{L2}(t) &= \rho^j(t) + c\Delta\delta^j + c[\delta^{iono}_{f_{L2}} + \delta^{trop}(t)] \end{aligned} \tag{13.16}$$

where $R^j_{L1}(t)$ and $R^j_{L2}(t)$ are the observed pseudoranges as computed with frequency L1 or L2 (f_{L1} or f_{L2}) from satellite j to the receiver, $\rho^j(t)$ the geometric range as defined in Equation (13.12) from the satellite to the receiver, c the velocity of light in a vacuum, $\delta^{trop}(t)$ the delay in the signal caused by the tropospheric refraction, and δ^{iono} the ionospheric delay for the L1 and L2 frequencies, respectively.

A similar expression can be developed for the carrier phase-shift model and is

$$\begin{aligned} \Phi^j_{L1} &= \frac{1}{\lambda_{L1}}\rho^j(t) + f_{L1}\Delta\delta^j + N_{L1} - f_{L1}\delta^{iono} + f_{L1}\delta^{trop} \\ \Phi^j_{L2} &= \frac{1}{\lambda_{L2}}\rho^j(t) + f_{L2}\Delta\delta^j + N_{L2} - f_{L2}\delta^{iono} + f_{L2}\delta^{trop} \end{aligned} \tag{13.17}$$

where Φ^j_{L1} and Φ^j_{L2} are the carrier phase-shift observations from satellite j using frequencies L1 and L2, respectively, N_{L1} and N_{L2} are integer ambiguities for the two frequencies L_1 and L_2, and the other terms are as previously defined in Equations (13.14) and (13.16) for each frequency. Note that once L5 signals are

[9] 1999 had a period of high solar activity. The next peak period of high solar activity should occur around 2010.

available on a sufficient number of satellites, then an additional equation can be written for Equations (13.16) and (13.17).

By taking observations on both the L1 and L2 signals, and employing either Equations (13.16) or (13.17), the atmospheric refraction can be modeled and mathematically removed from the data. This is a major advantage of *dual-frequency receivers* (those which can observe both L_1 and L_2 signals) over their single-frequency counterparts, and allows them to accurately observe baselines up to 150 km accurately. The linear combination of the L1 and L2 signals for the code pseudorange model, which is almost free of ionospheric refraction, is

$$R_{L1,L2} = R_{L1} - \frac{(f_{L1})^2}{(f_{L2})^2} R_{L2} \tag{13.18}$$

where $R_{L1,L2}$ is the pseudorange observation for the combined L_1 and L_2 signals.

The carrier-phase model, which is also almost free of ionospheric refraction, is

$$\Phi_{L1,L2} = \Phi_{L1} - \frac{f_{L2}}{f_{L1}} \Phi_{L2} \tag{13.19}$$

where $\Phi_{L1,L2}$ is the phase observation of the linear combination of the $L1$ and $L2$ waves. By their very nature, single-frequency receivers cannot take advantage of the two separate signals, and thus they must use ionospheric modeling data that is part of the broadcast message. This limits their effective range between 10 and 20 km, although, this limit is dependent on the space weather at the time of the survey.

The advantage in having the satellites at approximately 20,200 km above the Earth is that signals from one satellite going to two different receivers pass through nearly the same atmosphere. Thus, the atmosphere has similar effects on the signals and its affects can be practically eliminated using mathematical techniques as discussed in Sections 13.7 through 13.9. For long lines Equations (13.18) and (13.19) are typically used.

As can be seen in Figure 13.9, signals from satellites that are on the horizon of the observer must pass through considerably more atmosphere than signals coming from high above the horizon. Because of the difficulty in modeling the atmosphere at low altitudes, signals from satellites below a certain threshold angle, are typically omitted from the observations. The specific value for this angle (known as the satellite *mask angle*) is somewhat arbitrary. It can vary between 10° and 20° depending on the desired accuracy of the survey. This is discussed further in Chapter 14.

13.6.3 Other Error Sources

Several other smaller error sources contribute to the positional errors of a receiver. These include (1) satellite ephemeris errors; (2) multipathing errors; (3) errors in centering the antenna over a point; (4) errors in measuring antenna height above the point; and (5) errors due to satellite geometry.

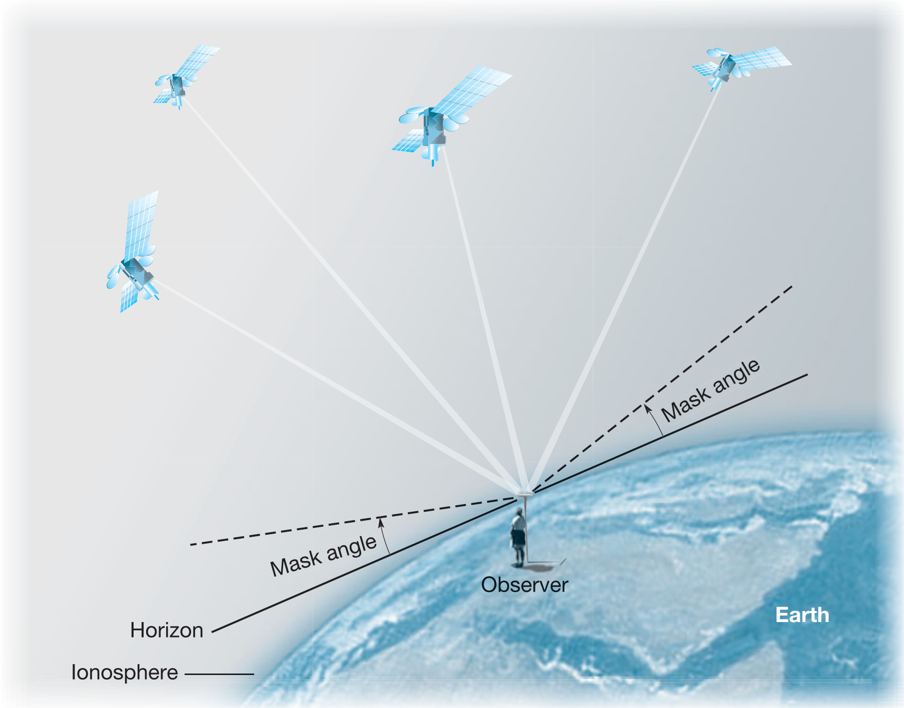


Figure 13.9
Relative positions of satellites, ionosphere, and receiver.

As noted earlier, the broadcast ephemeris predicts the positions of the satellites in the near future. However, because of fluctuations in gravity, solar radiation pressure, and other anomalies, these predicted orbital positions are always somewhat in error. In the code-matching method, these satellite position errors are translated directly into the computed positions of ground stations. This problem can be reduced by updating the orbital data using information obtained later, which is based on the actual positions of the satellites determined by the tracking stations. One disadvantage of this is the delay that occurs in obtaining the updated data. One of three updated postsurvey ephemerides are available: (1) ultra-rapid ephemeris, (2) the *rapid ephemeris*, and (3) the *precise ephemeris*. The ultra-rapid ephemeris is available twice a day; the rapid ephemeris is available within two days after the survey; the precise ephemeris (the most accurate of the three) is available two weeks after the survey. The ultra-rapid and rapid ephemerides are sufficient for most surveying applications.

As shown in Figure 13.10(a), *multipathing* occurs when a satellite signal reflects from a surface and is directed toward the receiver. This causes multiple signals from a satellite to arrive at the receiver at slightly different times. Vertical structures such as buildings and chain link fences are examples of reflecting surfaces that can cause multipathing errors. Mathematical techniques have been developed to eliminate these undesirable reflections, but, in extreme cases, they can cause a receiver to *lose lock* on the satellite—loss of lock is essentially a situation where the receiver cannot use the signals from the satellite. This can be caused not only by multipathing, but also by obstructions, or high ionospheric activity.

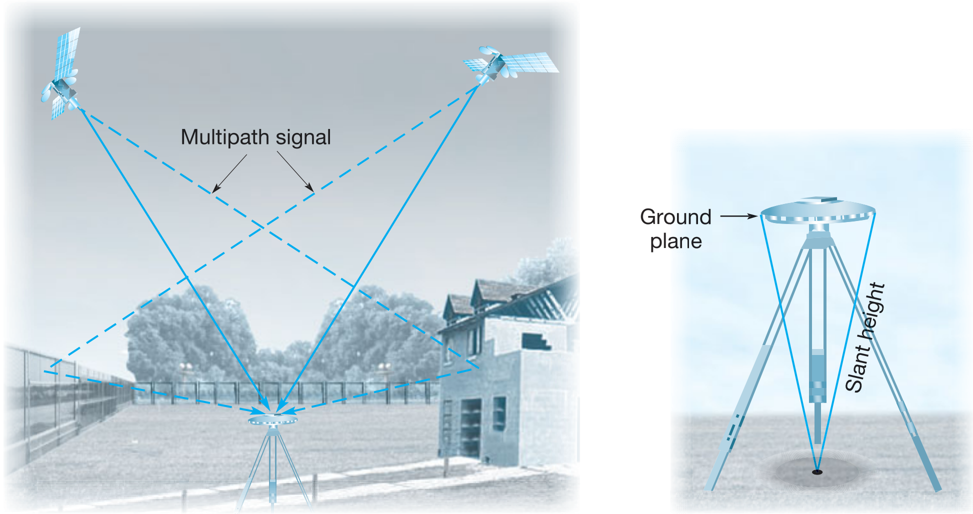


Figure 13.10 (a) Multipathing and (b) slant height measurements.

In satellite surveying, pseudoranges are observed to the receiver antennas. For precise work, the antennas are generally mounted on tripods, set up and carefully centered over a survey station, and leveled. Miscentering of the antenna over the point is another potential source of error. Set up and centering over a station should be carefully done following procedures like those described in Section 8.5. For any precise surveying work, including satellite surveys, it is essential to have a well-adjusted tripod, tribrach, and optical plummet. Any error in miscentering of the antenna over a point will translate directly into an equal-sized error in the computed position of that point.

Observing the height of the antenna above the occupied point is another source of error in satellite surveys. The ellipsoid height determined from satellite observations is determined at the phase center of the antenna. Therefore, to get the ellipsoid height of the survey station, it is necessary to measure carefully, and record the height of the antenna's phase center above the occupied point, and account for it in the data reduction. The distance shown in Figure 13.10(b) is known as the *slant height* and can be observed. The observations are made to the *ground plane* (a plane at the base of the antenna, which protects it from multipath signals reflecting from the ground). The slant height should be observed at several locations around the ground plane, and if the observations do not agree, the instrument should be checked for level. Software within the system converts the slant height to the antenna's vertical distance above the station. Mistakes in identifying and observing heights of phase centers have caused errors as great as 10 cm in elevation. In precise satellite surveys, many surveyors use fixed-height tripods and rods that provide a constant offset from the point to the *antenna reference point* (ARP)—typically set at 2 m.

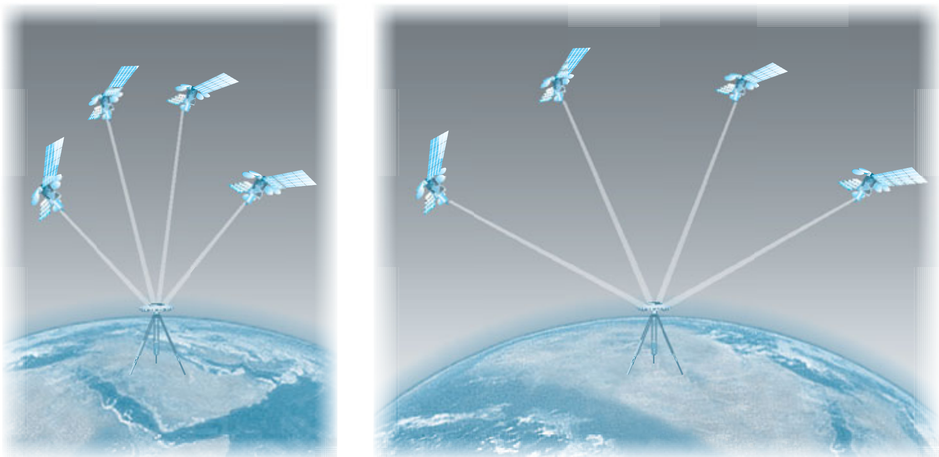
Additionally, the *phase center*, which is the electronic center of the antenna, varies with the orientation of the antenna, elevation of the satellites, and frequency of the signals. In fact, the physical center of the antenna seldom matches the phase center of the antenna. This fact is accounted for by *phase center offsets*, which are translations necessary to make the phase center and physical center of the antenna match.

For older antennas it is important to orient the antennas of multiple receivers in the same azimuth. This ensures the same orientation of the phase centers at all stations and eliminates a potential systematic error if the phase center is not precisely at the geometric center of the antenna. The same antenna should always be used with a given receiver in a precise survey, but if other antennas are used, their phase center offsets must be accounted for during post processing. Newer antennas are directionally independent. They no longer require azimuthal alignment.

The National Geodetic Survey calibrates GPS antennas with respect to satellite elevations. When processing GPS data (see Section 14.5), users should always include the NGS calibration data to account for varying offsets due to satellite elevations when processing baselines.

13.6.4 Geometry of Observed Satellites

An important additional error source in satellite surveying deals with the geometry of the visible satellite constellation at the time of observation. This is similar to the situation in traditional surveys, where the geometry of the network of observed ground stations affects the accuracies of computed positions. Figure 13.11 illustrates both weak and strong satellite geometry. As shown in Figure 13.11(a), small angles between incoming satellite signals at the receiver station produce weak geometry and generally result in larger errors in computed positions. Conversely, strong geometry, as shown in Figure 13.11(b), occurs when the angles between incoming satellite signals are large, and this usually provides an improved solution. Whether conducting a satellite survey or a traditional one, by employing least-squares adjustment in the solution, the effect of the geometry upon the expected accuracy of the results is determined.



Weak geometry

(a)

Strong geometry

(b)

Figure 13.11 Weak and strong satellite geometry.

TABLE 13.2 ERROR SOURCES AND SIZES THAT CAN BE EXPECTED IN OBSERVED GPS RANGES

Error Source	Current Sizes of Errors (m)	Anticipated Sizes of Errors with Two or More Coded Signals (m)
Clock and ephemeris errors	±2.3	±2.3
Ionospheric refraction	±7	±0.1
Tropospheric refraction	±0.2	±0.2
Receiver noise	±0.6	±0.6
Other (multipath, etc.)	±1.5	±1.5

Table 13.2 lists the various categories of errors that can occur in satellite positioning. For each category, the sizes of errors that could occur in observed satellite ranges if no corrections or compensations were made are given, for example, ±7.5 m could be expected as a result of ionospheric refraction, etc. But these error sizes assume ideal satellite geometry, that is, no further degradation of accuracy is included for weak satellite geometry. The anticipated size of these errors with the addition of the L2C and L5 signals is shown in the third column of Table 13.2. The L2C will be available to receivers as the satellites become available. The advantages of the L5 signal will not be apparent to users until a majority of the satellite constellation has been upgraded. It is anticipated that the entire satellite constellation will be upgraded with these new signals by 2030. By comparing the current errors with those anticipated with the inclusion of newer coded signals, it is obvious why the decision was made to fund the newer satellites. Using Equation (3.11), the total *User Equivalent Range Error* (UERE) is currently approximately ±7.5 m. It is anticipated that this error will drop to approximately ±2.8 m with the L2C and L5 signals.

As noted above, by employing least squares in the solution, the effect of satellite geometry can be determined. In fact, before conducting a satellite survey, the number and positions of visible satellites at a particular time and location can be evaluated in a preliminary least-squares solution to determine their estimated effect upon the resulting accuracy of the solution. This analysis produces so-called *Dilution Of Precision* (DOP) factors. The DOP factors are computed through error propagation (see Section 3.17). They are simply numbers, which when multiplied by the errors of Table 13.2, give the sizes of errors that could be expected based upon the geometry of the observed constellation of satellites. For example, if the DOP factor is 2, then multiplying the sizes of errors listed in Table 13.2 by 2 would yield the estimated errors in the ranges for that time and location. Obviously, the lower the value for a DOP factor, the better the expected precision in computed positions of ground stations. If the preliminary least-squares analysis gives a higher DOP number than can be tolerated, the observations should be delayed until a more favorable satellite constellation is available.

The DOP factors that are of most concern to surveyors are PDOP (dilution of precision in position), HDOP (dilution of precision in horizontal position), and VDOP (dilution of precision in height). For the best possible constellation of

TABLE 13.3 IMPORTANT CATEGORIES OF DILUTION OF PRECISION

Category of DOP	Stand. Dev. Terms	Equation	Acceptable Value (less than)*
PDOP, Positional DOP	σ in geocentric coordinates X, Y, Z	$\sqrt{\sigma_X^2 + \sigma_Y^2 + \sigma_Z^2}$	6
HDOP, Horizontal DOP	σ in local x, y coordinates	$\sqrt{\sigma_X^2 + \sigma_Y^2}$	3
VDOP, Vertical DOP	σ in height, h	σ_h	5

*These recommended values are general guides for average types of GPS surveys, but individual project requirements may require other specific values.

satellites, the average value for HDOP is under 2 and under 5 for PDOP. Other DOP factors such as GDOP (dilution of precision in geometry) and TDOP (dilution of precision in time) can also be evaluated, but are generally of less significance in surveying. Table 13.3 lists some important categories of DOP, explains their meanings in terms of standard deviations and equations, and gives maximum values that are generally considered acceptable for most surveys.

Multiplying the DOP factor by the UERE yields the positional error in code ranging using Equations (13.13). For example, the HDOP is typically about 1.5. Recall from Equation (3.8) that the 95% probable error is obtained using a multiplier of about 1.96. Using the error values from Table 13.2 and a HDOP of 1.5 the current 95% probable error in horizontal positioning is ± 22.5 m $(1.96 \times 1.5 \times 7.5)$. When the newer coded signals are available and used by receivers, the 95% horizontal positioning error will be approximately ± 8.5 m.

13.6.5 Selective Availability

Until May of 2000, GPS signals were degraded to intentionally reduce accuracies achievable using the code-matching method. The intent was to exclude the highest accuracy attainable with GPS from nonmilitary users, especially adversaries. Two different methods were used to degrade accuracy; the *delta* process, which dithered the fundamental frequency of the satellite clock, and the *epsilon* process, which truncated the orbital parameters in the broadcast message so that the coordinates of the satellites could not be computed accurately. The errors in the coordinates of the satellites roughly translated to similar ground positional errors. The combined effect of these errors was known as *selective availability* (SA), and resulted in a positional error of approximately 100 m in the horizontal, and 156 m in the vertical, at the 95% error level. However, this error could be removed by either differential or relative positioning techniques (see Sections 13.7 and 13.8).

When SA was initiated, the military and civilian communities were at odds about the need for it during times of peace. Initially, a plan was developed to turn SA off by 2006. But after agreement by the Departments of Transportation, Commerce and Defense, it was turned off at midnight on May 1, 2000 as the result of a Presidential Decision Directive (PDD). As was previously shown, with the removal of selective availability, code-based, real-time point positioning is about 20 m. Future satellites will not have the capability to implement SA.

13.7 DIFFERENTIAL POSITIONING

As discussed in the two preceding sections, accuracies of observed pseudoranges are degraded by errors that stem from clock biases, atmospheric refraction, and other sources. Because of these errors, positions of points determined by point positioning techniques using a single code-based receiver can be in error by 20 m or more. While this order of accuracy is acceptable for certain uses, it is insufficient for most surveying applications. *Differential GPS* (DGPS) on the other hand, is a procedure that involves the simultaneous use of two or more code-based receivers. It can provide positional accuracies to within a few meters, and thus the method is suitable for certain types of lower-order surveying work.

In DGPS, one receiver occupies a so-called *base station* (point whose coordinates are precisely known from previous surveying), and the other receiver or receivers (known as the *rovers*) are set up at stations whose positions are unknown. By placing a receiver on a station of known position, the pseudorange errors in the signal can be determined using Equation (13.16). Since this base station receiver and the rover are relatively close to each other (often less than a kilometer but seldom farther than a few hundred kilometers), the pseudorange errors at both the base station and at the rovers will have approximately the same magnitudes. Thus, after computing the corrections for each visible satellite at the base station, they can be applied to the roving receivers, thus substantially reducing or eliminating many errors listed in Table 13.2.

DGPS can be done in almost *real time* with a radio transmitter at the base station and compatible radio receivers at the rovers. This process is known as *real-time differential GPS* (RTDGPS). The radio transmissions to the rovers contain both *pseudorange corrections* (PRCs) for particular *epochs of time* (moments in time) and *range rate corrections* (RRCs)[10] so that they can interpolate corrections to signals between each epoch. Alternatively the errors can be eliminated from coordinates determined for rover stations during post processing of the data.

To understand the mathematics in the procedure, a review of Equation (13.13) is necessary. The various error sources presented in Section 13.6 cause the observed pseudorange $R_A^j(t_0)$ to be in error by a specific amount for any epoch, t_0. Letting this error at epoch t_0 be represented by $\Delta\rho_A^j(t_0)$, the *radial orbital error*, Equation (13.13) can be rewritten as

$$R_A^j(t_0) = \rho_A^j(t_0) + \Delta\rho_A^j(t_0) + c\delta^j(t_0) - c\delta_A(t_0) \tag{13.20}$$

where the other terms are as previously defined.

Because the coordinates of the base station are known, the geometric range $\rho_A^j(t_0)$ in Equation (13.20) can be computed using Equation (13.12). Also since the pseudorange $R_A^j(t_0)$ is observed, the difference in these two values will yield the necessary correction for this particular pseudorange. Since the error

[10]Pseudorange corrections (PRCs) are differences between measured ranges and ranges that are computed based upon the known coordinates of both the occupied reference station and those of the satellite. Because the satellites are moving, measured ranges to them are constantly changing. The rates of these changes per unit of time are the range rate corrections (RRCs).

conditions at each receiver are very similar, it can be assumed that the error in the pseudorange observed at the base station is the same as the error at the rovers. This error at the base station is known as the code *pseudorange correction* (PRC) for satellite j at reference epoch t_0, and is represented as

$$\begin{aligned} PRC^j(t_0) &= -R^j_A(t_0) + \rho^j_A(t_0) \\ &= -\Delta\rho^j_A(t_0) - c[\delta^j(t_0) - \delta_A(t_0)] \end{aligned} \tag{13.21}$$

Because computation of the correction and transmission of the signal make it impossible to assign the PRC to the same epoch at the rovers, a range rate correction (RRC) is approximated by numerical differentiation. This correction is used to extrapolate corrections for later epochs t. Thus, the pseudorange correction at any epoch t is given as

$$PRC^j(t) = PRC^j(t_0) + RRC^j(t_0)(t - t_0) \tag{13.22}$$

where $RRC^j(t_0)$ is the range rate correction for satellite j determined at epoch t_0.

Now this information can be used to correct the computed ranges at the roving receiver locations. For example, at a roving station B, the corrected pseudorange, $R^j_B(t)_{\text{corrected}}$, can be computed as

$$\begin{aligned} R^j_B(t)_{\text{corrected}} &= R^j_B(t) + PRC^j(t) \\ &= \rho^j_B(t) + [\Delta\rho^j_B(t) - \Delta\rho^j_A(t)] - c[\delta_B(t) - \delta_A(t)] \\ &= \rho^j_B(t) - c\Delta\delta_{AB}(t) \end{aligned} \tag{13.23}$$

where $\Delta\delta_{AB} = \delta_B(t) - \delta_A(t)$.

Notice that in the final form of Equation (13.23), it is assumed that the radial orbital errors at stations A and B, $\Delta\rho^j_A(t)$ and $\Delta\rho^j_B(t)$, respectively, are nearly the same, and thus are mathematically eliminated. Furthermore, the satellite clock bias terms will be eliminated. Finally, assuming the signals to the base and roving receivers pass through nearly the same atmosphere (which means they should be within a few hundred kilometers of each other), the ionospheric and tropospheric refraction terms are practically eliminated.

The U.S. Coast Guard maintains a system of beacon stations along the U.S. coast and waterways. Private agencies have developed additional stations. The correction signals described above are broadcast by modulation on a frequency between 285 and 325 kHz using the ***R****adio* ***T****echnical* ***C****ommission for* ***M****aritime Services* ***S****pecial* ***C****ommittee* 104 (RTCM SC-104) format. Among the data contained in this broadcast are C/A code differential corrections, delta differential corrections, reference station parameters, raw carrier phase measurements, raw code range measurements, carrier phase corrections, and code range corrections.

The *Wide Area Augmentation System* (WAAS) developed by the Federal Aviation Administration has a network of ground tracking base stations that collect GPS signals and determine range errors. These errors are transmitted to geosynchronous satellites that relay the corrections to rovers. GPS software typically allows users to access the WAAS system when performing RTK-GPS surveys (see Chapter 15). This option, called *RTK with infill*, accesses the WAAS corrections when base-station radio transmissions are lost. However, these corrections will provide significantly less accuracy than relative positioning techniques typically utilized by GPS receivers using carrier phase-shift measurements. In Europe, the *European Geostationary Navigation Overlay Service* (EGNOS) serves a similar role to WAAS. In Japan, the *Multifunctional Satellite Augmentation System* (MSAS) serves this purpose.

When the WAAS is combined with a *Local Area Augmentation System* (LAAS), it is anticipated that the system will enable aircraft to key in on their destinations, after which the navigation system would develop the necessary flight paths for making landings in zero visibility. This system is expected to provide centimeter-level, real-time accuracy when implemented. Corrections will be broadcast to aircraft using a very high frequency (VHF) radio wave. Private firms have created similar systems. These systems are available as a subscription service. The system is currently in its research and development stage.

13.8 KINEMATIC METHODS

Methods similar to DGPS can also be employed with carrier phase-shift measurements to eliminate errors. The procedure, called *Kinematic* surveying (see Chapter 15), again requires the simultaneous use of two or more receivers. All receivers must simultaneously collect signals from at least four of the same satellites through the entire observation process. Although single-frequency receivers can be used, kinematic surveying works best with dual-frequency receivers. The method yields positional accuracies to within a few centimeters, which makes it suitable for most surveying, mapping, and stakeout purposes.

As with DGPS, the fact that the base station's coordinates are known is exploited in *real-time kinematic* (RTK) surveys. Most manufacturers broadcast the observations at the base station to the rover. The roving receiver uses the relative positioning techniques discussed in Section 13.9 to determine the position of the roving receiver. However, it is possible to compute and broadcast pseudorange corrections (PRC). Once the pseudorange corrections are determined, they are used at the roving receivers to correct their pseudoranges. Multiplying Equation (13.14) by λ, and including the radial orbital error term, the carrier phase pseudorange at base station A for satellites j at epoch t_0 is

$$\lambda\Phi^j_A(t_0) = \rho^j_A(t_0) + \Delta\rho^j_A(t_0) + \lambda N^j_A + c[\delta^j(t_0) - \delta_A(t_0)] \tag{13.24}$$

where N^j_A is the initially unknown ambiguity, and all other terms were previously defined in Equation (13.20). Recalling that the base station is a point with known coordinates, the pseudorange correction at epoch t_0 is given by

$$\begin{aligned} PRC^j(t_0) &= -\lambda\Phi^j_A(t_0) + \rho^j_A(t_0) \\ &= -\Delta\rho^j_A(t_0) - \lambda N^j_A - c[\delta^j(t_0) - \delta_A(t_0)] \end{aligned} \tag{13.25}$$

and the pseudorange correction at any epoch t is

$$PRC^j(t) = PRC^j(t_0) + RRC^j(t_0)(t - t_0) \tag{13.26}$$

Using the same procedure as was used with code pseudoranges, the corrected phase range at the roving receiver for epoch t is

$$\lambda\,\Phi_B^j(t)_{\text{corrected}} = \rho_B^j(t) + \lambda\Delta N_{AB}^j - c\Delta\delta_{AB}(t) \qquad \textbf{(13.27)}$$

where $\Delta N_{AB}^j = N_B^j - N_A^j$ and $\Delta\delta_{AB}(t) = \delta_B(t) - \delta_A(t)$.

These equations can be solved as long as at least four satellites are continuously observed during the survey. The pseudorange corrections and the range rate corrections are transmitted to the receivers.

■ 13.9 RELATIVE POSITIONING

The most precise positions are currently obtained using relative positioning techniques. Similar to both DGPS and kinematic surveying, this method removes most errors noted in Table 13.2 by utilizing the differences in either the code or carrier phase ranges. The objective of relative positioning is to obtain the coordinates of a point relative to another point. This can be mathematically expressed as

$$\begin{aligned} X_B &= X_A + \Delta X \\ Y_B &= Y_A + \Delta Y \\ Z_B &= Z_A + \Delta Z \end{aligned} \qquad \textbf{(13.28)}$$

where (X_A, Y_A, Z_A) are the geocentric coordinates at the base station A, (X_B, Y_B, Z_B) are the geocentric coordinates at the unknown station B, and $(\Delta X, \Delta Y, \Delta Z)$ are the computed *baseline vector components* (see Figure 13.12).

Relative positioning involves the use of two or more receivers simultaneously observing pseudoranges at the endpoints of lines. Simultaneity implies that the receivers are collecting observations at the same time. It is also important that the receivers collect data at the same epoch rate. This rate depends on the purpose of the survey and its final desired accuracy, but common intervals are 1, 2, 5, 10, or 15 sec. Assuming that simultaneous observations have been collected,

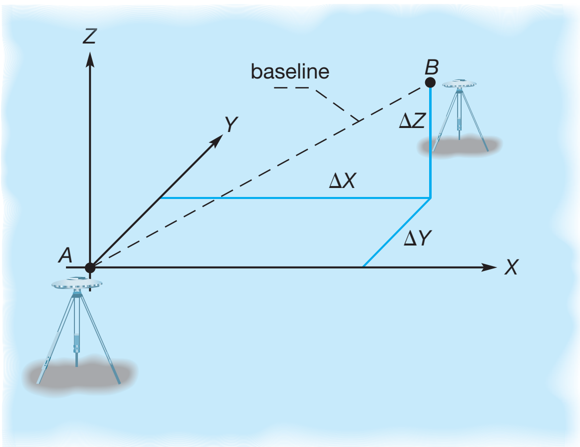


Figure 13.12
Computed baseline vector components.

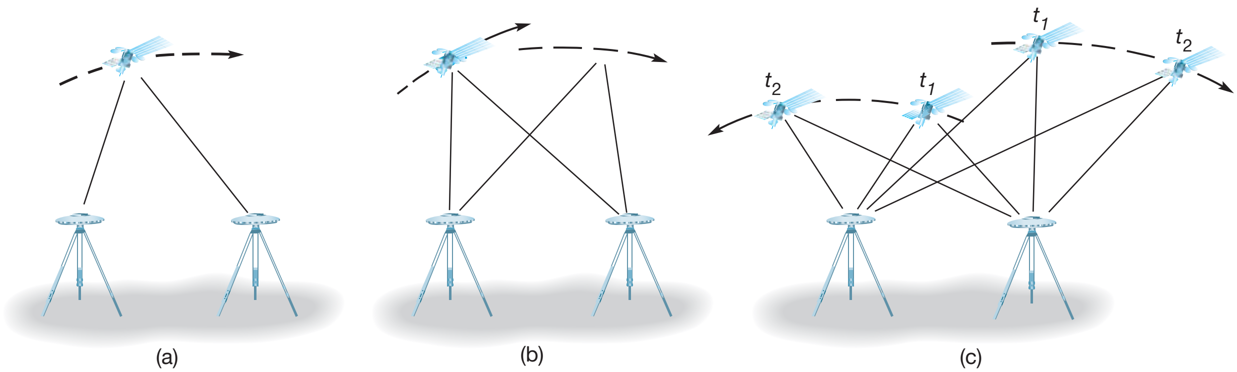


Figure 13.13 GPS differencing techniques: (a) single differencing, (b) double differencing, and (c) triple differencing.

different linear combinations of the equations can be produced, and in the process certain errors can be eliminated. Figure 13.13 shows three linear combinations and the required receiver-satellite combinations for each. These are described in the subsections that follow, and only carrier-phase measurements are considered.

13.9.1 Single Differencing

As illustrated in Figure 13.13(a), single differencing involves subtracting two simultaneous observations made to one satellite from two points. This difference eliminates the satellite clock bias and much of the ionospheric and tropospheric refraction from the solution. It would also eliminate the effects of *SA* if it were turned on. Following Equation (13.14), the phase equations for the two points are

$$\begin{aligned} \Phi_A^j(t) - f^j\delta^j(t) &= \frac{1}{\lambda}\rho_A^j(t) + N_A^j - f^j\,\delta_A(t) \\ \Phi_B^j(t) - f^j\delta^j(t) &= \frac{1}{\lambda}\rho_B^j(t) + N_B^j - f^j\,\delta_B(t) \end{aligned} \tag{13.29}$$

where the terms are as noted in Equation (13.14) for stations A and B. The difference in these two equations yields

$$\Phi_{AB}^j(t) = \frac{1}{\lambda}\rho_{AB}^j(t) + N_{AB}^j - f^j\,\delta_{AB}(t) \tag{13.30}$$

where the individual difference terms are

$$\begin{aligned} \Phi_{AB}^j(t) &= \Phi_B^j(t) - \Phi_A^j(t), \\ \rho_{AB}^j(t) &= \rho_B^j(t) - \rho_A^j(t), \\ N_{AB}^j &= N_B^j - N_A^j, \text{ and} \\ \delta_{AB}^j(t) &= \delta_B^j(t) - \delta_A^j(t). \end{aligned}$$

Note that in Equation (13.30), the satellite clock bias error, $f^j\delta^j(t)$ has been eliminated by this single differencing procedure.

13.9.2 Double Differencing

As illustrated in Figure 13.13(b), double differencing involves taking the difference of two single differences obtained from two satellites j and k. The procedure eliminates the receiver clock bias. Assume the following two single differences:

$$\begin{aligned} \Phi^j_{AB}(t) &= \frac{1}{\lambda}\rho^j_{AB}(t) + N^j_{AB} - f^j\,\delta^j_{AB}(t) \\ \Phi^k_{AB}(t) &= \frac{1}{\lambda}\rho^k_{AB}(t) + N^k_{AB} - f^k\,\delta^k_{AB}(t) \end{aligned} \tag{13.31}$$

Note that the receiver clock bias will be the same for observations on satellite j as it is for satellite k. Thus, by taking the difference between these two single differences, the following double difference equation is obtained, in which the receiver clock bias errors, $f^j\,\delta^j_{AB}(t)$ and $f^k\,\delta^k_{AB}(t)$ are eliminated.

$$\Phi^{jk}_{AB}(t) = \frac{1}{\lambda}\rho^{jk}_{AB}(t) + N^{jk}_{AB} \tag{13.32}$$

where the difference terms are

$$\begin{aligned} \Phi^{jk}_{AB}(t) &= \Phi^k_{AB}(t) - \Phi^j_{AB}(t) \\ \rho^{jk}_{AB}(t) &= \rho^k_{AB}(t) - \rho^j_{AB}(t) \\ N^{jk}_{AB} &= N^k_{AB} - N^j_{AB} \end{aligned}$$

13.9.3 Triple Differencing

The triple difference illustrated in Figure 13.13(c) involves taking the difference between two double differences obtained for two different epochs of time. This difference removes the integer ambiguity from Equation (13.32), leaving only the differences in the phase-shift observations and the geometric ranges. The two double-difference equations can be expressed as

$$\begin{aligned} \Phi^{jk}_{AB}(t_1) &= \frac{1}{\lambda}\rho^{jk}_{AB}(t_1) + N^{jk}_{AB} \\ \Phi^{jk}_{AB}(t_2) &= \frac{1}{\lambda}\rho^{jk}_{AB}(t_2) + N^{jk}_{AB} \end{aligned} \tag{13.33}$$

The difference in these two double differences yields the following triple difference equation, in which the integer ambiguities have been removed. The triple difference equation is

$$\Phi^{jk}_{AB}(t_{12}) = \frac{1}{\lambda}\rho^{jk}_{AB}(t_{12}) \tag{13.34}$$

In Equation (13.34) the two difference terms are

$$\begin{aligned} \Phi^{jk}_{AB}(t_{12}) &= \Phi^{jk}_{AB}(t_2) - \Phi^{jk}_{AB}(t_1) \\ \rho^{jk}_{AB}(t_{12}) &= \rho^{jk}_{AB}(t_2) - \rho^{jk}_{AB}(t_1) \end{aligned}$$

The importance of employing the triple difference equation in the solution is that by removing the integer ambiguities, the solution becomes immune to *cycle slips*. Cycle slips are created when the receiver *loses lock* during an observation session. The three main sources of cycle slips are (1) obstructions, (2) low *signal to noise ratio* (SNR), and (3) incorrect signal processing. Signal obstructions can be minimized by careful selection of receiver stations. Low SNR can be caused by undesirable ionospheric conditions, multipathing, high receiver dynamics, or low satellite elevations. Malfunctioning satellite oscillators can also cause cycle slips, but this rarely occurs. It should be noted that today's processing software rarely, if ever, uses triple differencing since the integer ambiguities are resolved using more advanced on-the-fly techniques, which are discussed in Section 15.2.

13.10 OTHER SATELLITE NAVIGATION SYSTEMS

Satellite positioning affects all walks of life including transportation, agriculture, data networks, cell phone technology, sporting events, and so on. In fact, the military and economic benefits of satellite positioning have been so great that other nations have or will be developing their own networks. This plethora of positioning satellites will greatly increase the utility and accuracy available from satellite positioning system. Other implemented or planned satellite positioning systems are discussed in the following subsections.

13.10.1 The GLONASS Constellation

The ***Glo****bal* ***Na****vigation* ***S****atellite* ***S****ystem* (GLONASS) is the Russian equivalent of GPS. The GLONASS constellation has 24 satellites equally spaced in three orbital planes making a 64.8° nominal inclination angle with the equatorial plane of the Earth. The satellites orbit at a nominal altitude of 19,100 km and have a period of 11.25 h. At least five are always visible to users. The system is free from selective availability, but does not permit public access to the P code. Each satellite broadcasts two signals with frequencies that are unique. The frequencies of the satellites are determined as

$$\begin{aligned} f^j_{L1} &= 1602.0000 \text{ MHz} + j \times 0.5625 \text{ MHz} \\ f^j_{L2} &= 1246.0000 \text{ MHz} + j \times 0.4375 \text{ MHz} \end{aligned} \tag{13.35}$$

where j represents the channel number assigned to the specific satellite,[11] and varies from 1 to 24, and $L1$ and $L2$ represent the broadcast bands.

As discussed in Section 13.3, GPS satellites broadcast their positions in every repetition of the broadcast message using the WGS84 reference system as the basis for coordinates. The GLONASS satellites only broadcast their positions every 30 min and use the PZ-90 reference ellipsoid as the basis for coordinates. Thus, GNSS receivers must extrapolate positions of the satellites for real-time reductions.

[11]Some antipodal satellites use the same frequencies.

The time reference systems used in GPS and GLONASS are also different. At the request of the international community, the timing of GLONASS satellites has moved toward the international standard as set by the *Bureau Internationale de l'Heure* (International Bureau of Time). This standard is based on the frequency of the atom Cesium 133 in its ground state.[12] This standard differs from the orbital period of the earth by approximately 1 sec every 6 months. To compensate, one leap second is periodically added to the atomic time (IAT) to create Universal Coordinated Time (UTC), which agrees with the solar day (see Section C.5). Currently, the GLONASS system clocks differ from Universal Coordinated Time by 3 h. In contrast, the GPS system clocks never account for the leap second, and differ from IAT by a constant of 19 sec. Thus, two GLONASS satellites must be visible to combine GLONASS and GPS satellites in a GNSS receiver.

13.10.2 The Galileo System

In 1998, the European Union decided to implement another satellite positioning system called Galileo. The Galileo system will offer five levels of service with subscriptions required for some of the services. The five levels of service are (1) *open* service (OS), (2) *commercial* service (CS), (3) *safety-of-life* service (SOL), (4) *public-regulated* service (PR), and (5) *search and rescue* (SAR) service. Open service will be a free offering positioning down to 1 m. The commercial service is an encrypted, subscription service, which will provide positioning at the cm level. The safety-of-life service will be a free providing both guaranteed accuracy and integrity messages to warn of errors. The public-regulated service will be available only to government agencies; which is similar to the current P-code. The search-and-rescue service will pickup distress beacon locations and be able to send feedback indicating that help is on the way.

The Galileo space segment will consist of 27 satellites plus 3 spares orbiting in three planes that are inclined to the equator at 56°. The satellites will have a nominal orbital altitude of 23,222 km above the Earth. The satellites will broadcast six navigation signals denoted as L1F, L1P, E6C, E6P, E5a, and E5b. The first Galileo experimental satellite was launched in December of 2005. After a failure in the second satellite, the second launch was delayed to late 2007. The *European Space Agency* (ESA) recently signed a contract to launch the first four operational satellites. These satellites will be used to validate the system. After validation, the remainder of the system will be launched over time. Galileo may offer greater accuracy than GPS with its commercial service providing meter-level point positioning. Like the modernized GPS satellites, the strength of its signals should allow work in canopy situations. The United States and European Union have agreed to make their systems interoperable. Thus future receivers will be able to use satellites from either system.

[12]One second is defined as 9,192,631,770 periods of the radiation of the ground state of the cesium 133 atom.

13.10.3 The Compass System

In 2006, China confirmed that it will create a fourth GNSS. Compass[13] will contain 35 satellites. Five of these satellites will be geostationary Earth orbit (GEO) satellites with the remaining 30 satellites at about 20,000 km. Compass will offer two levels of service—an open and commercial service with real-time positioning accuracy of 10 m. The announced completion date for the system is around 2020.

Even though the satellite constellations of the systems are not complete at the time of this writing, manufacturers of satellite receiver technology are building receivers that will utilize all GPS, GLONASS, and Galileo systems. The obvious advantage of using multiple systems is that many more satellites are available for observation by receivers. By combining these systems, the surveyor can expect improvements in increased speed and accuracy. Furthermore, the combination of systems will provide a viable method of bringing satellite positioning to difficult areas such as canyons, deep surface mines, and urban areas surrounded by tall buildings (urban canyons).

13.11 THE FUTURE

The overall success of satellite positioning in the civilian sector is well documented by the number and variety of enterprises that are using the technology. This has lead to increasing and improving GNSS constellations. In the near future, improvements will occur in signal acquisition and positioning. For example, signals from all of the satellite-positioning systems will be able to penetrate canopy situations and may provide satellite positioning capabilities from within buildings. The additional signals from within each system will improve both ambiguity resolution and atmospheric corrections. For example, in GPS with the addition of the L2C and L5 signals, real-time ionospheric corrections to the code pseudoranges will become possible by implementing Equations (13.18). Additionally, the addition of the L2C and L5 signal will enhance our ability to correctly and quickly determine the integer ambiguities for phase-shift observations. In fact, in theory ambiguities can be determined with a single epoch of data. It is anticipated that accuracies in the modernized system will be reduced to the millimeter level. In fact, it is anticipated that code-based solutions will be available at the centimeter level. The full implementation of the GLONASS system and Galileo system is only expected to enhance these capabilities. This will provide civil satellite positioning users with unprecedented real-time determination of highly accurate location anywhere on the Earth.

The use of satellites in the surveying (geomatics) community has continued to increase as the costs of the systems have decreased. This technology has and will undoubtedly continue to have considerable impact on the way data is collected and processed. In fact, as the new satellite technologies are developed, the use of conventional surveying equipment will decrease. This is due to the ease, speed, and achievable accuracies that satellite positioning technologies provide.

---

[13]The Chinese name for their system is Beidou, which stands for North Dipper. Compass is being used in English writings on this system.

As is the current trend, less field time will be required for the surveyor (geomatics engineer), and more time will be used to analyze, manage, and manipulate the large volumes of data that this technology and others provide. Those engaged in surveying (geomatics) in the future will need to be knowledgeable in the areas of information management and computer science and will provide products to clients that currently do not exist.

PROBLEMS

Asterisks (*) indicate problems that have partial answers given in Appendix G.

13.1 Define the line of apsides.

13.2 Briefly describe the orbits of the GLONASS satellites.

13.3 Why is a fully operational satellite positioning system designed to have at least four satellites visible at all time?

13.4* Discuss the purpose of the pseudorandom noise codes.

13.5 What is the purpose of the Consolidated Space Operation Center in GPS?

13.6 Describe the three segments of GPS.

13.7 Describe the content of the GPS broadcast message.

13.8 What is anti-spoofing?

13.9 What errors affect the accuracy of satellite positioning?

13.10 Define the terms "geodetic height," "geoid height," and "orthometric height." Include their relationship to each other.

13.11 Define PDOP, HDOP, and VDOP.

13.12 Define WAAS and EGNOS.

13.13 How can the receiver clock bias error be eliminated from carrier-phase measurements?

13.14 How can the satellite clock bias error be eliminated from carrier-phase measurements?

13.15 What is single differencing?

13.16 What is double differencing?

13.17 List and discuss the ephemerides.

13.18 Describe how the travel time of a GPS signal is measured.

13.19 If the HDOP during a survey is 1.53 and the UERE is estimated to be 1.65 m, what is the 95% horizontal point positioning error?

13.20 In Problem 13.19, if the VDOP is 3.3, what is the 95% point positioning error in geodetic height?

13.21* What are the geocentric coordinates in meters of a station in meters which has a latitude of 39°27′07.5894″ N, longitude of 86°16′23.4907″ W, and height of 203.245 m. (Use the WGS84 ellipsoid parameters.)

13.22 Same as Problem 13.21 except with geodetic coordinates of 45°26′32.0489″ N, longitude of 110°54′39.0646″ W, and height of 335.204 m?

13.23 Same as Problem 13.21 except with geodetic coordinates of 28°47′06.0936″ N, longitude of 75°52′35.0295″ W, and height of 845.678 m?

13.24* What are the geodetic coordinates in meters of a station with geocentric coordinates of (136,153.995, −4,859,278.535, 4,115,642.695)? (Use the WGS84 ellipsoid parameters.)

13.25 Same as Problem 13.24, except with geocentric coordinates in meters are (2,451,203.546, −4,056,568.907, 4,542,988.809)?

13.26 Same as Problem 13.24, except with geocentric coordinates in meters are (566,685.776, −4,911,654.896, −4,017,124.050)?

13.27 The GNSS determined height of a station is 588.648 m. The geoid height at the point is −28.45 m.

13.28* The GNSS determined height of a station is 284.097 m. The geoid height at the point is −30.052 m. What is the elevation of the point?

13.29 Same as Problem 13.28, except the height is 64.684 m and the geoid height is −28.968 m.

13.30* The elevation of a point is 124.886 m. The geoid height of the point is −28.998 m. What is the geodetic height of the point?

13.31 Same as Problem 13.30, except the elevation is 686.904 m, and the geoid height is −22.232 m.

13.32 The GNSS observed height of two stations is 124.685 m and 89.969 m, and their orthometric heights are 153.104 m and 118.386 m, respectively. These stations have model-derived geoid heights of −28.454 m and −28.457 m, respectively. What is the orthometric height of a station with a GNSS measured height of 105.968 m and a model-derived geoid height of −28.453 m?

13.33 Why are satellites at an elevation below 10° from the horizon eliminated from the positioning solution?

13.34 Research the Chinese satellite positioning system known as Compass and prepare a written report on the system.

13.35 Create a computational program that converts geocentric coordinates to geodetic coordinates.

13.36 Create a computational program that converts geodetic coordinates to geodetic coordinates.

13.37 Find at least two Internet sites that describe how GPS works. Summarize the contents of each site.

BIBLIOGRAPHY

Dodo, J. D., M. N. Kamarudin, and M. H. Yahya. 2008. "The Effect of Tropospheric Delay on GPS Height Differences along the Equator." *Surveying and Land Information Science* 68 (No. 3): 145.

Hofmann-Wellenhof, B., et al. 2004. *GPS Theory and Practice* 5th Ed. New York: Springer-Verlag.

Martin, D. J. 2003. "Around and Around with Orbits." *Professional Surveyor* 23 (No. 6): 50.

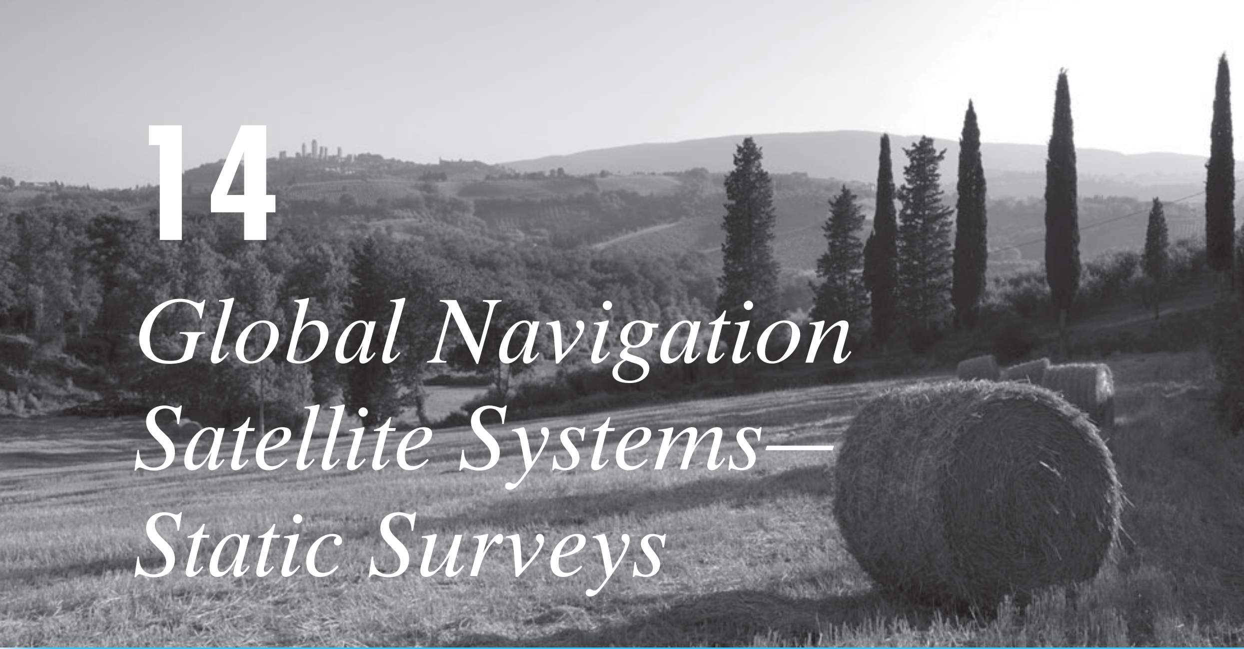
———. 2003. "Reaching New Heights in GPS, Part 3." *Professional Surveyor* 23 (No. 4): 42.

Reilly, J. 2003. "On Galileo, the European Satellite Navigation System." *Point of Beginning* 28 (No. 12): 46.

———. 2003. "On Geoid Models." *Point of Beginning* 29 (No. 12): 50.

Snay, R., et al. 2002. "GPS Precision with Carrier Phase Observations: Does Distance and/or Time Matter?" *Professional Surveyor* 22 (No. 10): 20.

Vittorini, L. D. and B. Robinson. 2003. "Optimizing Indoor GPS Performance." *GPS World* 14 (No. 11): 40.



14.1 INTRODUCTION

Many factors can have a bearing on the ultimate success of a satellite survey. Also there are many different approaches that can be taken in terms of equipment used and procedures followed. Because of these variables, satellite surveys should be carefully planned prior to going into the field. Small projects of lower-order accuracy may not require a great deal of preplanning beyond selecting receiver sites and making sure they are free from overhead obstructions. On the other hand, large projects that must be executed to a high order of accuracy will require extensive preplanning to increase the probability that the survey will be successful. As an example, a survey for the purpose of establishing control for an urban rapid transit project will command the utmost care in selecting personnel, equipment, and receiver sites. It will also be necessary to make a presurvey site visit to locate existing control, and identify possible overhead obstructions that could interfere with incoming satellite signals at all proposed receiver sites. In addition, a careful preanalysis should be made to plan optimum *observation session*[1] times, the durations of the sessions, and to develop a plan for the orderly execution of the sessions. The project will probably require ground communications to coordinate survey activities, a transportation analysis to ensure reasonable itineraries for the execution of the survey, and installation of monuments to permanently mark the new points that will be located in the survey. Consideration of these factors, and others, in planning and executing GNSS projects are the subjects of this chapter.

[1]An observation session denotes the period of time during which all receivers being employed on a project have been set up on designated stations and are simultaneously engaged in receiving satellite signals. When one session is completed, all receivers except one are generally moved to different stations and another observation session is conducted. The sessions are continued until all planned project observations have been completed.

Code-based receivers are used for positioning by people in all walks of life. They can be used by surveyors to gather details in situations not requiring typical survey accuracies. Examples are the approximate location of monuments, boundary or otherwise, to aid in later relocation, the collection of data to update small-scale maps in a geographic information system (GIS—see Chapter 28), and the navigation to monuments that are part of the National Spatial Reference System (see Chapter 19). The use of code-based receivers in nonsurveying applications includes the tracking of vehicles in transportation. The shipping industry uses code-based receivers for navigation. Likewise, surveyors may use the navigation functions of a code-based receiver to locate control monuments or other features where geodetic coordinates are known. Since the use of code-based receivers is so extensive and reaches far beyond the realm of the surveying community, their uses will not be covered in detail in this book.

This chapter concentrates on the use of carrier phase-shift receivers and the use of relative positioning techniques. This combination can provide the highest level of accuracy in determining the positions of points, and thus it is the preferred approach in surveying (geomatics) applications. But as noted in Chapter 13, the accuracy of a survey is also dependent on several additional variables. An important one is the type of carrier phase receiver used on the survey. As noted in Chapter 13, there are several types: *GNSS* receivers, which can utilize the multiple signals available from several different constellations; *dual-frequency receivers*, which can observe and process the multiple signals from the GPS constellation; and *single-frequency receivers*, which can observe only the $L1$ band. In precise surveys, GNSS and dual-frequency receivers are preferred for several reasons: they can (a) collect the needed data faster; (b) observe longer baselines with greater accuracy; and (c) eliminate certain errors, such as ionospheric refraction, and therefore yield higher positional accuracies. Receivers also vary by the number of channels. This controls the number of satellites that they can track simultaneously. As a minimum, carrier phase-shift receivers must have at least four channels, but some are capable of tracking as many as 30 satellites from the GPS, GLONASS, and Galileo constellations simultaneously using multiple frequency bands resulting in more than 60 channels. These receivers provide higher accuracies due to the increased number of satellites and increased strength in satellite geometry.

Other important variables that bear on the accuracy of a static survey include the (1) accuracy of the reference station(s) to which the survey will be tied, (2) number of satellites visible during the survey, (3) geometry of the satellites during the observation sessions, (4) atmospheric conditions during the observations, (5) lengths of observation sessions, (6) number and nature of obstructions at the proposed receiver stations, (7) number of redundant observations taken in the survey, and (8) method of reduction used by the software. Some of these factors are beyond control of the surveyor (geomatics engineer), and therefore it is imperative that observational checks be made. These are made possible by the redundant observations. This chapter will discuss these checks.

The use of satellites for specific types of surveys, for example, construction surveys, land surveys, photogrammetric surveys, etc., are covered in later chapters in this text. A GPS unit being used for a construction stakeout survey is shown in Figure 14.1. Use of satellite receivers for topographic surveys, and this application is covered in Section 17.9.5.

Figure 14.1
GPS receiver being used in construction stakeout. (Courtesy of Ashtech, LLC.)

14.2 FIELD PROCEDURES IN SATELLITE SURVEYS

In practice, field procedures employed on surveys depend on the capabilities of the receivers and the type of survey. Some specific field procedures currently being used in surveying include the *static, rapid static, pseudokinematic*, and *kinematic* methods. These are described in the subsections that follow. All are based on carrier phase-shift measurements and employ *relative positioning* techniques (see Section 13.9); that is, two (or more) receivers, occupying different stations and simultaneously making observations to several satellites. The vector (distance) between receivers is called a *baseline* as described in Section 13.9, and its ΔX, ΔY, and ΔZ coordinate difference components (in the geocentric coordinate system described in Section 13.4.2) are computed as a result of the observations.

14.2.1 Static Relative Positioning

For highest accuracy, for example geodetic control surveys, static surveying procedures are used. In this procedure, two (or more) receivers are employed. The process begins with one receiver (called the *base receiver*) being located on an existing control station, while the remaining receivers (called the *roving receivers*) occupy stations with unknown coordinates. For the first observing session, simultaneous observations are made from all stations to four or more satellites for a time period of an hour or more depending on the baseline length. (Longer baselines require greater observing times.) Except for one, all the receivers can be moved upon completion of the first session. The remaining receiver now serves as the base station for the next observation session. It can be selected from any of the receivers used in the first observation session. Upon completion of the second session, the process is repeated until all stations are

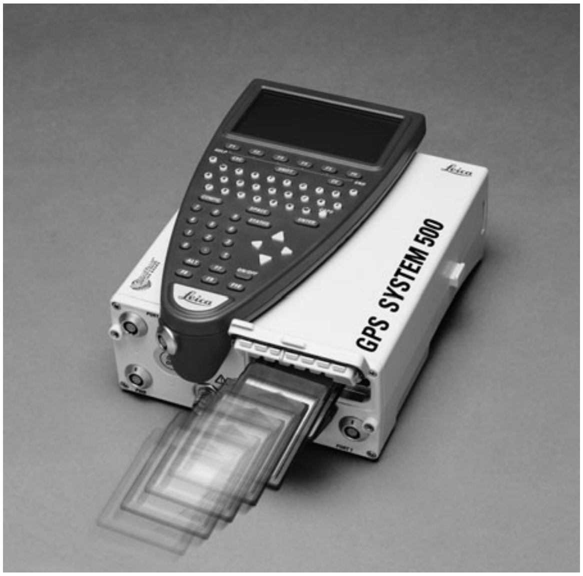


Figure 14.2 LEICA GPS System 500 with PCMCIA card. (Courtesy Leica Geosystems AG.)

occupied, and the observed baselines form geometrically closed figures. As discussed in Section 14.5, for checking purposes some repeat baseline observations should be made during the surveying process.

The value for the *epoch rate*[2] in a static survey must be the same for all receivers during the survey. Typically, this rate is set to 15 sec to minimize the number of observations and thus the data storage requirements. Most receivers either have internal memory capabilities or are connected to *controllers* that have internal memories for storing the observed data. After all observations are completed, data are transferred to a computer for post-processing. The receiver shown in Figure 14.2 has a PCMCIA card for downloading field data (see Section 2.12).

Relative accuracies with static relative positioning are about ±(3 to 5 mm + 1 ppm). Typical durations for observing sessions using this technique, with both single- and dual-frequency receivers, are shown in Table 14.1.

TABLE 14.1 TYPICAL SESSION LENGTHS FOR VARIOUS OBSERVATION METHODS

Method of Survey	Single Frequency	Dual Frequency
Static	30 min + 3 min/km	20 min + 2 min/km
Rapid Static	20 min + 2 min/km	10 min + 1 min/km

[2]GPS satellites continually transmit signals, but if they were continuously collected by the receivers, the volume of data and hence storage requirements would become overwhelming. Thus, the receivers are set to collect samples of the data at a certain time interval, which is called the epoch rate.

14.2.2 Rapid Static Relative Positioning

This procedure is similar to static surveying, except that one receiver always remains on a control station while the other(s) are moved progressively from one unknown point to the next. An observing session is conducted for each point, but the sessions are shorter than for the static method. Table 14.1 also shows the suggested session lengths for single- and dual-frequency receivers. The rapid static procedure is suitable for observing baselines up to 20 km in length under good observation conditions. Rapid static relative positioning can also yield accuracies on the order of about $\pm(3 \text{ to } 5 \text{ mm} + 1 \text{ ppm})$. However, to achieve these accuracies, optimal satellite configurations (good PDOP) and favorable ionospheric conditions must exist. This method is ideal for small control surveys. As with static surveys, all receivers should be set to collect data at the same epoch rate. Typically the epoch rate is set to 5 sec with this method.

14.2.3 Pseudokinematic Surveys

This procedure is also known as the *intermittent* or *reoccupation* method, and like the other static methods requires a minimum of two receivers. In pseudokinematic surveying, the base receiver always stays on a control station, while the rover goes to each point of unknown position. Two relatively short observation sessions (around 5 min each in duration) are conducted with the rover on each station. The time lapse between the first session at a station, and the repeat session, should be about an hour. This produces an increase in the geometric strength of the observations due to the change in satellite geometry that occurs over the time period. Reduction procedures are similar to those described in Section 13.9 and accuracies approach those of static surveying.

A disadvantage of this method, compared to other static methods, is the need to revisit the stations. This procedure requires careful presurvey planning to ensure that sufficient time is available for site revisitation, and to achieve the most efficient travel plan. Pseudokinematic surveys are most appropriately used where the points to be surveyed are along a road, and rapid movement from one site to another can be readily accomplished. During the movement from one site to another, the receiver can be turned off. Some projects for which pseudokinematic surveys may be appropriate include alignment surveys (see Chapters 24 and 25), photo-control surveys (see Chapter 27), lower-order control surveys, and mining surveys. Given the speed and accuracy of kinematic surveys, however, this survey procedure is seldom used in practice.

14.2.4 Kinematic Surveys

As the name implies, during kinematic surveys one receiver, the rover, can be in continuous motion. This is the most productive of the survey methods but is also the least accurate. The accuracy or a kinematic survey is typically in the range of $\pm(1 \text{ to } 2 \text{ cm} + 2 \text{ ppm})$. This accuracy is sufficient for many types of surveys and thus is the most common method of surveying. Kinematic methods are applicable for any type of survey that requires many points to be located, which makes it very appropriate for most topographic and construction surveys. It is also excellent for dynamic surveying, that is, where the observation station is in motion. The range of a kinematic survey is typically limited to the broadcast range of the base radio. However, real-time networks have made kinematic surveys possible over large regions. Chapter 15 explores the procedures used in kinematic surveying in greater detail. The remainder of this chapter is devoted to static surveying methods.

14.3 PLANNING SATELLITE SURVEYS

As noted earlier, small surveys generally do not require much in the way of project planning. However, for large projects and for higher-accuracy surveys, project planning is a critical component in obtaining successful results. The subsections that follow discuss various aspects of project planning with emphasis on control surveys.

14.3.1 Preliminary Considerations

All new high-accuracy survey projects that employ relative positioning techniques must be tied to nearby existing control points. Thus, one of the first things that must be done in planning a new project is to obtain information on the availability of existing control stations near the project area. For planning purposes, these should be plotted in their correct locations on an existing map or aerial photos of the area.

Another important factor that must be addressed in the preliminary stages of planning for projects is the selection of the new station locations. Of course, they must be chosen so that they meet the overall project objective. But in addition, terrain, vegetation, and other factors must be considered in their selection. If possible, they should be reasonably accessible by either the land vehicles or aircraft that will be used to transport the survey hardware. The stations can be somewhat removed from vehicle access points since hardware components are relatively small and portable. Also, the receiver antenna is the only hardware component that must be accurately centered over the ground station. It is easily hand carried and, when possible, can be separated from the other components by a length of cable, as shown in Figure 14.3. Once the preliminary station locations are selected, they should be plotted on the map or aerial photo of the area.

Another consideration in station selection is the assurance of an overhead view free of obstructions. This is known as *canopy restrictions*. Canopy restrictions may possibly block satellite signals, thus reducing observations and possibly adversely affecting satellite geometry. At a minimum, it is recommended that visibility be clear in all directions from a mask angle (altitude angle) of 10° to 20° from the horizon. In some cases, careful station placement will enable this visibility criterion to be met without difficulty; in other situations clearing around the stations may be necessary. Furthermore, as discussed in Section 13.6.3, potential sources that can cause interference and multipath errors should also be identified when visiting each site.

Selecting suitable *observation windows* is another important activity in planning surveys. This consists of determining which satellites will be visible from a given ground station or project area during a proposed observing

Figure 14.3
GPS antenna attached with cables to controller unit. (Courtesy of Trimble.)

period. To aid in this activity, azimuth and elevation angles to each visible satellite can be predetermined using almanac data for times within the planned observation period. Required computer input, in addition to observing date and time, includes the station's approximate latitude and longitude, and a relatively current satellite almanac. Additionally, the space weather should be checked for possible solar storms during the periods of occupation. Days where the solar radiation storm activity is rated from strong to extreme should be avoided. Section 15.2 discusses the effects of space weather on GNSS surveys in detail.

To aid in selecting suitable observation windows, a *satellite availability plot*, as shown in Figure 14.4, can be applied. The shaded portion of this diagram shows the number of satellites visible from station *PSU1* whose position is 41°18′00.00″ N latitude and 76°00′00.40″ W longitude. The diagram is applicable for August 4, 1999 between the hours of 8:00 and 17:00 EDT. A mask angle of 15° has been used. In addition to showing the number of visible satellites, the lines running through the plot depict the predicted PDOP, HDOP, and VDOP (see Section 13.6.4) for this time period. It should be noted that, for the day shown in Figure 14.4, only two short time periods are unacceptable for data collection. DOP *spikes* occur between 8:02 and 8:12 when only four satellites are above the

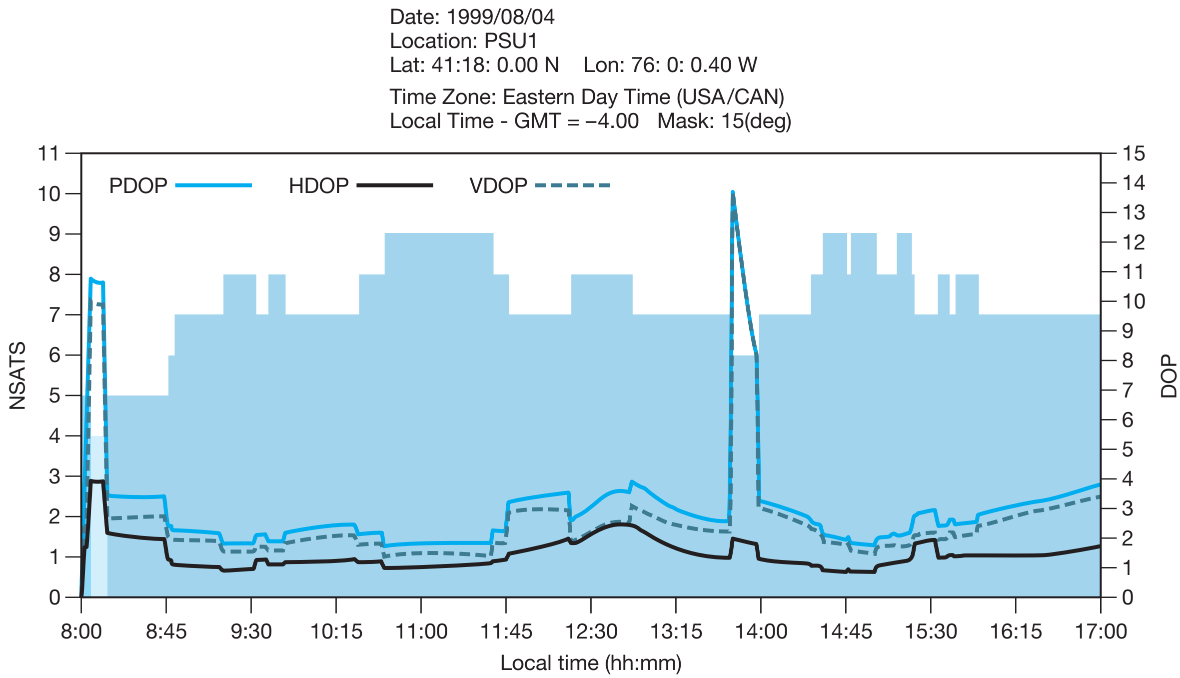


Figure 14.4 Satellite availability plot.

horizon mask angle, and between 13:45 and 14:00 when both VDOP and PDOP are unacceptable because of weak satellite geometry. However during this latter period, the HDOP is acceptable, indicating that a horizontal control survey could still be executed. Note also that one of the better times for data collection is between 10:40 and 11:30 when PDOP is below 2 because 9 satellites are visible during that time. However, if the receiver has less than 9 channels, then the satellite availability chart must be altered by disabling some of the usable satellites or by raising the mask angle.

Satellite visibility at any station is easily and quickly investigated using a *sky plot*. These provide a graphic representation of the azimuths and elevations to visible satellites from a given location. As illustrated in Figure 14.5(a) and (b), sky plots consist of a series of concentric circles. The circumference of the outer circle is graduated from 0° to 360° to represent satellite azimuths. Each successive concentric circle progressing toward the center represents an increment in the elevation angle with the radius point corresponding to the zenith.

For each satellite, the PRN number is plotted beside its first data point, which is its location for the selected starting time of a survey. Then arcs connect successive plotted positions for the given time increments after the initial time. Thus, travel paths in the sky of visible satellites are shown. Sky plots are valuable in survey planning because they enable operators to quickly visualize not only the number of satellites available during a planned observation period, but also their geometric distribution in the sky.

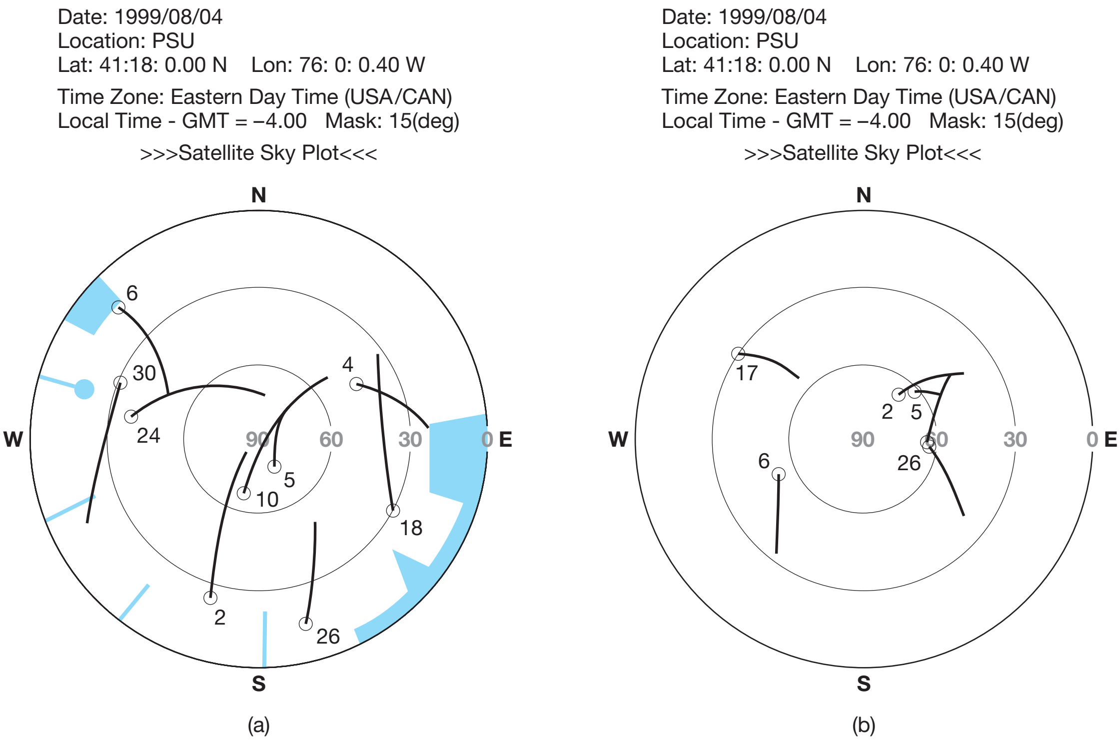


Figure 14.5 Sky plots for station PSU1 on August 4, 1999. Plot (a) shows the obstructing objects around the station from 10 to 12. Plot (b) shows the weak satellite geometry from 13:30 to 14.

In project planning, the elevations and azimuths of vertical obstructions near the station can be overlaid with the sky plot to form *obstruction diagrams*. The diagram will then show whether crucial satellites are removed by the obstructions and also indicate the best times to occupy the station to avoid the obstructions. As shown in Figure 14.5(a), a building will obscure satellite 6 briefly during the time period shown. Also note that signals from satellite 30 will experience a brief interruption caused by the presence of a nearby pole later in the session, and that satellite 4 will be lost near the end of the session because of a nearby building. None of these obstructions appear to be critical to the session.

The analysis of sky obstructions and satellite geometry is important for the highest accuracy in surveys. Recall from Section 13.6.4 that observations should be taken on groups of four or more widely spaced satellites that form a strong geometric intersection at the observing station. This condition is illustrated in Figures 13.11(b) and 14.5(a). Weaker geometry, as shown in Figures 13.11(a) and 14.5(b), should be avoided if possible as it will yield lower accuracy. The PDOP and VDOP spikes shown in Figure 14.4 between the times from 13:45 and 14:00 are caused by the closely clustered distribution of the satellites as shown in Figure 14.5(b).

It is important to note that the optimal observation times will repeat 4 min earlier for each day following the planning session. That is, in Figure 14.4, the same satellite visibility chart will apply for the period from 7:56 to 16:56 on August 5, and from 7:52 to 16:52 on August 6, etc. Of course, the periods of poor PDOP will also shift by 4 min each day. This shift occurs because sidereal days are approximately 4 min shorter than solar days. Modern receivers, with their built-in microprocessors, can derive sky plots, compute PDOP values in the field, and display the results on their screens.

14.3.2 Selecting the Appropriate Survey Method

As discussed in Section 14.2, several different methods are available with which to accomplish a survey. Each method provides a unique set of procedural requirements for field personnel. In high-accuracy surveys that involve long baselines, the static surveying method with GNSS receivers is the best solution. However in typical surveys limited to small areas, a single-frequency receiver using rapid static, pseudokinematic, or kinematic surveying methods (see Chapter 15) may be sufficient. Because of the variability in requirements and restrictions of surveys, the selection of the appropriate survey method is dependent on the (1) desired level of accuracy in the final coordinates, (2) intended use of the survey, (3) type of equipment available for the survey, (4) size of the survey, (5) canopy and other local conditions for the survey, and (6) available software for reducing the data, there is seldom only one method for accomplishing the work.

GNSS receivers will reduce the time required at each station in a static survey due to the increased number of visible satellites and the improved satellite geometry. Figure 14.6 is a schematic diagram that categorizes the various survey methods. The surveying community has used those shown on the left side of the diagram traditionally.

The selection of the appropriate survey method is dependent on the (1) desired level of accuracy in the final coordinates, (2) intended use of the survey, (3) type of equipment available for the survey, (4) size of the survey, (5) canopy and other local conditions for the survey, and (6) available software for reducing the data.

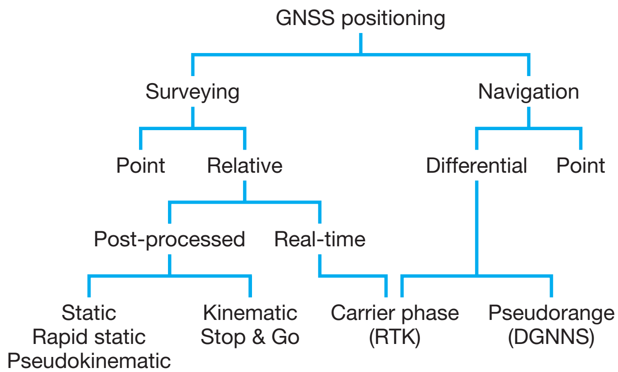


Figure 14.6
Schematic depicting GNSS positioning methods.

For mapping or inventory surveys where centimeter to submeter accuracy is sufficient, code-based receivers or an RTK survey (see Chapter 15) may provide the most economical product. However if the area to be mapped has several overhead obstructions, it may only be possible to use one of the static survey procedures to bring control into the region, and do limited kinematic mapping in small areas where clear overhead views to the satellites are available. Recognizing this, many manufacturers have developed equipment that allows the surveyor (geomatics engineer) to switch between a receiver and a total station instrument (see Chapter 8) using the same data collector or survey controller. This capability is useful in areas where GNSS surveys are not practical. A more in-depth discussion on this topic will be presented in Chapter 17. Restrictions in use of GNSS surveys due to canopy restrictions will be greatly reduced when modernized GNSS constellations are available.

The preferred approach in performing high-accuracy control surveys is the static method. Often a combination of the static methods will provide the most economical results for large projects. As an example, a static survey may be used to bring a sparse network of accurate control into a project area. This could be followed by a rapid static survey to densify the control. Finally, a kinematic survey could be used to establish control along specific alignments. In smaller areas with favorable viewing conditions, the best method for the survey may simply be the rapid static or kinematic methods. Available equipment, software, and experience often dictate the survey method of choice.

The NGS provides the *Online Positioning User Service* (OPUS). When using this service with a post-processed kinematic survey, the base station does not have to be positioned over a point with known coordinates. However, for this service to achieve its maximum accuracy, the base station must continuously collect data for a minimum of 2 h. If dual-frequency or GNSS receivers are used in the survey, however, the NGS provides a rapid static OPUS service that will determine the position of a receiver with as little as 15 min of data. Initialization of the roving receiver is obtained using the on-the-fly method of ambiguity resolutions. After the initialization, the roving receiver can proceed with the survey. Before post-processing the data from the roving receiver, the RINEX observation file from the base station is submitted to the NGS for processing over the Internet. In a matter of minutes, OPUS returns the coordinates of the base station using the data from three nearby CORS stations. This procedure removes the need for an initial known base station.

14.3.3 Field Reconnaissance

Once the existing nearby control points and new stations have been located on paper, a reconnaissance trip to the field should be undertaken to check the selected observation sites for (1) overhead obstructions that rise above 10° to 15° from the horizon, (2) reflecting surfaces that can cause multipathing, (3) nearby electrical installations that can interfere with the satellite's signal, and (4) other potential problems. If reconnaissance reveals that any preliminary selected station locations are unsatisfactory, adjustments in their positions should be made. For the existing control stations that will be used in the survey, ties can be made to nearby permanent objects, and also photos or *rubbings*[3] of the monument caps should be created. These items will aid crews in locating the stations later during the survey, reduce the amount of time spent at each station, and minimize station misidentifications. Web services such as GOOGLE EARTH can often be used to make preliminary decision about the suitability of a site for occupation by a GNSS receiver. However, a site visit is the only method to confirm its suitability.

Once final sites have been selected for the new stations, permanent monuments should be set, and the positions of these stations also documented with ties to nearby objects, photos, and rubbings. At this time, an accurate horizon plot of any surrounding obstructions can be prepared, and road directions and approximate driving times between stations recorded. There are several web services that can be used to obtain driving times and directions between stations. A valuable aid in identifying locations of stations is the use of a code-based receiver. These inexpensive devices will determine the geodetic coordinates of the stations with sufficient accuracy to enable plotting them on a map, navigation to the station, and project planning.

14.3.4 Developing an Observation Scheme

For surveying projects, especially those that employ relative positioning and are applied in control surveying, once existing nearby control points have been located and the new stations set, they and the observations that will be made comprise what is termed a *network*. Depending on the nature of the project and extent of the survey, the network can vary from only a few stations to very large and complicated configurations. Figure 14.7 illustrates a small network consisting of only two existing control points and four new stations.

After the stations are established, an observation scheme is developed for performing the work. The scheme consists of a planned sequence of observing

Approx. Distances

Line	Length (km)
AB	17 (between 2 control points)
AC	13
AE	7
AF	7
BC	11
BD	11
BF	11
FC	11
FE	7
FD	9
ED	9
CD	18

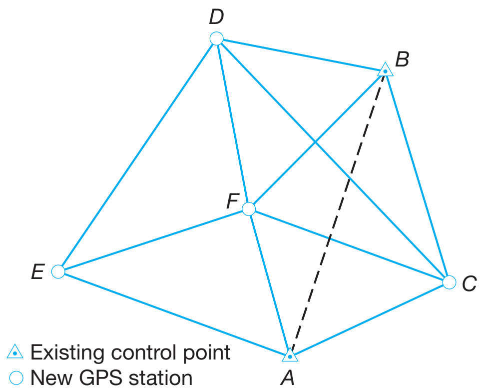


Figure 14.7
A GNSS network.

[3]Monuments used to mark stations generally have metal caps (often brass) that give the name of the point and other information about the station. This information is stamped into the cap, and by laying a piece of paper directly over the cap, and rubbing across the surface with the side of a pencil lead, an imprint of the cap is obtained. This helps to eliminate mistakes in station identification.

sessions that accomplishes the objectives of the survey in the most efficient manner. As a minimum, it must ensure that every station in the network is connected to at least one other station by a *nontrivial* (also called *independent*) baseline. (Nontrivial baselines are described in the following paragraph.) However, the plan should also include some redundant observations (i.e., baseline observations between existing control stations, multiple occupations of stations, and repeat observations of certain baselines) to be used for checking purposes, and for improving the precision and reliability of the work. Desired accuracy is the principal factor governing the number and type of redundant observations. The *Federal Geodetic Control Subcommittee* (FGCS) has developed a set of standards and specifications for GPS relative positioning (see Section 14.5.1) that specify the number and types of redundant observations necessary for accuracy orders *AA*, *A*, *B*, and *C*. Generally on larger, high-accuracy projects, these standards and specifications, or other similar ones, govern the conduct of the survey work and must be carefully followed.

In relative positioning, for any observing session the number of nontrivial baselines measured is one less than the number of receivers used in the session or

$$b = n - 1 \tag{14.1a}$$

where b is the number of nontrivial baselines, and n is the number of receivers being employed in the session. When only two receivers are used in a session, only one baseline is observed and it is nontrivial. If more than two receivers are used, both nontrivial and *trivial* (*dependent*) baselines will result. The total number of baselines can be computed as

$$T = \frac{n(n - 1)}{2} \tag{14.1b}$$

where T is the total number of baselines possible and n is as previously defined. The number of trivial baselines is

$$t = \frac{(n - 1)(n - 2)}{2} \tag{14.1c}$$

where t is the total number of trivial baselines. To differentiate between these two types of baselines, and to understand how they occur, refer to Figure 14.8. In this figure, an observing session involving three receivers *A*, *B*, and *C* observing four satellites *1*, *2*, *3*, and *4* is shown. Pseudoranges *1A*, *1B*, *2A*, *2B*, *3A*, *3B*, *4A*, and *4B* are employed to compute the baseline vector *AB*. Also pseudoranges *1B*, *1C*, *2B*, *2C*, *3B*, *3C*, *4B*, and *4C* are used in computing the baseline vector *BC*. Thus, all possible pseudoranges in this example have been used in calculating baselines *AB* and *BC*, and the computation of baseline *CA* would be redundant; that is, it would be based upon observations that have already been used. In this example, baselines *AB* and *BC* are considered nontrivial and *CA* trivial. This is reinforced with the fact that adding vectors *AB* and *BC* will result in vector *CA*, which demonstrates their mathematical dependence. However, the selection of the trivial baseline is arbitrary, that is, either *AB*, *BC*, or *AC* could be selected as the trivial baseline, dependent on which two baselines were selected as nontrivial. If four receivers are used in a session, six baselines will result: three nontrivial and three trivial. Students should verify this with a sketch. For meeting accuracy standards,

Figure 14.8
GNSS observation session using three receivers. In the case shown, *AB* and *BC* are considered nontrivial baselines. Thus, *AC* is a trivial baseline.

only nontrivial baselines can be considered, and thus distinguishing between them is important.

When possible, at least one baseline should be observed between existing control monuments of higher accuracy for every receiver-pair used on a project to check the performance of the equipment and the reliability of the control. Also as noted above, some baselines should be observed more than once. These repeat baselines should ideally be observed at or near the beginning and at the end of the project observations to check the equipment for repeatability. The analysis of these duplicate observations will be covered in Section 14.5.

For control surveys, the baselines should form closed geometric figures since they are necessary to perform closure checks (see Section 14.5). The simple network of baselines, shown in Figure 14.7, will be used as an example to illustrate survey planning. Assume that the project is in the area of control station *PSU1* and that the survey dates will be August 4 and 5, 1999. Thus, the satellite availability plot and sky plots of Figures 14.4 and 14.5, respectively, apply. Stations *A* and *B* in Figure 14.7 are existing control monuments, and a baseline observation between them will be planned to (1) verify the accuracy of the existing control and (2) confirm that the equipment is in proper working condition.

In the example of Figure 14.7, it is assumed that two dual-frequency receivers are available for the survey and that the rapid static method will be used. Following the minimum session lengths, as given in Table 14.1, of 10 min + 1 min/km, baseline *AB* would require $10 + 1 \times 17$, or 27, minutes of observation time. The remaining baseline observation times are listed in Table 14.2 using the same computational techniques.

Two two-person crews, each working individually with separate vehicles, are assumed for conducting the survey. It is also assumed that setup and teardown times at each station are both approximately 15 min. By rounding each minimal observation session up to a nearest 5-min interval, the following observation

TABLE 14.2 MINIMUM SESSION LENGTHS AND APPROXIMATE DRIVING TIMES BETWEEN STATIONS FOR BASELINES IN FIGURE 14.11

Line	Length (km)	Session Lengths (min)	Driving Time (min)
AB	17	27	15
AC	13	23	10
AE	7	17	8
AF	7	17	25
BC	11	21	15
BD	11	21	10
BF	11	21	20
FC	11	21	15
FE	7	17	15
FD	9	19	10
ED	9	19	15
CD	18	28	25

sessions and times were planned for a two-day data collection project. It is important to remember that a session involves the simultaneous collection of satellite data. Thus, a session does not start until all receivers involved in the session are set up and running on their respective stations. Thus communication between the field crews is crucial to a successful survey.

The observation plan for this example is given in Table 14.3. The table gives the itinerary for both field crews, allowing time for set up and teardown of the equipment, travel between stations, and collecting sufficient observations. The plan includes all lines in the network as baselines. These include those for checking purposes, an observation of the control baseline AB, and repeat observations of AF and BF. Note that the two unfavorable times for collecting observations shown in Figure 14.4 are not scheduled as times for collecting data, but are used for other ancillary operations. In the event that operations should actually run ahead of, or fall behind the planned schedule for some unforeseen reason, it is prudent to include a statement on the itinerary indicating that no baseline observations should be collected between the times of 8:00 to 8:15 and 13:40 to 14:00. Note that as indicated in the planned schedule, the crew with the stationary receiver should continue to collect data during the entire period of occupation of the station. This includes the periods of time during which the other crew is moving between stations. If a CORS site (see Section 14.3.5) is available, the data collected by the stationary receiver can be used to create stronger baseline link with the CORS station. Optionally, the longer sessions can be processed using OPUS to create additional control in the project. It is desirable to provide the field personnel with communication devices during the survey so that they can coordinate times for sessions, and handle unforeseen logistical problems that inevitably arise.

TABLE 14.3 OBSERVATION ITINERARY FOR FIGURE 14.11

DAY 1 (August 4, 1999)

Time	Crew 1	Crew 2	Baseline	Session
8:00–8:45	Drive to Station *A*	Drive to Station *C*		
9:00–9:25	Collect data	Collect data	*AC*	A1
9:40–9:55	Collect data	Drive to Station *F*		
9:55–10:15	Collect data	Collect data	*AF*	A2
10:30–10:45	Collect data	Drive to Station *E*		
11:00–11:20	Collect data	Collect data	*AE*	A3
11:35–11:50	Drive to Station *B*	Drive to Station *F*		
12:05–12:30	Collect data	Collect data	*BF*	A4
12:45–1:00	Collect data	Drive to Station *C*		
13:15–13:40	Collect data	Collect data	*BC*	A5
13:55–14:05	Collect data	Drive to Station *A*		
14:20–14:50	Collect data	Collect data	*AB*	A6
15:05–15:15	Drive to Station *D*	Drive to Station *C*		
15:30–16:00	Collect data	Collect data	*CD*	A7
16:00–17:00	Return to office	Download data		

DAY 2 (August 5, 1999)

Time	Crew 1	Crew 2	Baseline	Session
8:00–9:00	Drive to Station *A*	Drive to Station *F*		
9:15–9:35	Collect data	Collect data	*FA*	B1
9:50–10:00	Drive to Station *E*	Collect data		
10:15–10:30	Collect data	Collect data	*FE*	B2
10:45–11:00	Drive to Station *D*	Collect data		
11:00–11:20	Collect data	Collect data	*FD*	B3
11:35–11:45	Drive to Station *B*	Collect data		
12:00–12:25	Collect data	Collect data	*FB*	B4
12:40–12:55	Drive to Station *C*	Collect data		
13:10–13:35	Collect data	Collect data	*FC*	B5
13:50–14:05	Drive to Station *B*	Drive to Station *D*		
14:20–14:45	Collect data	Collect data	*BD*	B6
15:00–15:15	Drive to Station *D*	Drive to Station *E*		
15:30–15:50	Collect data	Collect data	*ED*	B7
15:50–17:00	Return to office	Download data		

Note: No baseline observations should be made between 8:00–8:15 and 13:40–1400 on August 4, and between 7:56–8:11 and 13:36–13:56 on August 5.

14.3.5 Availability of Reference Stations

As explained in Section 14.3.4, the availability of high-quality reference control stations is necessary to achieve the highest order of accuracy in positioning. To meet this need, individual states, in cooperation with the NGS, have developed ***H**igh **A**ccuracy **R**eference **N**etworks* (HARNs). The HARN is a network of control points that were precisely observed using GPS under the direction of the National Geodetic Survey (NGS). These HARN points are now available to serve as reference stations for surveys in their vicinity.

In recent years, the NGS with cooperation from other public and private agencies has created a national system of ***C**ontinuously **O**perating **R**eference **S**tations*, also called the *National CORS Network*. The location of stations in the CORS network as of 2003 is shown in Figure 14.9. These stations not only have their positions known to high accuracy[4] but also they are occupied by a receiver

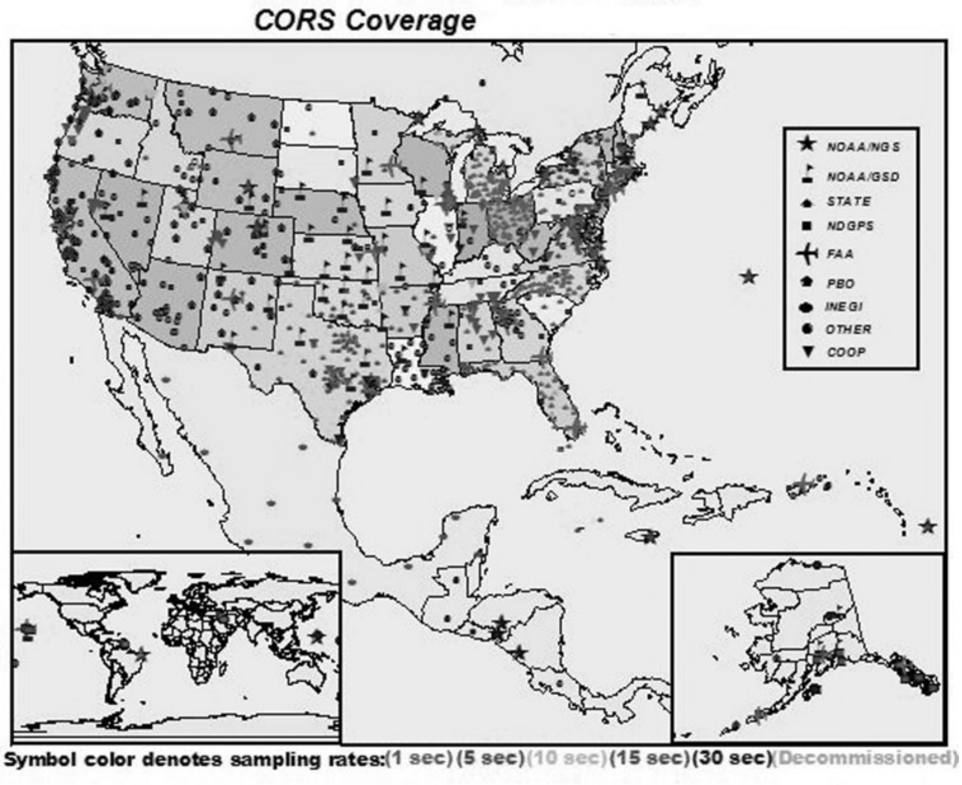


Figure 14.9 Locations that have continuously operating reference stations (CORS) and their coverages. (Courtesy National Geodetic Survey.)

[4]At the time of this writing, the CORS and HARN sites are computed using different reference coordinate systems, and thus do not represent homogenous sets of coordinates. The NGS is currently working on readjusting both networks of points so that this problem does not exist in the future. However, current surveys should not mix HARN and CORS stations until the NGS has completed the adjustment and published the results.

that continuously collects satellite data. The collected data is then downloaded and posted on the NGS Internet site at http://www.ngs.noaa.gov/CORS/. This information can be used as base station data to support roving receivers operating in the vicinity of the CORS station. The data is stored in a format known as ***R**eceiver **IN**dependent **EX**change* (RINEX2). This format is a standard that can be read by all post-processing software. This Internet site also provides coordinates for the stations, and the ultra-rapid, rapid, and precise ephemerides (see Section 13.6.3).

Because of plate tectonics and the differences in which the reference systems handle these motions, their velocity vectors accompany the ITRF00 coordinates for the stations. While motions in the eastern United States may be under 1 mm per year, these motions may be considerably more in the western United States. Thus, coordinates derived from control surveys should be accompanied by their reference frame (see Section 19.6.3) and epoch. The NGS has written software known as *horizontal time-dependent positioning* (HTDP) using 14 parameters to transform the ITRF00 positions with their velocity vectors into NAD83 (CORS96) equivalents (see Section 19.6.6). The WGS84(G1150) coordinates are close to ITRF00 coordinates. This coordinate transformation is demonstrated in the Mathcad® file *C14.XMCD*, which is on the companion website for this book at http://www.pearsonhighered.com/ghilani.

The CORS data files are easily downloaded using the *User-Friendly CORS* (UFCORS) option on the NGS website. This utility provides the user with an interactive form which requests the (1) starting date, (2) time, (3) duration of the survey, (4) site selection, and (5) the collection interval among other things. The request is interpreted by the server, and the appropriate data is sent, via the Internet, to the user within minutes.

Several factors can cause data not to be collected at a particular CORS site for short periods of time. These include local power outages, storm damage, and software and hardware failures. Thus, if a certain CORS station is planned for use on a particular survey, it is important to check the availability of the station prior to the beginning of a survey to ensure that it is functioning properly and that the necessary data will be available following the survey. However, given the number of CORS stations available in the conterminous United States (CONUS), it is always possible to select other CORS sites for processing when the desired station is unavailable.

14.4 PERFORMING STATIC SURVEYS

During a control survey using either the static or rapid static method, the field crew initially locates the control station and erects the antenna over the station so that it is level. To minimize setup errors, 2-m fixed-height poles or tripods are often used. Cables are connected to the antenna and controller, and the procedures specified by the manufacturer are followed to initiate data collection.

While the data is being logged, other ancillary information at the site can be collected and recorded. Typical ancillary data obtained during a survey includes (1) project and station names, (2) ties to the station, (3) a rubbing of the monument cap, (4) photos of the setup showing identifying background features, (5) potential obstructive or reflective surfaces, (6) date and session number, (7) start and stop times, (8) name of observer, (9) receiver and antenna serial number, (10) meteorological data, (11) PDOP value, (12) antenna height, (13) orientation of antenna, (14) rate of data collection (epoch rate), and (15) notations on any problems experienced. This ancillary data is typically entered onto a station log sheet. An example of a site log sheet is shown in Figure 14.10.

SITE LOG

Project: ______________________________ Station: ________

Session Number: ______ Observer: ______________ Date: ___ /___ /___

Start Time: ____ : ____ (AM) (PM) Stop Time: ____ : ____ (AM) (PM)

Latitude: ___ ° ____ ′ _____ ″ Latitude: ___ ° ____ ′ _____ ″ Height: _____

Temperature: ______ (F) (C) Pressure: ______ Relative Humidity: _______

PDOP: ________ Epoch Rate: ________ sec

Initial Slant Height: ______ (m) (ft) Final Slant Height: ______ (m) (ft)

Receiver S/N: __________ Antenna S/N: __________

Photo #: ________ Visible Satellites: ________________

Description of Monument *(include rubbing of surface)*:

Potential Problems:

Ties to Monument:

Comments:

____
Signature of Observer

Figure 14.10
Data log sheet for a static or rapid static survey.

Modern survey controllers can store the satellite observations internally, however, the receiver's internal memory may be used also. In the latter case, a survey controller may not even be necessary to collect data. At least once a day, and preferably twice a day, the data should be transferred to a computer. This can be done in the field with a laptop computer or at the end of the day in the office. File transfer is accomplished using software provided by the manufacturer. The GPS receiver shown in Figure 14.2 can transfer data from the controller to a PCMCIA card.

14.5 DATA PROCESSING AND ANALYSIS

The first step in processing the data is to transfer the observational files from the receiver to the computer. Typically, software provided by the hardware manufacturer performs this function. In this process, the software will download all of the observation files into one subdirectory on the computer. This directory is part of the project's directory. As the observation files are downloaded, special attention should be given to checking station information that is read directly from the file with the site log sheets. Catching incorrectly entered items such as station identification, antenna heights, and antenna offsets at this point can greatly reduce later problems during processing. Batch processing typically performs the reduction process. Figure 14.11 shows an individual processing satellite data with a PC.

There are three types of processing software: (1) *single baseline*, (2) *session processing*, and (3) *multipoint solutions*. The single baseline solution is most commonly used and will be discussed here. The session processing software simultaneously computes all nontrivial baselines for any particular session. The

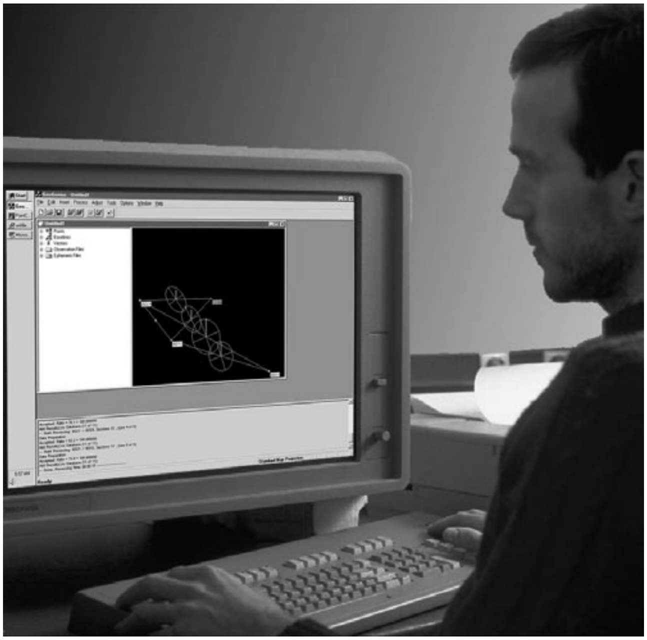


Figure 14.11 Processing GNSS data. (Courtesy of Ashtech, LLC.)

most common method of using multipoint solution software is to first use the single baseline software to isolate any problems in the observations because it provides better checks to isolate “bad” baselines. The multipoint solution software eliminates the need for the adjustment of the baselines as discussed in Section 14.5.5.

The initial processing can use the “broadcast” ephemeris, but it is recommended to use any of the “precise” ephemerides to achieve a higher level of accuracy since the precise ephemeris will remove orbital errors from the processing (see Section 13.6.3). When CORS sites are used for a project, the most precise ephemeris available at the time of the download can be requested. Proper use of the “user-friendly” option, which is available on the NGS website, will automatically result in interpolated values for the data at the specified epoch rate to match the survey. Additionally as mentioned in Section 13.6.3, NGS antenna calibration data should be used when determining the baseline vectors to account for changing antenna offsets due to changes in satellite elevation.

As shown in Figure 13.12, the baseline is computed as changes in X, Y, and Z between two stations. Thus if station A has known coordinates, then the coordinates of station B can be computed using Equations (13.28). The single baseline processing software will (1) generate orbit files, (2) compute the best-fit point positions from the code pseudoranges, (3) compute baseline components (ΔX, ΔY, ΔZ) using the double difference Equation (13.32), and (4) compute statistical information for the baseline components.

As each baseline is computed, any integer ambiguity problems should be identified and corrected. Integer ambiguity problems occur when a receiver loses lock on the satellite, which can happen because of obstructions, high solar activity, or multipathing. The only option typically provided by the software to correct this problem is to eliminate the periods of satellite data where the problem occurred. This may be done graphically or manually depending on the specific software being used.

After all baselines have been computed, their geometric closures can be analyzed. This step follows a series of procedures that check the data for internal consistency and eliminate possible blunders. No control points are needed for these analyses. Depending on the actual observations taken and the network geometry, these procedures may involve analyzing (1) differences between fixed and measured baseline components, (2) differences between repeated measurements of the same baseline components, and (3) loop closures. Procedures for making these analyses are described in the following subsections.

14.5.1 Specifications for Static Surveys

The *Federal Geodetic Control Subcommittee* (FGCS) has published the document entitled “Geometric Geodetic Accuracy Standards and Specifications for Using GPS Relative Positioning Techniques.”[5] The document specifies seven

[5]Available from the National Geodetic Information Center, NOAA, National Geodetic Survey, N/CG17, SSMC3 Station 09535, 1315 East West Highway, Silver Spring, MD 20910, telephone (301) 713-3242. Their email address is info_center@ngs.noaa.gov, and their website address is http://www.ngs.noaa.gov.

TABLE 14.4 GPS RELATIVE POSITIONING ORDERS OF ACCURACY

Order	Allowable Error Ratio	Parts Per Million (ppm)
AA	1:100,000,000	0.01
A	1:10,000,000	0.1
B	1:1,000,000	1.0
C-1	1:100,000	10
C-2-I	1:50,000	20
C-2-II	1:20,000	50
C-3	1:10,000	100

different orders of accuracy for relative positioning and provides guidelines on instruments and field and office procedures to follow to achieve them. Table 14.4 lists these orders of accuracy.

The FGCS document also makes recommendations concerning categories of surveys for which the different orders of accuracy are appropriate. Some of these recommendations include order AA for global and regional geodynamics and deformation measurements; order A for "primary" networks of the National Spatial Reference System (NSRS), and regional and local geodynamics; order B for "secondary" NSRS networks and high-precision engineering surveys; and the various classes of order C for mapping control surveys, property surveys, and engineering surveys. The allowable error ratios given in these standards imply the extremely high accuracies that are now possible with relative positioning techniques.

The National Geodetic Survey has created geodetic height accuracy standards that can be used for vertical surveys. These standards are based on the changes in geodetic heights between control stations. Following a correctly weighted, minimally constrained, least-squares baseline adjustment (see Section 14.5.5), the geodetic height order and class can be determined. This standard is based on the standard deviation (s) of the geodetic height difference (in millimeters) between two points obtained from the adjustment, and the distance (d) between the two control points (in kilometers). The ellipsoid height difference accuracy (b) is computed as

$$b = \frac{s}{\sqrt{d}} \tag{14.2}$$

Table 14.5 lists classifications versus value of b computed in Equation (14.2).

14.5.2 Analysis of Fixed Baseline Measurements

As noted earlier, job specifications often require that baseline observations be taken between fixed control stations. The benefit of making these observations is to verify the accuracy of both the observational process and the control being held fixed. Obviously smaller discrepancies between observed and known baseline

TABLE 14.5 NGS GEODETIC HEIGHT ORDER AND CLASS

Order	Class	Maximum Height Difference Accuracy (*b*)
First	I	0.5
First	II	0.7
Second	I	1.0
Second	II	1.3
Third	I	2.0
Third	II	3.0
Fourth	I	6.0
Fourth	II	15.0
Fifth	I	30.0
Fifth	II	60.0

lengths mean better precisions. If the discrepancies are too large to be tolerated, the conditions causing them must be investigated before proceeding further. Table 14.6 shows the computed baseline vector data from a survey of the network of Figure 14.7. Note that one fixed baseline (between control points A and B) was observed.

TABLE 14.6 OBSERVED BASELINE VECTORS FOR FIGURE 14.11

Obs. Baseline	ΔX(m)	ΔY(m)	ΔZ(m)
AC	11644.2232	3601.2165	3399.2550
AE	−5321.7164	3634.0754	3173.6652
BC	3960.5442	−6681.2467	−7279.0148
BD	−11167.6076	−394.5204	−907.9593
DC	15128.1647	−6286.7054	−6371.0583
DE	−1837.7459	−6253.8534	−6596.6697
FA	−1116.4523	−4596.1610	−4355.8962
FC	10527.7852	−994.9377	−956.6246
FE	−6438.1364	−962.0694	−1182.2305
FD	−4600.3787	5291.7785	5414.4311
FB	6567.2311	5686.2926	6322.3917
BF	−6567.2310	−5686.3033	−6322.3807
AF	1116.6883	4596.4550	4355.3008
AB	7683.6883	10282.4550	10678.3008

Assuming the geocentric coordinates of station A are (402.3509, −4,652,995.3011, 4,349,760.7775) and those of station B are (8086.0318, −4,642, 712.8474, 4,360, 439.0833), the analysis of the fixed baseline is as follows.

1. Compute the coordinate differences between stations A and B as

$$\begin{aligned}
\Delta X_{AB} &= 8086.0318 - 402.3509 &&= 7683.6809 \\
\Delta Y_{AB} &= -4{,}642{,}712.8474 + 4{,}652{,}995.3011 &&= 10{,}282.4537 \\
\Delta Z_{AB} &= 4{,}360{,}439.0833 - 4{,}349{,}760.7775 &&= 10{,}678.3058
\end{aligned}$$

2. Compute the absolute values of differences between the observed and fixed baselines as

$$\begin{aligned}
dX &= |7683.6883 - 7683.6809| &&= 0.0074 \\
dY &= |10{,}282.4550 - 10{,}282.4537| &&= 0.0013 \\
dZ &= |10{,}678.3008 - 10{,}678.3058| &&= 0.0050
\end{aligned}$$

3. Compute the length of the baseline as

$$AB = \sqrt{(7683.6809)^2 + (10{,}282.4537)^2 + (10{,}678.3058)^2} = 16{,}697.126 \text{ m}$$

4. Express the differences, as computed in Step 2, in parts per million (ppm) by dividing the difference by the length of the baseline computed in Step 3 as

$$\begin{aligned}
\Delta X\text{-ppm} &= 0.0074/16{,}697.126 \times 1{,}000{,}000 = 0.44 \\
\Delta Y\text{-ppm} &= 0.0013/16{,}697.126 \times 1{,}000{,}000 = 0.08 \\
\Delta Z\text{-ppm} &= 0.0050/16{,}697.126 \times 1{,}000{,}000 = 0.30
\end{aligned}$$

5. Check the computed values for the ppm against a known standard. Typically, the FGCS standard given in Section 14.5.1 is used.

14.5.3 Analysis of Repeat Baseline Measurements

Another procedure employed in evaluating the consistency of the observed data and in weeding out blunders is to make repeat observations of certain baselines. These repeat measurements are taken in different observing sessions and the results compared. Significant differences in repeat baselines indicate problems with field procedures or hardware. For example in the data of Table 14.6, baselines AF and BF were repeated. Table 14.7 gives comparisons of these observations using the same procedure that was used in Section 14.5.2. Column (1) lists the baseline vector components to be analyzed, columns (2) and (3) give the repeat baseline vector components, column (4) lists the absolute values of the differences in these two observations, and column (5) gives the computed ppm values that are computed similar to the procedure given in Step 4 of Section 14.5.2.

14.5.4 Analysis of Loop Closures

Static surveys consist of many interconnected closed loops typically. For example in the network of Figure 14.7, points $ACBDEA$ form a closed loop. Similarly, $ACFA$, $CFBC$, $BDFB$, etc. are other closed loops. For each closed loop, the algebraic sum of the ΔX components should equal zero. The same condition should

TABLE 14.7 ANALYSIS OF REPEAT BASELINE OBSERVATIONS

(1) Component	(2) First Observation	(3) Second Observation	(4) Difference	(5) ppm
ΔX_{AF}	1116.4577	1116.4523	0.0054	0.84
ΔY_{AF}	4596.1553	4596.1610	0.0057	0.88
ΔZ_{AF}	4355.9141	4355.9062	0.0079	1.23
ΔX_{BF}	−6567.2310	−6567.2311	0.0001	0.01
ΔY_{BF}	−5686.3033	−5686.2926	0.0107	1.00
ΔZ_{BF}	−6322.3807	−6322.3917	0.0110	1.02

exist for the ΔY and ΔZ components. An unusually large closure within any loop will indicate that either a blunder or a large error exists in one (or more) of the baselines of the loop. It is important not to include any trivial baselines (see Section 14.3.4) in these computations since they can yield false accuracies for the loop.

To compute loop closures, the baseline components are added algebraically for that loop. For example, the closure in X components for loop $ACBDEA$ is

$$cx = \Delta X_{AC} + \Delta X_{CB} + \Delta X_{BD} + \Delta X_{DE} + \Delta X_{EA} \tag{14.3}$$

where cx is the loop closure in X coordinates. Similar equations apply for computing closures in Y and Z coordinates. Substituting numerical values into Equation (14.3), the closure in X coordinates for loop $ACBDEA$ is

$$\begin{aligned} cx &= 11{,}644.2232 - 3960.5442 - 11{,}167.6076 - 1837.7459 + 5321.7164 \\ &= 0.0419 \text{ m} \end{aligned}$$

Similarly, closures in Y and Z coordinates for that loop are

$$\begin{aligned} cy &= 3601.2165 + 6681.2467 - 394.5204 - 6253.8534 - 3634.0754 \\ &= 0.0140 \text{ m} \end{aligned}$$

$$\begin{aligned} cz &= 3399.2550 + 7279.0148 - 907.9593 - 6596.6697 - 3173.6652 \\ &= -0.0244 \text{ m} \end{aligned}$$

For evaluation purposes, loop closures are expressed in terms of the ratios of resultant closures to the total loop lengths. They are given in ppm. For any loop, the resultant closure is

$$lc = \sqrt{cx^2 + cy^2 + cz^2} \tag{14.4}$$

where lc is the length of the misclosure in the loop.

Using the values previously determined for loop $ACBDEA$, the length of the misclosure is 0.0505 m. The total length of a loop is computed by summing its legs, each leg being computed from the square root of the sum of the squares of its observed ΔX, ΔY, and ΔZ components. For loop $ACBDEA$, the total loop length is 50,967 m and the closure ppm ratio is therefore (0.0505/50,967) 1,000,000 = 0.99 ppm. Again these ppm ratios can be compared against values given in the FGCS guidelines (Table 14.4) to determine if they are acceptable for the order or accuracy of the survey. As was the case with repeat baseline observations, the FGCS guidelines also specify other criteria that must be met in loop analyses besides the ppm values.

For any network, enough loop closures should be computed so that every baseline is included within at least one loop. This should expose any large blunders that exist. If a blunder does exist, its location often can be determined through additional loop closure analyses. For example, assume that the misclosure of loop *ACDEA* reveals the presence of a blunder. By also computing the closures of loops *AFCA*, *CFDC*, *DFED*, and *EFAE*, the exact baseline containing the blunder can be detected. In this example, if a large misclosure were found in loop *DFED* and all other loops appeared to be blunder free, the blunder would be in line *DE*, because that leg was also common to loop *ACDEA*, which contained a blunder as well.

14.5.5 Baseline Network Adjustment

After the individual baselines are computed, a least-squares adjustment (see Section 16.8) of the observations is performed. This adjustment software is available from the receiver manufacturer and will provide final station coordinate values and their estimated uncertainties. If more than two receivers are used in a survey, trivial baselines will be computed during the single baseline reduction. These trivial baselines should be removed before the final network adjustment.

The observations used in the baseline network adjustment should be part of an interconnected network of baselines. Initially, a minimally constrained adjustment should be performed (see Section 16.11). The adjustment results should be analyzed both for mistakes and large errors. As an example, antenna height mistakes, which are not noticeable during a single baseline reduction, will be noticeable during the network adjustment. After the results of a minimally constrained adjustment are accepted, a fully constrained adjustment should be performed. During a fully constrained adjustment, all available control is added to the adjustment. At this time, any scaling problems between the control and the observations will become apparent by the appearance of overly large residuals. Problems that are identified should be corrected and removed before the results are accepted.

Since these computations are performed in a geocentric coordinate system, the final adjusted values can be transformed into a geodetic coordinate system using procedures as outlined in Section 13.4.3, or into a plane coordinate system (see Chapter 20). Recall that geodetic elevations are measured from the ellipsoid and thus, as discussed in Section 13.4.3, the geoid height must be applied to these heights to derive orthometric elevations. The GEOID09 software can be used to determine the geoid height. This model is included in the software typically, and the user needs only to load the appropriate data file from the NGS for their region.

Finally the horizontal and vertical accuracy of the survey can be determined based on the FGCS or NGS horizontal and vertical accuracy standards (see Section 14.5.1).

14.5.6 The Survey Report

A final survey report is helpful in documenting the project for future analysis. At a minimum, the report should contain the following items.

1. The location of the survey and a description of the project area. An area map is recommended.
2. The purpose of the survey and its intended specifications.
3. A description of the monumentation used including the tie sheets, photos, and rubbings of the monuments.
4. A thorough description of the equipment used including the serial numbers, antenna offsets, and the date the equipment was last calibrated.
5. A thorough description of the software used including name and version number.
6. The observation scheme used, including the itinerary, the names of the field crew personnel, and any problems that were experienced during the observation phase.
7. Computation scheme used to analyze the observations and the results of this analysis.
8. A list of the problems encountered in the process of performing the survey, or its analysis including unusual solar activity, potential multipathing problems, or other factors that can affect the results of the survey.
9. An appendix containing all written documentation, original observations, and analysis. Since the computer can produce volumes of printed material in a typical survey, only the most important files should be printed. All computer files should be copied onto some safe backup storage. A CD or DVD provides an excellent permanent storage media that can be inserted into the back of the report.

■ 14.6 SOURCES OF ERRORS IN SATELLITE SURVEYS

As is the case in any project, observations are subject to instrumental, natural, and personal errors. These are summarized in the following subsections.

14.6.1 Instrumental Errors

Clock Bias. As mentioned in Section 13.6.1, both the receiver and satellite clocks are subject to errors. They can be mathematically removed using differencing techniques for all forms of relative positioning.

Setup Errors. As with all work involving tripods, the equipment must be in good adjustment (see Section 8.19). Careful attention should be paid to maintaining tripods that provide solid setups, and tribrachs with optical plummets that will center the antennas over the monuments. In GNSS work, tribrach adapters are often used that allow the rotation of the antenna without removing it from the tribrach. If these adapters are used, they should be inspected for looseness or "play" on a regular basis. Because of the many possible errors that can occur when using a standard tripod, special fixed-height tripods and rods are often used. The fixed-height rods can be set up using either a bipod or tripod with a rod on the point. They typically are set to a height of precisely 2 m from the antenna reference point (ARP).

Nonparallelism of the Antennas. Pseudoranges are observed from the phase center of the satellite antenna to the phase center of the receiver antenna. As with EDM observations, the phase center of the antenna may not be the geometric center of the antenna. Each antenna must be calibrated to determine the phase center offsets for both the $L1$ and $L2$ bands. For antennas, with phase center offsets, the antennas are aligned in the same direction. Generally, they are aligned according to local magnetic north using a compass.

Receiver Noise. When working properly, the electronics of the receiver will operate within a specified tolerance. Within this tolerance, small variations occur in the generation and processing of the signals that can eventually translate into errors in the pseudorange and carrier-phase observations. Since these errors are not predictable, they are considered as part of the random errors in the system. However, periodic calibration checks and tests of receiver electronics should be made to verify that they are working within acceptable tolerances.

14.6.2 Natural Errors

Refraction. Refraction due to the transit of the signal through the atmosphere plays a crucial role in delaying the signal from the satellites. The size of the error can vary from 0 to 10 m. Dual-frequency receivers can mathematically model and remove this error using Equations (13.18) and (13.19). With single-frequency receivers, this error must be modeled. For surveys involving small areas using relative positioning methods, the majority of this error will be removed by differencing. Since high solar activity affects the amount of refraction in the ionosphere, it is best to avoid these periods.

Relativity. GNSS satellites orbit the Earth in approximately 12 h. The speed of the satellites causes their atomic clocks to slow down according to the theories of relativity. The master control station computes corrections for relativity and applies these to the clocks in the satellites.

Multipathing. Multipathing occurs when the signal emitted by the satellite arrives at the receiver after following more than one path. It is generally caused by reflective surfaces near the receiver. As discussed in Section 13.6.3, multipathing can become so great that it will cause the receiver to lose lock on the signal. Many manufacturers use signal filters to reduce the problems of multipathing. However, these filters will not eliminate all occurrences of multipathing, and are susceptible to signals that have been reflected an even number of times. Thus, the best approach to reducing this problem is to avoid setups near reflective surfaces. Reflective surfaces include flat surfaces such as the sides of building, vehicles, water, and chain link fences.

14.6.3 Personal Errors

Tripod Miscentering. This error will directly affect the final accuracy of the coordinates. To minimize it, check the setup carefully before data collection begins and again after it is completed.

14.7 MISTAKES IN SATELLITE SURVEYS

A few of the more common mistakes in GNSS work are listed here.

Misreading of the Antenna Height. The height from the ground to the antenna ground plane or reference point should be read at several times. When measuring to a ground plane, it should be measured at several locations around the ground plane and the average recorded. To ensure that the tripod hasn't settled during the observation process and that the initial readings were correct, the slant height should be also measured at the end of the observation process. To avoid this problem, only fixed-height tripods should be used in the most precise surveys.

Incorrect Identification of the Station. This mistake can cause hours of wasted time in processing of the data, and will sometimes require that the survey be repeated. To limit this possibility, each station should be located from at least four readily visible permanent objects. Also if possible, rubbings of the monument caps should be obtained and photos taken of the area showing the location of the monument. During the observation phase of the survey, a second rubbing of the monument and a photo of the setup showing the surrounding area should be taken. This data should be correlated before baseline processing.

Processing of Trivial Baselines. This mistake can only occur when more than two receivers are used in a survey. While this mistake will not generate false coordinates, it will generate false accuracies for the survey. Care should be taken to remove all trivial baselines before a network adjustment is attempted.

Misidentification of Antennas. Since each antenna type has different phase center offsets, misidentifying an antenna will directly result in an error in the derived pseudorange. Using antennas from only one manufacturer can reduce this mistake or by correctly identifying and entering phase center offsets into the processing software.

PROBLEMS

Asterisks (*) indicate problems that have partial answers given in Appendix G.

14.1 Explain the differences between a static survey and rapid static survey.

14.2 When using the rapid static surveying method, what is the minimum recommended length of the session required to observe a baseline that is 10 km long for ***(a)** a dual-frequency receiver, **(b)** a single-frequency receiver?

14.3 What would be the recommended epoch rates for the surveys given in Problem 14.2?

14.4 For a 38-km baseline using a dual-frequency receiver, **(a)** what static surveying method should be used, **(b)** for what time period should the baseline be observed, and **(c)** what epoch rate should be used?

14.5 What variables affect the accuracy of a static survey?

14.6 Why are dual-frequency and GNSS receivers preferred for high-accuracy control stations?

14.7 Explain why canopy obstructions are a problem in a static survey.

14.8 Why is it recommended to use a precise ephemeris when processing a static survey?

14.9 What are the recommended rates of data collection in a **(a)** static survey, and a ***(b)** rapid static survey?

14.10 List the fundamental steps involved in planning a static survey.

14.11* How many nontrivial baselines will be observed in one session with four receivers?

14.12 What variables should be considered when selecting a site for a static survey?

14.13 Why should a control survey using static methods form closed geometric figures?

14.14 What is the purpose of an obstruction diagram in planning a static survey?

14.15 Explain why periods of high solar activity should be avoided when collecting satellite observations.

14.16 Why the survey vehicle should be parked at least 25 m from the observing station in a static survey?

14.17 When can a rapid static horizontal control survey with high PDOP continue?

14.18 When using four receivers, how many sessions will it take to independently observe all the baselines of a hexagon?

14.19 Plot the following ground obstructions on a obstruction diagram.
(a) From an azimuth of 65° to 73° there is a building with an elevation of 20°.
(b) From an azimuth of 355° to 356° there is a pole with an elevation of 35°.
(c) From an azimuth of 125° to 128° there is a tree with an elevation of 26°.

14.20* In Problem 14.19, which obstruction is unlikely to interfere with GPS satellite visibility in the northern hemisphere?

14.21 What items should be considered when deciding which method to use for a GPS survey?

14.22 What items should be included in a site log sheet?

14.23 What is a satellite availability chart and how is it used?

14.24* What order of accuracy does a survey with a standard deviation in the geodetic height difference of 15 mm between two control stations that are 5 km apart meet?

14.25 Do Problem 14.24 when the standard deviation in the geodetic height difference is 5 mm for two control points 15 km apart?

14.26 Use the NGS website to download the station coordinates for the nearest CORS station.

14.27 What are CORS and HARN stations?

14.28 Why should repeat baselines be performed in a static survey?

14.29 What is the purpose of developing a site log sheet for each session?

14.30 Using loop *ACFDEA* from Figure 14.7, and the data from Table 14.6, what is the
(a) Misclosure in the X component?
(b) Misclosure in the Y component?
(c) Misclosure in the Z component?
(d) Length of the loop misclosure?
(e) *Derived ppm for the loop?

14.31 Do Problem 14.30 with loop *BCFB*.

14.32 Do Problem 14.30 with loop *BFDB*.

14.33 List the contents of a typical survey report.

14.34 The observed baseline vector components in meters between two control stations is (3814.244, −470.348, −1593.650). The geocentric coordinates of the control stations are (1,162,247.650, −4,655656.054, 4,188,020.271) and (1,158,433.403, −4,655,185.709, 4,189,613.926). What are

(a)* ΔX ppm?

(b) ΔY ppm?

(c) ΔZ ppm?

14.35 Same as Problem 14.34 except the two control station have coordinates in meters of (−1,661,107.767, −4,718,275.246, 3,944,587.541) and (1,691,390.245, −4,712,916.010, 3,938,107.274), and the baseline vector between them was (30282.469, −5359.245, 6480.261).

14.36 List the various survey types that could be performed using static survey.

14.37 Employ the user-friendly button in the NGS CORS Internet site at http://www.ngs.noaa.gov/CORS/ to

(a) Download the navigation and observation files for station PSU1 between the hours of 10 and 11 local time for the Monday of the current week using a 5-sec data collection rate.

(b) Print the files and comment on the contents of them. (*Hint*: An explanation of the contents of the RINEX2 data file is contained at http://www.ngs.noaa.gov/CORS/Rinex2.html on the Internet in the CORS site.)

BIBLIOGRAPHY

Denny, M. 2002. "Surveying Little Egypt." *Point of Beginning* 27 (No. 8): 26.

Devine, D. 2002. "Mapping the CSS Hunley." *Professional Surveyor* 22 (No. 3): 6.

Fotopoulos, G. et al. 2003. "How Accurately Can We Determine the Orthometric Height Differences from GPS and Geoid Data?" *Journal of Surveying Engineering* 129 (No. 1): 1.

Hartzheim, P. 2002. "No Roads Untraveled—How GPS Has Eased the Tasks of the Wisconsin Department of Transportation." *Point of Beginning* 27 (No. 12): 14.

Henstridge, F. 2001. "The National Height Modernization Program." *Professional Surveyor* 21 (No. 6): 54.

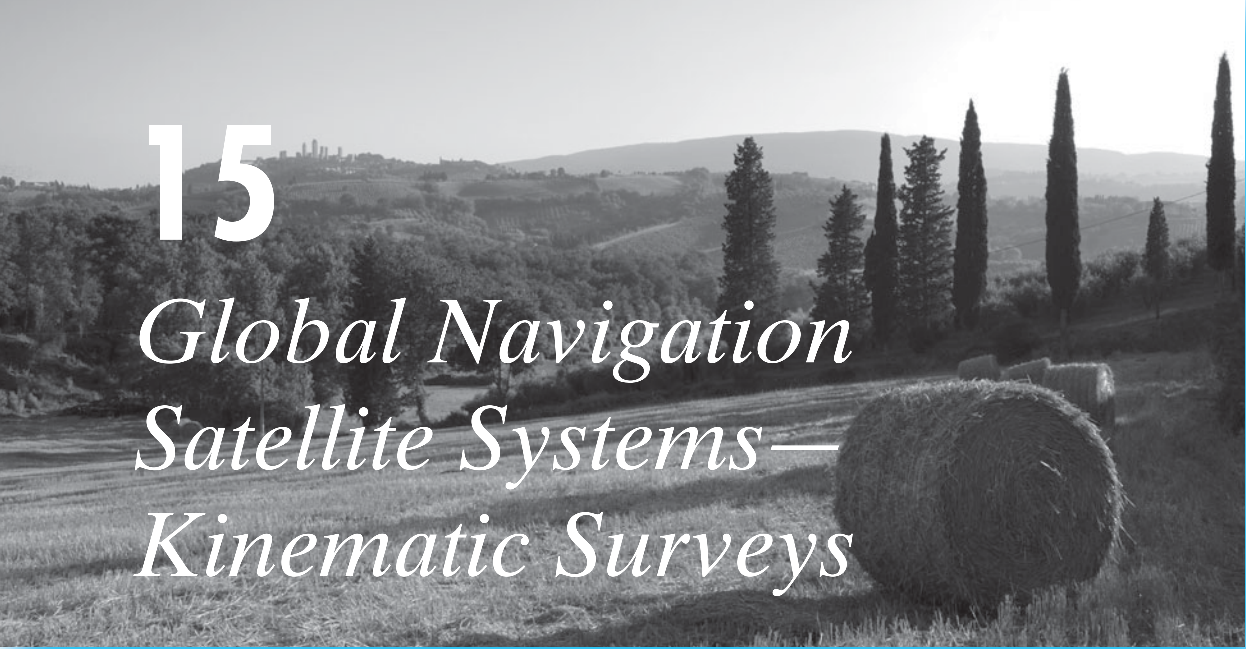
Kuang, S., et al. 2002. "GPS Control Densification Project for Illinois Department of Transportation." *Surveying and Land Information Science* 62 (No. 4): 225.

Licht, R. 2003. "A Step Ahead—Employees of a Minnesota Firm Take GPS One Step Further with the Application of Cell Phones." *Point of Beginning* 28 (No. 12): 32.

Mader, G. L., et al. 2003. "NGS Geodetic Tool Kit, Part II: The On-Line Positioning User Service (OPUS)." *Professional Surveyor* 23 (No. 5): 26.

Steinberg, G. and Even-Tzur, G. 2008. "Official GNSS-Derived Vertical Orthometric Height Control Networks." *Surveying and Land Information Science* 68 (No. 1): 29.

This page intentionally left blank



15 Global Navigation Satellite Systems—Kinematic Surveys

■ 15.1 INTRODUCTION

In many areas of surveying, speed and productivity are essential elements to success. In satellite surveying, the most productive form of surveying is kinematic surveying. It uses relative positioning techniques with carrier-phase observations as discussed in Sections 13.5.2 and 13.8. These surveys can provide immediate values to the coordinates of the points while the receiver is stationary or in motion. Its accuracy is typically less than that obtainable with static surveys, but is adequate for most forms of surveys. It has applications in many areas of surveying including mapping (Chapter 17), boundary (Chapters 21 and 22), construction (Chapters 23 and 24), and photogrammetry (Chapter 27). This chapter will look at both the post processed and real-time kinematic surveying methods.

Kinematic surveying can provide immediate results using the *real-time kinematic* (RTK) mode or in the office using the *post-process kinematic* (PPK) mode. Kinematic surveying provides positioning while the receiver is in motion. For example, kinematic surveys have been successfully used in positioning sounding vessels during hydrographic surveys (Section 17.13) and aerial cameras during photogrammetric surveys (Section 27.16). In large construction projects, it is used in machine control to guide earthwork operations. It is also useful in nonsurveying applications such as high-precision agriculture.

It shares many commonalities with static surveys. For example, a kinematic survey requires two receivers collecting observations simultaneously from a pair of stations with one receiver, the base, occupying a station of known position and another, the rover, collecting data on points of interest. It also uses relative positioning computational procedures similar to those used in static surveys. Thus, it requires that the integer ambiguities (see Section 13.5.2) be resolved before the survey is started. The main difference between static and kinematic surveying techniques is the length of time per session. In a kinematic survey, observations from a single epoch may be all that is used to determine position of the roving receiver. Establishment of control points using a static surveying method requires much longer sessions than are typically used in kinematic surveys.

As previously stated, the accuracy of kinematic surveys does not match that of static surveys typically. Some of the limiting factors are the lack of repeated observations and the length of the session. For example, a rapid static survey may use a 5-sec epoch rate to sample data over a 1-h session. This results in a total of 720 sets of observations per satellite. During the observation period, additionally, the satellite geometry also changes creating different solution geometries. The combination of a large set of observations with varying satellite geometry results in a better solution for the receiver coordinates. In a kinematic survey, the receiver may collect 600 observations per satellite using a 1-sec epoch rate over a 10-min interval. However, since the satellite geometry does not change significantly, the solution is often weaker than the static survey. Another accuracy degrading factor in RTK surveys includes the motion of the rover during data processing. Since the observations from the base receiver must be transmitted, received, and processed at the rover, any motion by the rover during this time will cause errors in its computed position. Since the motion of the rover is often small, the errors caused by this time difference, known as *data latency*, tend to be small and are generally adequate for the lower-accuracy surveys previously cited. However, they can be significant in cases involving fast moving rovers such as in the positioning of a camera station during a photogrammetric mission. Other factors that limit the positioning accuracy of kinematic surveys are its susceptibility to errors such as DOP spikes, atmospheric and ionospheric refraction, multipathing, and obstructions to satellite signals. Often the effect of these factors can be minimized with careful planning or by advanced processing techniques.

15.2 PLANNING OF KINEMATIC SURVEYS

All too often kinematic surveys are performed with little or no preplanning. Inevitably, when this is done an occasional survey will produce poor results. This happens due to several reasons including many of those discussed in Section 14.3. However, kinematic surveys are particularly vulnerable to poor observation conditions due to the relatively low number of observations taken at any location and the small changes that occur in satellite geometry at the time of these observations. For example, if a site has canopy problems or is susceptible to multipathing, a weak survey with poor results can occur. Since kinematic surveys are the most productive, they are used predominately in practice. Thus, the National Geodetic Survey is now in the process of developing guidelines for performing kinematic surveys.[1]

A typical point located using kinematic surveying methods may have as few as 1 epoch to a few minutes of observational data. Thus, canopy restrictions, solar activity, multipathing, DOP spikes, and many other sources of error can have drastic

[1]*National Geodetic Survey User Guidelines for Classical Real Time GNSS Positioning* by William Henning is available at http://www.ngs.noaa.gov/PUBS_LIB/NGSRealTimeUserGuidelines.v2.0.4.pdf.

effects on the determined locations of the receiver. For example, if a kinematic survey had been performed during the periods of DOP spikes shown in Figure 14.4, the resultant coordinates of these points would show larger errors when compared to others measured during periods of low PDOP. While a static survey is generally of sufficient length to "survive" the typical DOP spikes seen today, a kinematic survey is extremely vulnerable to them. Thus, it is important to be aware of the periods of DOP spikes and avoid them when performing a kinematic survey.

In addition to the additional noise that multipathing can cause in a receiver, it can also cause problems in the resolution of the integer ambiguities for each observed satellite. Multipathing is cyclical and can be modeled in the longer sessions typically present in static surveys by the post-processing software. However, the short duration of the typical kinematic session prevents similar modeling in a kinematic survey. A receiver in multipathing conditions will continue to display precise results even though the opposite is true. Thus, the base station for a kinematic survey should never be placed in a location that is susceptible to multipathing, and the rover should similarly avoid these conditions. Tall buildings, trees, fencing, vehicles, poles, other similar reflective objects should be avoided. These objects can typically be located using offset procedures found in most of today's survey controllers.

Except when using real-time networks (see Section 15.8), the software used in kinematic surveys assumes that both base and roving receivers are in the same atmospheric conditions. Thus, baselines should be kept under 20 km and surveys should be suspended when the base and roving receiver are not in similar conditions such as when a storm front moves through the project area.

Refraction caused by the free electrons in the ionosphere and by weather conditions in the troposphere can adversely affect the positioning results in a kinematic survey. During periods of high solar activity, the errors due to ionospheric refraction can be large. Since the ionosphere will remain charged for extensive periods of time, there will be some days when a satellite survey simply should not be attempted. In periods of very high solar activity, radio signals from the satellites can be interrupted. Additionally during these periods, radio communication between the base and roving receivers in an RTK survey can be interrupted. The National Oceanic and Atmospheric Administration (NOAA) Space Weather Prediction Center[2] provides forecasts for solar activity and its effect on radio communications. You can receive automatic updates from the Space Weather Center by registering with them. In particular, users should monitor geomagnetic storms, solar radiation storms, and radio blackouts. Geomagnetic storms may cause satellite orientation problems, increasing broadcast ephemeris errors, satellite communication problems and can lead to problems in initialization. Solar radiation storms may also create problems with satellite operations, orientation, and communications, which can cause increased noise at the receiver resulting in degraded positioning accuracy. Radio blackouts can cause intermittent loss of satellite and radio communications, which can increase noise at the receiver degrading positional accuracy. These are identified on the NOAA

[2]Space weather forecasts are available at http://www.swpc.noaa.gov/NOAAscales/index.html.

website in five categories from mild to extreme. In general, a satellite survey should not be attempted when any of the three is rated in the range from strong to extreme. It should be noted that in a post processed kinematic (PPK) survey as with static surveys, the effects of geomagnetic storms creating ephemeral errors can be alleviated by using one of the available precise ephemerides during processing of the observations. Additionally in a PPK survey, a radio link between the base and rover is not required, and thus radio blackouts are not a problem.

Of course, the equipment used in any survey should be calibrated. For example, fixed height poles and tripods should be checked for accuracy by measuring from the tip of the pole to the mounting plate of the receiver, which is also known as the antenna reference point (ARP). Poles should be checked for straightness and legs of tripods checked for tightness. Additionally, circular bubbles should be regularly checked to ensure that they are in adjustment.

■ 15.3 INITIALIZATION

To start a kinematic survey, the receivers must be *initialized*. This process includes determining the integer ambiguity (see Section 13.5.2) for each pseudorange observation. Following any of the methods described below can yield initialization of the receivers.

One procedure for initializing the receivers uses a baseline whose ΔX, ΔY, and ΔZ components are known. A very short static observing session is conducted with base and roving receivers occupying two stations with known positions simultaneously. Because the baseline coordinate differences are known, differencing of the observations will yield the unknown integer ambiguities. These differencing computations are performed in a post-processing operation using the data from both receivers. If only one control station is available, a second one can be set using the static or rapid static surveying methods described in Section 14.2.

An alternative initialization procedure, called *antenna swapping*, is also suitable if only one control station is available. Here receiver A is placed on the control point and receiver B on a nearby, unknown point. For convenience, the unknown point can be within 30 ft (10 m) of the control station. After a few minutes of data collection with both receivers, their positions are interchanged while keeping them running. In the interchange process, care must be exercised to make certain continuous tracking, or "lock" is maintained on at least four satellites. After a few more minutes of observations, the receivers are interchanged again, returning them to their starting positions. This procedure enables the baseline coordinate differences and the integer ambiguities to be determined, again by differencing techniques.

Finally, the most advanced techniques of initialization are known as *on-the-fly* (OTF) ambiguity resolution methods. These methods require five usable satellites during the initialization process and dual-frequency receivers. OTF, which involves the solution of a sophisticated mathematical algorithm, has resolved ambiguities to the centimeter level in 2 min for a 20-km line. However, longer sessions are sometimes necessary to resolve the ambiguities since ideal conditions are not always available. The typical period for ambiguity resolution is usually less than five minutes. It is not uncommon with current processing techniques to resolve ambiguities in under a minute. As mentioned in Section 13.11, when four satellites with the L5 frequencies are available, the ambiguities can be mathematically determined in a single epoch of data eliminating the need for the previous methods discussed herein.

15.4 EQUIPMENT USED IN KINEMATIC SURVEYS

The operator's body can be an obstruction when performing a kinematic survey. Thus, the antenna is often mounted on a fixed-height rod that is 2 m in length to avoid operator obstructions. Similarly, as shown in Figure 15.1, the base receiver is often mounted on fixed-height tripods. In any case, the advantages of fixed-height rods and tripods in all GNSS surveys are that they minimize measurement errors in the height of the receiver and help avoid operator caused obstructions.

Other equipment used in kinematic surveys includes traditional adjustable-height tripods and poles. However, the adjustments in these can often lead to errors in the measured heights to the antenna. Another factor to consider with traditional tripod equipment is the need of a tribrach for mounting of the antenna. As with traditional equipment, when tribrachs are used, it is extremely important to check the adjustment of the optical plummets (see Section 8.19.4). Similarly when using either fixed height or adjustable rods, it is important to regularly check and adjust the level bubbles (see Section 8.19.5).

For horizontal positioning, setup errors due to misleveling of the bubble can be minimized by using lower setups. The amount of setup error can be determined using Equation (8.1). For example, a fixed height tripod set at 2.000 m with

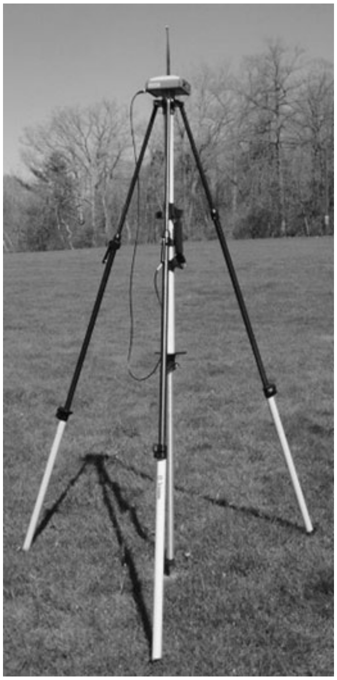


Figure 15.1
A GNSS receiver mounted on a fixed height tripod.

a misleveling of 2 min will cause a horizontal positioning error of 1.1 mm. At a height of 1.500 m, this same misleveling will result in horizontal positioning error of 0.9 mm. Since both of these errors are under the error achievable by satellite surveys, they are typically ignored. However, in kinematic surveys where the roving pole is often carried in a less careful manner, these errors can be significant. For example, a 2.000-m pole held within 5 min of level will result in a horizontal positioning error of 3 mm. This error is approaching the achievable precision in satellite surveying and thus leads to higher positional errors in kinematic surveys when compared to static methods. Thus, many fixed height tripods and some poles have several set positions for the mounting of the receiver. Typical settings are 1.500 m, 1.800 m, and 2.000 m. Fixed length extensions can be added to increase the height of the receiver; however, as the height of the receiver increases, the amount of error in horizontal positioning may also increase.

An analysis of error in misleveling on the derived height of a point can be done using simple trigonometry. For example, a 2.000-m pole that is held within 5 min of level will introduce a height error of only 0.002 mm $(2 - 2\cos 5')$. However, in some kinematic surveying work the tip of the pole is carried off the ground. This additional error is introduced into the derived heights. Even though these errors are greater than what is expected from a static survey, they are well below what is typically needed for the types of surveys where kinematic surveying is appropriate.

In kinematic surveys, overhead obstructions should be avoided at the base and rover stations. Additionally, the base receiver location should be free of reflective objects such as buildings, fences, and vehicles. Manufacturers sell 100-ft cables that enable the user to locate the vehicle away from the base receiver to avoid potential multipathing problems from the vehicle.

For the most RTK surveys, the radio antenna at the base receiver is often mounted on a nearby tripod. It is important to have the radio antenna vertical to match orientation of the antenna on the rover. Mounting the radio antenna high can increase the range of the base radio. However, repeater stations can also be used to extend the range of the base station radio in situations where necessary, as well as avoid obstructions.

Several factors may determine the "best" location for the base station in a RTK survey. Since the range of the radio can be increased with increasing height of the radio antenna, it is advantageous to locate the base station on a local high point. As previously stated, the base station should be located in an area that is free of multipathing conditions. Additionally, since the base station in an RTK survey requires the most equipment, it is also preferable to place the base station in an easily accessible location. The radios in an RTK survey are low power. Thus, it is wise to avoid sources of high electromagnetic activity such as power grid substations, high-tension power lines, or buildings containing large electric motors since these items generate substantial electromagnetic fields that can disrupt radio transmissions. Furthermore, the radio signals can interfere with the receiver antenna. Thus, the radio antenna should be placed a few meters from the receiver antenna.

Other equipment needed for RTK surveys include a radio and its power source. Typically, an external battery is used for the larger base receiver radio. When planning an RTK survey, it is important to provide some backup source of power for the inevitable event of a dead battery. A vehicle can serve as source for charging batteries and providing power for base station radios. Finally, a survey controller is required to control the collection of data from both the base and roving receivers.

■ 15.5 METHODS USED IN KINEMATIC SURVEYS

After initialization has been completed using one of the above techniques, the base receiver remains at the control station while the rover moves to collect data on features. Although the rover's positions can be determined at intervals as short as 0.2 sec, 1-sec epoch rates are typically used in kinematic surveys. The precision of intermediate points is in the range of $\pm(1 - 2 \text{ cm} + 2 \text{ ppm})$. In kinematic surveys, both receivers must maintain lock on at least four satellites throughout the entire session. If lock is lost, the receivers must be reinitialized. Thus, care must be taken to avoid obstructing the rover's antenna by carrying it close to buildings, beneath trees or bridges, shielding it with the operator's body, and so on. At the end of the survey, the rover should be returned to its initial control station, or another, as a check.

Kinematic surveys generally follow two forms of data collection. In *true kinematic* mode, data is collected at a specific rate. This method is useful for collecting points along an alignment, or grade elevations for topographic surveys. An alternative to the true kinematic mode is to stop for several epochs of data at each point of interest. This method, known as *semikinematic* or the *stop-and-go* mode, is useful for mapping and construction surveys where increased accuracy is desired for a specific feature. In the semikinematic mode, the antenna is positioned over points of interest and a point identifier is entered into the survey controller for each feature. Since multiple epochs of data are usually recorded at each point, the accuracy of this mode is greater than that obtainable in the true kinematic mode. In both surveys, the rate of data collection at the base station and rover is typically set to 1 sec.

As discussed in Section 15.2, in kinematic surveys the rover is never at a station long enough to survive a PDOP spike. For kinematic surveys, PDOP values should be less than four. Additionally, since high free electron counts in the ionosphere can affect satellite signals greatly, it is important to collect data only during periods of low solar activity. During periods of high solar activity, poor positioning results can be obtained and communications to the receiver and rover can be disrupted. Both high PDOP values and high solar activity periods can be avoided with careful project planning.

In post-processed kinematic (PPK) surveys, the collected data are stored on the survey controller or receiver until the fieldwork is completed. The data are then processed in the office using the same software and processing techniques used in static surveys. Data latency is not a problem in PPK surveys since the data is post processed. Other advantages of PPK surveys are that (1) precise ephemeris can be combined with the observational data to remove errors in the broadcast ephemeris and (2) the base station coordinates can be resolved after the fieldwork is complete. Thus, the base station's coordinates do not have to be known prior to the survey. The lack of data latency and use of a precise ephemeris results in PPK surveys having slightly higher accuracies than those obtainable from RTK surveys.

As discussed in Section 13.8, RTK surveying, as implied by its name, enables positions of points to be determined instantaneously as the rover occupies a

Figure 15.2
Note the antenna used in this stop-and-go RTK survey. (Courtesy Leica Geosystems AG.)

point. Like the other kinematic methods, RTK surveying requires that two (or more) receivers be operated simultaneously. The unique aspect of this procedure is that a radio is used to transmit the base receiver coordinates and its observations to the rover. At the rover, the observations from both receivers are processed in real time by the unit's on-board computer to produce a nearly immediate determination of its location according to Equation (13.27). Like PPK surveys, the processing techniques are similar to those used in static surveys. However, the epoch rate for data collection is typically set to 1 sec. Figure 15.2 shows a stop-and-go RTK survey in progress.

As shown in Figure 15.3, RTK surveys require compatible hardware at each end of the radio link. Normally, this equipment is purchased from one

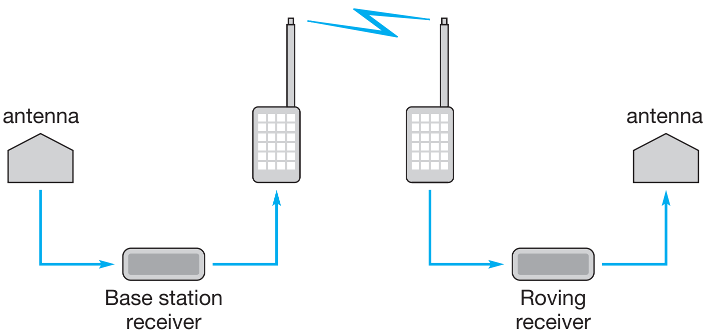


Figure 15.3
A base receiver and rover with compatible internal radios used in RTK surveys.

manufacturer. In North America and in other areas of the world, frequencies in the range of 150 to 174 MHz in the VHF radio spectrum, and from 450 to 470 MHz in the UHF radio spectrum can be used for RTK transmissions. Typically, the messages are updated at the rover every 0.5 to 2 sec. The data link for RTK receiver requires a minimum of 2400 baud or higher for operation. However, it is typically much higher with a baud rate of 38,400.

Even with higher baud rates, the application of the base receiver data to the rover data is delayed because of delays in transmitting the base receiver's observations and position to the rover and the additional time required to compute the rover's position. Typically, the data latency is between 0.05 and 1.0 sec. Data latency plays a role in the final accuracies of derived positions. However, these problems have been minimized since selective availability was turned off. Even with selective availability, these errors tended to be small and did not significantly affect the final quality of most surveys.

The radio link used with RTK can limit the distance between the base receiver and rover(s) to under 10 km, or about 6 mi. This distance can be increased with more powerful transmitters, or through the use of repeater stations as shown in Figure 15.4. A repeater station receives the signal from a transmitter

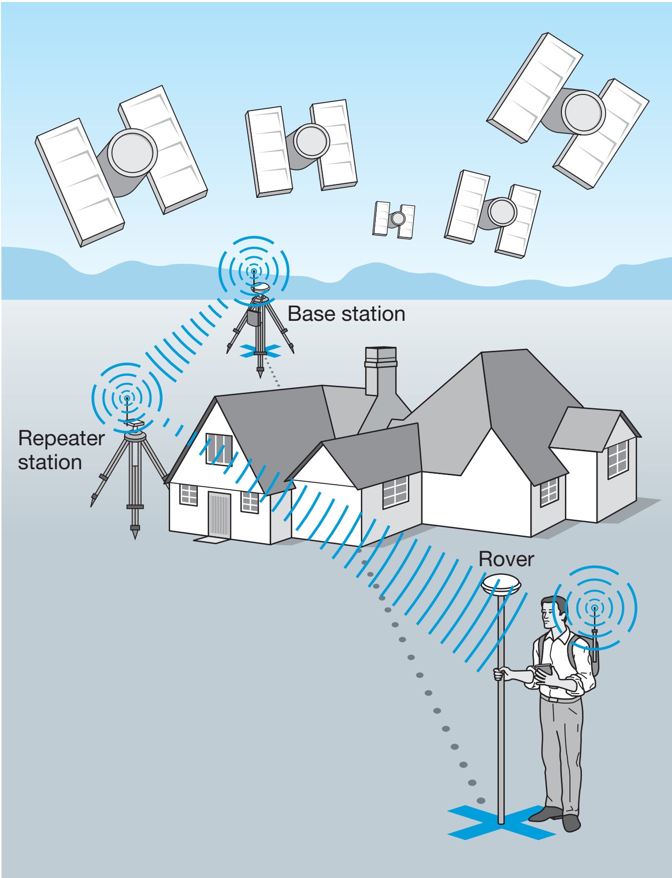


Figure 15.4 Use of a repeater radio to work around obstructions. (Courtesy Ashtech, Inc.)

such as the base radio and retransmits it. Some transmitters require a Federal Communication Commission (FCC) license to broadcast the data. With low-power radios, a line of sight between the transmitter and the receiver is required. An advantage of repeater stations is that they can be used to survey around obstructions and increase the range of the base radio. In areas where cellular coverage is available, data modems can also be used to broadcast data from the base receiver to the rover.

The advantages of RTK surveys over PPK surveys are the reduction in office time and the ability to verify observations in the field. When using RTK, the data can also be downloaded immediately into a GIS (see Chapter 28) or an existing surveying project. This increases the overall productivity of the survey.

15.6 PERFORMING POST PROCESSED KINEMATIC SURVEYS

A PPK survey requires a base receiver that is collecting data at the same epoch rate as the rover. The base receiver is usually set over a reference station established from a prior static survey. If a local CORS station is collecting data at the same rate, it can also be used as a base receiver. For example, if a local CORS station is gathering data at a 5-sec epoch rate and the rover is also set to a 5-sec epoch rate, the CORS station's observation files can be downloaded and used to reduce the rover's data. Some CORS stations have an epoch rate of 1 sec for this purpose.

When a local reference station is not immediately available for use as a base station, it is possible to establish the temporary coordinates of the base receiver by using its autonomous position. The autonomous position of the receiver can be derived from as little as a single epoch of data.[3] This position is determined from code-ranging and is not high in accuracy, but allows the PPK survey to continue. Then, prior to processing the roving receiver positions, accurate positioning of the base station can be achieved by post-processing the base receiver's data with an established reference network station. Ties to the base station receiver can be performed by collecting data from a local CORS station, or any other established network reference station including HARN stations or stations established by state Department of Transportations and local surveying (geomatics) engineering firms.

As an example, a CORS station is a readily available established reference control station. The National Geodetic Survey (NGS) has established a website known as *Online Processing User Service* (OPUS)[4] to perform this function for the surveyor. It uses three CORS stations near the base station receiver to determine the position of the base station receiver during the PPK survey. A CORS station has the additional advantage of not requiring another receiver to gather data at the network reference station. It is not important that data for this connection be gathered at the same epoch rate as the PPK survey. What is important is that the length of the observing session be sufficient in time to solve for the

[3]Due to the inaccuracy of the autonomous position of a GNSS receiver, it should never be used in RTK stakeout surveys.

[4]OPUS is available on the NGS website at http://www.ngs.noaa.gov/OPUS/.

connecting baseline vector. For example, the base receiver can gather data at a 1-sec epoch rate to support the PPK survey while the receiver on the network station collects data at a 5-sec epoch rate. What is important is that both receivers are collecting data in the same multiples. For example, if a receiver on the reference station is collecting data at a 3-sec epoch rate while the base station is collecting at a 2-sec rate, common data will only occur every 6 sec and would result in two out of every three base station observations being unusable. These situations should be avoided. Table 14.1 provides typical lengths of sessions for various baseline lengths. While the connection data are being collected, the rover can proceed with the PPK survey; however, the base station must collect sufficient data to determine its position later.

The NGS recommends a minimum session length of 2 h at the base receiver when using the static processing option. The static option has a maximum length of 24 h. However, the NGS provides a *rapid static service* (OPUS-RS) that can determine the position of the base station receiver with as little as 15 min of data under good observation conditions while the PPK survey is performed. After the survey is complete, the RINEX observation file (see Section 14.3.5) from the base receiver is sent to OPUS using the Internet. OPUS solves for the position of the base receiver using nearby CORS stations. Typically within a few minutes, the results of the computations are sent back to the user via e-mail. The OPUS derived position for the base station is then entered into the user's processing software to determine the locations of the rover during the simultaneous PPK survey. In order to achieve the best results, the model of the base receiver antenna and receiver antenna height in relation to the antenna reference point (see Section 13.6.3) must be known and entered with the observation file into OPUS.

PPK surveys are typically used to collect data for mapping surveys. They are especially useful for large surveys with minimal obstructions where the rover can be mounted on a vehicle. As mentioned in Section 15.5, features can be collected using the semikinematic or true kinematic mode. In areas where canopy obstructions are a problem, the semikinematic mode can be used to establish temporary, higher-accuracy reference points for later use by a total station. A minimum of two points is required. One station is the reference station for the total station while the other is its backsight which establishes the rotational position of the survey. As the new and stronger L2C and L5 signals become available, the problem caused by canopy obstructions could disappear significantly, thus removing the need for a total station to gather data in canopy-restricted areas.

When collecting data using the true kinematic mode, it is important to set a reasonable epoch rate based on the speed of motion of the rover. For example, if the rover is being hand carried, a 1-sec epoch rate would result in data being gathered every 5 or 6 ft. This is an excessive amount of data for the typical topographic survey. Furthermore, excessive data collection in lines would result in a weak triangulated irregular network. To avoid this, some survey controllers allow their users to set the data collection rate on a specified two- or three-dimensional distance. Section 17.8 discusses the importance of properly collecting data for topographic features. Section 17.12 discusses methods that are used to efficiently produce line work on maps.

As previously discussed, the receivers must be initialized before a kinematic survey can be started. Once this occurs, data collection can proceed as long as lock on four satellites is maintained. Thus, it is important to watch the number of visible satellites and the PDOP of the solution. If canopy restrictions obscure satellites that are crucial to an accurate solution, the displayed PDOP on the receiver will increase. In this event, the user needs to proceed to an area where the PDOP is sufficiently low and survey this area later with a total station. If the number of satellites drops below four, the rover must be reinitialized. The most common method of reinitialization is accomplished by moving the rover to a location where five satellites are visible. At this location, OTF will quickly reestablish lock on the satellites. OTF can reestablish lock on the satellites in less than 1 min in these situations. However, if this is not achievable, the user can move to a previously surveyed identifiable feature to reestablish lock on the satellites. Temporary control points can be established at the beginning of the survey to facilitate this solution. Since returning to a previously surveyed, identifiable feature can be time consuming, most users try to maintain lock on five or more satellites at all time and avoid situations where loss-of-lock problems can occur. As previously stated, when the number of Block IIF and III satellites becomes significant, loss-of-lock problems may disappear.

Since kinematic surveys use a small number of observations to establish the coordinates of points, a PDOP value that is less than four is recommended for most surveys. However, a value as high as 6 is acceptable for certain types of surveys—a mapping survey, for example. The user can also watch the PDOP as the survey proceeds. When weak satellite geometry is present, the PDOP value will rise. No data should be collected if the PDOP value is greater than 6. A sudden change in the PDOP value is usually caused by an obstruction that has removed a key satellite from the geometric solution of the point. As mentioned previously, when this happens the user should proceed to a location where the PDOP is reduced and continue the survey.

After the data is collected, it is loaded into the processing software. An advantage PPK surveys have over RTK surveys is that precise ephemeris can be used in the processing. As discussed in Section 13.6.3, this will result in a better solution for the positions of the surveyed points since it removes ephemeral errors from the solution. The base station coordinates should be established or entered before the rover's observation file is downloaded. If the base station's coordinates are not known, the position of the base receiver should be computed in the processing software or obtained using software such as OPUS. Having established the base station's coordinates before loading the rover's observation file ensures that the vectors to the rover will radiate from the base station. The processing of the baseline vectors to the rover is then performed. Since this is a radial survey, no checks are available on the resultant coordinate values. However, for critical features, it is possible to resurvey these points from a second base station location. This is similar to the radial traversing procedure discussed in Section 9.9.

As discussed in Section 13.4.3, the heights determined by satellite surveys are in the geodetic coordinate system. Typically, topographic maps are produced using a map coordinate system and orthometric heights. The conversion of geodetic coordinates to map coordinates is covered in Chapter 20. As shown in

Equation (13.8), the geoid height at each point must be applied to the geodetic height to determine its orthometric height. If requested by the user and a geoid model is available, the software can determine the orthometric height of the points surveyed. The current geoid model for the United States is GEOID09. This model has an accuracy of a few centimeters for most of the United States. Thus, the derived orthometric heights will be slightly worse. The software manufacturer usually supplies support files to upgrade both the controller and software to the current geoid model.

■ 15.7 COMMUNICATION IN REAL-TIME KINEMATIC SURVEYS

Roving receivers in RTK surveys require continuous communication with base receivers. These communications can be accomplished with radios, wireless Internet connections, or data modems. Using these devices, the base receiver transmits both corrections and raw data to the rover. The rover processes this data using procedures similar to those discussed in Section 13.9.

The most common form of communication between the base receiver and the rover is by low-powered radios. These radios are often an integral part of the receiver. The Federal Communications Commission (FCC) does not require a license for radios that broadcast in the range from 157 to 174 MHz. However, all other frequencies given in Section 15.4 do require an FCC license. The stronger, external radio transmitters typically use the frequencies in the 450 to 470 MHz range. These frequencies require an FCC license. Since, by FCC regulations, voice communication takes precedence over data communication, radio transmitters generally come with as many as 10 or more preset frequencies or channels. The operator must find a channel that is not in use already. Additionally using an unlicensed channel is a violation of FCC regulations, which can result in stiff fines. Thus, it is wise to license several of the channels that are available on the transmitter. The maximum power of the radios is typically 35 watts. This form of communication will work in all areas of the world although additional licensing to use the frequencies may be required. When using radios, it is important to connect the antenna to the radio before powering the transmitter to avoid equipment failure.

Another option for communication between the base receiver and rover is data modems. These require cellular coverage in the area being surveyed. When coverage is available, the data is transmitted via cellular technology to the rover. The cell provider charges a monthly service fee to use this option. Obviously, this form of communication is not available in areas that do not have cell coverage. Additionally, data latency with this form of communication will be greater than that experienced with radios.

In areas where wireless Internet connections are available, it is possible for the base receiver and rover to communicate over the Internet. This option requires that the base receiver have an Internet connection and the rover have a wireless connection. Again, data latency will be greater than that experienced with radios using this form of communication.

Several problems can occur with communication equipment. Cables often develop breaks near connectors resulting in intermittent transmission problems. In severe cases, the cables fail and communication is impossible. Also the power of the radio limits its range. When using receivers with internal radios, the range is often limited to small areas around the base station, which is less than 3 km typically. As discussed in Section 15.3, this range can be increased with repeater stations. With larger 35-watt external radios, the achievable range of the survey is maximized, but is generally limited to areas under 6 mi (10 km) in radius. Again, larger ranges can be achieved with repeater stations. In one survey in Alaska under ideal conditions, the range from the base radio to the rover was 38 mi! Obviously, this was not a typical situation.

■ 15.8 REAL-TIME NETWORKS

A base station requires additional receivers and personnel to perform the survey. If the base receiver could be used as a rover, the work could be performed in half the time. A real-time network provides this capability. An option that eliminates the need for a base receiver in a RTK survey is known as a *Real-Time Network* (RTN). Both the private and public sectors are implementing this technology. The RTN is a network of base stations that are connected to a central processing computer using the Internet. Using the known positions of the base receivers and their observational data, the central processor models errors in the satellite ephemerides, range errors caused by ionospheric and tropospheric refraction, and the geometric integrity of the network stations. *Virtual reference station* (VRS) and *spatial correction parameter* (FKP)[5] are examples of two methods used in modeling these errors.

Of course these systems may not work reliably in areas that are cellularly challenged. Since the entire system involves communication from multiple base receivers to a central processor and finally to a rover, high traffic volume on the Internet, multiple connections between network servers to the central processor, and time of transmissions in the cellular world can create greater data latency than much simpler base-to-rover radio connections. Some manufacturers wait for the corrections from the central processor before processing the data at the rover. Others extrapolate the modeled errors to process the rover's observational data at the time of reception. The application software typically stops survey operations if the data latency becomes greater than a specified time interval. This value may be as great as four seconds! For this reason, surveyors should use RTNs cautiously or not at all in machine control operations (see Section 15.9).

As shown in Figure 15.5, in real-time kinematic surveys, the accuracy of the position degrades as the rover moves farther from the base station. This is principally due to differences in the ionosphere between the base and the rover. Notice that this distance changes with respect to solar activity and its affect on the ionosphere. This is particularly true in the vertical component where errors are traditionally two to three times greater than horizontal errors. In RTNs these errors are modeled and thus substantially reduced. However, as shown in Figure 15.5, these errors do increase as the rover moves farther from the network reference station.

[5]FKP is an acronym for Flächenkorrekturparameter, which is German for spatial correction parameter.

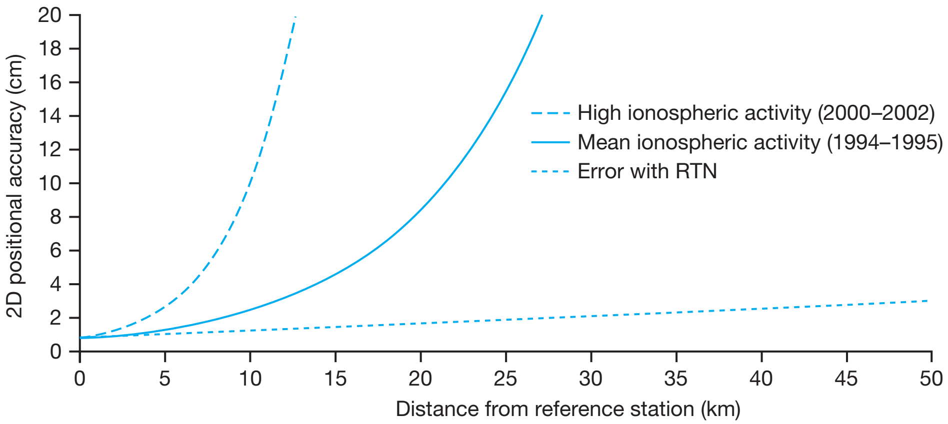


Figure 15.5 Comparison of horizontal errors during different periods of ionospheric activity. (Courtesy National Geodetic Survey.)

When the rover connects to the RTN, either the central processor or the rover interpolates the errors to a location at the survey site. A *virtual reference station* (VRS) is created that is used by the rover to determine its position. If the rover moves too far from the VRS, another virtual reference station is determined for the rover. When working with an RTN, the ppm error in surveying is removed resulting in better achievable accuracies than are present with a radio and single base receiver. The accuracy of positions determined using RTN is usually within 2 cm anywhere within a distance of 30 km from a reference station. Another advantage of using a RTN is that the coordinates obtained from the network are referenced to a common datum,[6] and thus results from many surveys will fit together seamlessly.

These RTN systems are sold usually as a subscription service. Users of this service save costs since they do not need a base receiver or the additional personnel to monitor the base receiver while performing a survey. The system should be periodically calibrated by locating a known position in the RTN system with the rover. HARN stations can serve as good reference stations. Caution should be exercised when using an RTN outside the bounds of the network since errors increase rapidly when extrapolation of the corrections occurs beyond the limits of the base stations.

■ 15.9 PERFORMING REAL-TIME KINEMATIC SURVEYS

As previously stated, the main difference between RTK surveys and PPK surveys is the fact that RTK surveys provide immediate results in the field. Thus, RTK surveys are used primarily in construction stakeout. Since RTK surveys provide immediate results, some form of communication as discussed in Section 15.6 must be established and maintained during the entire RTK survey. Similar to PPK surveys, the receivers must be initialized before the survey is started and initialization must be maintained during the entire survey. However, the process of surveying is similar to the methods used in a PPK survey.

[6]Section 19.6 discusses the reference frames that are currently used on the North American continent.

Stakeout surveys using RTK have some important distinctions from conventional surveys. One important difference is the reference frame (also called the datum). As discussed in Chapter 19, conventional surveys use some form of NAD 83 as their horizontal datum and NAVD 88 for their vertical datum typically. These reference frames are considered to be regional since they were developed using observations only on the North American continent. As discussed in Section 13.4.3, the broadcast ephemeris uses WGS 84, which is a worldwide reference frame. The current rendition of the WGS 84 reference frame closely approximates ITRF 2000. The difference in the origins of the NAD 83 and ITRF 2000 data is about 2.2 m. Thus, when performing a stakeout survey, the coordinates for stations produced by receivers can differ significantly from coordinates of the same stations in the regional reference frame that were used to perform the engineering design.

As discussed in Section 19.6.6, the surveyed coordinates of the points can be brought into the regional datum using a coordinate transformation. To do this, points having regional reference frame coordinate values must be established on the perimeter of the project area. A minimum of two horizontal control points and three vertical control points should exist. However, it is better to have a minimum of three horizontal control points and four vertical control points for the purpose of redundancy and checks. As discussed in Section 19.6.6, a minimal transformation considering only the translation factors between the reference data can be performed if only a single regional control point exists. However, this will result in a significant loss in accuracy to the survey. Furthermore as discussed in Section 19.6.6, a simplified two-dimensional coordinate transformation (see Equation 11.37) can be used to transform survey-derived positions into a local reference frame. It is important when performing these transformations to have control on the exterior and surrounding the project area to avoid extrapolation errors.

Survey controllers have this transformation built into their software. Depending on the vendor, this transformation is known as *localization* or *site calibration*. This procedure should be performed at the beginning of each project that requires local or arbitrary coordinates. The procedure involves occupying the control stations with the antenna. The controller then computes the transformation parameters and allows the user to view the errors. It is wise to perform this procedure whenever questions concerning the stability of the control arise to eliminate possible errors. With this in mind, an alignment survey should be planned with control points along the entire corridor to ensure their quick availability. Following this procedure, the stakeout of the design points can proceed. However, this procedure should be performed only once for any project, since errors in the systems can produce significantly different results if the procedure is repeated.

Since receivers create observation files during a RTK survey, it is possible to convert a RTK survey into a PPK survey in the office. This may be helpful when problems are experienced in the field. However, this would serve no purpose on a stakeout survey.

■ 15.10 MACHINE CONTROL

Traditionally construction projects were executed by placing stakes at key locations in the project (see Chapter 23) to establish the levels of materials and grades

Figure 15.6 A dozer and grader using machine control to create an intersection of roads. (Courtesy Topcon Positioning Systems, Inc.)

of finished work. However, with RTK surveying methods, it is possible to load the project design, digital terrain models (see Section 18.14), and site calibration parameters into a computer that guides the vehicle during the construction process. This technology known as *machine control* allows the machine operator to see their position in a construction project, cut and fill levels, and finished grades of the project, all in real time. As shown in Figures 15.6 and 15.7, this is achieved by placing RTK units on the construction equipment. One aspect of this technology is that the antenna must be calibrated with respect to the construction vehicle. For instance, the distance between the antenna reference point and the cutting edge of the blade on a machine must be observed and entered into the machine control system so that the height of the surface under the blade is accurately known.

To accurately achieve this level of automation, the surveyor must place sufficient horizontal and vertical control about the construction project area. A range of about 10 km (6 mi) from the base station is possible with a high-powered radio transmitter. Additionally in locations where obstructions may interfere with satellite transmissions, the surveyor must add sufficient control to support the use of a robotic total station. In these areas, a robotic total station can guide the construction equipment past the obstructions. The surveyor must also create

Figure 15.7 GNSS antenna mounted on a grader blade. (Courtesy Topcon Positioning Systems, Inc.).

a digital terrain model (see Section 18.14) of the existing terrain prior to the start of the project and a proposed three-dimensional surface model of the finished project. These two items are loaded into the machine control system along with site calibration parameters. As discussed in Section 15.8, the localization parameters are needed to transform the surveyed coordinates into the project datum. With an existing digital terrain model (DTM—see Section 18.14), final digital surface, and localization parameters loaded into the machine control system, the construction vehicle is guided by the machine control system through the project.

The accuracy of using RTK is about 1 cm in horizontal and 2 cm in vertical. This accuracy is sufficient for excavation purposes. However, finished surfaces need to have accuracies under 0.02 ft (5 mm). This accuracy is achieved by augmenting the machine control system with laser levels as shown in Figure 24.2 or robotic total stations. One manufacturer has incorporated a laser level into their machine control systems to achieve millimeter accuracies in all three dimensions. When using this equipment, sufficient control must be placed at the perimeters of the project area to guide the construction vehicles through the project. For example, robotic total stations have a working radius of about 1000 ft from the construction vehicle and laser levels have a working radius of about 1500 ft from the construction vehicle. Thus, control must be placed within the appropriate limits to provide sufficient guidance to support the machine control system. Again, RTNs should be used with caution in machine control since data latency can be large which can lead to significant real-time errors during the finishing process in a construction project. In fact, some manufacturers do not recommend the use of RTNs at all in machine control projects.

As shown in Figure 15.8, another area that is utilizing RTK is precision agriculture. This area does not require a surveyor's expertise, but is interesting nonetheless. In precision agriculture, crop yields are monitored with respect to positions of the harvester on the field. Additionally, soil samples are located and tested for fertility, drainage, and so on to provide the farmer with a complete picture of yields versus growing conditions. In the following year, this information is fed into a guidance system that controls tillage equipment, planters, sprayers, and fertilizer spreaders so that appropriate tillage and chemicals are used as required for various locations on the field. The end results are an economy in fuel and

Figure 15.8 A large tractor pulling a land leveler. (Courtesy Topcon Positioning Systems, Inc.).

chemicals that increase yields in crops. This is an example of surveying technology reaching into nonsurveying fields.

15.11 ERRORS IN KINEMATIC SURVEYS

Kinematic surveys suffer from some same error sources that are found in conventional surveys. These include:

1. Setup errors at the base station and rover.
2. Errors in reading the height to the base station or rover antenna.

Additionally, all satellite surveys have the following errors sources.

1. Ionospheric refraction
2. Tropospheric refraction
3. Errors in ephemerides
4. Base station coordinate errors
5. Weak satellite geometry

15.12 MISTAKES IN KINEMATIC SURVEYS

Some of the more common mistakes that can be made in kinematic surveys include:

1. Misidentification of stations
2. Incorrect identification of the station
3. Starting or proceeding with the survey before the integer ambiguities are resolved
4. Misidentification of the antenna
5. Surveying during high periods of solar activity
6. Incorrect settings for the radio or wireless connection
7. Surveying under overhead obstructions

PROBLEMS

Asterisks (*) indicate problems that have partial answers given in Appendix G.

15.1* What is the typical epoch rate for a kinematic survey?

15.2 What are the advantages of a PPK survey over an RTK survey?

15.3 What are the advantages of an RTK survey over a PPK survey?

15.4 Why are repeater stations used in an RTK survey?

15.5 How can ephemeral errors be eliminated in a kinematic survey?

15.6* How much error in horizontal position occurs if the antenna is mounted on a 2.000-m pole that is 10 min out of level?

15.7 Do Problem 15.6, but this time assume the level is 4 min out of level.

15.8 How much error in vertical position occurs with the situation described in Problem 15.6?

15.9* Why should the radio antenna at the base station be mounted as high as possible?

15.10 List two reasons why PPK surveys usually provide better results than RTK surveys.

15.11 What is OPUS and how can it be used in a PPK survey?

15.12 What are the available methods for initializing a receiver?

15.13 Discuss the differences between the stop-an-go and the true kinematic modes of surveying.
15.14 Discuss the appropriate steps used in processing PPK data.
15.15 Why is the use of a real-time network not recommended in machine control?
15.16 What is VRS?
15.17 What limitations occur in an RTK survey?
15.18* What frequencies found in RTK radios require licensure?
15.19 What is localization of a survey?
15.20 Why is it important to localize a survey?
15.21 Discuss how the coordinate differences between a regional datum and satellite-derived coordinates can be resolved.
15.22 What three surveying elements are needed in machine control?
15.23 A 5-mi stretch of road has numerous canopy restrictions. What is the minimum number of control stations required to support machine control in this part of the road if a robotic total station is used?
15.24 How are robotic total stations used in machine control?
15.25 How are finished grades determined in machine control?
15.26 What factors may determine the best location for a base station in an RTK survey?
15.27 What should be considered in planning a kinematic survey?
15.28* How many total pseudorange observations will be observed using a 1-sec epoch rate for a total of 10 min with eight usable satellites?
15.29 How many pseudorange observations will be observed using a 5-sec epoch rate for a total of 30 min with eight usable satellites?
15.30 The baseline vector between the base and roving receivers is 1000 m long. What is the estimated uncertainty in the length of the baseline vectors if an RTK survey is performed?
15.31 Discuss the importance of knowing the space weather before performing a kinematic survey.
15.32 Why must the antenna be calibrated to the cutting edge of the blade in a machine control system?
15.33 Where are the best locations for control used in localizing a project?
15.34 Why should a localization occur only once on a project?

BIBLIOGRAPHY

Asher, R. 2009. "Crossing the RTK Bridge." *Professional Surveyor* 29 (No. 6): 18.
Barr, M. 2006. "Real-Time Connection." *Point of Beginning* 31 (No. 4): 22.
Crawford, W. 2006. "What Are Your Tolerances?" *Point of Beginning* 32 (No. 3): 46.
Henning, W. 2006. "The New RTK—Changing Techniques for GPS Surveying in the USA." *Surveying and Land Information Science* 66 (No. 2): 107.
Mosby, M. 2006. "Advancing with Machine Control." *Point of Beginning* 32 (No. 3): 32.
Pugh, N. 2007. "The Specifics on Managing Network RTK Integrity." *Point of Beginning* 33 (No. 1): 34.
Schrock, G. 2006. "RTN-101: An Introduction to Network Corrected Real-Time GPS/GNSS (Part 1)." *The American Surveyor* 3 (No. 6): 28.
———. 2006. "RTN-101: An Introduction to Network Corrected Real-Time GPS/GNSS (Part 2)." *The American Surveyor* 3 (No. 7): 38.
———. 2006. "RTN-101: An Introduction to Network Corrected Real-Time GPS/GNSS (Part 3)." *The American Surveyor* 3 (No. 8): 38.
———. 2006. "RTN-101: On-Grid—An Initiative in Support of RTN Development (Part 4)." *The American Surveyor* 3 (No. 9): 39.

This page intentionally left blank

16 Adjustments by Least Squares

16.1 INTRODUCTION

The subject of errors in measurements was introduced in Chapter 3 where the two types of errors, systematic and random, were defined. It was noted that systematic errors follow physical laws, and if conditions producing them are observed, corrections can be computed to eliminate them. However, random errors exist in all observed values. Additionally as discussed in Chapter 3, observations can contain mistakes (blunders). Examples of mistakes are setting an instrument on the wrong station, sighting the wrong station, transcription errors in recording observed values, and so on. Mistakes should be removed when possible before the adjustment process. As further explained in Chapter 3, experience has shown that random errors in surveying follow the mathematical laws of probability and in any group of observations they are expected to conform to the laws of a *normal distribution*, as illustrated in Figure 3.3.

In surveying (geomatics), after eliminating mistakes and making corrections for systematic errors, the presence of the remaining random errors will be evident in the form of misclosures. Examples include sums of interior angles in closed polygons that do not total $(n - 2)180°$, misclosures in closed leveling circuits, and traverse misclosures in departures and latitudes. To account for these misclosures, adjustments are applied to produce mathematically perfect geometric conditions. Although various techniques are used, the most rigorous adjustments are made by the method of least squares, which is based on the laws of probability.

Although the theory of least squares was developed in the late 1700s, because of the lengthy calculations involved, the method was not used commonly prior to the availability of computers. Instead, arbitrary, or "rule of thumb," methods such as the compass (Bowditch) rule were applied. Now least-squares calculations are handled routinely and making adjustments by this method is rapidly becoming indispensable in modern surveying (geomatics). The method of least squares is currently being used to adjust all kinds of observations, including differences in elevation, horizontal distances, and horizontal and vertical angles. It has become essential in the adjustment of GNSS observations and is also widely used in adjusting photogrammetric data. Adjustments by the least-squares method have taken on added importance with the most recent surveying accuracy standards. These standards include the use of statistical quantities that result from least-squares adjustment. Thus in order to evaluate a survey for compliance with the standards, least-squares adjustments must first be performed.

Least-squares adjustments provide several advantages over other arbitrary methods. First of all, because the method is based upon the mathematical theory of probability, it is the most rigorous of adjustment procedures. It enables all observations to be simultaneously included in an adjustment, and each observation can be weighted according to its estimated precision. Furthermore, the least-squares method is applicable to any observational problem regardless of its nature or geometric configuration. In addition to these advantages, the least-squares method enables rigorous statistical analyses to be made of the results of the adjustment, that is, the precisions of all adjusted quantities can be estimated, and other factors investigated. The least-squares method even enables presurvey planning to be done so as to ensure that required precisions of adjusted quantities are obtained in the most economical manner.

A simple example can be used to illustrate the arbitrary nature of "rule of thumb" adjustments, as compared to least squares. Consider the horizontal survey network shown in Figure 16.1. If the compass rule was used to adjust the observations in the network, several solutions would be possible. To illustrate one variation, suppose that traverse *ABCDEFGA* is adjusted first. Then holding the adjusted values of points *G* and *E*, traverse *GHKE* is adjusted, and finally, holding the adjusted values on *H* and *C*, traverse *HJC* is adjusted. This obviously would yield a solution, but there are other possible approaches. In another variation, traverse *ABCDEFGA* could be adjusted followed by *GHJC*, and then

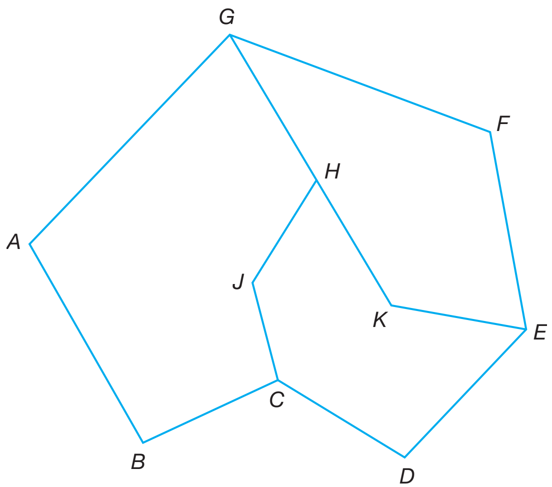


Figure 16.1 A horizontal network.

HKE. This sequence would result in another solution, but with different adjusted values for points *H*, *J*, and *K*. There are still other possible variations. This illustrates that the compass rule adjustment is properly referred to as an "arbitrary" method. On the other hand, the least-squares method simultaneously adjusts all observations and for a given set of weights there is only one solution—that which yields the most probable values for the given set of observations.

In the sections that follow, the fundamental condition enforced in least squares is described and elementary examples of least-squares adjustments are presented. Then systematic procedures for forming and solving least-squares equations using matrix methods are given and demonstrated with examples. The examples involving differential leveling, GNSS baselines, and horizontal networks are performed using the software WOLFPACK, which is on the companion website for this book at http://www.pearsonhighered.com/ghilani. For these examples, sample data files and the results of the adjustments are shown. A complete description of the data files is given in the *help* system provided with the software. For those wishing to see programming of these problems, Mathcad® worksheets that demonstrate the differential leveling, baseline vector, and plane survey adjustments are available on the companion website for this book.

16.2 FUNDAMENTAL CONDITION OF LEAST SQUARES

It was shown through the discussion in Section 3.12 and the normal distribution curves illustrated in Figures 3.2 and 3.3, that small errors (residuals) have a higher probability of occurrence than large ones in a group of normally distributed observations. Also discussed was the fact that in such a set of observations there is a specific probability that an error (residual) of a certain size will exist within a group of errors. In other words, there is a direct relationship between probabilities and residual sizes in a normally distributed set of observations. The method of least-squares adjustment is derived from the equation for the normal distribution curve. *It produces that unique set of residuals for a group of observations that have the highest probability of occurrence*.

For a group of equally weighted observations, the fundamental condition enforced by the least-squares method is that the *sum of the squares of the residuals is a minimum*. Suppose a group of m observations of equal weight were taken having residuals $v_1, v_2, v_3, \ldots, v_m$. Then, in equation form, the fundamental condition of least squares is

$$\sum_{i=1}^{m} v_i^2 = v_1^2 + v_2^2 + v_3^3 + \cdots + v_m^2 \rightarrow \text{minimum}^1 \qquad \textbf{(16.1)}$$

For any group of observed values, weights may be assigned to individual observations according to a *priori* (before the adjustment) estimates of their relative worth or they may be obtained from the standard deviations of the observations,

[1]Refer to Ghilani (2010) cited in the bibliography at the end of this chapter, for a derivation of this equation.

if available. An equation expressing the relationship between standard deviations and weights, given in Section 3.20 and repeated here, is

$$w_i = \frac{1}{\sigma_i^2} \tag{16.2}$$

In Equation (16.2), w_i is the weight of the ith observed quantity and σ_i^2 the variance of that observation. *This equation states that weights are inversely proportional to variances*.

If observed values are to be weighted in least-squares adjustment, then the fundamental condition to be enforced is that *the sum of the weights times their corresponding squared residuals is minimized* or, in equation form

$$\sum_i^m w_i v_i^2 = w_1 v_1^2 + w_2 v_2^2 + w_3 v_3^2 + \cdots + w_m v_m^2 \rightarrow \text{minimum} \tag{16.3}$$

Some basic assumptions underlying least-squares theory are that (1) mistakes and systematic errors have been eliminated, so only random errors remain in the set of observations; (2) the number of observations being adjusted is large; and (3) as stated earlier, the frequency distribution of the errors is normal. Although these basic assumptions are not always met, least-squares adjustments still provide the most rigorous error treatment available.

■ 16.3 LEAST-SQUARES ADJUSTMENT BY THE OBSERVATION EQUATION METHOD

Two basic methods are employed in least-squares adjustments: (1) the *observation equation* method and (2) the *condition equation* method. The former is most common and is the one discussed herein. In this method, "observation equations" are written relating observed values to their residual errors and the unknown parameters. One observation equation is written for each observation. For a unique solution, the number of equations must equal the number of unknowns. If redundant observations are made, the least-squares method can be applied. In that case, an expression for each residual error is obtained from every observation equation. The residuals are squared and added to obtain the function expressed in either Equation (16.1) or Equation (16.3).

To minimize the function in accordance with either Equation (16.1) or Equation (16.3), partial derivatives of the expression are taken with respect to each unknown variable and set equal to zero. This yields a set of so-called *normal equations*, which are equal in number to the number of unknowns. The normal equations are solved to obtain most probable values for the unknowns. The following elementary examples illustrate the procedures.

■ Example 16.1

Using least squares, compute the most probable value for the equally weighted distance observations of Example 3.1.

Solution

1. For this problem, as was done in Example 3.1, let $\overline{M}$ be the most probable value of the observed length. Then write the following observation equations that define the residual for any observed quantity as the difference between the most probable value and any individual observation:

$$\begin{aligned}
\overline{M} &= 538.57 + v_1 \\
\overline{M} &= 538.39 + v_2 \\
\overline{M} &= 538.37 + v_3 \\
\overline{M} &= 538.39 + v_4 \\
\overline{M} &= 538.48 + v_5 \\
\overline{M} &= 538.49 + v_6 \\
\overline{M} &= 538.33 + v_7 \\
\overline{M} &= 538.46 + v_8 \\
\overline{M} &= 538.47 + v_9 \\
\overline{M} &= 538.55 + v_{10}
\end{aligned}$$

2. Solve for the residual in each observation equation and form the function $\sum v^2$ according to Equation (16.1)

$$\begin{aligned}
\sum v^2 = {} & (\overline{M} - 538.57)^2 + (\overline{M} - 538.39)^2 + (\overline{M} - 538.37)^2 \\
& (\overline{M} - 538.39)^2 + (\overline{M} - 538.48)^2 + (\overline{M} - 538.49)^2 \\
& (\overline{M} - 538.33)^2 + (\overline{M} - 538.46)^2 + (\overline{M} - 538.47)^2 \\
& (\overline{M} - 538.55)^2
\end{aligned}$$

3. Take the derivative of the function $\sum v^2$ with respect to $\overline{M}$, set it equal to zero (this minimizes the function)

$$\begin{aligned}
\frac{\partial \sum v^2}{\partial \overline{M}} = 0 = {} & 2(\overline{M} - 538.57) + 2(\overline{M} - 538.39) + 2(\overline{M} - 538.37) \\
& + 2(\overline{M} - 538.39) + 2(\overline{M} - 538.48) + 2(\overline{M} - 538.49) \\
& + 2(\overline{M} - 538.33) + 2(\overline{M} - 538.46) + 2(\overline{M} - 538.47) \\
& + 2(\overline{M} - 538.55)
\end{aligned}$$

4. Reduce and solve for $\overline{M}$

$$\begin{aligned}
10\overline{M} &= 5384.50 \\
\overline{M} &= \frac{5384.50}{10} = 538.45
\end{aligned}$$

Note that this answer agrees with the one given for Example 3.1. Note also that this procedure verifies the statement given earlier in Section 3.10 that *the most probable value for an unknown quantity, measured repeatedly using the same equipment and procedures, is simply the mean of the observations*.

Example 16.2

In Figure 8.9(c), the three horizontal angles observed around the horizon are $x = 42°12'13''$, $y = 59°56'15''$, and $z = 257°51'35''$. Adjust these angles by the least-squares method so that their sum equals the required geometric total of 360°.

Solution

1. Form the observation equations

$$x = 42°12'13'' + v_1 \quad \text{(a)}$$
$$y = 59°56'15'' + v_2 \quad \text{(b)}$$
$$z = 257°51'35'' + v_3 \quad \text{(c)}$$

2. Write an expression that enforces the condition that the sum of the three adjusted angles total 360°.

$$x + y + z = 360° \quad \text{(d)}$$

3. Substitute Equations (a), (b), and (c) into Equation (d), and solve for v_3

$$(42°12'13'' + v_1) + (59°56'15'' + v_2) + (257°51'35'' + v_3) = 360°$$
$$v_3 = -3'' - v_1 - v_2 \quad \text{(e)}$$

(Because of the 360° condition, if v_1 and v_2 are fixed, v_3 is also fixed. Thus, there are only two independent residuals in the solution.)

4. Form the function Σv^2, which involves all three residuals but includes only the two independent variables v_1 and v_2

$$\sum v^2 = v_1^2 + v_2^2 + (-3'' - v_1 - v_2)^2 \quad \text{(f)}$$

5. Take partial derivatives of Equation (f) with respect to the variables v_1 and v_2, and set them equal to zero.

$$\frac{\partial \sum v^2}{\partial v_1} = 0 = 2v_1 + 2(-3'' - v_1 - v_2)(-1); \qquad 4v_1 + 2v_2 = -6'' \quad \text{(g)}$$

$$\frac{\partial \sum v^2}{\partial v_2} = 0 = 2v_2 + 2(-3'' - v_1 - v_2)(-1); \qquad 2v_1 + 4v_2 = -6'' \quad \text{(h)}$$

6. Solve Equations (g) and (h) simultaneously

$$v_1 = -1'' \text{ and } v_2 = -1''$$

7. Substitute v_1 and v_2 into Equation (e) to compute v_3

$$v_3 = -3'' + 1'' + 1'' = -1''$$

8. Finally substitute the residuals into Equations (a) through (c) to get the adjusted angles

$$
\begin{aligned}
x &= 42°12'13'' - 1'' &&= 42°12'12''\\
y &= 59°56'15'' - 1'' &&= 59°56'14''\\
z &= 257°51'35'' - 1'' &&= \underline{257°51'34''}\\
&\qquad\qquad\quad \sum &&= 360°00'00''\text{(Check!)}
\end{aligned}
$$

Note that this result verifies another basic procedure frequently applied in surveying (geomatics) that *for equally weighted angles observed around the horizon, corrections of equal size are applied to each angle*. The same result occurs when equally weighted interior angles in a closed polygon traverse are adjusted by least squares. That is, each receives an equal-size correction.

Examples 16.1 and 16.2 are indeed simple, hardly the type for which least squares is best suited. However, they do supply the basis for some commonly applied simple adjustments and also illustrate procedures involved in making least-squares adjustments without complicating the mathematics. The following example illustrates least-squares adjustment of distance observations that are functionally related.

Example 16.3

Adjust the three equally weighted distance observations taken (in feet) between points A, B, and C of Figure 16.2.

Solution

1. Let the unknown distances AB and BC be x and y, respectively. These two unknowns are related through the observations as follows:

$$
\begin{aligned}
x + y &= 393.65\\
x &= 190.40 \qquad\qquad \textbf{(i)}\\
y &= 203.16
\end{aligned}
$$

2. Values for x and y could be obtained from any two of these equations so that the remaining equation is redundant. However, notice that values

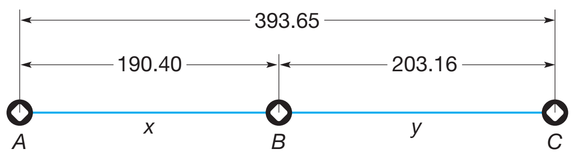


Figure 16.2 Equally weighted distance observations of Example 16.3.

obtained for x and y will differ, depending on which two equations are solved. It is therefore apparent that the observations contain errors. Equations (i) may be rewritten as observation equations by including residual errors as follows:

$$\begin{aligned} x + y &= 393.65 + v_1 \\ x &= 190.40 + v_2 \\ y &= 203.16 + v_3 \end{aligned} \tag{j}$$

3. To obtain the least-squares solution, the observation equations (j) are rearranged to obtain expressions for the residuals. These are squared and added to form the function given in Equation (16.1) as follows:

$$\sum_{i=1}^{m} v_i^2 = (x + y - 393.65)^2 + (x - 190.40)^2 + (y - 203.16)^2 \tag{k}$$

4. Function (k) is minimized, enforcing the condition of least squares, by taking partial derivatives with respect to the unknowns x and y and setting them equal to zero. This yields the following two normal equations:

$$\frac{\partial \sum v^2}{\partial x} = 0 = 2(x + y - 393.65) + 2(x - 190.40)$$

$$\frac{\partial \sum v^2}{\partial y} = 0 = 2(x + y - 393.65) + 2(y - 203.16)$$

5. Reducing the normal equations and solving yields $x = 190.43$ ft and $y = 203.19$ ft. The residuals can now be calculated by substituting x and y into the original observation equations (j)

$$\begin{aligned} v_1 &= 190.43 + 203.19 - 393.65 = -0.03 \text{ ft} \\ v_2 &= 190.43 - 190.40 = +0.03 \text{ ft} \\ v_3 &= 203.19 - 203.16 = +0.03 \text{ ft} \end{aligned}$$

■ 16.4 MATRIX METHODS IN LEAST-SQUARES ADJUSTMENT[2]

It has been noted that least-squares computations are quite lengthy, and therefore generally performed on a computer. Their solution follows a systematic procedure that is conveniently adapted to matrix methods. In general, any group of observation equations may be represented in matrix form as

$${}_{m}A^{n}\,{}_{n}X^{1} = {}_{m}L^{1} + {}_{m}V^{1} \tag{16.4}$$

[2]The balance of this chapter requires a basic understanding of matrix algebra. Students who do not have this background may consult any mathematics book with introductory coverage on matrices. Alternatively, a good primer on matrices can be found in Appendixes A and B of *Adjustment Computations: Spatial Data Analysis* by C. Ghilani, John Wiley & Sons, New York, 2010.

where A is the matrix of coefficients for the unknowns, X the matrix of unknowns, L the matrix of observations, and V the matrix of residuals. The detailed structures of these matrices are

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad L = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_m \end{bmatrix} \qquad V = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$$

The normal equations that result from a set of equally weighted observation equations [Equations (16.4)] are given in matrix form by

$$A^T A X = A^T L \tag{16.5}$$

In Equations (16.5), $A^T A$ is the matrix of normal equation coefficients for the unknowns. Premultiplying both sides of Equation (16.5) by $(A^T A)^{-1}$ and reducing yields

$$X = (A^T A)^{-1} A^T L \tag{16.6}$$

Equation (16.6) is the least-squares solution for equally weighted observations. The matrix X consists of most probable values for unknowns $x_1, x_2, x_3, \ldots, x_n$.

For a system of weighted observations, the following equation provides the $\boldsymbol{X}$ matrix:

$$X = (A^T W A)^{-1} A^T W L \tag{16.7}$$

In Equation (16.7) the matrices are identical to those of the equally weighted case, except that W is a diagonal matrix of weights defined as follows:[3]

$$W = \begin{bmatrix} w_1 & & & \\ & w_2 & & zeros \\ zeros & & \ddots & \\ & & & w_n \end{bmatrix}$$

If the observations in an adjustment are all of equal weight, Equation (16.7) can still be used, but the W matrix becomes an identity matrix. It therefore reduces exactly to Equation (16.6). Thus, Equation (16.7) is general and can be used for both the unweighted and weighted adjustments. It is readily programmed for computer solution.

[3]For a group of independent and *uncorrelated observations* (a case frequently encountered in surveying), the weight matrix is diagonal, that is, all off-diagonal elements are zeros. In certain cases, however, observations are correlated, that is, they are related to each other. An example occurs in GNSS baseline measurements, where the vector components result from least-squares adjustments and thus are correlated. As will be shown in Section 16.8, this yields off-diagonal elements in the W matrix.

Example 16.4

Solve Example 16.3 using matrix methods.

Solution

1. The observation equations of Example 16.3 can be expressed in matrix form as follows:

$$_3A^2\,{}_2X^1 = {}_3L^1 + {}_3V^1$$

where

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad X = \begin{bmatrix} x \\ y \end{bmatrix} \qquad L = \begin{bmatrix} 393.65 \\ 190.40 \\ 203.16 \end{bmatrix} \qquad V = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

2. Solving matrix Equation (16.6)

$$A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$(A^T A)^{-1} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \qquad A^T L = \begin{bmatrix} 584.05 \\ 596.81 \end{bmatrix}$$

$$X = (A^T A)^{-1} A^T L = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 584.05 \\ 596.81 \end{bmatrix} = \begin{bmatrix} 190.43 \\ 203.19 \end{bmatrix}$$

Note that this solution yields $x = 190.43$ ft and $y = 203.19$ ft, which are exactly the same values obtained through the algebraic approach of Example 16.3.

16.5 MATRIX EQUATIONS FOR PRECISIONS OF ADJUSTED QUANTITIES

The matrix equation for calculating residuals after adjustment, whether the adjustment is weighted or not, is

$$V = AX - L \tag{16.8}$$

The standard deviation of unit weight for an unweighted adjustment is

$$\sigma_0 = \sqrt{\frac{V^T V}{r}} \tag{16.9}$$

The standard deviation of unit weight for a weighted adjustment is

$$\sigma_0 = \sqrt{\frac{V^T W V}{r}} \tag{16.10}$$

In Equations (16.9) and (16.10), r is the *number of degrees of freedom* in an adjustment, which usually equals the number of observations minus the number of unknowns, or $r = m - n$.

Standard deviations of the individual adjusted quantities are

$$\sigma_{x_i} = \sigma_0 \sqrt{q_{x_i x_i}} \tag{16.11}$$

In Equation (16.11), σ_{x_i} is the standard deviation of the ith adjusted unknown x_i, that is the value in the ith row of the X matrix; σ_0 the standard deviation of unit weight as calculated by Equation (16.9) or (16.10); and $q_{x_i x_i}$ the diagonal element in the ith row and ith column of matrix $(A^T A)^{-1}$ in the unweighted case, or matrix $(A^T W A)^{-1}$ in the weighted case. The $(A^T A)^{-1}$ and $(A^T W A)^{-1}$ matrices are the so-called *covariance* matrices, and symbolized hereon by Q_{xx}.

Example 16.5

Calculate the standard deviation of unit weight and the standard deviations of the adjusted quantities x and y for the unweighted problem of Example 16.4.

Solution

1. By Equation (16.8), the residuals are

$$V = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 190.43 \\ 203.19 \end{bmatrix} - \begin{bmatrix} 393.65 \\ 190.40 \\ 203.26 \end{bmatrix} = \begin{bmatrix} -0.03 \\ 0.03 \\ 0.03 \end{bmatrix}$$

2. By Equation (16.9), the standard deviation of unit weight is

$$V^T V = [-0.03 \quad 0.03 \quad 0.03] \begin{bmatrix} -0.03 \\ 0.03 \\ 0.03 \end{bmatrix} = [0.0027]$$

$$\sigma_0 = \sqrt{\frac{0.0027}{3 - 2}} = \pm 0.052 \text{ ft}$$

3. Using Equation (16.11), the standard deviations of the adjusted values for x and y are

$$\sigma_x = \pm 0.052 \sqrt{\frac{2}{3}} = \pm 0.042 \text{ ft}$$

$$\sigma_y = \pm 0.052 \sqrt{\frac{2}{3}} = \pm 0.042 \text{ ft}$$

In part 3, the numbers 2/3 under the radicals are the elements in row 1, column 1, and row 2, column 2, respectively, of the $(A^T A)^{-1}$ matrix of Example 16.4. The interpretation of the standard deviations computed under step 3 of Example 16.5 is that a 68% probability exists that the adjusted values for x and y are within ± 0.042 ft of their true values. Note that for this simple example, the three residuals calculated in step 1 are equal, and the standard deviations of x and y are equal in step 3. This is caused by the symmetric nature of this particular problem (illustrated in Figure 16.2), but it is not generally the case with more complex problems.

16.6 LEAST-SQUARES ADJUSTMENT OF LEVELING CIRCUITS

When control leveling is being done to establish new benchmarks, for example, benchmarks for a construction project, it is common practice to create a network like that illustrated in Figure 16.3. This enables each new benchmark to benefit from redundant observations and least-squares adjustment. In Figure 16.3, A and B are two new project benchmarks being established near a construction site. Each could be set by running a single loop, such as from BM 1 to A and back to establish A, and from BM 3 to B and back to set B. To build redundancy into the survey, and to increase the precisions of the new benchmarks, additional lines from other nearby benchmarks can be run. Thus in Figure 16.3, five loops are run rather than the minimum of two needed to establish A and B. All observations within this leveling network can be adjusted simultaneously using the least-squares method to obtain most probable adjusted values for the two benchmarks.

In adjusting level networks, the observed difference in elevation for each course is treated as one observation containing a single random error. Observation equations are written that relate these observed elevation differences and their residual errors to the unknown elevations of the benchmarks involved. These can then be processed through the matrix equations given in Sections 16.4 and 16.5 to obtain adjusted values for the benchmarks and their standard deviations. The procedure is illustrated with the following example.

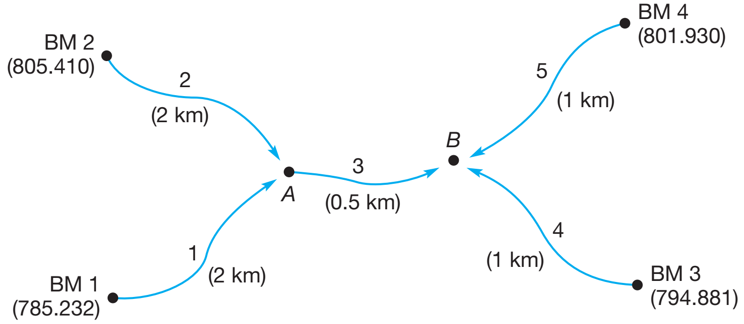


Figure 16.3 Level net for Example 16.6.

Example 16.6

Adjust the level net of Figure 16.3 by weighted least squares, and compute precisions of the adjusted benchmarks. In the figure, the benchmark elevations (in meters) and course lengths (in kilometers) are shown in parentheses. Observed elevation differences for courses 1 through 5 (given in order) are +10.997, −9.169, +3.532, +4.858, and −2.202 m. Arrows on the courses in the figure indicate the direction of leveling. Thus for course 1 having a length of 2 km, leveling proceeded from BM 1 to *A* and the observed elevation difference was 10.970 m.

Solution

1. Observation equations are written relating each line's observed elevation difference to its residual error and the most probable values for unknown elevations *A* and *B* as follows:

$$
\begin{aligned}
A &= \text{BM }1 + 10.997 + v_1 \\
A &= \text{BM }2 - 9.169 + v_2 \\
B &= A + 3.532 + v_3 \\
B &= \text{BM }3 + 4.858 + v_4 \\
B &= \text{BM }4 - 2.202 + v_5
\end{aligned}
\qquad \text{(l)}
$$

2. Substituting the elevations of BM 1, BM 2, BM 3, and BM 4 into Equations (l) and rearranging gives

$$
\begin{aligned}
A \phantom{{}+B} &= 796.229 + v_1 \\
A \phantom{{}+B} &= 796.241 + v_2 \\
-A + B &= 3.532 + v_3 \\
B &= 799.739 + v_4 \\
B &= 799.728 + v_5
\end{aligned}
$$

3. The *A*, *X*, *L*, and *V* matrices for this adjustment are

$$
A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ -1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}
\qquad X = \begin{bmatrix} A \\ B \end{bmatrix}
\qquad L = \begin{bmatrix} 796.229 \\ 796.241 \\ 3.532 \\ 799.739 \\ 799.728 \end{bmatrix}
\qquad V = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix}
$$

4. Weights in differential leveling are inversely proportional to course lengths. Thus after inverting the lengths, the course weights are 0.5, 0.5, 2, 1, and 1, respectively, and the weight matrix is

$$
W = \begin{bmatrix} 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}
$$

5. The weighted matrix solution for the most probable values according to Equation (16.7) is

$$A^TW = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 & -2.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 2.0 & 1.0 & 1.0 \end{bmatrix}$$

$$A^TWA = \begin{bmatrix} 0.5 & 0.5 & -2.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 2.0 & 1.0 & 1.0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ -1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -2 & 4 \end{bmatrix}$$

$$Q_{xx} = (A^TWA)^{-1} = \frac{1}{8}\begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}$$

$$A^TWL = \begin{bmatrix} 0.5 & 0.5 & -2.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 2.0 & 1.0 & 1.0 \end{bmatrix} \begin{bmatrix} 796.229 \\ 796.241 \\ 3.532 \\ 799.739 \\ 799.728 \end{bmatrix} = \begin{bmatrix} 789.171 \\ 1606.531 \end{bmatrix}$$

$$X = (A^TWA)^{-1}A^TWL = \frac{1}{8}\begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 789.171 \\ 1606.531 \end{bmatrix} = \begin{bmatrix} 796.218 \\ 799.742 \end{bmatrix}$$

Thus, the adjusted benchmark elevations are $A = 796.218$ m and $B = 799.742$ m.

6. The residuals by Equation (16.8) are

$$V = AX - L = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 796.218 \\ 799.742 \end{bmatrix} - \begin{bmatrix} 796.229 \\ 796.241 \\ 3.532 \\ 799.739 \\ 799.728 \end{bmatrix} = \begin{bmatrix} -0.011 \\ -0.023 \\ -0.008 \\ 0.003 \\ 0.014 \end{bmatrix}$$

7. Utilizing Equation (16.10), the estimated standard deviation of unit weight is

$$V^TWV = [-0.011 \quad -0.023 \quad -0.008 \quad 0.003 \quad 0.014] \begin{bmatrix} 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -0.011 \\ -0.023 \\ -0.008 \\ 0.003 \\ 0.014 \end{bmatrix}$$

$$= [0.00066]$$

Thus, the standard deviation of unit weight is $\sigma_0 = \sqrt{\dfrac{0.00066}{5-2}} = \pm 0.015$ m.

8. By Equation (16.11), the estimated standard deviations of adjusted benchmark elevations of A and B are

$$\sigma_A = \sigma_0\sqrt{q_{AA}} = \pm 0.015\sqrt{\frac{4}{8}} = \pm 0.010 \text{ m}$$

$$\sigma_B = \sigma_0\sqrt{q_{BB}} = \pm 0.015\sqrt{\frac{3}{8}} = \pm 0.009 \text{ m}$$

Note in these calculations that terms in the radicals are the diagonal elements, (1, 1) and (2, 2), of the Q_{xx} matrix. Note also that B has a lower standard deviation than A, indicating that its precision is better. A check of Figure 16.3 reveals that this should be expected, because the benchmarks closest to A are both 2 km away, and those closest to B are only 1 km away. Thus, the weights of level circuits coming into B are higher than those coming into A, giving B a higher precision.

The software WOLFPACK can be used to adjust the data of Example 16.6. The format for the data file is shown in Figure 16.4. Note that the lengths of the level lines, the number of setups, or the standard deviations of the observations can weight the observations. In this example, course lengths were used. The

```
Differential leveling example of Section 15.6   {title line}
4 6 5                                            {Number of benchmarks; number of stations; number
                                                 of elev. diff.}
1 785.232                                        {BM identifier and elevation}
2 805.410
3 794.881
4 801.930
1 A 10.997 2                                     {from, to, elevation difference, [dist.|setups|σ}
2 A -9.169 2
A B 3.532 0.5
3 B 4.858 1
4 B -2.202 1
```

Figure 16.4 WOLFPACK data file for Example 16.6.

```
******************************
Adjusted Elevation Differences
******************************
      From        To     Elevation Difference      V       σ
=============================================================
         1         A         10.986             -0.011   0.0105
         2         A         -9.192             -0.023   0.0105
         A         B          3.524             -0.008   0.0091
         3         B          4.861              0.003   0.0091
         4         B         -2.188              0.014   0.0091

*******************
Adjusted Elevations
*******************
   Station      Elevation        σ
===================================
         A      796.218      0.0105
         B      799.742      0.0091

Standard Deviation of Unit Weight: 0.015
```

Figure 16.5 Results from adjustment of data in Example 16.6 using WOLFPACK.

results of the adjustment are shown in Figure 16.5. The programming behind this problem is contained in a Mathcad® worksheet *llsq.xmcd* on the companion website for this book. On the companion website for this book at http://www.pearsonhighered.com/ghilani are instructional videos that downloaded. This example problem is solved using MATRIX in the video *LSQ I.mp4*. WOLFPACK, MATRIX, and the Mathcad worksheet are also available for download on this website.

16.7 PROPAGATION OF ERRORS

In Section 3.17, the propagation of errors in functions using independent observations was discussed. At the completion of a least-squares adjustment, the unknowns are no longer independent as evidenced by the off-diagonal terms in the covariance matrix discussed in Section 16.5. When observations are not independent, errors propagate as

$$Q_{\ell\ell} = AQ_{xx}A^T \tag{16.12}$$

where Q_{xx} *equals* the matrix $(A^TA)^{-1}$ in the unweighted case, or the matrix $(A^TWA)^{-1}$ in the weighted case.

The standard deviations in the computed observations are

$$\sigma_{\ell_i} = \sigma_0\sqrt{q_{\ell_i\ell_i}} \tag{16.13}$$

In Equation (16.13), σ_{ℓ_i} is the standard deviation of the *i*th adjusted observation; σ_0 the standard deviation of unit weight as calculated by Equation (16.9) or (16.10); and $q_{\ell_i\ell_i}$ *i*th diagonal element of the $Q_{\ell\ell}$ matrix of Equation (16.12).

Example 16.7

Compute the adjusted elevation differences and their standard deviations for Example 16.6.

Solution

By Equation (16.12) the $Q_{\ell\ell}$ matrix is

$$Q_{\ell\ell} = \frac{1}{8}\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ -1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} = \frac{1}{8}\begin{bmatrix} 4 & 4 & -2 & 2 & 2 \\ 4 & 4 & -2 & 2 & 2 \\ -2 & -2 & 3 & 1 & 1 \\ 2 & 2 & 1 & 3 & 3 \\ 2 & 2 & 1 & 3 & 3 \end{bmatrix}$$

A tabulation of the observations, their residuals, adjusted values, and standard deviations are

From	To	Obs. Δ Elev	V	Adj. Δ Elev	σ
BM 1	*A*	10.997	−0.011	10.986	±0.010
BM 2	*A*	−9.169	−0.023	−9.192	±0.010
A	*B*	3.532	−0.008	3.524	±0.009
BM 3	*B*	4.858	0.003	4.861	±0.009
BM 4	*B*	−2.202	0.014	−2.188	±0.009

The contents of the third line in the above table are explained to clarify the calculations. The adjusted elevation difference for the third observation, which consisted of leveling from *A* to *B*, is obtained by adding the observed elevation difference (see the data for Example Problem 16.6) and its residual (see step 6 of Example Problem 16.6), as follows:

$$\text{Adj. } \Delta\text{Elev} = 3.532 - 0.008 = 3.524$$

Also the standard deviation, *S*, for the third observation is computed as

$$\sigma = \pm 0.015\sqrt{\frac{3}{8}} = \pm 0.009$$

where 3/8 is the third diagonal element of the $Q_{\ell\ell}$, and 0.015 is σ_0 as determined in Example 16.6. The adjusted elevations and standard deviations for the remaining observations are computed in similar fashion.

16.8 LEAST-SQUARES ADJUSTMENT OF GNSS BASELINE VECTORS

It was previously noted in Section 14.5.5 that the least-squares method is essential in adjusting GNSS observations. It is applied in this work in two different stages: (1) for adjusting the massive quantities of redundant data that result after several receivers have made repeated observations on multiple satellites over a period of time (this yields baseline components ΔX, ΔY, and ΔZ) and (2) it is applied in adjusting redundant observations of these baseline components to make them consistent in static networks. The reduction software, which can be purchased, is programmed to perform the first stage of these two applications and thus its development is not covered in this text. However, the second application is within the scope of this text and is illustrated with an example—the adjustment of the network of Figure 14.10.

Since the baseline vector data are the results of the first least-squares adjustment noted above, each baseline has it own 3×3 covariance matrix. This matrix not only contains terms along the diagonal, but it also has elements in the off-diagonal locations. The covariance terms depict the amount of correlation between the adjusted ΔX, ΔY, and ΔZ values. For each baseline, the carrier-phase reduction software will list the baseline components and their covariance terms. For example for baseline AC in Figure 14.10, the vector and its covariance terms were listed as

ΔX	11,644.2232	$9.8E-4$	$-9.6E-6$	$9.5E-6$
ΔY	3601.2165		$9.4E-4$	$-9.5E-6$
ΔZ	3399.2550			$9.8E-4$

Note that only the upper-triangular portion of the covariance matrix is shown to the right of the vector components. This is because the covariance matrix is symmetric and thus the elements of the lower-triangular portion mirror their upper-triangular values and need not be repeated. Since the baseline vectors are not independent, the weight matrix is computed as the inverse of the covariance matrix, or in matrix symbology is

$$W = \sum\nolimits^{-1} \qquad \textbf{(16.14)}$$

where $\sum$ is the covariance matrix for the baseline vectors, and W their weight matrix. It can be shown that this equation is also valid for independent observations and thus is a general equation for weighting observations.

The baseline vector components (ΔX, ΔY, ΔZ), and their covariance terms for the survey of the network of Figure 14.10 are shown in Table 16.1. In the column labeled (1) of this table, the (1,1) element of the covariance matrix is listed; column (2) lists the (1,2) element of the covariance matrix; column (3) the (1,3) element; column (4) the (2,2) element; column (5) the (2,3) element; and column (6) the (3,3) element of the covariance matrix. In the survey, two HARN stations (see Section 14.3.5) were held fixed. These stations and their coordinates are listed in Table 16.2.

From the known X, Y, and Z coordinates of stations A and B, and the observed ΔX, ΔY, and ΔZ components, coordinates of new stations C, D, E, and F can be calculated. However, an adjustment is necessary because redundant observations exist. In applying least squares to this problem, observation equations are written that relate the unknown adjusted coordinates of new stations C, D, E, and F to the observed ΔX, ΔY, and ΔZ values and their residual errors. As shown in Figure 14.10, excluding the check observation of fixed baseline AB,

TABLE 16.1 **OBSERVED BASELINE VECTORS AND THEIR COVARIANCE MATRIX VALUES FOR FIGURE 14.10**

Baseline	ΔX	ΔY	ΔZ	(1)	(2)	(3)	(4)	(5)	(6)
AC	11644.2232	3601.2165	3399.2550	9.8E-4	−9.6E-6	9.5E-6	9.4E-4	−9.5E-6	9.8E-4
DC	15128.1647	−6286.7054	−6371.0583	1.5E-4	−1.4E-6	1.3E-6	1.6E-4	−1.4E-6	1.3E-4
AE	−5321.7164	3634.0754	3173.6652	2.2E-4	−2.1E-6	2.2E-6	1.9E-4	−2.1E-6	2.0E-4
BC	3960.5442	−6681.2467	−7279.0148	2.3E-4	−2.2E-6	2.1E-6	2.5E-4	−2.2E-6	2.2E-4
BD	−11167.6076	−394.5204	−907.9593	2.7E-4	−2.8E-6	2.8E-6	2.7E-4	−2.7E-6	2.7E-4
DE	−1837.7459	−6253.8534	−6596.6697	1.2E-4	−1.2E-6	1.2E-6	1.3E-4	−1.2E-6	1.3E-4
FA	−1116.4523	−4596.1610	−4355.9062	7.5E-5	−7.9E-7	8.8E-7	6.6E-5	−8.1E-7	7.6E-5
FC	10527.7852	−994.9370	−956.6246	2.6E-4	−2.2E-6	2.4E-6	2.2E-4	−2.3E-6	2.4E-4
FE	−6438.1364	−962.0694	−1182.2305	9.4E-5	−9.2E-7	1.0E-6	1.0E-4	−8.9E-7	8.8E-5
FD	−4600.3787	5291.7785	5414.4311	9.3E-5	−9.9E-7	9.0E-7	9.9E-5	−9.9E-7	1.2E-4
FB	6567.2311	5686.2926	6322.3917	6.6E-5	−6.5E-7	6.9E-7	7.5E-5	−6.4E-7	6.0E-5
BF	−6567.2310	−5686.3033	−6322.9807	5.5E-5	−6.3E-7	6.1E-7	7.5E-5	−6.3E-7	6.6E-5
AF	1116.4577	4596.1553	4355.9141	6.6E-5	−8.0E-7	9.0E-7	8.1E-5	−8.2E-7	9.4E-5

TABLE 16.2 HARN STATION GEOCENTRIC COORDINATES

Station	X	Y	Z
A	402.3509	−4,652,995.3011	4,349,760.7775
B	8086.0318	−4,642,712.8474	4,360,439.0833

there are 11 different baselines. However, two of these, AF and BF, were repeated giving a total of 13 baseline observations. The following observation equations are written for the first two baselines:

$$
\begin{aligned}
X_C &= X_A + \Delta X_{AC} + v_1 \\
Y_C &= Y_A + \Delta Y_{AC} + v_2 \\
Z_C &= Z_A + \Delta Z_{AC} + v_3 \\
X_C - X_D &= \Delta X_{DC} + v_4 \\
Y_C - Y_D &= \Delta Y_{DC} + v_5 \\
Z_C - Z_D &= \Delta Y_{DC} + v_6
\end{aligned}
\tag{16.15}
$$

Similar equations can be written for the other 11 baseline observations, giving a total of 39 observation equations. These observation equations can be expressed in matrix form according to Equation (16.4). To illustrate the contents of the matrices, the partial matrices that result from the observation equations of Equations (16.15) are

$$
{}_{39}A^{12} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \\ 0 & 1 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 1 & 0 & 0 & 0 & \\ 1 & 0 & 0 & -1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & -1 & 0 & \\ 0 & 0 & 1 & 0 & 0 & -1 & \\ & & & \vdots & & & \end{bmatrix};\ {}_{12}X^{1} = \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ X_D \\ Y_D \\ Z_D \\ \vdots \end{bmatrix};
$$

$$
{}_{39}L^{1} = \begin{bmatrix} X_A + \Delta X_{AC} \\ Y_A + \Delta Y_{AC} \\ Z_A + \Delta Z_{AC} \\ \Delta X_{DC} \\ \Delta Y_{DC} \\ \Delta Z_{DC} \\ \vdots \end{bmatrix};\ {}_{39}V^{1} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ \vdots \end{bmatrix}
\tag{16.16}
$$

To complete the above A matrix, the coefficients for the remaining observation equations are entered expanding the dimensions of this matrix to 39 rows and 12 columns. Note that there are three unknowns for each of the four new ground stations, thus the X matrix has a total of 12 elements. The L and V matrices each have 39 elements, one for each observation equation.

The partial covariance matrix that results from the first two baseline observations is

$$
{}_{39}\Sigma^{39} = \begin{bmatrix}
9.8E-4 & -9.6E-6 & 9.5E-6 & 0 & 0 & 0 & & \\
-9.6E-6 & 9.4E-4 & -9.5E-6 & 0 & 0 & 0 & & \\
9.5E-6 & -9.5E-6 & 9.8E-4 & 0 & 0 & 0 & & \\
0 & 0 & 0 & 1.5E-E & -1.4E-6 & 1.3E-6 & \dots & \dots \\
0 & 0 & 0 & -1.4E-6 & 1.6E-4 & -1.4E-6 & & \\
0 & 0 & 0 & 1.3E-6 & -1.4E-6 & 1.3E-4 & & \\
 & & & \vdots & & & & \ddots
\end{bmatrix}
$$

The Σ matrix, when completed, has 39 rows and 39 columns. Its formation follows the same procedure for each observed baseline as is demonstrated above. As can be seen, each observed baseline creates a 3×3 submatrix within Σ. Thus the structure of this matrix is *block diagonal*, that is, each individual submatrix on the diagonal is 3×3, and all other elements are zeros. Thus the first 3 rows and 3 columns are nonzero elements that pertain to the first baseline, the next 3 rows and 3 columns contain nonzero elements that pertain to the second baseline, and so on.

The weight matrix is obtained by inverting the Σ matrix according to Equation (16.14). Once all matrices are formed, the solution for the unknown station coordinates and their standard deviations can be determined using Equation (16.7). After the adjustment, the resulting geocentric coordinates can be used to determine geodetic coordinates using the procedures discussed in Section 13.4.3 and if requested state plane coordinates can be derived from the geodetic coordinates using procedures discussed in Chapter 20.

For this particular problem, the software program WOLFPACK was used to adjust the baselines using the least-squares method. The input file is shown in Figure 16.6, and the results of the adjustment are given in Figure 16.7. Note that a standard deviation was computed for each baseline, and that the precision of

```
Example of Section 15.8                {title line}
2 13                                   {Number of control stations; Number of baselines)
A 402.3509 -4652995.3011 4349760.7775  {identifier, X, Y, Z coordinates of control station}
B 8086.0318 -4642712.8474 4360439.0833
A C 11644.2232 3601.2165 3399.2550 9.8E-4 -9.6E-6 9.5E-6 9.4E-4 -9.5E-6 9.8E-4    {baselines}
A E -5321.7164 3634.0754 3173.6652 2.2E-4 -2.1E-6 2.2E-6 1.9E-4 -2.1E-6 2.0E-4
B C 3960.5442 -6681.2467 -7279.0148 2.3E-4 -2.2E-6 2.1E-6 2.5E-4 -2.2E-6 2.2E-4
B D -11167.6076 -394.5204 -907.9593 2.7E-4 -2.8E-6 2.8E-6 2.7E-4 -2.7E-6 2.7E-4
D C 15128.1647 -6286.7054 -6371.0583 1.5E-4 -1.4E-6 1.3E-6 1.6E-4 -1.4E-6 1.3E-4
D E -1837.7459 -6253.8534 -6596.6697 1.2E-4 -1.2E-6 1.2E-6 1.3E-4 -1.2E-6 1.3E-4
F A -1116.4523 -4596.1610 -4355.9062 7.5E-5 -7.9E-7 8.8E-7 6.6E-5 -8.1E-7 7.6E-5
F C 10527.7852 -994.937 -956.6246 2.6E-4 -2.2E-6 2.4E-6 2.2E-4 -2.3E-6 2.4E-4
F E -6438.1364 -962.0694 -1182.2305 9.4E-5 -9.2E-7 1.0E-6 1.0E-4 -8.9E-7 8.8E-5
F D -4600.3787 5291.7785 5414.4311 9.3E-5 -9.9E-7 9.0E-7 9.9E-5 -9.9E-7 1.2E-4
F B 6567.2311 5686.2926 6322.3917 6.6E-5 -6.5E-7 6.9E-7 7.5E-5 -6.4E-7 6.0E-5
B F -6567.2310 -5686.3033 -6322.3807 5.5E-5 -6.3E-7 6.1E-7 7.5E-5 -6.3E-7 6.6E-5
A F 1116.4577 4596.1553 4355.9141 6.6E-5 -8.0E-7 9.00E-7 8.1E-5 -8.2E-7 9.4E-5
```

Figure 16.6 Input file for least-squares adjustment problem in Section 16.8.

```
                    Degrees of Freedom = 27
                  Reference Variance = 0.5010
           Standard Deviation of Unit Weight = ±0.71

*************************
Adjusted Distance Vectors
*************************
      From        To          dX            dY            dZ         Vx        Vy        Vz
=============================================================================================
         A         C    11644.2232     3601.2165     3399.2550   0.00665   0.00219   0.03197
         A         E    -5321.7164     3634.0754     3173.6652   0.02665   0.00579   0.01212
         B         C     3960.5442    -6681.2467    -7279.0148   0.00475   0.01169  -0.00403
         B         D   -11167.6076     -394.5204     -907.9593  -0.00730  -0.00137  -0.00065
         D         C    15128.1647    -6286.7054    -6371.0583  -0.00084  -0.00783  -0.00058
         D         E    -1837.7459    -6253.8534    -6596.6697  -0.00985   0.00267   0.00117
         F         A    -1116.4523    -4596.1610    -4355.9062   0.00197   0.00527  -0.00773
         F         C    10527.7852     -994.9370     -956.6246  -0.00568  -0.00004  -0.00236
         F         E    -6438.1364     -962.0694    -1182.2305  -0.00368  -0.00514  -0.00611
         F         D    -4600.3787     5291.7785     5414.4311  -0.00563  -0.00230   0.00083
         F         B     6567.2311     5686.2926     6322.3917  -0.00053   0.00537   0.00017
         B         F    -6567.2310    -5686.3033    -6322.3807   0.00043   0.00533  -0.01117
         A         F     1116.4577     4596.1553     4355.9141  -0.00737   0.00043  -0.00017

*****************************************************
Advanced Statistical Values
*****************************************************
      From        To      ±σ   Vector Length     Prec
======================================================
         A         C  0.0105      12,653.538 1,206,000
         A         E  0.0091       7,183.255   794,000
         B         C  0.0105      10,644.668 1,015,000
         B         D  0.0087      11,211.408 1,282,000
         D         C  0.0107      17,577.670 1,641,000
         D         E  0.0097       9,273.836   960,000
         F         A  0.0048       6,430.015 1,344,000
         F         C  0.0104      10,617.871 1,019,000
         F         E  0.0086       6,616.111   770,000
         F         D  0.0083       8,859.035 1,066,000
         F         B  0.0048      10,744.075 2,246,000
         B         F  0.0048      10,744.075 2,246,000
         A         F  0.0048       6,430.015 1,344,000

********************
Adjusted Coordinates
********************
   Station           X                  Y                   Z         σx       σy       σz
=================================================================================================
         A        402.3509    -4,652,995.3011    4,349,760.7775
         B      8,086.0318    -4,642,712.8474    4,360,439.0833
         C     12,046.5808    -4,649,394.0824    4,353,160.0645  0.0061   0.0061   0.0059
         E     -4,919.3388    -4,649,361.2199    4,352,934.4548  0.0052   0.0053   0.0052
         D     -3,081.5831    -4,643,107.3692    4,359,531.1234  0.0049   0.0051   0.0051
         F      1,518.8012    -4,648,399.1454    4,354,116.6914  0.0027   0.0028   0.0028
```

Figure 16.7 Results of adjustment of data file in Figure 16.6 using WOLFPACK.

each baseline was determined as the ratio of the standard deviation, σ over the vector length times one million. For those wishing to explore the programming of this problem, the Mathcad® worksheet *GPS.xmcd* demonstrates the programming of this network.

16.9 LEAST-SQUARES ADJUSTMENT OF CONVENTIONAL HORIZONTAL PLANE SURVEYS

Traversing (described in Chapters 9 and 10), and trilateration and triangulation (discussed in Chapter 19), are traditional ground-surveying methods for conducting *horizontal surveys* [those surveys that establish either X and Y coordinates, usually in some grid system such as state plane coordinates (see Chapter 20), or geodetic latitudes and longitudes of points (see Section 19.4)]. The basic observations that are made in traversing, trilateration, and triangulation are horizontal angles and horizontal distances. As with other types of surveys, they are most appropriately adjusted by the method of least squares.

To adjust horizontal surveys by the least-squares method, it is necessary to write observation equations for the horizontal distance and horizontal angle observations. These observation equations are nonlinear, and to facilitate solving them they are linearized using a first-order Taylor series expansion. The procedure is described in the following subsection.

16.9.1 Linearizing Nonlinear Equations

The general form of the Taylor series linearization of a nonlinear equation is

$$\begin{aligned} F(x_1, x_2, \ldots, x_n) = F(x_{1_0}, x_{2_0}, \ldots, x_{n_0}) &+ \left(\frac{\partial F}{\partial x_1}\right)_0 dx_1 \\ &+ \left(\frac{\partial F}{\partial x_2}\right)_0 dx_2 + \cdots + \left(\frac{\partial F}{\partial x_n}\right)_0 dx_n + R \end{aligned} \tag{16.17}$$

where $F(x_1, x_2, \ldots, x_n)$ is a nonlinear function in terms of the unknowns $x_1, x_2, \ldots, x_n$, which represents a measured quantity; $x_{1_0}, x_{2_0}, \ldots, x_{n_0}$ are approximate values for the unknowns $x_1, x_2, \ldots, x_n$; $(\partial F/\partial x_1)_0, (\partial F/\partial x_2)_0, \ldots, (\partial F/\partial x_n)_0$ are the partial derivatives of the function F with respect to $x_1, x_2, \ldots, x_n$ evaluated using the approximate values of $x_{1_0}, x_{2_0}, \ldots, x_{n_0}$; $dx_1, dx_2, \ldots, dx_n$ are corrections to the approximate values of $x_{1_0}, x_{2_0}, \ldots, x_{n_0}$, such that $x_1 = x_{1_0} + dx_1$; $x_2 = x_{2_0} + dx_2, \ldots, x_n = x_{n_0} + d_{x_n}$; and R is a remainder. In Equation (16.17) the only unknowns are $dx_1, dx_2, \ldots, dx_n$, and R. The term R is also nonlinear, but if the values assigned for $x_{1_0}, x_{2_0}, \ldots, x_{n_0}$ are close to the true values of the unknowns, then R is small, and is dropped, which linearizes the equation. However, this makes the equation an approximation and thus the solution must be obtained *iteratively*, that is, the corrections $dx_1, dx_2, \ldots, dx_n$ are computed repetitively until their sizes become negligible.

After dropping R and rearranging Equation (16.17), the following general linear form of the equation is obtained:

$$\begin{aligned} &\left(\frac{\partial F}{\partial x_1}\right)_0 dx_1 + \left(\frac{\partial F}{\partial x_2}\right)_0 dx_2 + \cdots + \left(\frac{F}{\partial x_n}\right)_0 \\ &\qquad = F(x_1, x_2, \ldots, x_n) - F(x_{1_0}, x_{2_0}, \ldots, x_{n_0}) \end{aligned} \tag{16.18}$$

The subscript zeros attached to the coefficients on the left side of Equation (16.18) indicate that these coefficients are simply numbers obtained by substituting the approximate values $x_{1_0}, x_{2_0}, \ldots, x_{n_0}$ into those partial derivative functions. Also, the right side of Equation (16.18) is the observed value, $[F(x_1, x_2, \ldots, x_n)]$, minus the computed value obtained by substituting the initial approximations into the original function, $[F(x_{1_0}, x_{2_0}, \ldots, x_{n_0})]$.

The process of solving a pair of nonlinear equations using the Taylor's series will be illustrated. Suppose that the following two nonlinear functions, $F(x, y)$ and $G(x, y)$ express the relationship between observed values 115 and 75, respectively, and the unknowns x and y:

$$\begin{aligned} F(x, y) &= x^2 + 3y = 115 \\ G(x, y) &= 5x + y^2 = 75 \end{aligned} \tag{a}$$

Partial derivatives of the functions with respect to the unknowns are

$$\frac{\partial F}{\partial x} = 2x; \quad \frac{\partial F}{\partial y} = 3; \quad \frac{\partial G}{\partial x} = 5; \quad \frac{\partial G}{\partial y} 2y \tag{b}$$

A. First Iteration

Assume that through either estimation or preliminary calculations based upon one or more observations values of 9 and 4 are selected as initial estimates for the unknowns, x_0 and y_0. Then using the functions of Equations (a) and substituting these initial approximations and the partial derivatives from (b) into Equation (16.18), the following two linearized equations are obtained:

$$\begin{aligned} (2 \times 9)dx + 3dy &= 115 - [9^2 + 3(4)] \\ 5dx + (2 \times 4)dy &= 75 - [5(9) + 4^2] \end{aligned} \tag{c}$$

Equations (c) are now in linear form and contain only two unknowns, dx and dy. The solution of this pair of equations yields the following corrections: $dx = 1.04$, and $dy = 1.10$.

B. Second Iteration

Using corrections from the first iteration, new approximations x_0 and y_0 are computed as

$$\begin{aligned} x_0 &= 9.00 + 1.04 = 10.04 \\ y_0 &= 4.00 + 1.10 = 5.10 \end{aligned} \tag{d}$$

These new approximations are now used to repeat the solution. Substituting again into Equation (16.18), the following linearized equations result:

$$\begin{aligned} (10.04)^2 + 3 \times 5.10 + 2(10.04)dx + 3dy &= 115 \\ 5 \times 10.04 + (5.10)^2 + 5dx + 2(5.10)dy &= 75 \end{aligned} \tag{e}$$

Solving Equations (e) for the unknown parameters yields: $dx = -0.08$ and $dy = -0.08$.

C. Third Iteration

The corrections from the second iteration are used to get updated values for the coordinates as

$$\begin{aligned} x_0 &= 10.04 - 0.08 = 9.96 \\ y_0 &= 5.10 - 0.08 = 5.02 \end{aligned} \tag{f}$$

New linearized equations are formed by substituting these initial approximations into Equations (16.18) as follows:

$$
\begin{aligned}
(9.96)^2 + 3 \times 5.02 + 2(9.96)dx + 3dy &= 115 \\
5 \times 9.96 + (5.02)^2 + 5dx + 2(5.02)dy &= 75
\end{aligned} \tag{g}
$$

Solving Equations (g) for the unknown corrections gives $dx = 0.04$ and $dy = -0.02$. The updated values for x and y are therefore

$$
\begin{aligned}
x_0 &= 9.96 + 0.04 = 10.00 \\
y_0 &= 5.02 - 0.02 = 5.00
\end{aligned}
$$

A fourth iteration (not shown) yields zeros for both dx and dy and thus the solution has converged. The final answers are $x = 10.00$ and $y = 5.00$.

Although only two unknowns existed in the above example, the first-order Taylor's series expansion is applicable to linearizing and solving nonlinear equations with any number of unknowns.

All that is necessary is to select an initial approximation for each unknown and take partial derivatives of the function with respect to each unknown, as indicated in Equation (16.18). As is discussed in the subsections that follow, in least-squares adjustments of horizontal surveys, up to four unknowns can appear in distance observation equations, and up to six unknowns can appear in angle observation equations. In more advanced types of geodesy and photogrammetry problems, many more unknowns can appear in the nonlinear equations used.

On the companion website for this book at http://www.pearsonhighered.com/ghilani are instructional videos that can be downloaded. This example problem is solved using a spreadsheet and MATRIX in the video *LSQ II.mp4*. MATRIX is also available for download on this website.

16.9.2 The Distance Observation Equation

The observation equation for an observed distance is expressed in terms of the X and Y coordinates of its end points, and includes the residual error. Referring to Figure 16.8, the following distance observation equation can be written for the line whose end points are identified by I and J:

$$L_{IJ} + v_{IJ} = \sqrt{(X_J - X_I)^2 + (Y_J - Y_I)^2} \tag{16.19}$$

In Equation (16.19), L_{IJ} is the observed length of the line IJ; v_{IJ} the residual error in the observation; X_I, Y_I, X_J, and Y_J the most probable values for stations I and J, respectively. By applying Taylor's series [Equation (16.18)] to this nonlinear distance equation, the following linearized form of the equation results:[4]

$$
\begin{aligned}
&\left(\frac{X_I - X_J}{IJ}\right)_0 dX_I + \left(\frac{Y_I - Y_J}{IJ}\right)_0 dY_J + \left(\frac{X_J - X_I}{IJ}\right)_0 dX_J \\
&\qquad + \left(\frac{Y_J - Y_I}{IJ}\right)_0 dY_J = k_{IJ} + v_{IJ}
\end{aligned} \tag{16.20}
$$

[4]The complete development of this linearized equation, and the linearized equations for azimuths and horizontal angles, which are given in the following subsections, are presented in Ghilani (2010) cited in the bibliography at the end of this chapter.

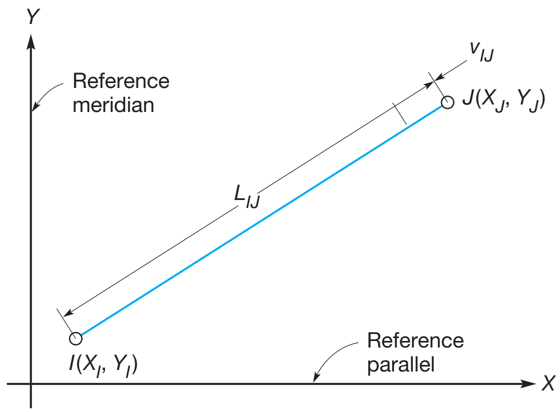


Figure 16.8 Observed distance expressed in terms of coordinates.

In Equation (16.20), $k_{ij} = L_{ij} - (IJ)_0$, where L_{ij} is the observed length of the line, and $(IJ)_0$ the length based upon initial approximations for the coordinates of points I and J and computed as

$$(IJ)_0 = \sqrt{(X_{J_0} - X_{I_0})^2 + (Y_{J_0} - Y_{I_0})^2} \tag{16.21}$$

For a specific distance such as AB of Figure 16.1, Equation (16.21) would be written as

$$\left(\frac{X_A - X_B}{AB}\right)_0 dX_A + \left(\frac{Y_A - Y_B}{AB}\right)_0 dY_A + \left(\frac{X_B - X_A}{AB}\right)_0 dX_B + \left(\frac{Y_B - Y_A}{AB}\right)_0 dY_B = k_{AB} + v_{AB} \tag{16.22}$$

In Equation (16.22), each coefficient of the unknown dx and dy corrections is evaluated using initial approximations selected for the unknown coordinates of stations A and B; k_{AB} is $L_{AB} - (AB)_0$, where L_{AB} is the observed distance, and $(AB)_0$ the distance computed using Equation (16.21) and the approximate coordinates; and v_{AB} the residual error in the distance.

Example 16.8

Write the linearized observation equation for the distance AB whose observed length is 132.823 m. Assume the approximate coordinates for station A and B are (1023.151, 873.018) and (1094.310, 985.163), respectively.

Solution

Step 1: Compute the appropriate coordinate differences

$$(X_B - X_A)_0 = 1094.310 - 1023.151 = 71.159 \text{ m}$$
$$(Y_B - Y_A)_0 = 985.163 - 873.018 = 112.145 \text{ m}$$

Step 2: Compute $(AB)_0$

$$(AB)_0 = \sqrt{71.159^2 + 112.145^2} = 132.816 \text{ m}$$

Step 3: Substitute the appropriate values into Equation (16.22) to develop the linearized observation equation as

$$\left(\frac{-71.159}{132.816}\right) dX_A + \left(\frac{-112.145}{132.816}\right) dY_A + \left(\frac{71.159}{132.816}\right) dX_B + \left(\frac{112.145}{132.816}\right) dY_B = (132.823 - 132.816) + v_{AB}$$

Reducing:

$$-0.53577\ dX_A - 0.84436\ dY_A + 0.53577\ dX_B + 0.84436\ dY_B = 0.007 + v_{AB}$$

16.9.3 The Azimuth Observation Equation

Equation (10.11) expressed the azimuth of a line in terms of the X and Y coordinates of the line's end points. That expression is written here in observation equation form, and to make it general, the line designation has been changed from AB to IJ, and thus subscripts I and J have replaced A and B:

$$Az_{IJ} + v_{IJ} = \tan^{-1}\left(\frac{X_J - X_I}{Y_J - Y_I}\right) + C \tag{16.23}$$

In Equation (16.23), Az_{IJ} is the observed azimuth of line IJ, v_{IJ} the residual error in the observation and the coordinates in the expression on the right side of the equation are most probable values of the line's end points. The value of the constant C depends on the direction of the line. If the azimuth of the line is between 0° and 90°, the value of C is 0°. If the azimuth of the line is between 90° and 270°, C is 180° and if the line's azimuth is between 270° and 360°, C is 360°. Equation (16.23) is nonlinear, but again by applying Equation (16.18) the following linearized form of this equation is obtained:

$$\rho\left(\frac{Y_I - Y_J}{IJ^2}\right)_0 dX_I + \rho\left(\frac{X_J - X_I}{IJ^2}\right)_0 dY_I + \rho\left(\frac{Y_J - Y_I}{IJ^2}\right)_0 dX_J + \rho\left(\frac{X_I - Y_J}{IJ^2}\right)_0 dY_J = k_{IJ} + v_{IJ} \tag{16.24}$$

As with the linearized distance observation equation, the coefficients of the unknown dx and dy terms in Equation (16.24) are obtained by using initial approximations for the coordinates of the end points of the line. The IJ^2 terms in the denominators of the coefficients are simply the squares of the line lengths as computed using initial approximations for the coordinates in Equation (16.21). The right side of Equation (16.24), which includes the constant term k_{IJ} and the residual, is given in seconds. Thus to make the units consistent on both sides of the equation, the coefficients on the left are multiplied by rho (ρ), which is 206,265 sec/rad. The constant term k_{IJ} is computed as follows:

$$k_{IJ} = Az_{IJ} - \tan^{-1}\left(\frac{X_{J_0} - X_{I_0}}{Y_{J_0} - X_{I_0}}\right) + C \qquad \textbf{(16.25)}$$

In Equation (16.25), Az_{IJ} is the observed azimuth, the arc tangent function is the computed azimuth based upon initial approximations for the coordinates and C is the constant, as previously described.

16.9.4 The Angle Observation Equation

As illustrated in Figure 16.9, an angle can be expressed as the difference between the azimuths of two lines. Thus angle BIF in the figure is simply azimuth IF minus azimuth IB. The nonlinear observation equation for angle, BIF, is therefore

$$\theta_{BIF} + v_{BIF} = \tan^{-1}\left(\frac{X_F - X_I}{Y_F - Y_I}\right) - \tan^{-1}\left(\frac{X_B - X_I}{Y_B - Y_I}\right) + D \qquad \textbf{(16.26)}$$

In Equation (16.26), θ_{BIF} is the observed value for angle BIF and v_{BIF} the residual error in the observation. The right side of this equation is simply the difference in the azimuths IF and IB, where these azimuths are expressed in the manner of Equation (16.23). The constant term $D = C_{IF} - C_{IB}$, where C was defined for Equation (16.23), and C_{IF} and C_{IB} apply to the azimuths of IF and IB, respectively. By applying Equation (16.18), the following linearized form of Equation (16.26) is obtained:

$$\begin{aligned}
&\rho\left(\frac{Y_I - Y_B}{IB^2}\right)_0 dX_B + \rho\left(\frac{X_B - X_I}{IB^2}\right)_0 dY_B + \rho\left(\frac{Y_B - Y_I}{IB^2} - \frac{Y_F - Y_I}{IF^2}\right)_0 dX_I \\
&\quad + \rho\left(\frac{X_I - X_B}{IB^2} - \frac{X_I - X_F}{IF^2}\right)_0 dY_I + \rho\left(\frac{Y_F - Y_I}{IF^2}\right)_0 dX_F \qquad \textbf{(16.27)} \\
&\quad + \rho\left(\frac{X_I - X_F}{IF^2}\right)_0 dY_F = k_{BIF} + v_{BIF}
\end{aligned}$$

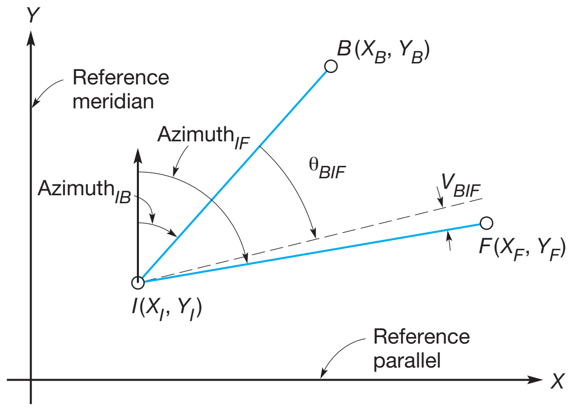


Figure 16.9 Measured angle expressed in terms of coordinates. (Note that an angle is simply the difference between the two azimuths.)

In Equation (16.27), the term k_{BIF} is

$$k_{BIF} = \theta_{BIF} - \tan^{-1}\left(\frac{Y_F - Y_I}{X_F - X_I}\right)_0 - \tan^{-1}\left(\frac{X_B - X_I}{Y_B - Y_I}\right)_0 + D$$

where θ_{BIF} is the observed value for the angle. As with the distance and azimuth equations, the coefficients of the unknown corrections dX and dY in Equation (16.27) are developed using approximate coordinates for stations B, I, and F. In this equation, B represents the backsight station, I the instrument station, and F the foresight station for clockwise angle BIF. In Equation (16.27), it should again be noted that the coefficients on the left side of the equation are multiplied by ρ (206, 265″/rad), so that the units on both sides are seconds. For a specific angle, such as angle GAB of Figure 16.1, after substitution of corresponding subscripts into Equation (16.27), the following linearized angle observation equation results:

$$\begin{aligned}
&\rho\left(\frac{Y_A - Y_G}{AG^2}\right)_0 dX_G + \rho\left(\frac{X_G - X_A}{AG^2}\right)_0 dY_G + \left(\frac{Y_G - Y_A}{AG^2} - \frac{Y_B - Y_A}{AB^2}\right)_0 dX_A \\
&+ \left(\frac{X_A - X_G}{AG^2} - \frac{X_A - X_B}{AB^2}\right)_0 dY_A + \rho\left(\frac{Y_B - Y_A}{AB^2}\right)_0 dX_B \\
&+ \rho\left(\frac{X_A - X_B}{AB^2}\right)_0 dY_B = k_{GAB} + v_{GAB}
\end{aligned} \tag{16.28}$$

Example 16.9

Angle GAB was observed as 107°29′40″. The backsight station G, instrument station A, and foresight station B had the following approximate X and Y coordinates, respectively: (578.741, 1103.826); (415.273, 929.868); and (507.934, 764.652). (Note that all coordinate values are given in units of meters.) Write the linearized observation equation for this angle.

Solution

Step 1: Compute the appropriate coordinate differences for substitution into Equation (16.28).

$$\begin{aligned}
(X_G - X_A)_0 &= 578.741 - 415.273 = -163.468 \text{ m} \\
(Y_G - Y_A)_0 &= 1103.826 - 929.868 = 173.958 \text{ m} \\
(X_B - X_A)_0 &= 507.934 - 415.273 = 92.661 \text{ m} \\
(Y_B - Y_A)_0 &= 764.652 - 929.868 = -165.216 \text{ m}
\end{aligned}$$

Step 2: Compute the distances AG and AB, and angle GAB_0

$$\begin{aligned}
(AG)_0 &= \sqrt{(163.468)^2 + (173.958)^2} = 238.711 \text{ m} \\
(AB)_0 &= \sqrt{(92.661)^2 + (-165.216)^2} = 189.726 \text{ m}
\end{aligned}$$

$$GAB_0 = \tan^{-1}\left(\frac{92.661}{-165.216}\right) - \tan^{-1}\left(\frac{163.468}{173.958}\right) + 180° = 107°29'42''$$

Step 3: Substitute the appropriate values into Equation (16.28).

$$
\begin{aligned}
&\rho\left(\frac{-173.958}{238.711^2}\right)dX_G + \rho\left(\frac{163.468}{238.711^2}\right)dY_G \\
&+ \rho\left(\frac{173.958}{238.711^2} - \frac{-165.216}{189.726^2}\right)dX_A \\
&+ \rho\left(\frac{-16.468}{238.711^2} - \frac{-92.661}{189.726^2}\right)dY_A + \rho\left(\frac{-165.216}{189.726^2}\right)dX_B \\
&+ \rho\left(\frac{-92.661}{189.726^2}\right)dY_B \\
&= (107^\circ 29'40'' - 107^\circ 29'42'') + v_{GAB}
\end{aligned}
$$

Reducing:

$$
\begin{aligned}
&-629.684''\ dX_G + 591.713''\ dY_G + 1579.405''\ dX_A \\
&-59.065''\ dY_A - 949.721''\ dX_B - 532.649''\ dY_B = -2'' + v_{GAB}
\end{aligned}
$$

16.9.5 A Traverse Example Using WOLFPACK

Traverse adjustments by the least-squares method involve distance and angle observation equations, and sometimes they include azimuth observation equations as well. Because of the lengthy calculations involved in forming and solving the observation equations, and because the solution is iterative which requires repetitive computations, the least-squares adjustment of even small traverses should be done by computers. To perform a least-squares adjustment by computer, a data file must be prepared in which all observations and their identity (given by the end stations of lines for distances and azimuths, and by backsight, instrument and foresight stations for angles), must be entered. The computer can be programmed to calculate initial approximations for the coordinates of the unknown stations by using some limited amount of the observed data. In a traverse like that of Figure 11.1(a) for example, the angles at stations A, B, C, and D, and distances AB, BC, CD, and DE could be used to compute coordinates of the four unknown stations B, C, D, and E. Once the initial approximations have been calculated, a computer can easily determine the coefficients of the unknowns in the observation equations, as well as the constant terms. By employing Equation (16.2), relative weights can be determined from the standard deviations of the observed quantities. Thus the computer is able to prepare the $\boldsymbol{A}$, $\boldsymbol{W}$, and $\boldsymbol{L}$ matrices, so that Equation (16.7) can be solved.

The WOLFPACK software, which is available on the companion website for this book, has been used to adjust the traverse network of Figure 16.1. For this adjustment, station A with coordinates of $x = 415.273$ m and $y = 929.868$ m, was held fixed. Also azimuth AB was held fixed at $150^\circ 42'51''$ by applying a heavy weight (assigning it a standard deviation of $\pm 0.001''$). The coordinates of station A fix the survey in position while the azimuth AB fixes the survey in rotation. The observed data for the traverse network, which includes distance observations and their standard deviations, and angle observations and their standard deviations, are listed in Tables 16.3 and 16.4, respectively.

TABLE 16.3 DISTANCE OBSERVATIONS FOR NETWORK SHOWN IN FIGURE 16.1

From	To	Distance (m)	σ (m)
A	B	189.436	0.007
B	C	122.050	0.007
C	D	121.901	0.007
D	E	145.256	0.007
E	F	168.180	0.007
F	G	231.021	0.007
G	A	238.714	0.007
G	H	143.780	0.007
H	K	119.631	0.007
K	E	114.695	0.007
H	J	96.036	0.007
J	C	85.908	0.007

TABLE 16.4 ANGLE OBSERVATIONS FOR NETWORK SHOWN IN FIGURE 16.1

Backsight Station	Instrument Station	Foresight Station	Angle	σ
G	A	B	107°29′40″	8.9″
A	B	C	94°44′24″	11.7″
B	C	D	235°09′26″	13.7″
C	D	E	104°08′40″	12.7″
D	E	F	124°27′36″	11.2″
E	F	G	121°37′08″	9.5″
F	G	A	112°23′00″	8.3″
F	G	H	38°25′46″	9.9″
G	H	J	243°15′20″	14.6″
H	J	C	135°08′30″	18.0″
J	H	G	116°44′44″	14.6″
J	H	K	296°44′38″	15.0″
H	K	E	131°16′30″	14.3″
K	E	F	68°40′36″	12.3″

Figure 16.10 shows the format and order of preparing the data file for input to the least-squares adjustment program of the WOLFPACK software. A printout giving the results of the adjustment is shown in Figure 16.11. Note that this latter table lists the adjusted coordinates of all new stations, as well as adjusted

```
Example 15.9.4 data for least squares adjustment    {Title line}
12 14 1 1 10          {Number of distances, angles, azimuths, control and total stations}
A 415.273 929.868     {Control station: Identification, X, Y}
B 507.934 764.652     {Unknown station: Identification, X, Y}
C 618.952 815.353
D 723.852 753.287
E 826.128 856.438
F 794.659 1021.655
G 578.741 1103.826
H 652.221 980.245
J 600.595 899.272
K 713.362 877.418
A B 189.436 0.007     {Distance obs: Stations: From, To, Measured Distance, and Std Dev}
B C 122.050 0.007
C D 121.901 0.007
D E 145.256 0.007
E F 168.180 0.007
F G 231.021 0.007
G A 238.714 0.007
G H 143.780 0.007
H K 119.631 0.007
K E 114.695 0.007
H J 96.036 0.007
J C 85.908 0.007
G A B 107 29 40 8.9   {Angle obs: Stations: B, I, F, Measured Angle, and Std Dev}
A B C 94 44 24 11.7
B C D 235 09 26 13.4
C D E 104 08 40 12.7
D E F 124 27 36 11.2
E F G 121 37 08 9.5
F G A 112 23 00 8.3
F G H 38 25 46 9.9
G H J 243 15 20 14.6
H J C 135 08 30 18
J H G 116 44 44 14.6
J H K 296 44 38 15.0
H K E 131 16 30 14.3
K E F 68 40 36 12.3
A B 150 42 51 0.001   {Azimuth obs: Stations: I, F, Observed Azimuth, and Standard Deviation}
```

Figure 16.10 Data file for adjustment of Figure 16.1.

values for all distance and angle observations (obtained by adding the residuals to their corresponding observed values). The adjusted azimuth of line *AB* is also listed, which is the same as its input value. This is expected since that azimuth was held in the adjustment by assigning it a large weight.

These output results contain some of the quantities that are needed to meet newer surveying accuracy standards as was noted in Section 16.1. For those who wish to see problem programmed, it is also demonstrated in the Mathcad® worksheet *hlsq.xmcd*, which is on the companion website for this book.

16.10 THE ERROR ELLIPSE

Error ellipses give a two-dimensional representation of the uncertainties of the adjusted coordinates of points as determined in a least-squares adjustment. They can be plotted at enlarged scales directly on scaled diagrams showing the

****
Adjusted stations
****

Station	Northing	Easting	σn	σe	σu	σv	t
B	764.645	507.938	0.0038	0.0021	0.0044	0.0000	150.06°
C	815.350	618.955	0.0049	0.0046	0.0050	0.0045	159.09°
D	753.286	723.867	0.0069	0.0064	0.0074	0.0058	36.95°
E	856.441	826.133	0.0092	0.0053	0.0093	0.0052	7.55°
F	1,021.654	794.661	0.0086	0.0058	0.0091	0.0049	156.30°
G	1,103.827	578.746	0.0045	0.0058	0.0060	0.0042	111.69°
H	980.245	652.226	0.0061	0.0049	0.0063	0.0047	157.64°
J	899.270	600.599	0.0058	0.0050	0.0058	0.0050	176.23°
K	877.418	713.370	0.0073	0.0056	0.0073	0.0056	176.91°

****
Adjusted Distance Observations
****

Station Occupied	Station Sighted	Distance	V	σ
A	B	189.434	0.0016	0.0044
B	C	122.048	0.0022	0.0045
C	D	121.895	0.0055	0.0044
D	E	145.257	-0.0005	0.0044
E	F	168.184	-0.0040	0.0042
F	G	231.024	-0.0027	0.0042

****
Adjusted Angle Observations
****

Station Backsighted	Station Occupied	Station Foresighted	Angle	V	σ
G	A	B	107°29'39"	0.8"	5.0"
A	B	C	94°44'17"	6.5"	6.4"
B	C	D	235°09'20"	5.8"	7.9"
C	D	E	104°08'39"	1.2"	7.1"
D	E	F	124°27'46"	-9.6"	5.9"
E	F	G	121°37'16"	-7.9"	5.1"
E	G	A	112°23'03"	-2.8"	4.4"

****
Adjusted Azimuth Observations
****

Station Occupied	Station Sighted	Azimuth	V	σ
A	B	150°42'51"	0.0"	0.001"

-----Standard Deviation of Unit Weight = 0.697667-----

Figure 16.11 Abbreviated results from adjustment of data file in Figure 16.10 from WOLFPACK.

points in the horizontal survey. When plotted in this manner, their sizes and appearances enable a quick visual analysis to be made of the overall relative precisions of all adjusted points. As discussed later in this section, this is useful in planning surveys and in analyzing the results of surveys for acceptance or rejection.

On the output listing of Figure 16.11, the adjusted coordinates for the stations in Figure 16.1 are listed, and to their right are columns titled σ_U, σ_V, and t. Respectively, these contain the semimajor axes, semiminor axes, and clockwise rotation angle from the Y-axis to the semimajor axis of the ellipse computed at each station. To compute these three terms, values from the Q_{xx} matrix (see Section 16.5), and the standard deviation of unit weight [see Equations (16.9) and (16.10)] are used with the following formulas.[5]

1. Rotation angle, t

$$\tan(2t) = \frac{2q_{xy}}{q_{yy} - q_{xx}} \tag{16.29}$$

In Equation (16.29), the values of q_{xx} and q_{yy} are the diagonal elements from the Q_{xx} matrix, and q_{xy} is the covariance off-diagonal element in the Q_{xx} matrix for a particular station. When computing t, it is important to establish its quadrant before dividing by 2.

2. *Semimajor axis:* $\sigma_U = \sigma_0\sqrt{q_{xx}\sin^2(t) + 2q_{xy}\cos(t)\sin(t) + q_{yy}\cos^2(t)}$ **(16.30)**

3. *Semiminor axis:* $\sigma_V = \sigma_0\sqrt{q_{xx}\cos^2(t) - 2q_{xy}\cos(t)\sin(t) + q_{yy}\sin^2(t)}$ **(16.31)**

To demonstrate these calculations, part of the Q_{xx} and X matrices that were generated in the least-squares adjustment of the horizontal network in Section 16.9.4 are listed below. (*Note*: These were not shown in the abbreviated output listing of Figure 16.11.) Only those parts of the matrices that pertain to stations B and C are shown. The elements in the rows of the X matrix indicate the order of the unknowns, and identify the elements of Q_{xx} that apply in computing error ellipses. The upper left 2×2 submatrix of Q_{xx} contains the elements that apply to station B. Because X_B is in row 1 of the X matrix, q_{xx} for point B occupies the row 1 column 1, or (1, 1) position. The Y_B coordinate is in row 2 of the X matrix and thus q_{yy} is located in the (2, 2) position of X. Also, the (1, 2) element (or the (2, 1) element which is the same because of symmetry) contains q_{xy}. The following computations yield the error ellipse data for station B.

$$_{18}[(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}]^{18} = \begin{bmatrix} \mathbf{0.00009440} & \mathbf{-0.00001683} & 0.00000739 & -0.00001304 & \cdots \\ \mathbf{-0.00001683} & \mathbf{0.00003001} & -0.00001318 & 0.00002325 & \cdots \\ 0.00000739 & -0.00001318 & 0.0004332 & -0.00000294 & \cdots \\ -0.00001304 & 0.00002325 & -0.00000294 & 0.00004989 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$_{18}X^1 = \begin{bmatrix} \boldsymbol{X_B} \\ \boldsymbol{Y_B} \\ X_C \\ Y_C \\ \vdots \end{bmatrix}$$

[5]For the derivations of these equations, see Ghilani (2010), which is cited in the bibliography at the end of this chapter.

Step 1: By Equation (16.29), compute rotation angle, t.

$$\tan(2t) = \frac{2(-0.00001683)}{0.00003001 - 0.00000944} = -1.59261$$

Since the numerator is negative and the denominator is positive, angle $2t$ is in the fourth quadrant[6] and thus angle t is

$$t = \frac{1}{2}\,[\tan^{-1}(-1.59261) + 360°] = \frac{1}{2}[302.1247164°] = 151°03'44''$$

Step 2: Compute the semimajor axis using Equation (16.30). In these computations, the standard deviation of unit weight, σ_0, is taken from the bottom of the printout given in Figure 16.11, and has been rounded to 0.70.

$$\begin{aligned}\sigma_U &= 0.70\sqrt{0.00000944\ \sin^2(t) + 2(-0.00001683)\ \cos(t)\ \sin(t) + 0.00003001\ \cos^2(t)}\\ &= 0.70\sqrt{0.000039447}\\ &= 0.004\ \text{m}\end{aligned}$$

Step 3: Compute the semiminor axis using Equation (16.31).

$$\begin{aligned}\sigma_V &= 0.70\sqrt{0.00000944\ \cos^2(t) - 2(-0.00001683)\ \cos(t)\ \sin(t) + 0.00003001\ \sin^2(t)}\\ &= 0.70\sqrt{0.000000002}\\ &= 0.000\ \text{m}\end{aligned}$$

Note again that only the upper left 2×2 submatrix of the full $(A^{TWA})^{-1}(Q_{xx})$ matrix was necessary for the computations. To compute the error ellipse for station C, only the (3,3), (3,4), and (4,4) elements of Q_{xx} are needed. This pattern is continued for each station.

In the computed error ellipse for station B of this example, the semiminor axis is nearly zero. Also note that the rotation angle of the semimajor axis closely matches that of the azimuth of the line AB. This could be predicted since the azimuth AB was held fixed during the adjustment by giving it a large weight. This shows both the power and the danger of weights. In this example, the large weight is necessary to fix the horizontal network rotationally in azimuth. If the azimuth had not been fixed by weighting, the network would have been free to rotate about station A and no solution could have been found. However, inappropriate application of weights can cause unwanted corrections in the adjustment. *In least-squares adjustments, it is very important to weight the observations according to their estimated uncertainties*.

---

[6]In Equation (16.29), $\tan 2t = \sin 2t/\cos 2t$. Thus the sines and cosines enable determining the quadrant of $2t$. If the sine and cosine are both plus, that is, the numerator and denominator of Equation (16.29) are both plus, then $2t$ is in the first quadrant (between 0° and 90°). Similarly, if the numerator is plus and the denominator minus, $2t$ is in the second quadrant; and if the numerator and denominator are both minus, $2t$ is in the third quadrant.

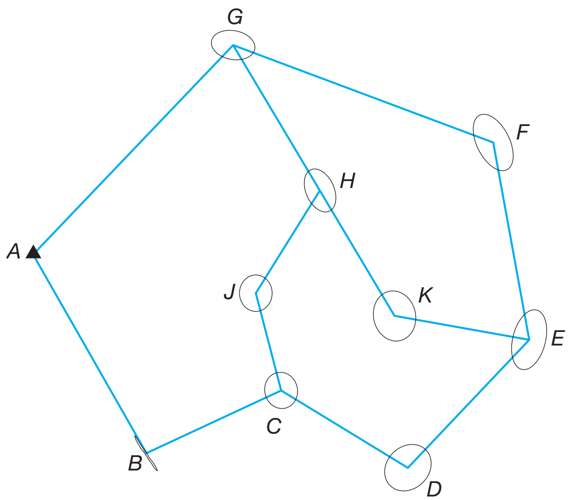


Figure 16.12
Error ellipses plotted at 200 times their actual sizes.

As noted above, when an adjustment is completed, it is often informative to view a graphic of the error ellipses. The error ellipses for this example, shown in Figure 16.12, have had their semiaxes magnified 200 times to make their relatively small values easily visible on the plot. If an error ellipse approximates a circle, this would indicate that point is of approximately equal precision in all directions. Long and slender error ellipses indicate low precisions in their long directions (along their semimajor axes) and high precisions in their narrow directions (along their semiminor axes). In Figure 16.12, it can be seen that the directions of the semimajor axes for stations D, E, F, and G are aligned on circles with their centers approximately at A. This shows the rotational instability in the observational system, that is, the held azimuth of line AB alone fixes the rotation in the network. This instability could be improved by observing an azimuth on another line such as EF, or by connecting either stations E or F to another neighboring control point (assuming one is available). In Figure 16.12 it can also be seen that the largest uncertainties exist on the stations that are farthest from the control station. This is the usual manner that errors propagate in observational systems—that is, the farther the unknown stations are from the control, the larger the estimated errors in their coordinates. With this in mind, the sizes of the errors at all stations, but especially E and F, could be reduced by connecting either of these stations to a neighboring control station (assuming one is available).

The analysis of plotted error ellipses is very useful in scrutinizing results of adjusted horizontal surveys, as illustrated by the simple analyses given in the previous paragraph. As demonstrated, both the sizes and shapes of error ellipses give an immediate visual impression about the relative accuracies of the points in a network and enable the most efficient plan to be developed for making additional observations to strengthen the observational scheme.

16.11 ADJUSTMENT PROCEDURES

Regardless of the nature of the specific adjustment problem, certain procedures should be followed. For example, before the adjustment is undertaken, all data must be carefully analyzed for blunders. Mistakes such as station misidentifications, transcription mistakes, reading blunders, and others must be identified and corrected. Failure to remove them will result in either an unsatisfactory adjustment or no adjustment at all. In several types of surveys, performing loop closures on the data can identify blunders. This is true in leveling, in GNSS networks, and in horizontal surveys, including traversing. Also, in traverses the methods discussed in Section 10.16 can be employed to detect blunders.

The minimum amount of control required for making adjustments varies with the type of problem. In differential leveling only one benchmark is needed, and in a network of GNSS baseline observations, only one station with known coordinate values is necessary. For horizontal surveys such as traverses and networks, one station with known coordinates and one course with known direction must be available. If more than the minimum amount of control is present, the adjustment should be performed in two stages as a further means of detecting blunders. The first adjustment, called a *minimally constrained adjustment*, should contain only the minimum amount of control. This should then be followed by a *constrained adjustment*, in which all available control is used.

The minimally constrained adjustment provides checks on the geometric closures and the consistency of the observations. After the adjustment is completed, the residuals should be analyzed. Any unusually large residuals, or a preponderance of minus or plus residuals (a group of random errors should contain approximately equal numbers of positive and negative values), will be keys to the existence of blunders in the data. However, even though a minimally constrained adjustment will check the data for *internal consistency*, it may fail to identify the presence of systematic errors and mistakes. For instance, assume during the observation of the leveling circuit shown in Figure 16.13 that a mistake of +1 m occurred during the observation of line 1 and that a −1 m error occurred during the observation of line 7. Even though every observed benchmark would have a 1-m error, the loop misclosure would not uncover this mistake and the observations would geometrically close. However, in that figure, if an additional line of differential levels is observed from BM 2 to *D*, the presence of these mistakes would become apparent.

Upon acceptance of the minimally constrained adjustment, all additional available control should be added to the data and the constrained adjustment

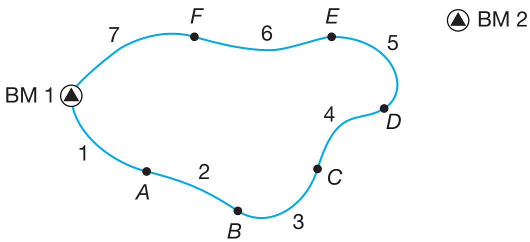


Figure 16.13
Differential leveling loop.

performed. This will aid in identifying compensating or systematic errors in the data and blunders in control. For instance, suppose that the coordinate values for control station A in Figure 16.12 are state plane coordinate values, but the distances are not properly reduced to the state plane grid (see Chapter 20). In this situation, the geometric closures of the adjustment could appear to be fine, but, because of the distance scaling error, the computed coordinates for the unknown stations would all be incorrect. In this case, the minimally constrained adjustment would fail to catch the scaling error. However, if observations had been made to connect the network to a second control station with known state plane coordinates, the scaling error would become apparent in the constrained adjustment. Similarly, a blunder in the control coordinates or benchmarks will not become apparent until the constrained adjustment is performed. Thus, it is important to perform a minimally constrained adjustment to obtain geometric checks on the data and a constrained adjustment to find possible compensating errors, systematic errors, and control blunders. Therefore it follows logically that every survey should have redundant observations to provide geometric checks and isolate mistakes caused by careless work.

Upon acceptance of the constrained adjustment, the post-adjustment statistics, as given in Sections 16.5, 16.7, and 16.10, should be computed. When possible, these statistical values should be compared against published accuracy standards. If the survey fails to meet the required standards, additional observations should be taken or the work repeated using more precise equipment.

16.12 OTHER MEASURES OF PRECISION FOR HORIZONTAL STATIONS

Since error ellipses are part of a bivariate distribution, their probability level is approximately 39%. Generally, surveyors prefer to state their results at a much higher level of confidence. For the semiminor and semimajor axes of the error ellipse, this is accomplished by using a multiplier that is based on critical values taken from the F distribution. This distribution is a function of the number of *degrees of freedom* (number of redundant observations) that existed in the adjustment. Some of the critical values from the F distribution are shown in Table 16.5. The multipliers for the semimajor and semiminor axes of an error ellipse expressed at other probabilities levels are determined from

$$c = \sqrt{2(F_{\alpha,2,\text{ degrees of freedom}})} \tag{16.32}$$

where the semimajor and semiminor axes would be computed as

$$\sigma_{U\%} = c\sigma_u \tag{16.33}$$

$$\sigma_{V\%} = c\sigma_v \tag{16.34}$$

Other measures of precision that can be used include the *circular error probable* (CEP), which is a function of the computed standard deviations for a horizontal station, or

$$\text{CEP} = 0.5887(\sigma_X + \sigma_Y) \tag{16.35}$$

TABLE 16.5 $F_{\alpha,2,\text{degrees of freedom}}$ **Critical Statistic Values for Selected Probabilities**

Degrees of Freedom	90%	Probability 95%	99%
1	49.50	199.50	4999.50
2	9.00	19.00	99.00
3	5.46	9.55	30.82
4	4.32	6.94	18.00
5	3.78	5.79	13.27
9	3.01	4.26	8.02
10	2.92	4.10	7.56
15	2.70	3.68	6.36
20	2.59	3.49	5.85
30	2.49	3.32	5.39

The 90% region of the CEP is called the *circular map accuracy standard* (CMAS) and is computed as

$$\text{CMAS} = 1.8227\ \text{CEP} \tag{16.36}$$

The *distance root mean square* (DRMS) error is another measure of precision. It can be computed as

$$\text{DRMS} = \sqrt{\sigma_X^2 + \sigma_Y^2} \tag{16.37}$$

Example 16.10

What are the 95% probability values for the semimajor and semiminor axes of station B in Section 16.10?

Solution

The adjustment of Figure 16.1 had 12 distance observations, 14 angle observations, and 1 azimuth observation for a total of 27 observations. There were 9 stations having 2 unknowns each for a total of 18 unknowns. Thus the number of degrees of freedom are 27 − 18, or 9. From Table 16.5, the appropriate F value for 9 degrees of freedom is 4.26. By Equation (16.32), the c-multiplier is

$$c = \sqrt{2(4.26)} = 2.92$$

From Section 16.10, the values for σ_U and σ_V were 0.004 and 0.000 m, respectively. Thus by Equations (16.33) and (16.34), the 95% values for the semimajor and semiminor axes are

$$\sigma_{U95\%} = 2.92\ (0.004) = \pm 0.012\ \text{m}$$
$$\sigma_{V95\%} = 2.92\ (0.000) = \pm 0.000\ \text{m}$$

16.13 SOFTWARE

In a typical surveying office, software is used to perform the adjustments discussed in this chapter. On the companion website for this book at http://www.pearsonhighered.com/ghilani is the WOLFPACK program, which incorporates some of the adjustments discussed in this chapter. The help file, which accompanies the software, describes the formats for each type of data file supported. For those wishing to program these adjustments in a higher-level programming language, several Mathcad® worksheets are available on the companion website also. In particular the worksheet *LLSQ.XMCD* demonstrates a differential leveling least-squares adjustment, *HLSQ.XMCD* demonstrates a horizontal survey least-squares adjustment, and *GPS.XMCD* demonstrates a GNSS network least-squares adjustment. Additionally, the Excel spreadsheet *C16.XLS* on the companion website demonstrates the computations in Examples 16.4, 16.5, 16.6, and the solution of nonlinear equations presented in Section 16.9.

16.14 CONCLUSIONS

As discussed in Chapter 3, the presence of errors in observations is inevitable. However, if the method of least squares is employed, the sizes of the errors can be assessed and if they are within acceptable limits, the observations can be adjusted to determine the most probable values for the unknowns. If some of the observations contain unacceptable errors, these observations must be repeated before the final adjustment is made. The advantages of the least-squares method over other adjustment techniques are many. Some of these are that it: (1) conforms to the laws of probability, (2) provides the most probable solution for a given set of observations, (3) allows individual weighting of observations, (4) forces geometric closures on the observations, (5) simultaneously adjusts all observations, (6) provides a single unique solution for a set of data, and (8) yields estimated precisions of adjusted quantities. The least-squares method is readily programmed for computer solution and the data is easily prepared for making adjustments. Because of these advantages and the fact that data from least-squares adjustments are now necessary for assessing compliance of surveys with modern standards such as the FGCS standards and specifications for GNSS Relative Positioning (see Section 14.5.1) and the ALTA-ACSM Land Title Survey Standards (see Section 21.10), all surveying offices should employ the method.

In this chapter, the basic theory of least-squares adjustment has been presented and its application to common surveying observations demonstrated. For further information on least squares, the reader is directed to the references in the bibliography.

PROBLEMS

Asterisks (*) indicate problems that have partial answers given in Appendix G.

16.1 What fundamental condition is enforced by the method of weighted least squares?

16.2 What are the advantages of adjusting observations by the method of least squares?

16.3 What are the advantages of the method of least squares over methods of adjustment?

16.4* What is the most probable value for the following set of 10 distance observations in meters? 532.688, 532.682, 532.682, 532.684, 532.689, 532.686, 532.690, 532.684, 532.686, 532.686.

16.5 What is the standard deviation of the adjusted value in Problem 16.4?

16.6 Three horizontal angles were observed around the horizon of station A. Their values are 165°07′54″, 160°25′36″, and 34°26′36″. Assuming equal weighting, what are the most probable values for the three angles?

16.7 What are the standard deviations of the adjusted values in Problem 16.6?

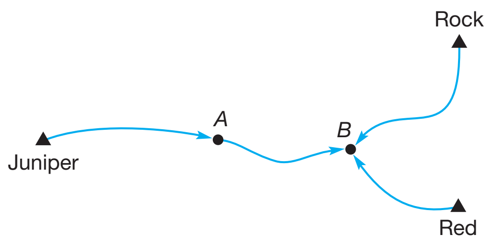
16.8 In Problem 16.6, the standard deviations of the three angles are ±1.5″, ±3.0″, and ±4.9″, respectively. What are the most probable values for the three angles?

16.9* Determine the most probable values for the x and y distances of Figure 16.2, if the observed lengths of AC, AB, and BC (in meters) are 294.081, 135.467, and 158.607, respectively.

16.10* What are the standard deviations of the adjusted values in Problem 16.9?

16.11 A network of differential levels is run from existing benchmark Juniper through new stations A and B to existing benchmarks Red and Rock as shown in the accompanying figure. The elevations of Juniper, Red, and Rock are 685.65, 696.75, and 705.27 ft, respectively. Develop the observation equations for adjusting this network by least squares, using the following elevation differences.

From	To	Elev. Diff. (ft)	(ft)
Juniper	A	−40.58	0.021
A	B	8.21	0.010
B	Red	43.44	0.020
B	Rock	51.96	0.026



Problem 16.11

16.12 For Problem 16.11, following steps outlined in Example 16.6 perform a weighted least-squares adjustment of the network. Determine weights based upon the given standard deviations. What are the:

(a) Most probable values for the elevations of A and B?
(b) Standard deviations of the adjusted elevations?
(c) Standard deviation of unit weight?
(d) Adjusted elevation differences and their residuals?
(e) Standard deviations of the adjusted elevation differences?

16.13 Repeat Problem 16.12 using distances for weighting. Assume the following course lengths for the problem.

From	To	Distance (ft)
Juniper	*A*	1500
A	*B*	300
B	Red	1200
B	Rock	2300

16.14 Use WOLFPACK to do Problem 16.12 and 16.13 and compare the solutions for *A* and *B*.

16.15 Repeat Problem 16.12 using the following data.

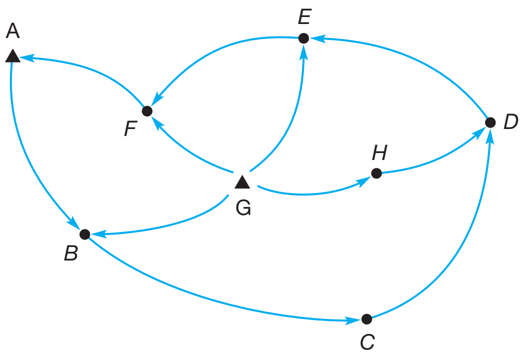
From	To	Elev. Diff. (ft)	σ(ft.)
Juniper	*A*	−40.62	0.027
A	*B*	8.19	0.015
B	Red	43.48	0.029
B	Rock	51.95	0.027

16.16 A network of differential levels is shown in the accompanying figure. The elevations of benchmarks *A* and *G* are 435.235 and 465.643 m, respectively. The observed elevation differences and the distances between stations are shown in the following table. Using WOLFPACK, determine the

(a) Most probable values for the elevations of new benchmarks *B*, *C*, *D*, *E*, *F*, and *H*?

From	To	Elev. Diff (m)	S (m)
A	*B*	−19.411	0.127
B	*C*	5.007	0.180
C	*D*	24.436	0.154
D	*E*	10.414	0.137
E	*F*	−17.974	0.112
F	*A*	−4.797	0.106
G	*F*	−25.655	0.112
G	*H*	−7.810	0.112
H	*D*	−13.011	0.112
G	*B*	−49.785	0.122
G	*E*	−10.477	0.117

(b) Standard deviations of the adjusted elevations?
(c) Standard deviation of unit weight?
(d) Adjusted elevation differences and their residuals?
(e) Standard deviations of the adjusted elevation differences?



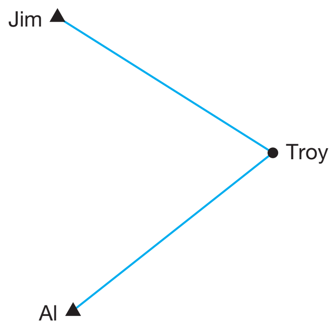
Problem 16.16

16.17 Develop the observation equations for Problem 16.16.

16.18 A network of GNSS observations shown in the accompanying figure was made with two receivers using the static method. Known coordinates of the two control stations are in the geocentric system. Develop the observation equations for the following baseline vector components.

Station	***X*(m)**	***Y*(m)**	***Z*(m)**
Jim	34.676	−4,497,514.405	4,507,737.916
Al	−268.920	−4,588,971.649	4,415,199.130

Jim to Troy				**Al to Troy**			
76,399.646	3.35E-03	2.79E-04	1.55E-03	76,703.377	3.37E-03	2.17E-04	5.24E-07
−45,109.020		3.18E-03	7.93E-05	46,348.167		3.45E-03	1.62E-05
−45,766.394			3.32E-03	46,772.390			3.38E-03



Problem 16.18

16.19 For Problem 16.18, construct the A and L matrices.

16.20 For Problem 16.18, construct the covariance matrix.

16.21 Use WOLFPACK to adjust the baselines of Problem 16.18.

16.22 Convert the geocentric coordinates obtained for station Troy in Problem 16.21 to geodetic coordinates.

16.23 A network of GNSS observations shown in the accompanying figure was made with two receivers using the static method. Use WOLFPACK to adjust the network, given the following data.

Station	***X*(m)**	***Y*(m)**	***Z*(m)**
Bonnie	−1,660,596.783	−4,718,761.893	3,944,402.433
Tom	−1,622,711.656	−4,733,328.952	3,942,760.219

Bonnie to Ray

54,807.272	3.48E-03	2.89E-04	1.62E-03
−65,078.175		3.61E-03	6.16E-05
−55,773.186			3.39E-03

Bonnie to Herb

80,093.477	3.07E-03	7.87E-05	5.88E-05
14,705.261		3.13E-03	1.36E-04
49,804.206			3.06E-03

Tom to Ray

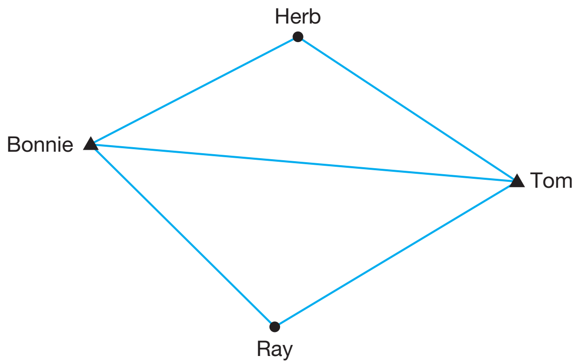
16,922.045	1.93E-03	4.49E-05	7.39E-05
−50,511.098		1.98E-03	−3.91E-06
−54,130.990			1.99E-03

Tom to Herb

42,208.529	1.79E-03	1.12E-05	4.48E-05
29,272.300		1.78E-03	5.65E-05
51,446.338			1.80E-03

Bonnie to Tom (Fixed line—Don't use in adjustment.)

37,885.108	5.55E-04	1.31E-05	1.21E-05
−14,567.048		5.55E-04	8.04E-06
−1,642.215			5.54E-04



Problem 16.23 through 16.28

16.24 For Problem 16.23, write the observation equations for the baselines "Bonnie to Ray" and "Tom to Herb."

16.25 For Problem 16.23, construct the A and L matrices for the observations.

16.26 For Problem 16.23, construct the covariance matrix.

16.27* After completing Problem 16.23, convert the geocentric coordinates for station Ray and Herb to geodetic coordinates. (*Hint*: See Section 13.4.3.)

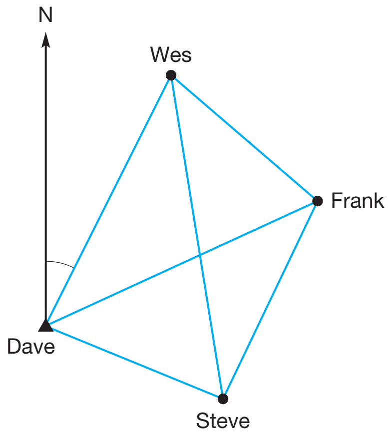
16.28 Following the procedures discussed in Section 14.5.2, analyze the fixed baseline from station Bonnie to Tom.

16.29 For the horizontal survey of the accompanying figure, determine initial approximations for the unknown stations. The observations for the survey are

Station	***X*(ft)**	***Y*(ft)**		**From**	**To**	**Azimuth**	***S***
Dave	2515.62	1941.12		Dave	Wes	23°43′46″	0.001″

From	To	Distance (ft)	σ(ft)
Dave	Steve	2049.59	0.016
Steve	Frank	2089.17	0.016
Frank	Wes	2639.06	0.017
Wes	Dave	3179.50	0.018
Dave	Frank	3181.90	0.018
Steve	Wes	3759.69	0.019

Backsight Station	Instrument Station	Foresight Station	Angle	σ
Wes	Dave	Frank	49°01′12″	3.1″
Frank	Dave	Steve	40°12′36″	3.2″
Dave	Steve	Wes	57°44′15″	3.2″
Wes	Steve	Frank	42°45′08″	3.2″
Steve	Frank	Dave	39°17′59″	3.2″
Dave	Frank	Wes	65°26′35″	3.2″
Frank	Wes	Steve	32°30′13″	3.2″
Steve	Wes	Dave	33°01′59″	3.1″



Problem 16.29 through 16.32

16.30* Using the data in Problem 16.29, write the linearized observation equation for the distance from Steve to Frank.

16.31 Using the data in Problem 16.29, write the linearized observation equation for the angle Wes-Dave-Frank.

16.32 Assuming a standard deviation ±0.001″ of for the azimuth line Dave-Wes, use WOLFPACK to adjust the data in Problem 16.29.

16.33* Given the following inverse matrix and a standard deviation of unit weight of 1.23, determine the parameters of the error ellipse.

$$(A^TWA)^{-1} = \begin{bmatrix} q_{xx} & q_{xy} \\ q_{xy} & q_{yy} \end{bmatrix} = \begin{bmatrix} 0.00023536 & 0.00010549 \\ 0.00010549 & 0.00033861 \end{bmatrix}$$

16.34 Compute S_x and S_y in Problem 16.33.

16.35 Given the following inverse matrix and a standard deviation of unit weight of 1.45, determine the parameters of the error ellipse.

$$(A^TWA)^{-1} = \begin{bmatrix} q_{xx} & q_{xy} \\ q_{xy} & q_{yy} \end{bmatrix} = \begin{bmatrix} 0.0004894 & 0.0000890 \\ 0.0000890 & 0.0002457 \end{bmatrix}$$

16.36 Compute S_x and S_y in Problem 16.35.

16.37 The well-known observation equation for a line is $mx + b = y + v_y$. What is the slope and y-intercept of the best fit line for a set of points with coordinates of (478.72, 3517.64), (1446.81, 2950.40), (2329.79, 2432.66), (3345.74, 1837.13), (4382.98, 1229.16)?

16.38 Use WOLFPACK and the following standard deviations for each observation to do a least-squares adjustment of Example 10.4, and describe any differences in the solution. What advantages are there to using the least-squares method in adjusting this traverse?

Stations	Angle ± S	Station	Distance ± S
E-A-B	100°45′37″ ± 16.7″	*AB*	647.25 ± 0.027
A-B-C	231°23′43″ ± 22.1″	*BC*	203.03 ± 0.026
B-C-D	17°12′59″ ± 21.8″	*CD*	720.35 ± 0.027
C-D-E	89°03′28″ ± 10.2″	*DE*	610.24 ± 0.027
D-E-A	101°34′24″ ± 16.9″	*EA*	285.13 ± 0.026
AZIMUTH *AB*	126°55′17″ ± 0.001″		

BIBLIOGRAPHY

Bell, J. 2003. "MOVE3." *Professional Surveyor* 23 (No. 11): 46.

Ghilani, C. D. 2003. "Statistics and Adjustments Explained—Part 1: Basic Concepts." *Surveying and Land Information Science* 63 (No. 2): 73.

_____. 2003. "Statistics and Adjustments Explained—Part 2: Sample Sets and Reliability." *Surveying and Land Information Science* 63 (No. 3): 141.

_____. 2004. "Statistics and Adjustments Explained—Part 3: Error Propagation." *Surveying and Land Information Science* 64 (No. 1): 23.

_____. 2010. *Adjustment Computations: Spatial Data Analysis*. New York, NY: Wiley.

Schwarz, C. R. 2005. "The Effects of Unestimated Parameters." *Surveying and Land Information Science* 65 (No. 2): 87.

Tan, W. 2002. "In What Sense a Free Net Adjustment?" *Surveying and Land Information Science* 62 (No. 4): 251.

17 Mapping Surveys

17.1 INTRODUCTION

Mapping surveys are made to determine the locations of *natural* and *cultural* features on the Earth's surface and to define the configuration (*relief*) of that surface. Once located, these features can be represented on maps. Natural features normally shown on maps include vegetation, rivers, lakes, oceans, etc. Cultural (*artificial*) features are the products of people and include roads, railroads, buildings, bridges, canals, boundary lines, etc. The relief of the Earth includes its hills, valleys, plains, and other surface irregularities. Lines and symbols are used to depict features shown on maps. Names and legends are added to identify the different objects shown.

Two different types of maps, *planimetric* and *topographic*, are prepared as a result of mapping surveys. The former depicts natural and cultural features in the plan (X-Y) views only. Objects shown are called *planimetric features*. Topographic maps also include planimetric features, but in addition they show the configuration of the Earth's surface. Both types of maps have many applications. They are used by engineers and planners to determine the most desirable and economical locations of highways, railroads, canals, pipelines, transmission lines, reservoirs, and other facilities; by geologists to investigate mineral, oil, water, and other resources; by foresters to locate access- or haul-roads, fire-control routes, and observation towers; by architects in housing and landscape design; by agriculturists in soil conservation work; and by archeologists, geographers, and scientists in numerous fields. Maps are used extensively in geographic information system (GIS) applications (see Chapter 28). Conducting the surveys necessary for preparing maps and the production of the maps from the survey data are the mainstay of many surveying businesses.

Relief is shown on maps by using various conventions and procedures. For topographic maps, *contours* are most commonly used and are preferred by surveyors and engineers. *Digital elevation models* (DEMs) and *three-dimensional perspective models* are newer methods for depicting relief, made possible by computers. *Color, hachures, shading*, and *tinting* can also be used to show relief, but these methods are not quantitative enough and thus are generally unsuitable for surveying and engineering work. Contours, digital elevation models, and three-dimensional perspective models are discussed in later sections of this chapter and in Chapter 18.

Traditionally, maps were prepared using manual drafting methods. Now however, as described in Chapter 18, the majority of maps are produced using computers, computer-aided drafting (CAD) software, and data collectors. Currently, some data collectors include drafting software so that field personnel can display their data in the field to check for mistakes and missing elements. This chapter discusses procedures for collecting planimetric and topographic mapping data.

17.2 BASIC METHODS FOR PERFORMING MAPPING SURVEYS

Mapping surveys are conducted by one of two basic methods: *aerial* (photogrammetric) or *ground* (field) techniques, but often a combination of both is employed. Refined equipment and procedures available today have made photogrammetry very accurate and economical. Hence, almost all mapping projects covering large areas now employ this method. However, as discussed in Section 27.18, airborne laser mapping systems may also be used. Ground surveys are still commonly used in preparing large-scale maps of smaller areas. Even when photogrammetry or airborne laser mapping is utilized, ground surveys are necessary to establish control and to field-check mapped features for accuracy. This chapter concentrates on ground methods and describes several field procedures for locating topographic features, both horizontally and vertically. Photogrammetry and airborne laser mapping are discussed in Chapter 27.

17.3 MAP SCALE

Map scale is the ratio of the length of an object or feature on a map to the true length of the object or feature. Map scales are given in three ways: (1) by *ratio* or *representative fraction*, such as 1:2000 or 1/2000; (2) by an *equivalence*, for example, 1 in. = 200 ft; and (3) by graphically using either a bar scale or labeled grid lines spaced throughout the map at uniform distances apart. Graphic scales permit accurate measurements to be made on maps, even though the paper upon which the map is printed may change dimensions.

An equivalence scale of 1 in./100 ft indicates that 1 in. on the map is equivalent to 100 ft on the object. In giving scale by ratio or representative fraction, the same units are used for the map distance and the corresponding object distance, and thus 1:1200 could mean 1 in. on the map is equivalent to 1200 in. on the object, but any other units would also apply. Obviously, it is possible to convert from an equivalence scale to a ratio, and vice versa. As an example, 1 in. = 100 ft is converted to a ratio by multiplying 100 ft by 12, which converts it to inches and gives a ratio of 1:1200. Those engaged in surveying (geomatics) and engineering generally prefer an equivalence scale and grid lines on their maps, while geographers often utilize a representative fraction and bar scale.

Choice of scale depends on the purpose, size, and required precision of the finished map. Dimensions of a standard map sheet, type and number of topographic symbols used, and accuracy requirements for scaling distances from the map are some additional considerations. Maps produced using the English system of units usually have their scales selected to be compatible with one of the standard graduations on engineer's scales. These standard graduations have 10, 20, 30, 40, 50, or 60 units per inch. Thus, scales of 1 in. = 100 ft and 1 in. = 1000 ft are compatible with the 10 scale; 1 in. = 200 ft and 1 in. = 2000 ft are consistent with the 20 scale, and so on. In the metric system, ratios or representative fractions such as 1:1000, 1:2000, 1:5000, and so on are usually employed.

Map scales may be classified as *large*, *medium*, and *small*. Their respective scale ranges are as follows:

Large scale, 1 in. = 200 ft (1:2400) or larger

Medium scale, 1 in. = 200 ft to 1 in. = 1000 ft (1:2400 to 1:12,000)

Small scale, 1 in. = 1000 ft (1:12,000) smaller

Large-scale maps are applied where relatively high accuracy is needed over limited areas; for example, in subdivision design and the design of engineering projects like roads, dams, airports, and water and sewage systems. Medium scales are often used for applications such as general preliminary planning where larger areas are covered but only moderate accuracy is needed. Applications include mapping the general layout of potential construction sites, proposed transportation systems, and existing facilities. Small-scale maps are commonly used for mapping large areas where a lower order of accuracy will suffice. They are suitable for general topographic coverage, applications in site-suitability analysis, preliminary layout of expansive proposed construction projects, and for special applications in forestry, geology, environmental impact and management, etc.

Maps in graphic form can have their scales enlarged or reduced photographically or by converting the maps to digital form and enlarging or reducing by computer processing. The enlargement ratios possible by either of these methods are virtually unlimited. *However, enlargements must be produced with caution since any errors in the original maps or digital data are also magnified, and the enlarged product may not meet required accuracy standards*.

The scale at which a map will be plotted directly affects the choice of instruments and procedures used in performing the mapping survey. This is because the accuracy with which the position of an object is depicted on a map is related to the map's scale, which in turn dictates the accuracy with which features must be surveyed. Consider, for example, a map plotted at a scale of 1 in. = 20 ft. If distances and locations can be scaled from the map to within say 1/50th in., this represents a scaling error of $(1/50)20 = \pm 0.4$ ft. To ensure that the accuracy of the surveyed data does not limit the accuracy with which information can be scaled from a map, features must be located on the map to an accuracy better than ±0.4 ft for this map. As a safety factor, many surveying and mapping agencies apply a rule of thumb in which they require features to be located in the field to at least twice the scaling accuracy, which in this instance would require accuracy to within ±0.2 ft or better. Following this same reasoning, if map scale is 1 in. = 200 ft, then ground features should be located to an accuracy of ±2 ft so as not to limit the accuracy of the map. Another consideration regarding map scale that affects surveying accuracy is the thicknesses of lines used to plot features. Assume, for example, that line widths on a map with a scale of 1:2000 are 0.3 mm. This means that each line represents 0.3(2000) = 600 mm = 0.6 m on the object. Therefore, to accurately depict an object on a map with this line width, the survey needs to be accurate to at least half the line width, or ±0.3 m. Obviously, the equipment and procedures used for the mapping work must be selected so that these accuracies are met.

While scaling factors such as those discussed above must be taken into account for each specific mapping project, it is important to also consider the possible future use of the map data being collected. Thus, even though the first map produced for a particular project may be a small-scale reconnaissance map, it is possible that as the project progresses, medium-scale planning maps and large-scale design maps will be needed, and that some or all of the data collected could also be used for these maps. Thus, even though relaxed accuracies may suffice for the reconnaissance map, for efficiency, the data should be collected to accuracy suitable for other maps that may follow.

17.4 CONTROL FOR MAPPING SURVEYS

Whether the mapping is done by ground or aerial methods, the first requirement for any project is good control. As discussed in Chapter 19, control is classified as either horizontal or vertical.

Horizontal control for a mapping survey is provided by two or more points on the ground, permanently or semipermanently monumented, and precisely fixed in position horizontally by distance and direction or coordinates. It is the basis for locating map features. Horizontal control can be established by the traditional ground surveying methods of *traversing, triangulation*, or *trilateration* (see Section 19.12), or by using *GNSS* surveys (see Chapters 14 and 15). For large areas, a sparse network of horizontal (and vertical) control can be densified using *photogrammetry* (see Chapter 27). For small areas, horizontal control for mapping surveys is generally established by traversing. Until recently, triangulation and trilateration were the most economical procedures available for establishing basic control for mapping projects extending over large areas such as a state or the entire United States. These techniques have now given way to GNSS surveys, which is not only highly accurate but also very efficient. Monuments whose positions have been established through higher-order control surveys and referenced in the state plane coordinate systems (see Chapter 20) are used to initiate surveys of all types, but unfortunately more are needed in most areas. However, GNSS surveys can bring state plane coordinates into any region where satellite visibility is available.

Vertical control is provided by benchmarks in or near the tract to be surveyed and becomes the foundation for correctly portraying relief on a topographic map. Vertical control is usually established by running lines of differential levels starting from and closing on established benchmarks (see Chapters 4 and 5). Project benchmarks are established throughout the mapping area in strategic locations and their elevations determined by including them as turning points in the differential leveling lines. In rugged areas, trigonometric leveling with total station instruments is practical and frequently used to establish vertical control for mapping. GNSS surveys are also suitable for establishing vertical control for topographic mapping, but the ellipsoidal heights derived from GNSS surveys must first be converted to orthometric heights using Equation (13.8). The latter two methods are of sufficient accuracy to support most mapping surveys.

Regardless of the methods used in conducting the control surveys for mapping projects, specified maximum allowable closure errors for both horizontal and vertical control should be determined prior to the fieldwork, then used to guide it. The locations of the features, which comprise the map (often also called *map details*), are based upon the framework of control points whose positions and elevations are established. Thus, any errors in the surveyed positions or elevations of the control points will result in erroneous locations of the details on the map. Therefore, it is advisable to run, check, and adjust the horizontal and vertical control surveys before beginning to locate map details, rather than carry on both processes simultaneously. The method selected for locating map details will govern the speed, cost, and efficiency of the survey. In later sections of this chapter, the different basic field procedures and the varying equipment that can be used are described.

■ 17.5 CONTOURS

As stated earlier, surveyors and engineers most often use contours to depict relief. The reason is that they provide an accurate quantitative representation of the terrain. Because planimetric features and contours are located simultaneously in most field topographic surveys, it is important to understand contours and their characteristics before discussing the various field procedures used to position them.

A *contour is a line connecting points of equal elevation*. Since water assumes a level surface, the shoreline of a lake is a visible contour, but in general, contours cannot be seen in nature. On maps, contours represent the planimetric locations of the traces of level surfaces for different elevations (see the plan view of Figure 17.1). Contours are drawn on maps by interpolating between points whose positions and elevations have been observed and plotted. As noted earlier, computerized mapping and contouring systems are replacing manual plotting methods, but the principles of plotting terrain points and of interpolating contours are still basically the same in either method.

The vertical distance between consecutive level surfaces forming the contours on a map (the elevation difference represented between adjacent contours) is called the *contour interval*. For the small-scale U.S. Geological Survey *quadrangle maps* (plotted at 1:24,000 scale), depending on the nature of the terrain one of the following contour intervals is used: 5, 10, 20, 40, or 80 ft. For larger-scale

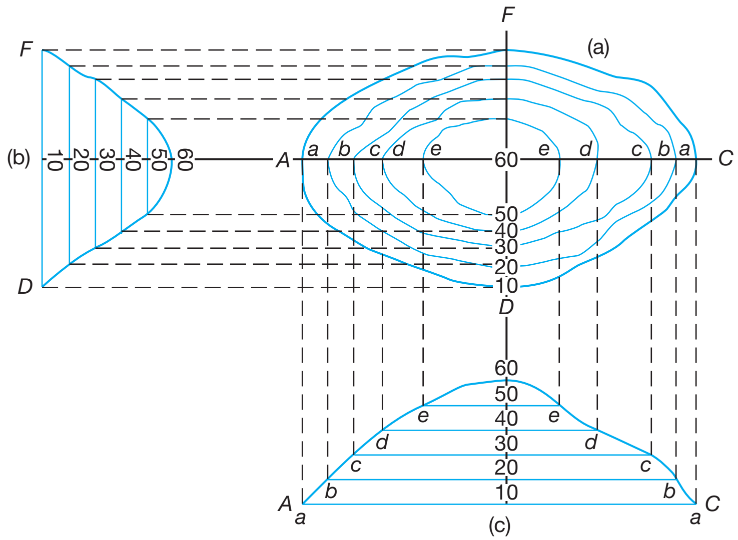


Figure 17.1
(a) Plan view of contour lines, (b) and (c) profile views.

maps used in engineering design, in the English system of units contour intervals of 1, 2, 5, or 10 ft are commonly used. In the metric system, a contour interval of 0.5, 1, 2, 5, or 10 m is generally selected. Figure 17.2 is a topographic map having 10-ft contours.

The contour interval selected depends on a map's purpose and scale, and upon the diversity of relief in the area. As examples, on a map to be used for designing the streets and water and sewer systems for a subdivision, a contour interval of 1 or 2 ft would perhaps be necessary, whereas a 10- or 20-ft contour interval may be suitable for mapping a large ravine to determine the reservoir capacity that would result from constructing a dam. Also, a smaller contour interval will normally be necessary to adequately depict gently rolling terrain with only moderate elevation differences, while rugged areas with large elevation differences normally require a larger contour interval so that the contours do not become too congested on the map. In general, reducing the contour interval requires more costly and precise fieldwork. In regions where both flat coastal areas and mountainous terrain are included in a map, supplementary contours, at one half or one fourth the basic contour interval, are often drawn (and shown with dashed lines).

Spot elevations are used on maps to mark unique or critical points such as peaks, potholes, valleys, streams, and highway crossings. They may also be used in lieu of contours for defining elevations on relatively flat terrain that extends over a large area.

Topographic mapping convention calls for drawing only those contours that are evenly divisible by the contour interval. Thus, for the 10-ft contour interval on the map in Figure 17.2, contours such as the 1100, 1110, 1120, and 1130 are shown. Elevations are shown in breaks in the contour lines, and to avoid confusion, at least every fifth contour is labeled. To aid in reading topographic maps, every fifth contour (each that is evenly divisible by five times the contour interval) is drawn using a heavier line. Thus, in Figure 17.2, the 1100, 1150, 1200, and so on contours are drawn more heavily.

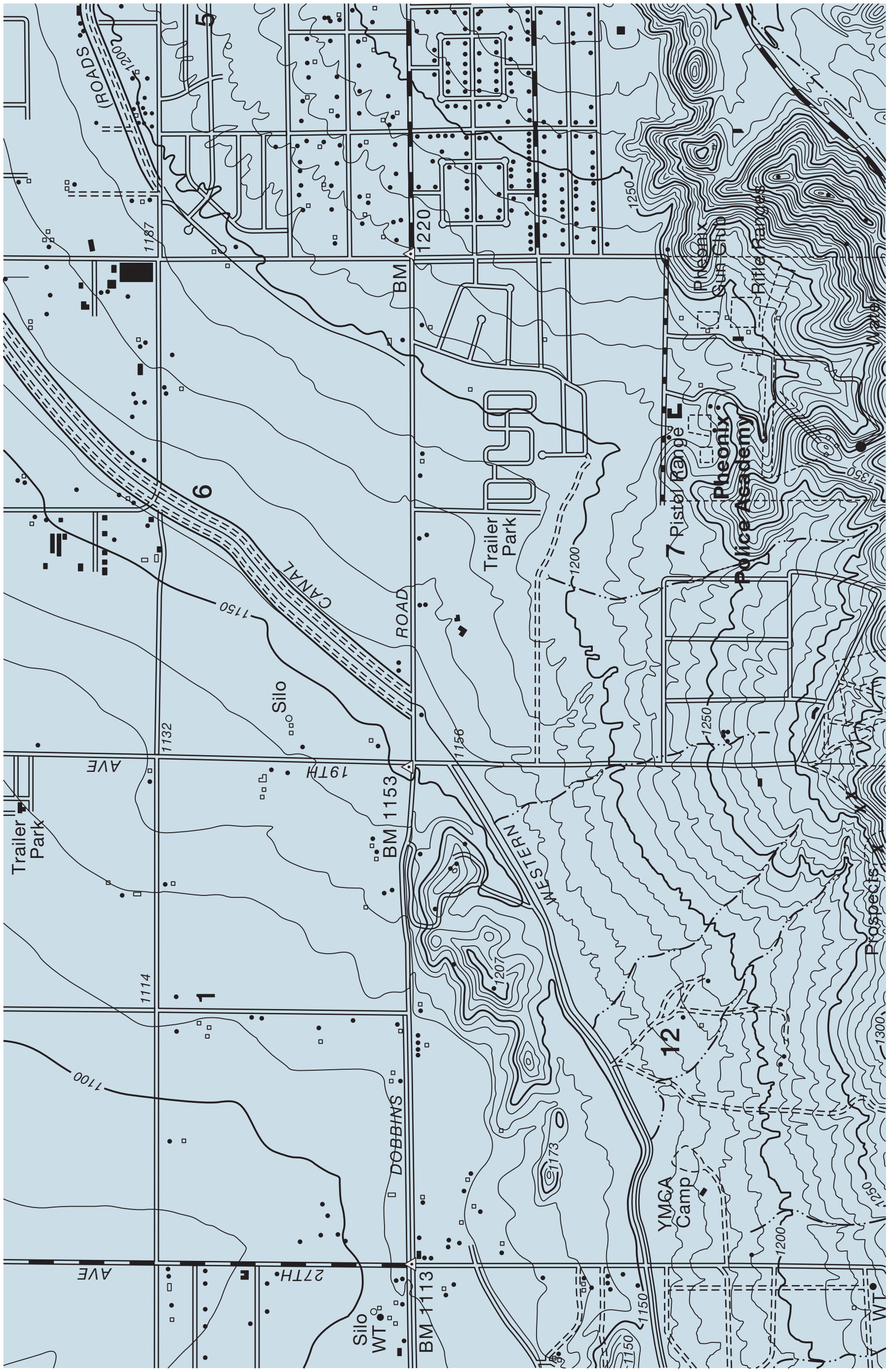


Figure 17.2 Part of U.S.G.S. Lone Butte quadrangle map. (Courtesy U.S. Geological Society.)

17.6 CHARACTERISTICS OF CONTOURS

Although each contour line in nature has a unique shape, all contours adhere to a set of general characteristics. Important ones, fundamental to their proper field location and correct plotting, are listed.

1. Contour lines must close on themselves, either on or off a map. They cannot dead end.
2. Contours are perpendicular to the direction of maximum slope.
3. The slope between adjacent contour lines is assumed to be uniform. (Thus, it is necessary that breaks (changes) in grade be located in topographic surveys.)
4. The distance between contours indicates the steepness of a slope. Wide separation denotes gentle slopes; close spacing, steep slopes; even and parallel spacing, uniform slope.
5. Irregular contours signify rough, rugged country. Smooth lines imply more uniformly rolling terrain.
6. Concentric closed contours that increase in elevation represent hills. A contour forming a closed loop around lower ground is called a depression contour (Spot elevations and hachures inside the lowest contour and pointing to the bottom of a hole or sink with no outlet make map reading easier.)
7. Contours of different elevations never meet except on a vertical surface such as a wall, cliff, or natural bridge. They cross only in the rare case of a cave or overhanging shelf. Knife-edge conditions are never found in natural formations.
8. A contour cannot branch into two contours of the same elevation.
9. Contour lines crossing a stream point upstream and form V's; they point down the ridge and form U's when crossing a ridge crest.
10. Contour lines go in pairs up valleys and along the sides of ridge tops.
11. A single contour of a given elevation cannot exist between two equal-height contours of higher or lower elevation. For example, an 820-ft contour cannot exist alone between two 810- or two 830-ft contours.
12. Cuts and fills for earth dams, levees, highways, railroads, canals, etc., produce straight or geometrically curved contour lines with uniform, or uniformly graduated spacing. Contours cross sloping or crowned streets in typical V- or U-shaped lines.

Keeping these characteristics in mind will (1) make it easier to visualize contours when looking at an area, (2) assist in selecting the best array of points to locate in the field when conducting a topographic survey, and (3) prevent serious mistakes in drawing contours.

17.7 DIRECT AND INDIRECT METHODS OF LOCATING CONTOURS

Contours can be established by either the direct method (*trace-contour method*) or the indirect method (*controlling-point method*). The controlling-point method is generally more convenient and faster, and therefore it is most often selected. It is also the most frequent choice when data is entered into a computer for automated contouring. These two methods are described in the subsections that follow.

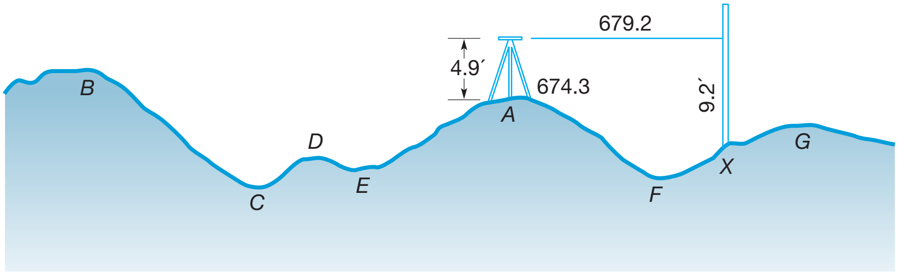


Figure 17.3
Direct method of locating contours.

17.7.1 Direct Method

This method can be performed using a total station instrument. After the instrument is set up, the *HI* is established, and the telescope oriented horizontally. Then for the existing *HI*, the rod reading (foresight) that must be subtracted to give a specific contour elevation is determined. The rodperson selects trial points expected to give this minus sight and is directed uphill or downhill by the instrument operator until the required reading is actually secured (to within perhaps 0.1 to 0.5 ft, depending on the allowable discrepancy which is dictated by the nature of the terrain, contour interval, and specified map accuracy).

In Figure 17.3, the instrument is set up at point *A*, elevation 674.3 ft, *hi* 4.9 ft, and *HI* 679.2 ft. If 5-ft contours are being located, a reading of 4.2 or 9.2 with the telescope level will place the rod on a contour point. For example, in Figure 17.3, the 9.2-ft rod reading means that point *X* lies on the 670-ft contour. After the point which gives the required rod reading has been located by trial, the horizontal position of the point is determined by observing its horizontal distance and direction from the instrument. Distances are observed electronically with total station instruments (see Chapter 6, Part III), but stadia (see Section 5.4) or taping (see Chapter 6, Part II) can be used. The process is repeated until the entire area has been surveyed. Work speed is increased by using a piece of red plastic flag that can be moved up and down on the level rod to mark the required reading and eliminate searching for a number. The maximum distance between contour points located in this method is determined by the terrain and accuracy required. Beginners have a tendency to take more sights than necessary in ordinary terrain. Contours are sketched by connecting located points having equal elevations. This is usually done as part of the office work, but they may also be drawn in the field book to clarify unusual conditions.

The direct method is suitable in gently rolling country, but generally not practical in undulating or rough, rugged terrain. Neither is it convenient for obtaining data to be used in computer-driven automated contouring systems.

17.7.2 Indirect Method

In the indirect method, the rod is set on “controlling points” that are critical to the proper definition of the topography. They include high and low points on the terrain, and locations where changes in ground slope occur, such as *B, C, D, E, F*, and *G* in Figure 17.3. Channels of drainage features and ridgelines must be included. Elevations are determined on these points using a total station instrument

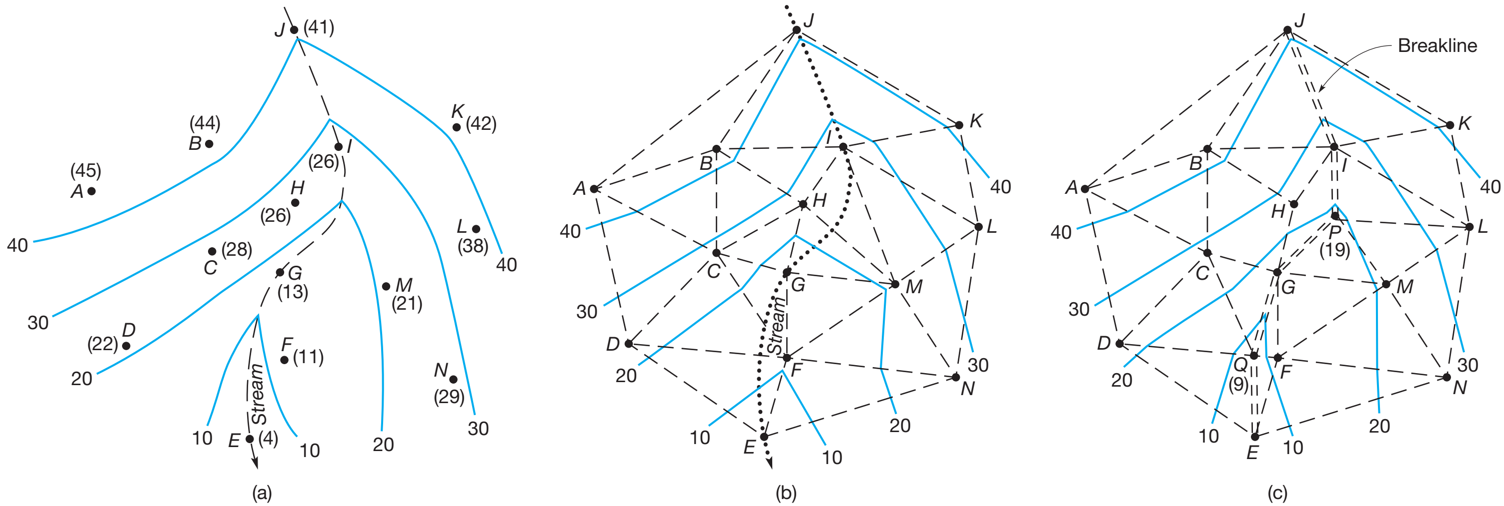


Figure 17.4 (a) Contours compiled by hand from controlling points *A* through *N*. (b) TIN model (dashed lines) constructed from data of (a), and contours (solid lines) derived from TIN model. The stream is shown with a dotted line. Note the striking differences between the 10- and 20-ft contours of (a) and (b). (c) TIN model (dashed lines) constructed from data of (a) but with the addition of two points, *P* and *Q*, and the designation of lines *EQ*, *QG*, *GP*, *PI*, and *IJ* as breaklines. Contours shown with solid lines were derived from this TIN model. Note the agreement of these contours with those of (a).

employing trigonometric leveling (see Section 4.5.4) or by using GNSS receivers with a geoid model. Horizontal distances and azimuths are also observed to locate the points. The positions of controlling points are then plotted and contours interpolated between elevations of adjacent points.

Figure 17.4(a) illustrates a set of controlling points labeled *A* through *N* that have been plotted according to their surveyed horizontal positions. Observed elevations (to the nearest foot) of the points are given in parentheses. Contours having a 10-ft interval have been sketched freehand between adjacent points by interpolation. It is improper to interpolate along lines that cross controlling features such as gullies, streams, rivers, ridge lines, roads, etc. Thus, to properly draw the contours of Figure 17.4(a) with the stream located on the map, elevations were first interpolated along its thread between surveyed points *E, G, I*, and *J*. Interpolations were then made from the stream to points on either side. As an example, it would have been incorrect to interpolate across the stream between points *D* and *F*. Rather the elevation of the stream on the line between *D* and *F* (approximately 9 ft) was used to interpolate both ways from the stream to points *D* and *F*. Note in Figure 17.4(a) that the gently curved contours tend to duplicate the naturally rolling topography of nature. Note also that contours crossing the stream form V's pointing upstream.

Numerous controlling points may be needed to locate a contour in certain types of terrain. For example, in the unusual case of a nearly level field that is at or close to a contour elevation, the exact location of that contour would be time consuming or perhaps impossible to determine. In these situations, a uniform distribution of *spot elevations* can be determined in the field and plotted on the map to convey the area's relief.

■ 17.8 DIGITAL ELEVATION MODELS AND AUTOMATED CONTOURING SYSTEMS

Data for use in automated contouring systems are collected in arrays of points whose horizontal positions are given by their *X* and *Y* coordinates and whose elevations are given as *Z* coordinates. Such three-dimensional arrays provide a *digital* representation of the continuous variation of relief over an area and are known as *digital elevation models* (DEMs). Alternatively, the term *digital terrain model* (DTM) is sometimes used.

Two basic geometric configurations are normally used in the field for collecting DEM data: the *grid method*, and the *irregular method*, although often a combination of the two methods is employed. In the grid method, elevations are determined on points that conform to a regular square or rectangular grid. The procedure is described in Section 17.9.3 and sample field notes are given in Plate B.2 of Appendix B. From the array of grid data, the computer interpolates between points along the grid lines to locate contour points and then draws the contour lines. The major disadvantage of this method is that critical high and low points and changes in slope do not generally occur at the grid intersections, and thus they are missed in the data collection process, which results in inaccurate relief portrayal.

The irregular method is simply the controlling-point method, but additional information (to be described later) is included. As previously noted, the controlling-point method involves determining the elevations of all high and low points and points where slopes change. Of course, this produces a DEM with an irregularly spaced configuration of surveyed points.

The first step taken by computerized contouring systems that utilize irregularly spaced spot elevations is to create a so-called *triangulated irregular network* (TIN) model of the terrain from the spot elevations. It is very important to understand the TIN model concept to ensure that an appropriate array of controlling points is selected and surveyed in the field if an automated contouring system is to be used. A TIN model is constructed by connecting points in the array to create a network of adjoining triangles. The dashed lines in Figure 17.4(b) show a TIN model created for the data in Figure 17.4(a). Various criteria can be used in the development of TIN models from an array of surveyed points, but one commonly used standard creates the "most equilateral network" of triangles.

Automated contouring systems generally make two assumptions concerning TIN models: (1) all triangle sides have a constant slope and (2) the surface area of any triangle is a plane. Based on these assumptions, elevations of contour crossings are interpolated along triangle edges, and contours are constructed such that they change direction only at triangle boundaries. Contours derived in this manner from the TIN model of Figure 17.4(b) are shown in the figure as solid lines. Note the disparities between the hand-drawn contours of Figure 17.4(a) and those derived from the TIN model of Figure 17.4(b). Differences are particularly obvious between the 10- and 20-ft contours. These occur because (1) the computer did not interpret the curved thread of the stream [shown as a dotted line in Figure 17.4(b)], and (2) in creating the network of triangles, several sides were constructed that cross the stream, resulting in improper interpolation across the stream.

From this example it is apparent, as noted earlier, that additional information must be provided for computer-driven systems to depict contours accurately. That important added information is the identification of controlling features, also more often called *breaklines*, or *fault lines* in modern computer mapping terminology. Breaklines are linear topographic features that delineate the intersection of two surfaces that have uniform slopes, and thus define changes in grade. *Automated mapping algorithms use these lines to define sides of the triangles that form the TIN model, and thus elevations are interpolated along them*. Streams, lake shores, roads, railroads, ditches, ridgelines, etc. are examples of controlling features or breaklines. Curved breaklines such as streams must have enough data points so that when adjacent ones are connected with straight lines, they adequately define the feature's alignment.

The dashed lines of Figure 17.4(c) represent the TIN model constructed from the same data set as in Figure 17.4(b), except that the stream (shown with a double dashed line) has now been identified as a breakline, and two additional points, P and Q, have been added to better approximate the curvature of the stream. In this figure, contours derived from the TIN model are shown. Note that these, now very nearly, duplicate the hand-drawn contours.

The important lesson of the foregoing is that if an automated contouring system is used, field points must be selected carefully, breaklines identified, and the data properly input to meet the system's assumptions. As indicated by this example, a few more controlling points may have to be surveyed, but the benefits of automated contouring systems make it worthwhile.

In order to avoid missing significant data during topographic surveys, it is usually best to collect features in groups. That is, data should be gathered first for (1) planimetric features, followed by (2) breaklines, (3) significant controlling points of elevation, and finally (4) sufficient *grade points* (those remaining points surveyed only to enable accurate depictions to be made of slopes and grades between the other types of points). Grade points are often most efficiently collected in a grid pattern throughout the entire area to be mapped. This grid should be sufficiently dense to avoid triangles in the TIN that are geometrically weak; that is, long and slender figures with one small angle. Varying grid sizes can be used, with larger spacing applied in areas of gradual slopes, and more dense patterns employed as the terrain becomes more undulating.

■ 17.9 BASIC FIELD METHODS FOR LOCATING TOPOGRAPHIC DETAILS

Objects to be located in a mapping survey can range from single points or lines to meandering streams and complicated geological formations. The process of tying mapping details to the control net is called *detailing*. Regardless of their shape, all objects can be located by considering them as composed of a series of connected straight lines, with each line being determined by two points. Irregular or curved lines can be assumed straight between points sufficiently close together; thus detailing becomes a process of locating points.

Location of planimetric features and contours can be accomplished by one of the following field procedures: (1) radiation by total station instrument, (2) coordinate squares or "grid" method, (3) offsets from a reference line, (4) use of portable GNSS units, or (5) a combination of these methods. An explanation of each system and a discussion on their uses, advantages, and disadvantages follow.

17.9.1 Radiation by Total Station

In the radiation method, illustrated in Figure 17.5, with a total station instrument set up at a control point, the zenith angle, slope distance, and direction are observed to each desired item of mapping detail. From the zenith angle and slope distance, the elevation of the point can be determined, and by incorporating the direction, its horizontal position can be computed. These computations are often performed by the internal computer in a total station or by the data collector. As shown in the figure, the sights to details radiate from the occupied station, hence the name for the procedure. This method is especially efficient if a data collector (see Sections 2.12 through 2.15) is used to record the point identities and their associated descriptions, vertical distances, horizontal distances, and directions. The data collector permits downloading the observations directly into a computer for processing through an automated mapping system. The field procedure of radiation with a total station can be made most efficient if the instrument is placed at a good vantage point (on a hill or ridge) that overlooks a large part or all of the area to be surveyed. This permits more and longer radial lines and reduces the number of setups required.

Table 17.1 is an example illustrating the use of a total station with data collector for topographic mapping. The example relates to Figure 17.5. In Figure 17.5,

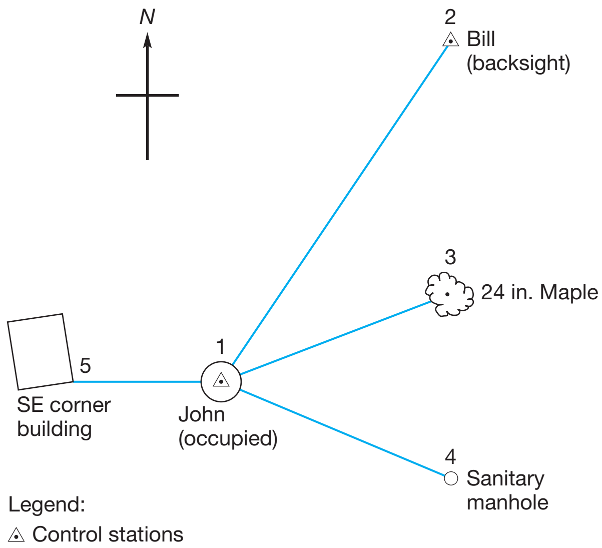


Figure 17.5
Radiation survey from control traverse.

TABLE 17.1 EXCERPT OF AUTOMATIC DATA COLLECTOR FIELD NOTES OF A RADIAL SURVEY FOR TOPOGRAPHIC DETAILS

Entry	Explanation
AC:SS	(Activity: Sideshot/keyboard entry by operator)
PN:3	(Point number: 3/keyboard entry by operator)
PD:24 IN MAPLE	(Point description: 24-in. maple/keyboard entry)
HZ:16.3744	(Horizontal angle: 16°37′44″/by total station)
VT:90.2550	(Vertical "zenith" angle: 90°25′50″/by total station)
DS:565.855	(Distance: 565.855 ft/by total station)
AC:SS	(Activity: Sideshot/keyboard entry by operator)
PN:4	(Point number: 4/keyboard entry by operator)
PD:SAN MH	(Point description: Sanitary manhole/keyboard entry)
HZ:70.3524	(Horizontal angle: 70°35′24″/by total station)
VT:91.1548	(Vertical "zenith" angle: 91°15′48″/by total station)
DS:463.472	(Distance: 436.472 ft/by total station)
AC:SS	(Activity: Sideshot/keyboard entry by operator)
PN:5	(Point number: 4/keyboard entry by operator)
PD:SE COR BLDG	(Point description: SE corner building/keyboard entry)
HZ:225.1422	(Horizontal angle: 225°14′22″/by total station)
VT:88.3036	(Vertical "zenith" angle: 88°30′36″/by total station)
DS:265.934	(Distance: 265.934 ft/by total station)

Source: Courtesy Michael J. Dercks, President, ABACUS, A Division of Calculus, Inc.

a total station instrument was set up at control station 1 (John) and oriented in azimuth with a backsight on control station 2 (Bill). Observations of azimuth, zenith angle, and distance, respectively, were then taken to points 3, 4, and 5, which are *sideshots* to mapping details. From these sideshots, two- or three-dimensional coordinates can be computed that are used to locate the points on a map sheet.

When using an automatic data collector in this process, initial data for the setup are first entered in the unit via the keyboard. These include the X and Y coordinates of stations John and Bill, the elevation of John, and the heights of both the total station instrument and the reflector. The left-hand column of Table 17.1 illustrates an actual set of field notes recorded by a data collector during the process of taking sideshots on points 3, 4, and 5. Six entries were recorded per point. On each line, the entry to the left of the colon was automatically supplied by the computer and appeared on the data collector display at the time of observation to prompt the operator. Entries to the right of the colon were supplied by the operator, either manually via the keyboard or by pressing the proper button on the total station instrument. Explanations to assist students in interpreting the data in Table 17.1 are given in parentheses.

As shown in Figure 17.6, details that have width such as trees are located with two separate observations. The first observation locates the azimuth to the object by observing an angle from a reference line to the front center of the object. The second shot measures the distance to the side center of the object. Using the azimuth of the first shot and the distance from the second shot, coordinates near the center of the object can be determined. Data collectors have various names for this collection routine such as *separate distance and angle* (SDA). This procedure should only be used when the diameter of the object is sufficiently large to cause a plotting error on the map. For smaller objects where the diameter will not noticeably displace the center of the object on the map, this procedure is unnecessary. Thus, the use of this method is dependent on the scale of the map and size of the object.

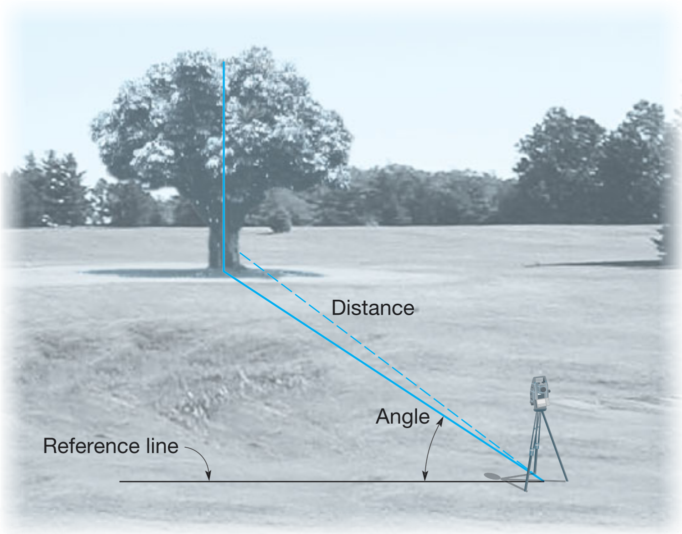


Figure 17.6 Proper location of objects such as trees.

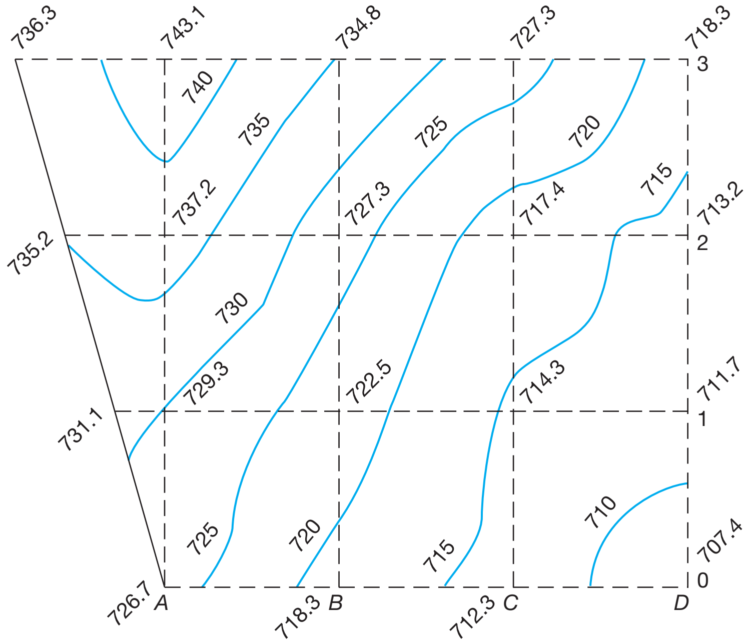


Figure 17.7
Coordinate squares.

17.9.2 Coordinate Squares or "Grid" Method

The method of coordinate squares (grid method) is better adapted to locating contours than planimetric features, but can be used for both. The area to be surveyed is staked in squares 10, 20, 50, or 100 ft (5, 10, 20, or 40 m) on a side, the size depending on the terrain and accuracy required. A total station instrument can be used to lay out control lines at right angles to each other, such as *AD* and *D*3 in Figure 17.7. Grid lengths are marked and the other corners staked and identified by the number and letter of intersecting lines.

Elevations of the corners can be obtained by differential or trigonometric leveling. Contours are interpolated between the corner elevations (along the sides of the blocks) by estimation or by calculated proportional distances. Elevations obtained by interpolation along the diagonals will generally not agree with those from interpolation along the four sides because the ground's surface is not a plane. Except for plotting contours, this is the same procedure as that used in the borrow-pit problem in Section 26.10. In plotting contours by the grid method, a widely spaced grid can be used for gently sloping areas, but it must be made denser for areas where the relief is rolling or rugged.

A drawback of the method is that no matter how dense the grid, critical points (high and low spots and slope changes) will not generally occur at grid locations, thus degrading the accuracy of the resulting contour map.

17.9.3 Offsets from a Reference Line

This procedure is most often selected for mapping long linear features and for performing surveys necessary for route locations. After the reference line or centerline has been staked and stationed, planimetric details are located by observing

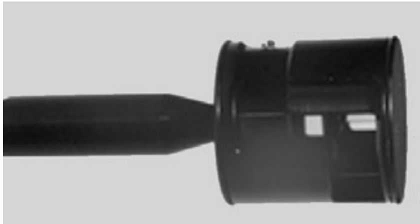


Figure 17.8 Double pentagonal prism (on side) for establishing perpendiculars to a reference line. (Courtesy Leica Geosystems AG.)

perpendicular offsets from it and noting the stations from which the offsets were taken. Features such as streams, trails, fences, buildings, utilities, trees, etc. can be located in this manner. Elevations for determining contour locations can also be determined by *cross-sectioning* (observing ground profiles along lines perpendicular to the reference line) as described in Section 26.3.

Shorter offset distances are generally most easily and quickly observed by taping, but longer lengths may be more efficiently obtained by electronic distance measurement using a total station instrument. Where steep slopes run transverse to the reference line, better accuracy and efficiency can often be obtained using a total station instrument. Perpendiculars to a reference line can be quickly established using a *pentagonal prism* (see Figure 17.8). This device works well when offset distances are being taped. While standing on the reference line at a location where a perpendicular is desired, the operator holds the instrument upright and looks through it sighting either ahead or back on the reference line. By means of the prisms, perpendicular views both left and right can be simultaneously seen. Alternatively, the total station instrument can be set up on the reference line at the location of a desired offset, oriented by sighting ahead or back on the line, and then used to turn a 90° angle.

Figures 17.9(a) and (b) illustrate examples of using offsets for mapping. In Figure 17.9(a), the location of Crooked Creek was determined by observing distances to the edge of the creek at intervals from the reference line. The offsets can be taken at regular intervals as illustrated in the figure, or can be spaced at distances that permit the curved nature of the stream to be considered as a series of straight segments between successive offsets. Figure 17.9(b) illustrates an example of locating planimetric features along a road right-of-way. This type of survey would be useful to locate buildings, utilities, trees and other features along a road for highway design, or for excavation to install an underground utility. After locating at least two corners of a building by observing their offset distances, for example, x_1 and x_2, in the figure, and their plusses on the reference line, its remaining dimensions can usually be quickly obtained by taping, for example, a and b in the figure.

In both of these examples, it would be convenient to include a sketch in the field book, and to record the observations directly on the sketch. Since data collected by this method are based purely on distances, it is difficult to merge them with data collected by the radiation method and is seldom used when employing computer drafting techniques.

17.9.4 Topographic Detailing with GNSS

GNSS receivers for topographic work are specially designed, small, and portable, and are interfaced with a keyboard for system control and entry of codes to identify

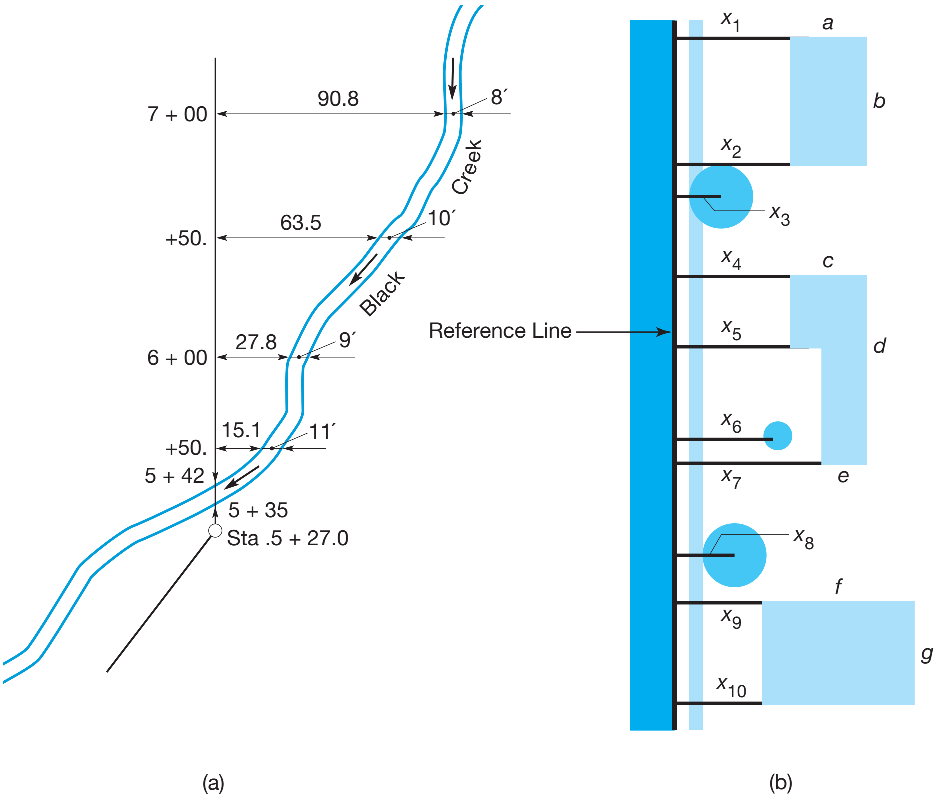


Figure 17.9
(a) Location of creek by perpendicular offsets from a reference line.
(b) Locating objects by offsets from a reference line.

features surveyed. The unit shown in Figure 14.1 is suitable for topographic work. These receivers can determine (in real time) the coordinates of locations where the receiver antenna is placed and can store data for the points in files. The files are then directly downloaded to a computer for further processing, which could include automatic map drafting. These systems make topographic data collection a simple and very fast one-person operation. The kinematic surveying methods discussed in Chapter 15 are most often used for mapping surveys; however, static surveys (Chapter 14) are sometimes used to establish control in the project area.

The stop-and-go method (semikinematic) has the advantage over the true kinematic method in that the operator can stop and collect multiple epochs of data for a point to increase positional accuracies. The semikinematic method generally results in a file size smaller than those using the true kinematic method. Additionally the stop-and-go method can be used to locate features or establish lower-order control points for occupation or sighting by conventional instruments such as total stations.

True-kinematic surveys collect data points at an operator-selected epoch (usually 1–5 sec interval), and thus this method can be used to quickly obtain cross-sections and locate linear features such as breaklines, roads, and streams. However, since the operator has no control over the actual instance of data collection, this method does not provide a convenient means for surveying key topographical and planimetric features.

During a mapping survey, it is possible to switch from a true kinematic to semikinematic, or static survey (see Chapter 14). Therefore, if the conditions within a mapping project change, the operator can select the survey method that is most appropriate for the task at hand. Regardless of the method selected, it is essential that the antennas have clear visibility to the satellites. Thus, GNSS surveys are generally not suitable for direct location of large trees, buildings, or other objects that could obscure the view of the satellites. In these situations, nearby temporary control stations can be established using the static GNSS survey method and traditional data collection methods with total stations used to locate these map details. However, when the modernized GPS constellation is fully implemented, this limitation is expected to disappear due to increased signal strength.

As discussed in Chapter 2, many data collectors work with GNSS receivers as well as total station instruments. However, since orthometric heights are normally required for the vertical component of a survey, it is important to transform GNSS-derived ellipsoid heights to their orthometric values using procedures discussed in Sections 13.4.3 and 19.5. Typically, survey control software provides an option to perform this conversion using an appropriate geoid model.

Code-based GNSS receivers can also be applied in topographic mapping, but their use is generally limited to lower-order work such as those found in GIS applications. These receivers are very affordable and currently have post-processed accuracies under five meters. Even though this is a relatively high positional uncertainty, these units can be used to do surveys for maps with scales smaller than 1:20,000 since the plotting errors become negligibly small. As additional channels are added to the GPS satellites and other satellite systems become mature, it is anticipated that the accuracies of code-based GNSS will be reduced to the meter level and possibly decimeter level in the differential positioning mode.

17.9.5 Laser Scanning

Laser scanners automate digital angle observation with reflectorless, laser EDM technologies. They can quickly produce grids of three-dimensional coordinates for user-specified scenes. A combination of rotating mirrors allows the instruments to observe distances and orthogonal angles in precise grid patterns. These instruments vary in capabilities to match the varying requirements of jobs. Characteristics of an instrument are defined by the *number of observations* they can observe per second, the observable distance from the instrument known as its *range*, the minimum spacing between observations or its *resolution, accuracy*, and *field of view*. In general, the higher the number of observations per second, the faster the data acquisition. Instruments vary from less than a hundred observations per second to 500,000 observations per second. The range of instruments can vary from several meters to several kilometers. Instrument resolutions can vary from a few millimeters to a several centimeters. Many ground-based instruments can produce finer resolutions on targets. However, it must be remembered that the higher the resolution, the larger the data files. Generally, ground-based laser scanners have distance measurement accuracies of a few millimeters. The Leica HDS3000 shown in Figure 1.5 has a range accuracy of ±6 mm at 50 m and angular accuracies of

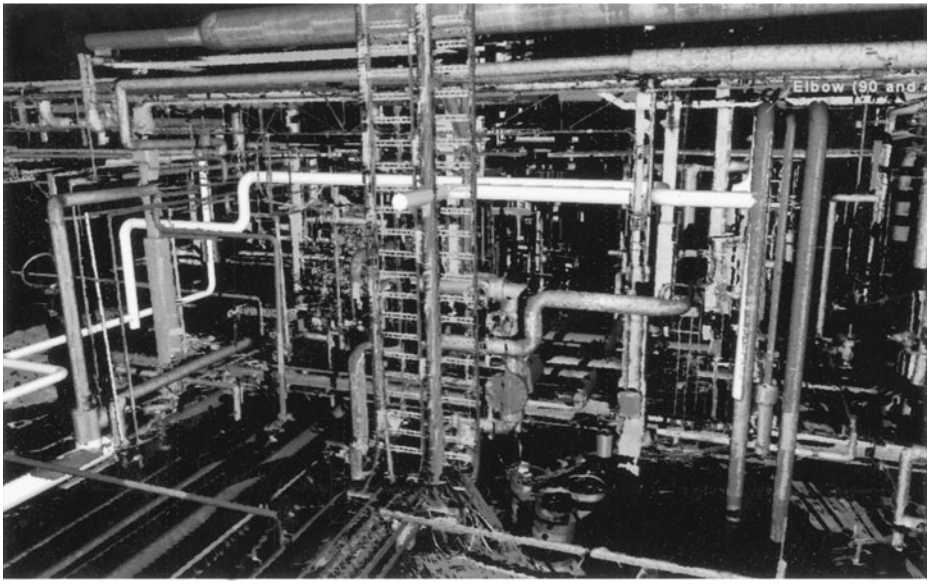


Figure 17.10 Laser-scanned image of refinery showing designed pipe alignment in white. (Courtesy of Christopher Gibbons, Leica Geosystems AG.)

$\pm 60\ \mu$rad.[1] The field of view dictates the area that a laser scanner can observe in a single setup. Some instruments such as the Leica HDS3000 can rotate 360° in horizontal and 270° in vertical allowing it to survey the entire scene surrounding the instrument. In general, users can specify the field of view to match the area of interest. Several manufacturers have added scanning capabilities to selected robotic total stations. Thus, the surveyor has the option of performing a traditional radiation survey or scanning the object. This is especially useful in rugged areas with large vertical relief and in industrial settings.

The resulting grid of scanned, three-dimensional points can be so dense that a visual image of the scene is formed. This so-called "point-cloud" image differs from a photographic image in that every point has a three-dimensional coordinate assigned to it. These coordinates can be used to obtain dimensions between any two observed points in the scene. In Figure 17.10, a point-cloud image of piping at a refinery provided engineers with the information needed to design a new pipe addition shown in white. The detail in this image would be difficult to recreate using other surveying processes. Because of the high density of the observed points in a scene, it is sometimes referred to as "high-definition" surveying. Some instruments also capture a digital image of the scene. The digital image can be integrated with the scanned points to create a three-dimensional image having color and texture. This process was used for the bridge shown in Figure 23.14. Laser scanning can play a significant role in as-built surveys, archeology, and mapping of artifacts. Section 27.19 discusses the use of airborne laser mapping

[1]A microradian equals 0.000001 radians, which is about 0.2″.

Figure 17.11 Point cloud from the IP-S2 3D mobile mapping system. (Courtesy Topcon Positioning Systems.)

known as LiDAR. The Florida Department of Transportation has mapped the entire state using this technology.

The original point cloud in many systems is determined in an arbitrary three-dimensional coordinate system. If it is necessary to have coordinates in a project-based coordinate system, a traditional survey can be used to establish coordinates on targets in the scene. Control must be strategically located at the edges of each scene. A minimum of three control points per scene is required. However, additional control is often used to provide redundancy. Multiple scenes can be connected using common control targets. After determining the project coordinates of the control, a three-dimensional conformal coordinate transformation, discussed in Section 17.10, is performed to transform points from the arbitrary coordinate system to the project coordinate system.

Figure 17.11 depicts a scene that was captured by a IP-S2 3D mobile mapping system shown in Figure 1.4. The IP-S2 incorporates multiple LiDAR scanners, a GNSS receiver, inertial measurement unit, high-quality digital camera, and odometry data. The system is capable of scanning objects within 30 m of the vehicle as it moves at highway speeds. It collects up to 1.3 million points per second. In Figure 17.11, the dots running down the street corridor are the GNSS position fixes of the system. The inertial measurement unit provides positioning for the system when canopy prohibits the GNSS fixes. The larger sphere in the center of the figure is a window into the high-resolution digital image of the scene. This digital image overlays the point cloud allowing the viewer to see a picture quality image of the scene. From this point, items can be identified and their positions determined or distances between objects computed. The IP-S2 system georeferences the point cloud. Thus, additional surveying and ground control are not required. However, higher accuracies can be obtained by post-processing the GNSS fixes against a GNSS base station. (See Chapter 14.)

■ 17.10 THREE-DIMENSIONAL CONFORMAL COORDINATE TRANSFORMATION

The three-dimensional conformal coordinate transformation transfers points from one three-dimensional coordinate system (xyz) into another (XYZ). This transformation is similar to the two-dimensional coordinate transformation covered in Section 11.8. However, the three-dimensional conformal coordinate transformation involves seven unknown parameters (three rotation angles, three translation factors, and one scale factor). The development of the rotations is covered in Section 19.16. The mathematical model for the transformation is:

$$X = S(m_{11}x + m_{21}y + m_{31}z) + T_X \tag{17.1a}$$
$$Y = S(m_{12}x + m_{22}y + m_{32}z) + T_Y \tag{17.1b}$$
$$Z = S(m_{13}x + m_{23}y + m_{33}z) + T_Z \tag{17.1c}$$

where S is a scale factor, T_X, T_Y, and T_Z are the translations in x, y, and z, respectively, and m_{11} through m_{33} are elements of the combined rotation matrix.

$$\begin{aligned}
m_{11} &= \cos(\theta_y)\cos(\theta_z)\\
m_{12} &= \sin(\theta_x)\ \sin(\theta_y)\cos(\theta_z) + \cos(\theta_x)\ \sin(\theta_z)\\
m_{13} &= -\cos(\theta_x)\ \sin(\theta_y)\cos(\theta_z) + \sin(\theta_x)\ \sin(\theta_z)\\
m_{21} &= -\cos(\theta_y)\ \sin(\theta_z)\\
m_{22} &= -\sin(\theta_x)\ \sin(\theta_y)\ \sin(\theta_z) + \cos(\theta_x)\cos(\theta_z)\\
m_{23} &= \cos(\theta_x)\ \sin(\theta_y)\ \sin(\theta_z) + \sin(\theta_x)\cos(\theta_z)\\
m_{31} &= \sin(\theta_x)\\
m_{32} &= -\sin(\theta_x)\cos(\theta_y)\\
m_{32} &= \cos(\theta_x)\cos(\theta_y)
\end{aligned}$$

θ_x, θ_y, and θ_z are the counterclockwise rotation angles about the X, Y, and Z axes, respectively, as viewed from their positive ends. To solve this set of nonlinear equations, methods similar to those discussed in Section 16.9.1 are employed. The linearized set of equations for this transformation given in matrix form are

$$\begin{bmatrix} \left(\frac{\partial X}{\partial S}\right)_0 & 0 & \left(\frac{\partial X}{\partial \theta_y}\right)_0 & \left(\frac{\partial X}{\partial \theta_z}\right)_0 & 1 & 0 & 0 \\ \left(\frac{\partial Y}{\partial S}\right)_0 & \left(\frac{\partial Y}{\partial \theta_x}\right)_0 & \left(\frac{\partial Y}{\partial \theta_y}\right)_0 & \left(\frac{\partial Y}{\partial \theta_z}\right)_0 & 0 & 1 & 0 \\ \left(\frac{\partial Z}{\partial S}\right)_0 & \left(\frac{\partial Z}{\partial \theta_x}\right)_0 & \left(\frac{\partial Z}{\partial \theta_y}\right)_0 & \left(\frac{\partial Z}{\partial \theta_z}\right)_0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} dS \\ d\theta_x \\ d\theta_y \\ d\theta_z \\ dT_X \\ dT_Y \\ dT_Z \end{bmatrix} = \begin{bmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{bmatrix} \tag{17.2}$$

where X_0, Y_0, and Z_0 are determined using Equations (17.1a) through (17.1c), respectively, with approximations for the unknown parameters. The coefficients from the linearized equations are

$$\frac{\partial X}{\partial S} = m_{11}x + m_{21}y + m_{31}z$$

$$\frac{\partial X}{\partial \theta_y} = S(-x\sin\theta_y\cos\theta_z + y\sin\theta_y\sin\theta_z + z\cos\theta_y)$$

$$\frac{\partial X}{\partial \theta_z} = S(m_{21}x - m_{11}y)$$

$$\frac{\partial Y}{\partial S} = m_{12}x - m_{22}y + m_{33}z$$

$$\frac{\partial Y}{\partial \theta_x} = -S(m_{13}x + m_{23}y + m_{33}z)$$

$$\frac{\partial Y}{\partial \theta_y} = S(x\sin\theta_x\cos\theta_y\cos\theta_z - y\sin\theta_x\cos\theta_y\sin\theta_z + z\sin\theta_x\sin\theta_y)$$

$$\frac{\partial Y}{\partial \theta_z} = S(m_{22}x - m_{12}y)$$

$$\frac{\partial Z}{\partial S} = m_{13}x + m_{23}y + m_{33}z$$

$$\frac{\partial Z}{\partial \theta_x} = S(m_{12}x + m_{22}y + m_{32}z)$$

$$\frac{\partial Z}{\partial \theta_y} = S(-x\cos\theta_x\cos\theta_y\cos\theta_z + y\cos\theta_x\cos\theta_y\sin\theta_z - z\cos\theta_x\sin\theta_y)$$

$$\frac{\partial Z}{\partial \theta_z} = S(m_{23}x - m_{13}y)$$

As outlined in Section 16.9.1, Equation (17.2) is formed using approximate values for the unknowns and iterated until the corrections to the unknown parameters become negligibly small. This process is demonstrated in a Mathcad worksheet *3DC.xmcd* on the companion website for this book at http://www.pearsonhighered.com/ghilani. The complete mathematical development of the transformation is covered in several books listed in the bibliography at the end of this chapter.

■ 17.11 SELECTION OF FIELD METHOD

Selection of the field method to be used on any topographic survey depends on many factors, including (1) purpose of the survey, (2) map use (accuracy required), (3) map scale, (4) contour interval, (5) size and type of terrain involved, (6) costs, (7) equipment and time available, and (8) experience of the survey personnel.

Items (1) to (5) are interdependent. The cost, of course, will be a minimum if the most suitable method is selected for a project. On large projects, personnel costs rather than equipment investment will usually govern. However, the equipment owned may govern the method chosen by a private surveyor making a topographic survey of 50 or 100 acres.

■ 17.12 WORKING WITH DATA COLLECTORS AND FIELD-TO-FINISH SOFTWARE

Surveying instruments equipped with data collectors can record and store field notes for electronic transmission to computers, plotters, and other office equipment for processing. Such systems, called "field to finish," rely on sophisticated software for their operation. Their use can increase productivity tremendously in surveying and mapping.

In using field-to-finish systems for topographic surveys, the data collector will store a point identifier and the *NEH* coordinates for each point located. However, in addition, ancillary descriptive information can accompany each surveyed point. For instance, in Figure 17.12, points 1, 2, 14, and 15 are the corners of a building; points 5, 6, 7, and 8 mark the corners of a sidewalk; points 4, 9, 10, 11, and 12 are points along a property line, with points 10 and 11 being on a curve; and points 3 and 13 are deciduous trees. It is possible to add and store this descriptive information in the data collector through the use of *notes*. If the notes are entered in a manner that is understandable by the field-to-finish software and the points are collected in a manner that is consistent with and supports the drafting system, the software will use appropriate symbols for plotting each feature, draw and close polygons, and create a complete and finished drawing, while some software requires that different features be placed on different layers to control plotting features. No matter the method employed, correctly entering the field data at the time of collection will greatly reduce the time in generating the final

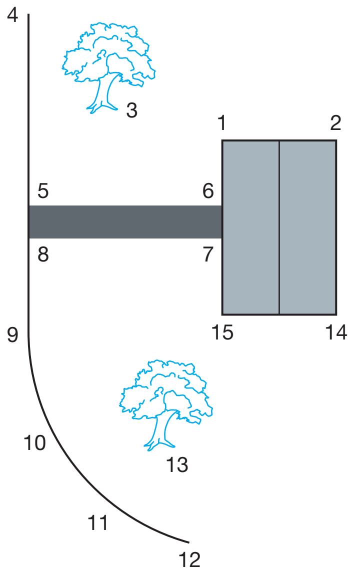


Figure 17.12 Example survey showing line work for planimetric features.

TABLE 17.2 DRAWING DESIGNATORS

Designator	System Commands
Line	. (period)
Curve	− (minus)
Close	+ (plus)
Join last	* (asterisk)
Bearing Close	# (pound sign)
Cross Section	= (equal sign)
Stop	! (exclamation mark)
Insert	* (asterisk)

map product. However, in order for the system to operate properly, the field personnel collecting the data must understand the requirements of the plotting software and implement these requirements during data collection.

While the field-to-finish software of the various vendors use somewhat different techniques in reducing field data to a finished map, the basic concepts of these systems can be discussed. Typically, drawing designators are used to construct the line work from the field data (Table 17.2). By entering appropriate designators in the notes for each point, as it is located, the field personnel instruct office personnel and field-to-finish software on how to draw the lines. Specific *codes* for drawing symbols can also be entered in the notes in the field as work progresses. For example, the identifier DTREE could indicate the deciduous trees shown in Figure 17.12. Using the proper drawing designators and codes, the notes for the planimetric survey of Figure 17.12 are shown in Table 17.3.

As shown in Table 17.3, the note for each line element begins with a period (.) symbol. Also, the fact that points 5 and 8 are common to two lines [i.e., they are part of sidewalk (SW1) and right-of-way (RW1)], is indicated in the table by the insertion of an additional period. Since point 10 is on the curve of the right-of-way, it is followed by a minus (−) symbol. The deciduous trees, points 3 and 13, have *DTREE* as their descriptive code, and this note will direct the computer to use the appropriate symbol for plotting these trees. Finally, note that the plus symbol (+) following point 15 indicates that lines B1 (designating building No. 1) close back on point 1. The notes for points 6 and 8 end with an exclamation point (!) that indicates that SW1 stops at this location.

It is important to include the appropriate designators and notes when collecting data. It is also necessary that successive shots on any object be collected and numbered in the correct order. For example, if the field personnel had collected the building data in the order of 1, 2, 15, and 14, the line work for the building would cross, creating an hourglass shape. Similarly, it was important to collect point 4 before point 5 so that the line work for the right-of-way would be drawn in a linear fashion.

TABLE 17.3 NOTES FOR DRAWING FEATURES IN FIGURE 17.11

Point	Notes
1	.B1
2	.B1
3	DTREE
4	.RW
5	.RW1.SW1
6	.SW1!
7	.SW1
8	.RW1.SW1!
9	.RW1
10	.RW1−
11	.RW1−
12	.RW1!
13	DTREE
14	.B1
15	.B1+

Since data collector viewing screens are typically small, many data collectors allow users to place field shots on different layers. These layers can be toggled off and on as required. By doing this, the field personnel can easily identify their progress on collecting data for various features in their project site. For example, all building points can be placed on a single layer. These points will later define the boundary of the building as well as a boundary for contouring in topographic mapping. By placing them on a single layer, the field personnel can turn off unnecessary layers so that they can easily view their progress on collecting the relevant features of the object. Grade shots, which are points located simply for later contouring, can be placed on a single layer while the breaklines are placed on another. All other layers can be turned off. Again, by doing so, the field personnel can identify topographic features that are missing or may require additional data for proper definition.

With the complexities of collecting data, selecting the locations of points, and properly noting the features, it is easy to see that some orderly plan for data collection should be developed before the instruments are taken from their cases. Also, there must be coordination between the field and drafting personnel. While each organization may develop their own procedures, some guidelines for collecting data suggest collecting planimetric feature data first, paying special attention to the sequence in which the data is collected. It is often most efficient to collect data for one feature type before beginning another; that is, locating all buildings, then the roadways, then vegetation, and so on in an orderly fashion at each instrument setup. Again, if these various features are placed on separate layers, the field personnel can more easily identify planimetric features that are missing or require more points for proper location. For topographic surveys, all controlling points may be collected, followed by break lines, and finally sufficient grade shots to allow accurate interpolation of the contours. (See Chapter 18 for a discussion on interpolation of contours.)

From the foregoing, it should be obvious that it is essential for field personnel to understand not only what data to collect, but also the order and manner in which to collect it so that proper plotting will occur. It should be noted here that different coding requirements are necessary for the various software used in practice. However, all field-to-finish systems have some drawing conventions that must be understood by both the field and office personnel. Correctly performed, a field-to-finish survey will greatly reduce the time it takes to create a finished and correct map.

■ 17.13 HYDROGRAPHIC SURVEYS

Hydrographic surveys determine depths and terrain configurations of the bottoms of water bodies. Usually the survey data are used to prepare hydrographic maps. In navigation and dredging, they may be recorded in electrical formats for real-time analysis. Bodies of water surveyed include rivers, reservoirs, harbors, lakes, and oceans.

Hydrographic surveys and maps are used in a variety of ways. As examples, engineers employ them for planning and monitoring harbor and river dredging operations, and to ascertain reservoir capacities for flood control and water supply systems; petroleum engineers use them to position offshore drilling facilities and locate underwater pipelines; navigators need them to chart safe passageways and avoid reefs, bars, and other underwater hazards; biologists and conservationists find them helpful in their study and management of aquatic life; and anglers use them to locate structures where fish is likely to be located.

Field procedures for hydrographic surveys are similar to those for topographic work; hence, the subject is discussed in this chapter. There are some basic differences in procedures used by surveyors since the land area being mapped cannot be seen, and the depth measurements must be made in water.

Two basic tasks involved in hydrographic surveys are (1) making *soundings* (measuring depths) from the water surface to bottom, and (2) locating the positions where soundings were made. Techniques used to perform these tasks vary depending on the water body's size, accuracy required, type of equipment to be used, and number of personnel available. The subsections that follow briefly describe procedures for mapping small to moderate-sized water bodies.

17.13.1 Equipment for Making Soundings

The size of a water body and its depth control the type of equipment used to measure depths. For shallow areas of limited size, a *sounding pole* can be used. This is usually a wooden or fiberglass staff resembling a level rod. It is perhaps 15 ft long, graduated in feet or tenths of feet, with a metal shoe on the

Figure 17.13
Depth sounder used in small lake hydrographic survey. (Sea Floor Systems, Inc.)

bottom. Direct depth measurements are made by lowering the pole vertically into the water until it hits the bottom, and then reading the graduation at the surface.

Lead lines can be used where depths are greater than can be reached with a sounding pole. These consist of a suitable length of stretch-resistant cord or other material, to which a heavy lead weight (perhaps 5–15 lb.) is attached. The cord is marked with foot graduations, and these should be checked frequently against a steel tape for their accuracy. In use, the weight is lowered into the water, being careful to keep the cord vertical. The graduation at the surface is read when the weight hits the bottom.

In deep water, or for hydrographic surveys of appreciable extent, electronically operated sonic depth recorders called *echo sounders* are used to measure depths. These devices, an example of which is shown in Figure 17.13, transmit an acoustic pulse vertically downward and measure the elapsed time for the signal to travel to the bottom, be reflected, and return. The travel time is converted to depth, and displayed in either digital or graphic form. A graphic profile of the depths, such as that shown in Figure 17.13 can be displayed on a computer screen. This graphic plot can be referred to repeatedly for plotting and checking.

Sounding poles and lead lines yield spot depths and are restricted for use in relatively shallow water. However, electronic depth sounders provide continuous profiles of the surface beneath the boat's path and can be used in water of virtually any depth. For example, in the profile of Figure 17.14, the chart's vertical range was set to 40 ft and profile depths shown vary from 10 to 24 ft.

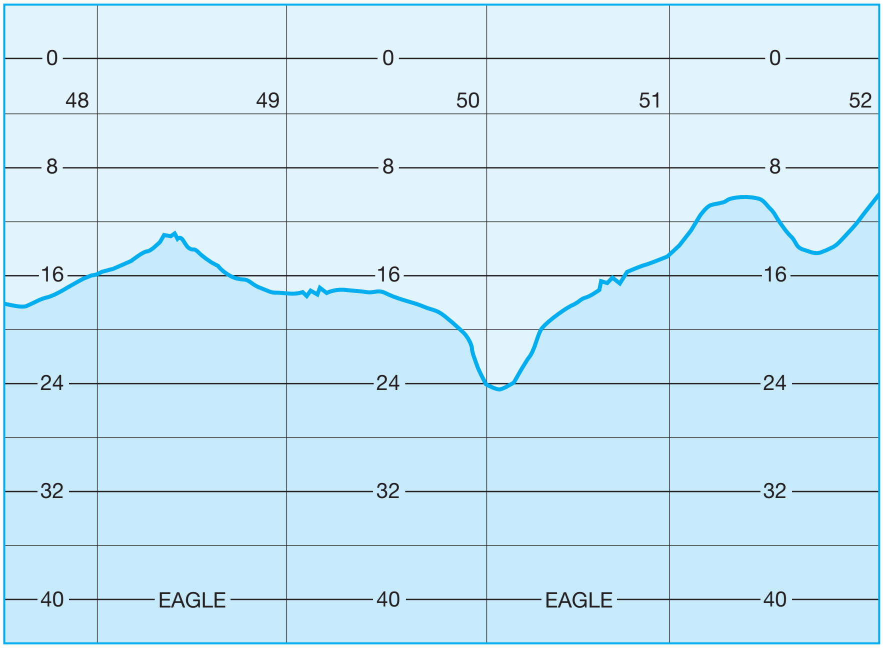


Figure 17.14 Bottom profile produced by electronic depth sounder.

The reference plane from which depth soundings are measured is the water surface. Because of surface fluctuations, its elevation or *stage* at the time of survey must be determined with respect to a fixed datum—for example, orthometric height. Running a level circuit to the water from a nearby benchmark can do this. In situations where soundings are repeated at regular intervals, a graduated staff can be permanently installed in the water so that its stage, in feet above the datum, can be read directly each time soundings are repeated.

17.13.2 Locating Soundings

Any of the traditional ground-surveying procedures can be used to locate positions where soundings are taken. In addition to these techniques, other methods have also been applied in hydrographic surveys, for example, GNSS receivers. If ground-surveying techniques are used, some horizontal control must first be established on shore. Ideal locations for control stations are on peninsulas or in open areas that afford a wide unobstructed view of the water body for tracking a sounding boat. The coordinate positions of the control points can be established by traverse, but triangulation and trilateration are also well suited for this work.

Among the various boat-positioning methods, *radiation* and *angle intersection* are usually selected if total station instruments are used. Radiation is particularly efficient, especially if a total station instrument is used, because only one person on shore is needed to track the boat. After setting up on one control station and back sighting another, angles and distances are measured to locate each boat position. Special total station instruments and reflectors are manufactured for this work to facilitate sighting and observing distances electronically to a

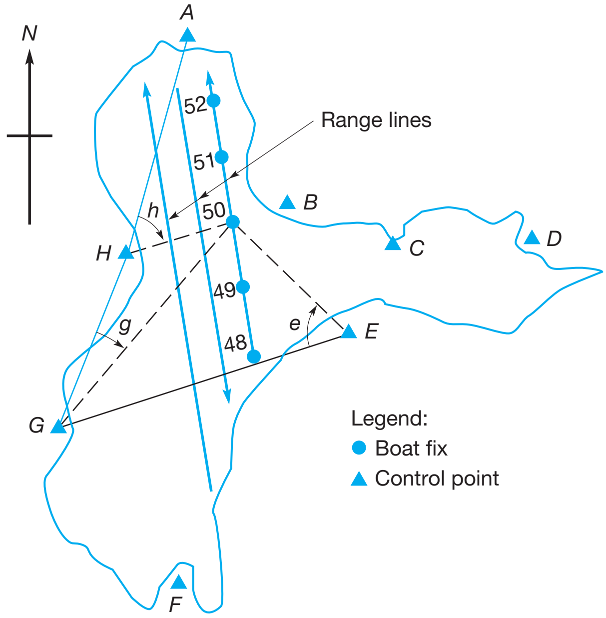


Figure 17.15 Angle intersection procedure for locating boat fixes along range lines.

moving target. From angles and distances, which are automatically read, the total station's computer determines the boat's coordinates. These can either be stored in an automatic data collector for later office use in mapping, or transmitted by radio to the boat if real-time positioning is required, as in controlling ongoing dredging.

Figure 17.15 illustrates the use of angle intersections in the hydrographic survey of a lake. Here the boat travels back and forth along *range lines* while the depth sounder continuously records bottom profiles. At regular intervals, *fixes* are taken by observing angles to the boat from shore stations. Two angles establish the boat's position, but three or more provide redundancy and a check. For example, in Figure 17.15, angle observations *e*, *g*, and *h*, for fix number 50 (indicated by dashed lines) have been made from shore stations *E*, *G*, and *H*, respectively. Prior to observing angles, the total stations or theodolites were oriented by backsighting on another visible control station, as at station *G* from *E*.

Flag or radio signals are given from the boat to coordinate fixes and ensure that angles from all shore stations are observed simultaneously. At the precise moment of any fix, the profile is also marked and the fix number noted. For example, in Figure 17.14, fixes 48 through 52 are identified and marked on the profile of Figure 17.13. This correlates bottom depths with specific locations in the water body—a necessity for mapping.

If the boat is driven back and forth along parallel range lines to cover the area of interest, and then the area traversed again with perpendicular courses, a grid of profiles results from which contours can be drawn. In larger bodies of water a compass is valuable to assist in keeping the range lines parallel. Required accuracy dictates the spacing between range lines, with closer spacing yielding more accurate results. Various other boat-positioning systems can be used, depending on circumstances. One that works well for hydrographic surveys of rivers or other relatively narrow water bodies consist in laying out uniformly spaced reference lines, which cross the water. Placing tall painted stakes on the bank on either side marks the lines. Then, fixes can be taken as the sounding boat navigates along the reference lines. However, to position each fix along the lines, either a distance must be observed from one reference point, or an azimuth to the boat from an independent control point. When the boat is moving perpendicular to the marked lines, its passage across projections between stakes locates fixes, but again a distance or angle is needed to complete the fix position.

Kinematic surveying methods (see Chapter 15) are ideal for establishing sounding locations for hydrographic surveys and also for guiding the boat along planned range lines on larger water bodies. With its many advantages, it has replaced other hydrographic positioning techniques. In this case, the transducer is located by measuring offset from the antenna reference point. Often the antenna is mounted on the sounding pole well above the top of the boat and directly over the transducer. In this case, only the distance from the antenna reference point to the transducer needs to be measured.

17.13.3 Hydrographic Mapping

Procedures for preparing hydrographic maps do not differ appreciably from those used in topographic mapping discussed in Chapter 18. Basically, depths are plotted in their surveyed positions and contours drawn. If an echo sounder is used, depths are interpolated from the profiles and plotted between fix locations. In addition to depth contours, the shoreline and other prominent features are also located on hydrographic maps. This is especially important for navigation and fishing maps, as the features are the means by which users line in and locate themselves on the water body. Planimetric features are most often located photogrammetrically (see Chapter 27), but the techniques for topographic mapping discussed in this chapter can also be used.

Modern hydrographic surveying systems utilize sophisticated electronic positioning and depth-recording devices. These, coupled with computers interfaced with plotters, enable rapid automated production of hydrographic maps in near real time. But the basic principles discussed here still apply.

17.14 SOURCES OF ERROR IN MAPPING SURVEYS

Some sources of error in planimetric and topographic surveys are:

1. Instrumental errors, especially an index error that affects vertical and zenith angles.
2. Errors in reading instruments.
3. Control not established, checked, and adjusted before beginning to collect details.
4. Control points too far apart and poorly selected for proper coverage of an area.
5. Sights taken on detail points which are too far away.
6. Poor selection of points for contour delineation.

17.15 MISTAKES IN MAPPING SURVEYS

Some typical mistakes in planimetric and topographic surveys are:

1. Unsatisfactory equipment or field method for particular survey and terrain conditions.
2. Mistakes in instrument reading and data recording.
3. Failure to periodically check azimuth orientation when many detail points are located from one instrument station.
4. Too few (or too many) contour points taken.
5. Failure to collect some mapping details.
6. Mistakes in entering point identifiers, drawing designators, and symbols when using field-to-finish surveying and mapping systems.

PROBLEMS

Asterisks (*) indicate problems that have partial answers given in Appendix G.

17.1 Describe what features are located on a topographic map.

17.2 What is the purpose of a mapping survey?

17.3 What factors must be considered when selecting the contour interval to be used for a given topographic map.

17.4 List the different methods that can be used for a ground survey to perform a mapping survey.

17.5 Why are spot elevations placed on a map?

17.6* On a map sheet having a scale of 1 in. = 360 ft, what is the smallest distance (in feet) that can be plotted with an engineer's scale? (Minimum scale graduations are 1/60th in.)

17.7 What ratio scales are suitable to replace the following equivalent scales: 1:1200, 1:2400, 1:3600, 1:4800, and 1:7200?

17.8 A topographic map has a contour interval of 1 ft and a scale of 1:600. If two adjacent contours are 0.5 in. apart, what is the average slope of the ground between the contours?

17.9* On a map whose scale is 1 in. = 50 ft, how far apart (in inches) would 2-ft contours be on a uniform slope (grade) of 2%?

17.10* On a map drawn to a scale of 1:1000, contour lines are 16 mm apart at a certain place. The contour interval is 1 m. What is the ground slope, in percent, between adjacent contours?

17.11 Similar to Problem 17.10, except for a 5-m interval, 20-mm spacing, and a map scale of 1:5000.

17.12 Sketch at a scale of 10 ft/in., the general shape of contours that cross a 20-ft wide street having a +4.00% grade, a 6-in. parabolic crown, and a 6-in. high curb.

17.13 Same as Problem 17.12, except a four-lane divided highway including a 30-ft wide median strip, two 12-ft wide lanes of pavement each side of the median having a +4.00% grade with a 1.50% side slope; and 10-ft wide shoulders both sides with a 2% side slope. The median slopes 10.0% toward the center.

17.14 When should points be located for contours connected by straight lines? When by smooth curves?

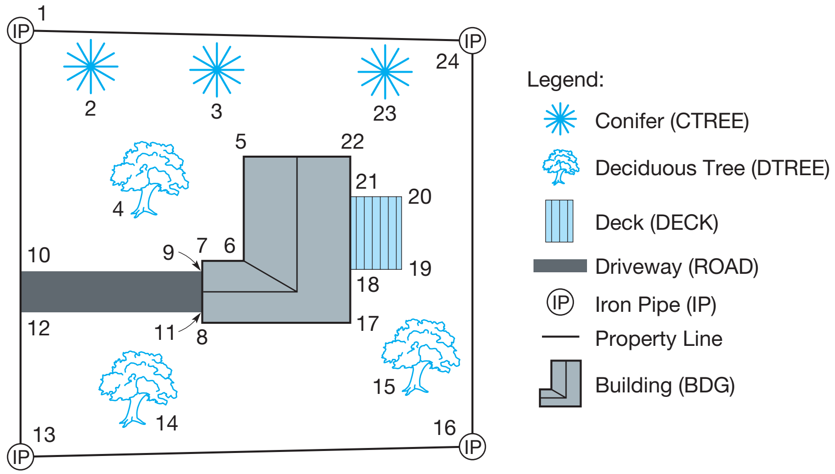
17.15* What conditions in the field need to exist when using kinematic satellite survey?

17.16 What is a digital terrain model?

17.17 Discuss why it is important to locate breaks in grade with "breaklines" in the field if contours will be drawn using a computerized automated contouring system.

17.18 What considerations should be given to a mapping survey using GNSS satellites?

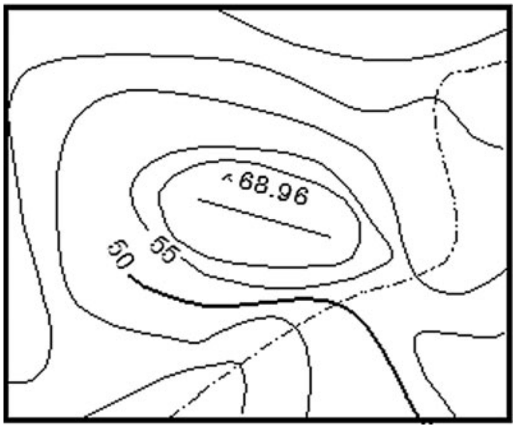
17.19 Using the labels given in parentheses in the legend of the accompanying figure create a set of notes using the drawing designators listed in Table 17.2.



Problem 17.19

17.20 How could GNSS survey methods be used where the area of interest has some overhead obstructions?

17.21 Using the rules of contours, list the contouring mistakes that are shown in the accompanying figure.



Problem 17.21

17.22 Discuss how a data collector with a total station instrument can be combined with satellite surveying methods to collect data for a topographic map.

17.23 Prepare a set of field notes to locate the topographic details in Figure 17.5. Scale additional distances and angles if necessary.

17.24 Why is it dangerous to run the control traverse at the same time as planimetric data is being collected?

17.25 Cite the advantages of locating topographic details by radial methods using a total station instrument with a data collector.

17.26 What does the term "point cloud" describe in laser scanning?

17.27 List the methods for establishing control to support a topographic mapping project.

For Problems 17.28 through 17.31, calculate the X, Y, and Z coordinates of point B for radial readings taken to B from occupied station A, if the backsight azimuth at A is $25°32'48''$, the elevation of $A = 610.098$ m, and $hi = 1.45$ m. Assume the XY coordinates of A are (10,000.000, 5000.000).

17.28* Clockwise horizontal angle $= 55°37'42''$, zenith angle $= 92°34'18''$, slope distance $= 435.098$ m, $hr = 2.000$ m.

17.29 Clockwise horizontal angle $= 272°42'22''$, zenith angle $= 92°28'16''$, slope distance $= 58.905$ m, $hr = 1.500$ m.

17.30 Clockwise horizontal angle $= 55°15'06''$, zenith angle $= 88°35'24''$, slope distance $= 103.023$ m, $hr = 1.500$ m.

17.31 Clockwise horizontal angle $= 307°56'52''$, zenith angle $= 87°17'40''$, slope distance $= 304.902$ m, $hr = 1.500$ m.

17.32 Describe how the arbitrary coordinates of a point cloud are transformed into a conventional coordinate system.

17.33 List various equipment used for making hydrographic depth soundings, and discuss the limitations, advantages, and disadvantages of each.

17.34* On a map having a scale of 200 ft/in. the distance between plotted fixes 49 and 50 of Figure 17.14 is 3.15 in. From measurements on the profile of Figure 17.13, determine how far from fix 50 the 20-ft contour (existing between fixes 49 and 50) should be plotted on the map.

17.35 Similar to Problem 17.34, except locate the 16-ft contour between fixes 50 and 51 if the corresponding map distance is 2.98 in.

17.36 Why is it important to show the shoreline and some planimetric features for navigation hydrographic maps?

17.37 Create a computational program that performs a three-dimensional conformal coordinate transformation as described in Section 17.10.

BIBLIOGRAPHY

Andelin, E. 2009. "On the Right Track." *Point of Beginning* 35 (No. 3): 12.

ASPRS. 1987. *Large Scale Mapping Guidelines*. Bethesda, MD: American Society for Photogrammetry and Remote Sensing.

Bennett, T. D. 2009. "BIM and Laser Scanning for As-Built and Adaptive Reuse Projects." *The American Surveyor* 6 (No. 6): 32.

Brinkman, B. and B. Stevens. 2009. "The Magic Bullet." *Point of Beginning* 35 (No. 1): 36.

Caneves, E. P., et al. 2009. "The Caves of Naica–Laser Scanning in Extreme Underground Environments." *The American Surveyor* 6 (No. 2): 8.

Cheves, M. 2009. "ASTM E57: 3D Imagining Systems." *The American Surveyor* 6 (No. 6): 44.

Fenicle, J. D. 2009. "Ground Penetrating Radar Holds Promise as a Practical Land Surveying Tool." *Point of Beginning* 34 (No. 7): 12.

Gardner, N. 2007. "LiDAR on a Stick." *Professional Surveyor* 27 (No. 2): 6.

Garret, J. 2007. "Reservoir of Lessons Learned." *Professional Surveyor* 27 (No. 2): 18.

Goucher, S. and B. L. Sheive. 2009. "Refined Dimensions." *The American Surveyor* 6 (No. 3): 24.

Jacobs, G. 2009. “3D Scanning: Accuracy of Scan Points.” *Professional Surveyor* 29 (No. 8): 24.

———. 2009. “3D Scanning: Laser Scanner Versatility Factors, Part 1.” *Professional Surveyor* 29 (No. 10): 34.

———. 2010. “3D Scanning: Laser Scanner Versatility Factors, Part 2.” *Professional Surveyor* 30 (No. 1): 32.

Longstreet, B. 2009. “Laser Scanning Brings New Asset to Accident Investigations and Surveyors.” *The American Surveyor* 6 (No. 7): 19.

Pesci, A. D. Conforti and M. Bacciocchi. 2007. “Morphing Mount Vesuvius.” *Professional Surveyor* 27 (No. 2): 12.

Rameriz, J. R. 2006. “A New Approach to Relief Representation.” *Surveying and Land Information Science* 66 (No. 1): 19.

Stewart, P. and P. Canter. 2009. “Creating a Seamless Model.” *Professional Surveyor* 29 (No. 8): 18.

Wagner, M. Jo. 2009. “Scanning the Horizon.” *Point of Beginning* 35 (No. 2): 24.

This page intentionally left blank

18 Mapping

18.1 INTRODUCTION

Maps are visual expressions of portions of the Earth's surface. Features are depicted using various combinations of points, lines, and standard symbols. Maps have traditionally been produced in *graphic*, or "hardcopy," form, that is, printed on paper or a stable-base plastic material. However, today most mapping data are collected in digital form and are then processed using *Computer Aided Drafting and Design* (CADD) systems to develop "softcopy" maps. Softcopy maps are stored within a computer, can be analyzed, modified, enlarged or reduced in scale, and have their contour intervals changed while being viewed on the monitors of CADD systems. Different types or "layers" of information can also be extracted from digital maps to be represented and analyzed separately, and softcopy maps can be transferred instantaneously to other offices or remote locations electronically. Of course they can also be printed in hardcopy form if desired. Softcopy maps are indispensable in the development and operation of modern *Land Information Systems* (LISs) and *Geographic Information Systems* (GISs) (see Chapter 28).

Throughout the ages, maps have had a profound impact on human activities and today the demand for them is perhaps greater than ever. They are important in engineering, resource management, urban and regional planning, management of the environment, construction, conservation, geology, agriculture, and many other fields. Maps show various features—for example, topography, property boundaries, transportation routes, soil types, vegetation, land ownership, and mineral and resource locations. Maps are especially important in engineering for planning project locations, designing facilities, and estimating contract quantities.

As noted above, maps are essential in the development and operation of LISs and GISs. These systems for spatial data analysis and management use the computer to store, retrieve, manipulate, merge, analyze, display, and disseminate information by means of digital maps (see Chapter 28). GISs have applications in virtually every field of endeavor. Spatial databases to support the systems are generally developed either by digitizing existing graphic maps or by generating new digital maps in the computer based upon digitized ground survey or photogrammetric data. Maps of various types needed to create spatial databases for LISs and GISs include topographic maps, which show the natural and cultural features and relief in an area; cadastral maps, which give boundaries of land ownership; natural resource maps, which provide the location and distribution of forest and water resources, wetlands, soil types, etc.; facilities maps, which show existing transportation networks, water and sewer mains, and distribution lines for electric power; and land-use maps, which show the various activities of humans related to the land. Applications of GIS technology have been expanding at a phenomenal rate and these activities will impose a heavy demand for high-quality maps of various types and scales in the future.

Cartography, the term applied to the overall process of map production, includes map design, preparing or compiling manuscripts, final drafting, and reproduction. These processes, which apply whether the maps are graphic or digital, are described in this chapter.

18.2 AVAILABILITY OF MAPS AND RELATED INFORMATION

Maps for a variety of different purposes, prepared at scales varying from large to small, and in both graphic and digital form, are prepared by private surveying and engineering companies, industries, public utilities, cities, counties, states, and agencies of the federal government. Unfortunately, with such a wide range in organizations and agencies involved, some duplication of effort has occurred because mapping activities generally have not been coordinated. Also, the existence of available maps and related information is often unknown to potential users. However, steps have been taken to improve this situation. The U.S. Geological Survey (USGS) now coordinates all mapping activities at the federal level. They offer nationwide information and sales service for map products and Earth science publications. The USGS provides information about topographic, land use, geologic and hydrologic maps, books and reports; Earth science and map data in digital format and related applications software; aerial, satellite, and radar images and related products; and geodetic data.[1]

Several states have established state cartographers' offices. One of their functions is the dissemination of local maps and related products and information to surveyors, engineers, cartographers, and the general public.

[1]The U.S. Geological Survey can be reached by telephone at (888)ASK-USGS, [(888)275-8747]. Information can also be obtained, and selected maps and other products can be downloaded at the following website: http://www.usgs.gov/pubprod/. Contact by mail can be made to the U.S. Department of the Interior, U.S. Geological Survey, 12201 Sunrise Valley Drive, Reston, VA 20192.

■ 18.3 NATIONAL MAPPING PROGRAM

The *National Mapping Program* was established to provide maps and other cartographic products needed by the citizens of the United States. This is the responsibility of the *National Mapping Division* of the USGS. The USGS began publishing topographic maps in 1886 as an aid to scientific studies. It now produces a variety of topographic maps at differing scales; however, its standard series has a scale of 1:24,000. In this series, individual sheets cover *quadrangles* of 7 1/2-min in both latitude and longitude. Each quadrangle map is named, usually according to the most prominent feature within its bounds. Except for Alaska, the entire United States is covered at the 1:24,000 scale and over 57,000 maps are involved. [Maps covering 15′ quadrangles at a scale of 1:63,360 (1 in./mi) are standard for Alaska.] On the quadrangle maps, cultural features are shown in black, contours in brown, water features in blue, urban regions in red, and woodland areas in green. Topographic coverage of the United States is also available at scales of 1:50,000 (county maps), 1:62,500 (the older 15′ quadrangles produced until about 1950), 1:100,000, and 1:250,000. The USGS has also published a state map series. Most are at a scale of 1:500,000, but a few are at 1:1,000,000 or other scales.

As noted in Section 18.1, requirements for digital cartographic data are growing rapidly to support LISs and GISs. To meet these needs, the U.S. Geological Survey has developed two very useful types of digital data: (1) *digital line graphs* (DLGs) and (2) *digital elevation models* (DEMs). Primarily digitizing existing maps and other cartographic products is generating these products. The digital line graphs contain only linear features or planimetry in an area. Included are political boundaries, hydrography, transportation networks, and the subdivision lines of the U.S. Public Land Survey System (see Chapter 22). The digital elevation models are arrays of elevation values, produced in grids of varying dimensions, depending on the source of the information. The horizontal positions of points in the DEMs are X and Y coordinates referenced to the Universal Transverse Mercator coordinate system (see Section 20.12). A grid of 30 m is used for DEMs generated from 7 1/2′ quadrangles, with larger spacings being used for those generated from smaller scale maps.

In addition to topographic maps, digital line graphs, and digital elevation models, a variety of other special maps and related products are published as a part of the National Mapping Program. As noted in the preceding section, information on all of their available maps and other related products is obtained through the U.S. Geological Survey.

■ 18.4 ACCURACY STANDARDS FOR MAPPING

To provide a set of uniform standards for guiding the production of maps, and to protect consumers of maps, the United States Bureau of the Budget developed the *National Map Accuracy Standards* (NMAS). These standards, first published in 1941 and revised in 1947, provide specifications governing both the horizontal and vertical accuracy with which features are depicted on maps. Published maps meeting these accuracy requirements may have the following note in their legends: "This map complies with National Map Accuracy Standards," thereby providing assurance that the map meets these specified accuracy levels.

To meet the NMAS horizontal position specification, for maps produced at scales larger than 1:20,000, not more than 10% of well-defined points tested shall be in error by more than 1/30 in. (0.8 mm). Accordingly, on a map plotted to a scale of 1 in. = 100 ft, point positions would have to be correctly portrayed to within ±3.3 ft to meet this specification. On smaller scale maps, the limit of horizontal error is 1/50 in. (0.5 mm), or approximately ±40 ft on the ground at a map scale of 1:24,000. These limits of accuracy apply to positions of well-defined points only, such as monuments, benchmarks, highway intersections, and building corners.

The NMAS vertical accuracy requirements specify that not more than 10% of elevations tested shall be in error by more than one half the contour interval, and none can exceed the interval. To meet this requirement, contours may be shifted by distances up to the horizontal positional tolerance (discussed above), if necessary.

The accuracy of any map can be tested by comparing the positions of points whose locations or elevations are shown on it with corresponding positions determined by surveys of a higher order of accuracy. Plotted horizontal positions of objects are checked by running an independent traverse or other survey to points selected by the person or organization for which the map was made. To check vertical accuracy, elevations obtained from field profile surveys are compared with elevations taken from profiles made from plotted contours. These procedures provide a check on both fieldwork and map drafting.

When the NMAS were developed, maps were being produced in hardcopy form. But as noted in Section 18.1, softcopy maps are now most common. To accommodate this change, the *Federal Geographic Data Committee* (FGDC) drafted a more current set of accuracy standards called the *Geospatial Positioning Accuracy Standards*.[2] The FGDC is composed of representatives from 19 Federal agencies and was established to coordinate policies, standards, and procedures for producing and sharing geographic information. The new Geospatial Positioning Accuracy Standards are completed, and like NMAS, the document specifies accuracies in separate horizontal and vertical components. But unlike NMAS, accuracies are specified in terms of coordinates of points, ground distances, and elevations at the 95% confidence level. Thus these new standards are applicable to all types and scales of maps, including those in digital form. The test for maps intended to meet this standard involves checking a set of at least 20 well-defined points against information obtained from an independent source of higher accuracy. Root mean square errors are computed and converted to the 95% confidence level by using appropriate multipliers (see Section 3.16). A digital planimetric map that passes at the 1-m level, for example, would contain the

[2]Information on the status of the **GEOSPATIAL POSITIONING ACCURACY STANDARDS** can be obtained from the Federal Geographic Data Committee by telephone at (703)648-5514, or at the following website: http://www.fgdc.gov. Copies of the current standards can also be downloaded at this website. Contact can also be made by mail at The Federal Geodetic Data Committee, U.S. Geological Survey, 590 National Center, Reston, VA 20192.

statement "Tested 1-meter horizontal accuracy at 95% confidence level" in its legend. A similar statement can be included that applies to a map's vertical accuracy.

The USGS has developed its own standards to govern the production of the maps and other products it provides through the National Mapping Program.[3] Standards have been developed not only for their hardcopy maps, but also for their digital products including digital elevation models, digital line graphs, digital orthophotos (see Section 27.15), and others.

The *American Society for Photogrammetry and Remote Sensing* (ASPRS) has also adopted its own standards to govern photogrammetric production of large-scale maps. It specifies standards for three levels of accuracy, classes 1, 2, and 3. For a map to meet its class 1 standards, the root mean square (rms)[4] error in both X and Y coordinates of well-defined points must not exceed ± 0.01 in. at map scale. Thus for a map scale of 500 ft/in., the allowable rms error in X and Y coordinates is ± 5.0 ft. Vertical accuracy is specified in terms of the map's contour interval (CI). For class 1, the rms error of well-defined points must not exceed $\pm(\text{CI}/3)$. These horizontal and vertical standards are both relaxed by factors of 2 and 3 for class 2 and class 3 maps, respectively.

The *American Society of Civil Engineers* (ASCE) has also developed a set of standards for topographic mapping that are aimed primarily at large-scale engineering maps. In addition to suggesting accuracies for various map scales, they also provide standards for contouring, map symbols, abbreviations, lettering, and other factors important in mapping.

18.5 MANUAL AND COMPUTER-AIDED DRAFTING PROCEDURES

As previously stated, maps may be drafted manually or produced with CADD systems. Manual procedures utilize standard drafting tools such as scales, protractors, compasses, triangles, and T-squares. CADD systems employ computers programmed with special software and interfaced with electronic plotting devices. With either approach, after deciding on scale and other factors that control overall map design, a manuscript is prepared. When completed, final drafting is performed.

In manual drafting, the manuscript is usually compiled in pencil. It should be prepared carefully to locate all features and contours as accurately as possible and be complete in every detail, including placement of symbols and letters. Lettering on the manuscript need not be done with extreme care, for its major purpose at this stage is to ensure good overall map design and proper placement. A well-prepared manuscript goes a long way toward achieving a good-quality final map.

[3]For information on the USGS standards, visit the following website: http://nationalmap.gov/gio/standards/. Contact can also be made by telephone at (888) ASK-USGS [(888) 275-8747], or by email at ask@usgs.gov.

[4]The rms error is defined as the square root of the average of squared discrepancies for points tested. Discrepancies are the differences between coordinates and elevations of points taken from the map, and their values as determined by check surveys.

The completed version of the manually compiled manuscript is drafted in ink, or *scribed*. If inked, the manuscript is placed on a light table, and features are traced on a stable-base transparent overlay material. Lettering is usually done first; then planimetric features and contours are made. Scribing is performed on sheets of transparent stable-base material coated with an opaque emulsion. Manuscript lines are transferred to the coating in a laboratory process. Lines representing features and contours are then made by cutting and scraping to remove the coating. Special scribing tools are used to vary line weights and make standard symbols. Scribing is generally easier and faster than inking. Reproductions are made from the finished inked or scribed product.

In drafting maps with CADD, a softcopy manuscript is compiled in the computer and displayed on its screen as work progresses. CADD software provides instructions to the computer, which basically duplicate manual drafting functions. A file containing coordinates of points, as well as specific instructions on how to plot them, must be input in preparation for mapping. An operator interactively designs and compiles the map by entering commands into the computer's keyboard or using a mouse to activate functions on a menu. Points, lines of various types, and a variety of symbols are available to the operator. Letters of differing sizes and styles can also be selected. When the manuscript is completely finished, simply activating the electronic plotter will draw the final map. Map production using CADD systems has many advantages over manual methods and therefore it has rapidly become the method of choice by most mapping firms. However, it is still important to learn the basics of manual map drafting techniques, since these are often duplicated in the CADD processes. CADD systems are described in more detail in Section 18.14.

18.6 MAP DESIGN

Before beginning the design of a map, the following two basic questions should be answered: (1) What is the purpose of the map? and (2) Who is the map intended to serve? All maps have a purpose, which in turn dictates the information that the map must convey. Once the purpose of the map is fixed, emphasis should be placed on achieving the design that best meets its objectives and conveys the necessary information clearly to its users.

Maps typically depict many different types and classes of details in portraying natural and cultural features and, if properly designed, they can convey an enormous amount of information. On the other hand, maps that are carelessly designed can be confusing, difficult to read, understand, or interpret. To achieve maximum effectiveness in map design, the following elements or factors should be considered: (1) clarity, (2) order, (3) balance, (4) contrast, (5) unity, and (6) harmony. Definitions of these six elements in relation to map design and explanations of their interdependence are discussed below.

1. *Clarity* relates to the ability of a map to convey its intended information completely and unambiguously. It can only be achieved after fully examining the objectives of the map and then emphasizing the features necessary to carry out those objectives. Maps should not be overloaded with details, as

this can cause congestion and confusion. If considerable detail must be included on a map, the information could be placed in a table. Other alternatives consist of preparing larger-scale *inset maps* of areas that contain dense detail, or creating an overlay to display some of the detail. The proper use of textual elements is very important in achieving clarity. (The subject of map lettering is covered in Section 18.11.)

2. *Order* refers to the logic of a map, and relates to the path that a user's eye would follow when looking at one. A design should be adopted that first draws the user's attention to the subject area of the map, then the map title, and then to any notes. Never let auxiliary elements such as bar scales and directional arrows dominate the map. A common mistake made by beginners is to make bar scales and north arrows so large and bold that they attract attention away from the subject of the map.
3. All elements on a map have weight, and they should be distributed uniformly around the "visual center" of the map to create good overall *balance*. The visual center is slightly above the geometrical center of the map sheet. In general, the weight of an element is affected by factors such as size, color, font, position, and line width. Map elements that appear at the center have less weight than those on the edges. Elements in the top or right half of the map will appear to have more weight than those in the bottom or left half of the map. Also map elements identified with thicker line widths will appear to have heavier weights than their slimmer counterparts. Colors such as red appear heavier than blue or yellow. Two examples illustrating balance are shown in Figure 18.1. In Figure 18.1(a), the map appears to be too heavily weighted to the left, and thus has poor overall balance. A redistribution of the map features, as shown in Figure 18.1(b), produces a more visually balanced product. The use of thumbnail sketches can often help to

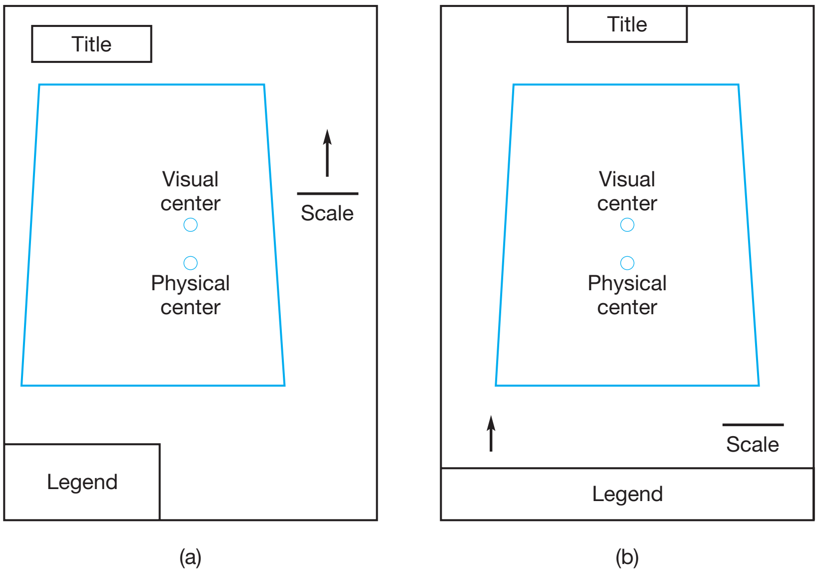


Figure 18.1
(a) The layout of a poorly balanced map sheet. (b) A better layout for the same map.

achieve a balanced layout for a map. It is important to place highest weights on those elements that enhance the purpose of the map.

4. *Contrast* relates primarily to the use of different line weights and fonts of varying sizes. Contrast can be used to enhance balance, order, and clarity. For example, the title of the map should be displayed in a larger font than the other textual elements. This will attract the viewer's attention, thereby enhancing the order and clarity of the map. Various fonts can also be used to provide balance with other elements on the map. Another example where contrast supports the clarity of a map is in contouring. Here *index contours* (every fifth contour) should be drawn with a heavier line than the other contours. This enhances the map's clarity and facilitates the determination of elevations.
5. *Unity* refers to the interrelationships between the backgrounds, shading, and colors on a map. Again these items can enhance clarity, balance, and contrast. They can also detract from these same items. For example, yellow lettering on a white background is difficult to see and often overlooked by the reader. However, this same yellow lettering on a black background will stand out and appear emphasized. A map with good unity is visualized as a unit and not as an assemblage of individual elements.
6. *Harmony* relates to the interrelationships between all elements on the map. If a map has good harmony, the elements work together. Common errors are the use of too many fonts, a north arrow that is too fancy or large, or a bar scale that is too large.

In designing maps, it is important to remember that different audiences may require different maps. For example, it would be difficult for a layperson to read and understand a map produced for an engineering project. Accordingly, maps that are developed for design professionals are not generally suitable for public hearings. In fact because laypeople often have no training in map reading, it may be best to develop specialized three-dimensional maps or models that depict relief, boundaries, proposed buildings, landscaping, and so on.

Sometimes when designing a map, certain elements of map design will be in conflict. When this happens, priorities must be established that provide a reasonable solution to the conflict. A perfect map rarely, if ever, exists, and there are generally several equally acceptable designs that could be adopted. Often there are design conflicts that cannot be resolved and a compromising solution may have to be accepted. Map creation is often subjective and the production of a well-designed map requires a combination of skill, art, and patience.

18.7 MAP LAYOUT

In general, the subject area of the map should be plotted at the largest scale that will enable it to fit neatly within its borders without producing overcrowding. It should also be centered on the map sheet and, if possible, should be aligned so that the edges of the map sheet coincide with the cardinal directions. If this is not done, users may experience some confusion when viewing the map. Accordingly, the size and shape of the map sheet, the size and shape of the area to be mapped,

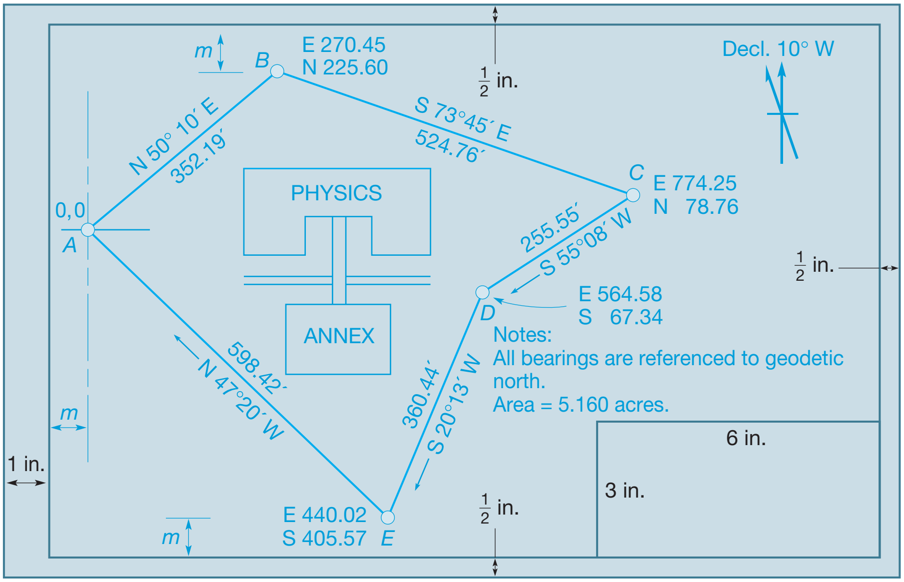


Figure 18.2
Map layout.

the orientation of the subject area on the map sheet, and map scale, must be jointly considered in map layout.

To illustrate, consider the example of Figure 18.2, which is a simple traverse from a planimetric survey. Before any plotting is done, the proper scale for a sheet of given size must be selected. Assume in this example that an 18- by 24-in. sheet will be used, with a 1-in. border on the left (for possible binding) and 1/2-in. borders on the other three sides. A borderline somewhat heavier than all other lines can be drawn to outline this area. If the most westerly station (*A* in the example) has been chosen as the origin of coordinates, then divide the total departure to the most easterly point *C* by the number of inches available for plotting in the east-west direction. The maximum scale possible in Figure 18.2 is 774.25 divided by 22.5, or 1 in. = 34 ft. The nearest standard scale that will fit is 1 in. = 40 ft.

This scale must be checked in the *Y* direction by dividing the total difference in *Y* coordinates, 225.60 + 405.57 = 631.17, by 40 ft, giving 15.8 in. required in the north-south direction. Since 17 in. are usable, a scale of 1 in. = 40 ft is satisfactory, although a smaller scale would yield a larger border margin. If a scale of 1 in. = 40 ft is not suitable for the map's purpose, a sheet of different size should be selected, or alternatively more than one sheet employed to map the required area.

In Figure 18.2, the traverse is centered between the borderlines in the *Y* direction by making each distance *m* equal to 1/2(17 − 631.17/40), or 0.61 in. The same 0.61 in. can be used for the left side. Weights of the title, notes, and north arrow compensate for the traverse being to the left of the sheet center and leave ample space for including the necessary auxiliary elements[5] of the map.

[5]Auxiliary elements include the map title block, notes, legend, bar scale, and north arrow. These elements are described in Section 18.12.

If a compromising choice must be made between map scale and the sizes of auxiliary elements, it is better to maximize the map's scale and minimize the size of the auxiliary elements. Beginners should avoid using oversized auxiliary elements to use up available or leftover space since doing so detracts from the map's order and balance.

18.8 BASIC MAP PLOTTING PROCEDURES

Map plotting may be done either manually or by employing automated CADD systems. Regardless of which method is used, the procedure consists fundamentally of plotting individual points. Lines are then drawn from point to point to portray features. Although this process may seem simple in principle, accurate work requires skill, patience, and care. While points could be laid out using angles and distances, or lines scaled and plotted directly, the most convenient method for laying out points and drafting maps involves plotting points by coordinates. Plotting by this procedure is also consistent with today's modern data collection systems, that is, total station instruments and portable GNS units, because those devices provide coordinates directly. The following subsections describe manual and CADD coordinate plotting procedures.

18.8.1 Manually Plotting by Coordinates

To plot points by coordinates, the map sheet is first laid out precisely in a grid pattern with unit squares of appropriate size. Squares of 2, 4, or 5 in. are commonly used, and depending on map scale, they may represent 100, 200, 400, 500, or 1000 ft, or 50, 100, 200, or 500 m. The grids are constructed using a sharp, hard pencil, and are checked by carefully measuring diagonals. The grid lines are labeled with coordinate values, making sure that the range of coordinates covered on the map will accommodate the most extreme X and Y coordinates to be plotted.

Initially, coordinates for all features to be mapped must be determined. These points are then plotted by laying off their X and Y coordinates from the ruled grid lines. Mistakes in plotting can be detected by comparing scaled lengths (and directions) of lines with their measured or computed values. Since each point is plotted independently, a mistake in one will not affect the others, and that point can simply be corrected.

Many mapping elements such as bar scales, legends, north arrows, etc. are prepared on separate map sheets. These items along with the map are then cut to their dimensions and manually located for optimal presentation. After the desired arrangement is achieved, the entire layout is copied onto a more stable medium as discussed in Section 18.5.

18.8.2 Plotting Using CADD

Fundamentally, CADD systems plot points and lines in a manner similar to manual drafting techniques. However, compared to manual map drafting, computer-assisted mapping offers advantages of increased accuracy, speed, flexibility, and reduced cost. Computers are capable of quickly performing many drafting chores that are tedious and time consuming if done by manual methods, for example, drawing complicated line types and symbols, and performing lettering. With CADD systems, lettering reduces to simply choosing letter sizes and styles and selecting and monitoring placement. Since these systems can often read files of coordinates, such as those from data collectors, the plotting process can become almost totally automated (see Section 17.11). For example, many common features of a map such as bar scale, north arrow, legend, and title block can be created as *blocks* and imported into any map with varied scales. This process simplifies the entire map production process and creates a standardized look for a mapping agency or company. Additionally, the digital environment of a CADD system allows for the easy arrangement of the mapping elements, which simplifies the process of map design and enables colors to be readily selected and changed.

In preparing the topographic data for computer drafting, it is advantageous to develop files that group similar categories of features in separate layers. As an example, individual layers can be created for buildings, vegetation, transportation routes, utilities, hydrology, contours, and so on. By developing the data structure in this manner, various types of maps at different scales can be produced from the same original topographic data file. This is particularly advantageous in mapping for LISs and GISs. This feature also enables specialized products to be created such as *ephemeral maps*—that is, those produced only for current conditions and then redrawn as the conditions change. Examples of these types of maps are those used in guidance, where a map is generated showing the location of a vehicle at a particular moment in time, and as the vehicle moves, the map is instantaneously redrawn on the monitor to replace its earlier version.

In order that digital map files developed during CADD drafting processes are suitable for importing into GISs, it is important that all closed features are actually "physically" closed in the mapping files. As shown in Figure 18.3, a frequent mistake in CADD mapping is the failure to close polygons that appear on the screen to be "visually" closed, but which in fact are not. Since GIS software packages use polygons to represent features, the visually closed but physically open features could appear as simply a series of random lines, or even be viewed as errors in the drawing when imported into the GIS package.

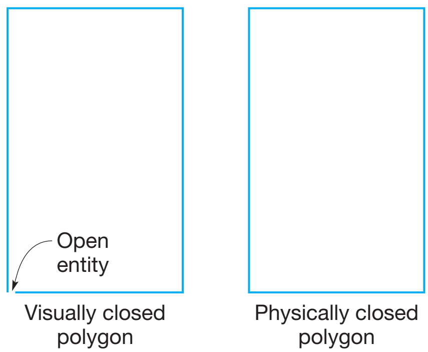


Figure 18.3
A visually closed feature versus a physically closed feature.

18.9 CONTOUR INTERVAL

As noted in Section 17.5, the choice of contour interval to be used on a topographic map depends on the map's intended use, required accuracy, type of terrain, and scale. If, according to National Map Accuracy Standards (see Section 18.4), elevations can be interpolated from a map to within one half the contour interval, then if elevations taken from the map must be accurate to within ± 1 ft, a 2-ft maximum interval is necessary. However, if only 10-ft accuracy is required, a 20-ft contour interval will suffice.

Terrain type and map scale combine to regulate the contour interval needed to produce a suitable density (spacing) of contours. Rugged terrain requires a larger contour interval than gently rolling country and flat ground mandates a relatively small one to portray the surface adequately. Also if the map scale is reduced, the contour interval must be increased; otherwise, lines are crowded, confuse the user, and possibly obscure other important details.

For average terrain, the following large and medium map scales and contour interval relationships generally provide suitable spacing:

English System		Metric System	
Scale (ft/in.)	**Contour Interval (ft)**	**Scale**	**Contour Interval (m)**
50	1	1:500	0.5
100	2	1:1000	1
200	5	1:2000	2
500	10	1:5000	5
1000	20	1:10,000	10

18.10 PLOTTING CONTOURS

In plotting contours, points used in locating them are first plotted on the map following techniques described in Section 18.8. Contours found by the *direct method* (see Section 17.7.1) are sketched through the plotted points. Interpolation between plotted points is necessary for the *indirect method* (see Section 17.7.2).

Interpolation to find contour locations between points of known elevation can be done in several ways:

1. Estimating.
2. Scaling the distance between points of known elevation and locating the contour points by proportion.
3. Using special devices called *variable scales*, which contain a graduated spring. The spring may be stretched to make suitable marks fall on the known elevations.
4. Using a triangle and scale, as indicated in Figure 18.4. To interpolate for the 420-ft contour between point *A* at elevation 415.2 and point *B* at elevation 423.6, first set the 152 mark on any of the engineer's scales opposite *A*. Then, with one side of the triangle against the scale and the 90° corner at 236, the scale and triangle are pivoted together around *A* until the

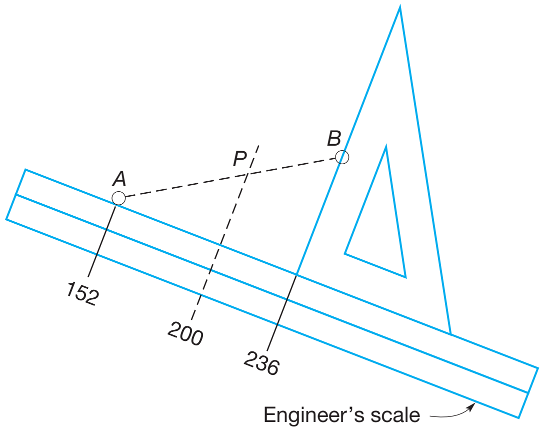


Figure 18.4 Interpolating using engineer's scale and triangle.

perpendicular edge of the triangle passes through point *B*. The triangle is then slid to the 200 mark and a dash drawn to intersect the line from *A* to *B*. This is the interpolated contour point *P*.

Contours are drawn only for elevations evenly divisible by the contour interval. Thus for a 20-ft interval, elevations of 800, 820, and 840 are shown, but 810, 830, and 850 are not. To improve legibility, every fifth line (those evenly divisible by five times the contour interval) is made heavier. So for a 20-ft interval, the 800, 900, and 1000 lines would be heavier.

■ 18.11 LETTERING

An important part of the contents of any map is its textual information. The title and all feature names, coordinate values, contour elevations, and other items must be clearly identified. To produce a professional looking drawing and one that clearly conveys the intended information, a suitable style of lettering must be selected. That style should be used consistently throughout the map, but the size varied in accordance with the importance of each particular item identified. Lettering that is too big or bold should not be used, but the letters must be large enough to be readable without difficulty.

Lettering should be carefully placed so that it is clearly associated with the item it identifies and so that letters do not interfere with other features being portrayed. Typically, the best balance results if names are centered in the objects being identified. Also both appearance and clarity are generally improved by aligning letters parallel with linear objects that run obliquely, as has been done with the traverse lengths and bearings of Figure 18.2. For ease in map reading, letters should be placed so that the map can be read from either the bottom or its right side.

Text should take precedence over line work. If necessary, lines should be broken where text is placed, as this improves clarity. An example of this is in the labeling of contours, where the lines are preferably broken and the contour elevation inserted in the break. It is best to select straight, or nearly straight, sections of contours for labeling. Contours should not be labeled around tight turns since this will remove valuable topographic information expressed by the contours. When manually drafting a map from a manuscript, the text should be lettered before the line work, and then during drafting the lines can be broken where text is encountered. When using automated drafting techniques, manuscripts must be carefully examined to make sure that the text and lines do not overwrite each other and any observed overwrites corrected.

Because of the importance of lettering to a map's overall appearance and utility, even when drafting is done manually, the text is seldom hand-lettered. Rather, mechanical lettering devices that produce uniform sizes and styles of inked letters, or special machines that print letters on adhesive-backed transparent tape are usually used. With the latter device, a variety of fonts and sizes are available. After the letters are printed, they are pasted onto the map, but can be lifted and moved later if necessary.

In computer-assisted mapping, lettering is greatly simplified. A wide choice of fonts and sizes are available and letters can be easily placed, aligned, rotated, and moved. However, to ensure a good-quality final product, the same rules stated above for manual lettering should be followed when using CADD. A common mistake by some in CADD is to use too many fonts.

■ 18.12 CARTOGRAPHIC MAP ELEMENTS

Notes, legends, bar scales, meridian arrows, and title blocks are essential cartographic elements included on maps. *Notes* cover special features pertaining specifically to a particular map. The following are examples:

> All bearings are geodetic (or magnetic or grid or record/deed number).
> Coordinates are based on the NAD83 Pennsylvania North Zone State Plane Coordinates.
> Datum for elevations is the NAVD88.
> Area by calculation is X acres (or hectares).

Notes must be in a prominent place where they are certain to be seen upon even a cursory examination of the map. The best location is near the title block. This is suggested because the user of a map will find and identify a specific plot by its title, and then presumably also check any special notes beside the title before examining the drawing.

Cartographic symbols and different line types are commonly used to represent and portray different topographic features on maps, and *legends* are employed to explain the meaning of those symbols and lines. Figure 18.5 gives some of the hundreds of symbols and line types employed in topographic mapping. The symbols shown in the legend should be replicas of those used on the map. In a CADD environment, it is usually expedient to copy the element to the legend from the map, and reduce its scale to match the size of the font. Often legend symbols are created as blocks in CADD, and recalled for later use in both the map and legend. Any symbol that is not self-explanatory should appear in the legend. Often the legend can be used to balance other map elements. Sometimes, especially if there are an unusually large number of elements in the legend, it is best to create a separate legend sheet for the map.

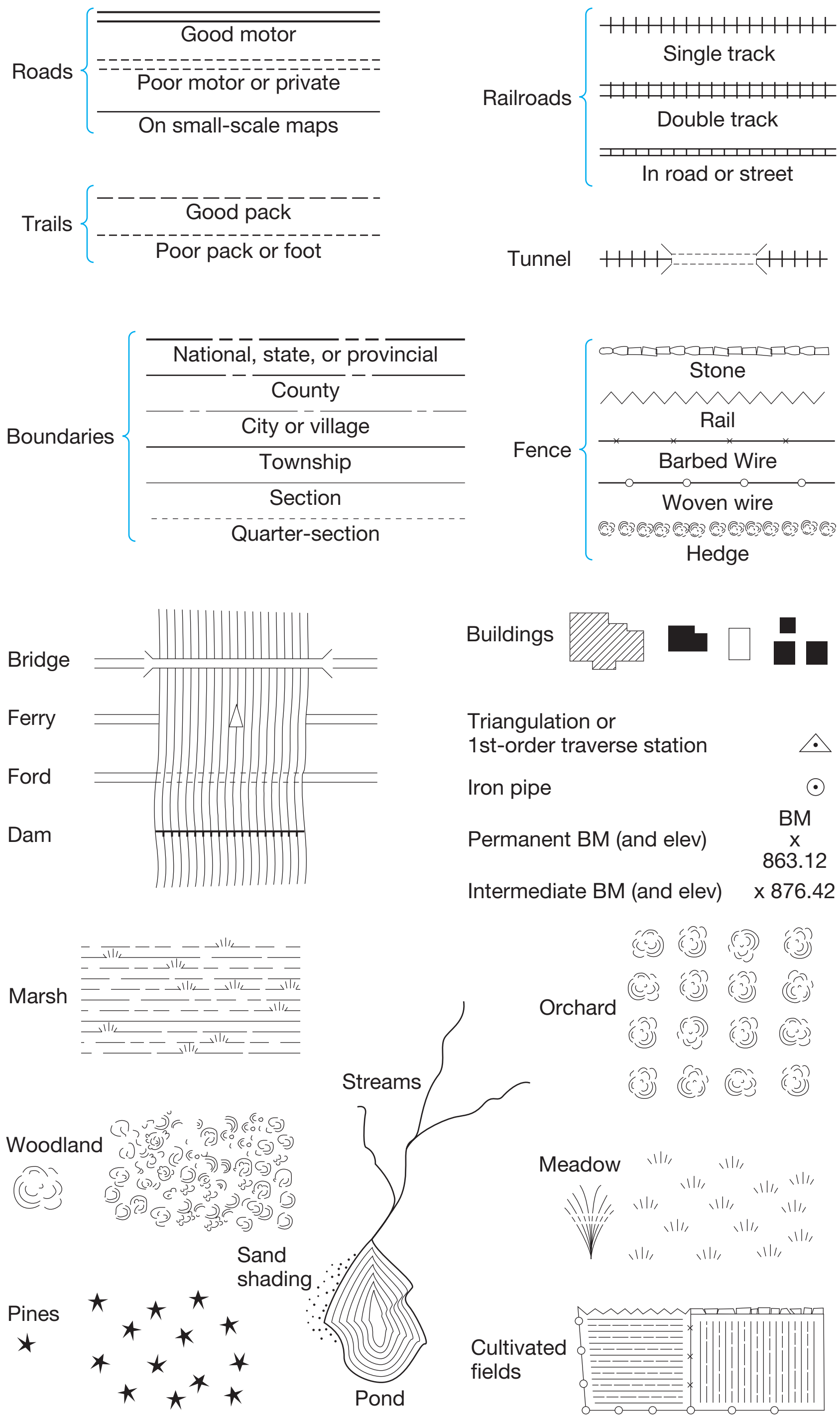


Figure 18.5 Topographic symbols.

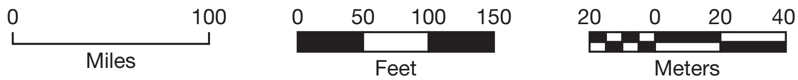


Figure 18.6 Typical graphical scales in mapping.

The *scale* of the map should preferably be presented as both a representative fraction and a graphical element. A few typical examples of graphical scales are shown in Figure 18.6. Note that the units are associated with each scale. If a map sheet is enlarged or reduced in a reproduction process, the graphical scale will change accordingly, and thus the original scale of the map will be preserved on the reproduction. When designing a bar scale, it is important to maintain a narrow bar. If the bar becomes too wide, it will draw undue attention.

Every map must display a *meridian arrow* for orientation purposes. However, the arrow should not be so large or elaborate that it becomes the focal point of a sheet. Geodetic, grid, or magnetic north (or all three) may be shown. Often the true-meridian arrow is identified by a full head and full feather; and a grid and/or magnetic arrow by a half-head and half-feather. The half-head and half-feather are put on the side away from the true north arrow to avoid touching it. The identity of the reference meridian used should be noted above or below the arrow in text to define the reference system. When magnetic directions are shown, the declination at the time of the survey should be indicated on the map. If a grid meridian is used, the grid system should be referenced. If an assumed meridian is used, information that will allow the reader to recreate the meridian in the field should be provided.

The *title block* should state the type of map, name of property or project and its owner or user, location or area, date completed, scale, contour interval, horizontal and vertical reference systems (datum) used, and for property surveys, the name of the surveyor with his or her license number. Additional data may be required on special-purpose maps. The title block may be placed wherever it will best balance the sheet, but it always should be kept outside of the subject area. Searching for a particular map in a bound set of maps, or loose pile of drawings, is facilitated if all titles are in the same location. Since sheets are normally filed flat, bound along the left border, or hung from the top, the lower right-hand corner is the most convenient position. Lettering within the title block should be simple in style rather than ornate and conform in size with the individual map sheet. Emphasis should be placed on the most important parts of the title block by increasing letter size or using uppercase (capital) letters for them. Perfect symmetry of outline about a vertical centerline is necessary since the eye tends to exaggerate any defection. An example title block is given in Figure 18.7. No part of a map

The Pennsylvania State University
Surveying Program

SURVEY OF CAMPUS

Scale: 1:480	Date: 16 Sept. 2000
Survey by: P. Dills and J. Tills	Map by: S. Smith

Figure 18.7 Title arrangement.

better portrays the artistic ability of the compiler than a neat, well-arranged title block. Today many companies and government agencies use sheets with preprinted title forms to be filled in with individual job data, or with CADD systems standard title blocks are stored, retrieved, and modified as appropriate for each new project.

18.13 DRAFTING MATERIALS

Polyester film and tracing and drawing papers are the materials commonly used for preparing maps in surveying and engineering offices. Polyesters such as Mylar are by far most frequently employed because they are dimensionally stable and are also strong, durable, and waterproof. In addition, they take pencil, ink, and stickup items, and withstand erasing, so they are ideal for manual drafting. Tracing papers are available in a variety of grades, and good ones also are stable, take pencil, ink, and stickups, and endure some erasing. Both Mylar and tracing paper are transparent, so blueprints can be made from them.

Papers of different types and grades are used for printing maps made with CADD systems. When CADD is used, the paper quality can be relaxed somewhat because erasing, stickup lettering, etc., will not be necessary. *If accurate measurements are to be extracted from the maps, then material with good dimensional stability should be used*.

18.14 AUTOMATED MAPPING AND COMPUTER-AIDED DRAFTING SYSTEMS

Digital computers have had a profound impact in all areas of life and surveying and mapping is certainly no exception. *Automated mapping* (AM) and *computer-aided drafting and design* (CADD) systems have now become commonplace in surveying and engineering offices throughout the world. Generic CADD systems developed for general drafting and engineering work are widely used for map drafting. In addition, special AM systems have been designed specifically for surveying, mapping, and GIS work.

The hardware necessary for AM and CADD systems varies, but as a minimum it will include a computer with a hard drive, at least one disk drive, and a high-resolution monitor; an input device such as a digitizer and/or mouse; and a plotting device. The most important component of any CADD system is its software. This enables an operator to interact with the computer and activate the system's various functions.

CADD systems enable operators to design and draw manuscript maps in real time using the computer. A visual display of the manuscript can be examined on the monitor as it is being compiled, and any additions, deletions, or changes can be made as needed. Lines can be added, deleted, or their styles altered; placement of symbols and lettering modified; and lettering sizes and styles varied. Parts of the drawing may be "picked up" and moved to other areas to resemble a "cut and paste" operation that is handy for subdivision design or placement of frequently occurring symbols. A zoom feature allows more complicated or crowded parts of the manuscript to be magnified for better viewing. In the end, the map can be checked for completeness and accuracy, and when the operator is satisfied that all requirements are met and the design is optimum, the final product plotted.

Required input to a computer for automated mapping includes a set of specific mapping instructions and a file of point locations and elevations. The instructions will include map scale, contour interval, line styles, lettering sizes and styles, symbols, and other items of information. Point locations are usually entered as a file of X, Y, Z coordinates, but angle and distance data can be entered and the coordinates computed. Special CADD systems for mapping with data collected by total station instruments and GNSS receivers have been developed.

As explained in Section 17.8, most automated mapping systems draw contours after constructing a *triangulated irregular network* (TIN) model. These are networks of nonoverlapping triangles, which represent the individual facets of the terrain. The computer interpolates contour crossings along the edges of the triangles and then draws the contours. A portion of a TIN model from an actual mapping project is illustrated in Figure 18.8(a), and the contours constructed from it are shown in Figure 18.8(b).

To understand the process of automatically constructing a contour line using a TIN, imagine contouring the TIN shown in Figure 18.9. To begin the procedure, a line on the edge of the TIN is randomly selected for contouring; in this instance line 1 has been chosen. Assume that the X, Y, Z coordinates of the end points of line 1 are (5401.08, 4369.79, 865.40) and (5434.90, 4456.90, 868.30), respectively, and that the 868-contour line is to be drawn. From the X and Y coordinate values, and applying Equations (10.11) and (10.12), the length and azimuth of line 1 are 93.45 ft and 21°13′06″, respectively. Also, from the Z coordinates, the elevation difference is $(868.30 - 865.40) = 2.9$ ft. The elevation difference from the first end point of line 1 to the 868-contour line

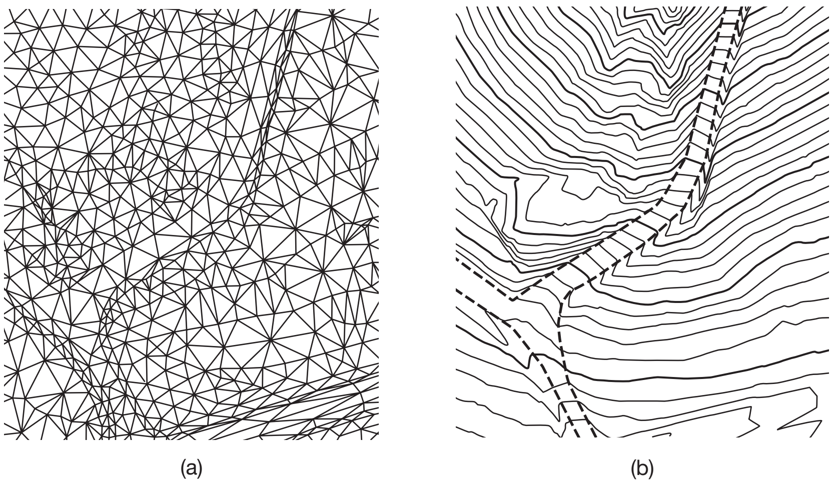


Figure 18.8 (a) Triangulated irregular network (TIN) model derived from digital-elevation model. (b) Contours derived by automated mapping system from TIN model of (a). Note the roadway edges were defined by breaklines in (a). (Courtesy Wisconsin Department of Transportation.)

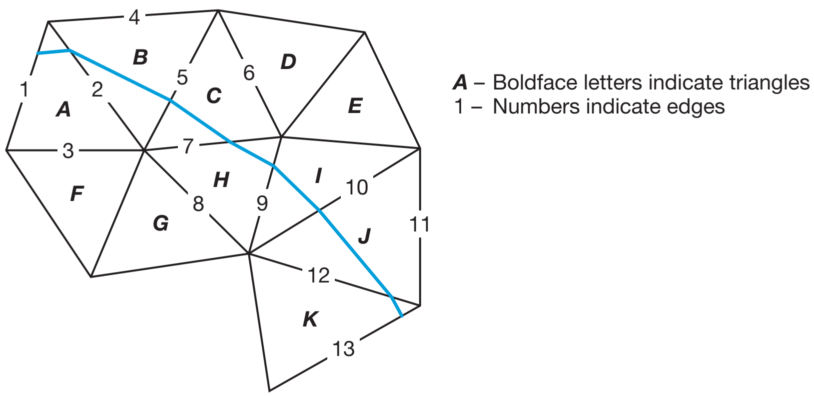


Figure 18.9 Automated contouring of TIN example.

is (868.00 − 865.40) = 2.6 ft. Now the following equivalent ratios can be formed:

$$\frac{2.6}{2.9} = \frac{\Delta L}{93.45}$$

$$\Delta L = 93.45 \times \frac{2.6}{2.9} = 83.78 \text{ ft}$$

This computation indicates that the distance (ΔL) from the first end point of the line to the 868-contour line is 83.78 ft. Using the previously derived azimuth for line 1, and Equation (10.7), the X, Y, Z coordinate values for the intersection of line 1 and the 868-contour are (5431.40, 4447.89, 868.0), respectively. This is the initial point of the 868-ft contour. Now a search algorithm checks the elevations of the end points of lines 2 and 3 (the other two sides of triangle A) for the continuation of the 868-contour. Once it determines that line 2 contains the continuation of the line, it again uses the same linear interpolation procedure to determine the coordinates of the intersection of the 868-contour with line 2, and draws the contour line from line 1 to 2. The contour is now ready to enter triangle B. It proceeds to check the elevations of the endpoints of lines 4 and 5, and determines that the 868-contour intersects with line 5. Then it again uses linear interpolation to determine the coordinate values for the intersection and continues with the drawing of the 868-contour to line 5. It continues to draw straight-line segments for the 868-contour as it passes through each triangle until it finally exits triangle K on line 13. From the preceding, it can be seen that this algorithm is repetitive and ideally suited for computer solution.

Contours compiled automatically should be carefully edited for correctness. In certain areas, *breaklines* (see Section 17.8) may need to be added or changed to obtain the proper terrain representation. Incorrect interpolation often occurs along the outer edges of automatically contoured areas, so these areas require special processing and additional field data. Thus, it is good practice to carry the field survey somewhat beyond the area of interest and "trim" the edges of the map.

Once all of the contours have been drawn as straight-line segments, the software then uses a *smoothing* algorithm to round the corners created at each intersection. The operator can usually control the amount of smoothing with a single entry called a *smoothing factor*. The higher the smoothing factor selected by the user, the smoother the intersections become, but the farther the contours depart from their computed values. Thus, the operator must choose a value for the smoothing factor carefully to ensure that the lines do not excessively depart from their original positions.

With terrain information stored in the computer in the form of TIN models, profiles and cross sections along selected lines can be derived automatically and plotted if desired. By including grade lines and design templates, earthwork computations can be made and stakeout information automatically derived for projects such as highways, railroads, and canals.

A TIN is a *digital terrain model* (DTM), which is also commonly called a *digital elevation model* (DEM), of the earth's surface. A DTM shows only topographic features of the earth and is devoid of any vegetation, or structures that lie on the surface. A DTM can be created with a TIN or by locating spot elevations in a grid, although, the former is more prevalent in practice. A DTM supports projects in machine control (see Section 15.9), flood modeling, highway design (see Chapters 24 and 25), as well as other construction projects (see Chapter 23). Figure 18.10 shows a DTM of a highway alignment in a three-dimensional perspective view.

The *three-dimensional perspective grid* is an alternative form of terrain representation. It can also be produced by the computer from TIN models. The example shown in Figure 18.10 illustrates its important advantage—it gives a very vivid impression of relief.

Figure 18.11 shows a CADD workstation. An operator controls the system by making entries on the keyboard, or by selecting instructions from a command board (*menu*) with a cursor.

Figure 18.12 is a portion of a topographic map for an engineering design project created with a CADD system, and Figure 18.13 is a subdivision plat, also

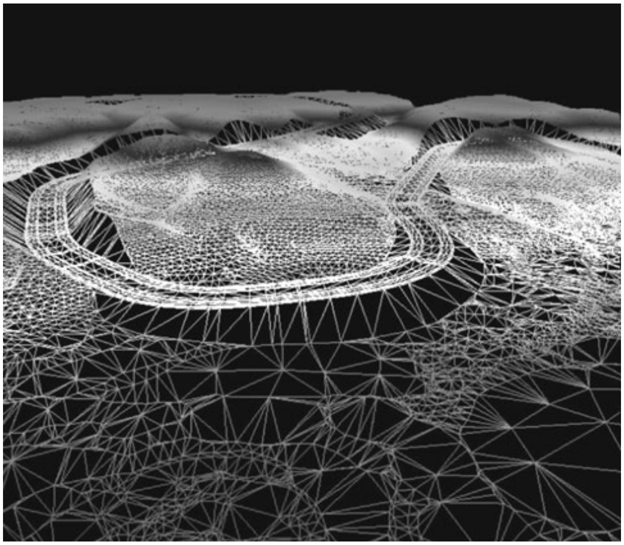


Figure 18.10 Three-dimensional perspective grid. (Courtesy of Ashtech, LLC.)

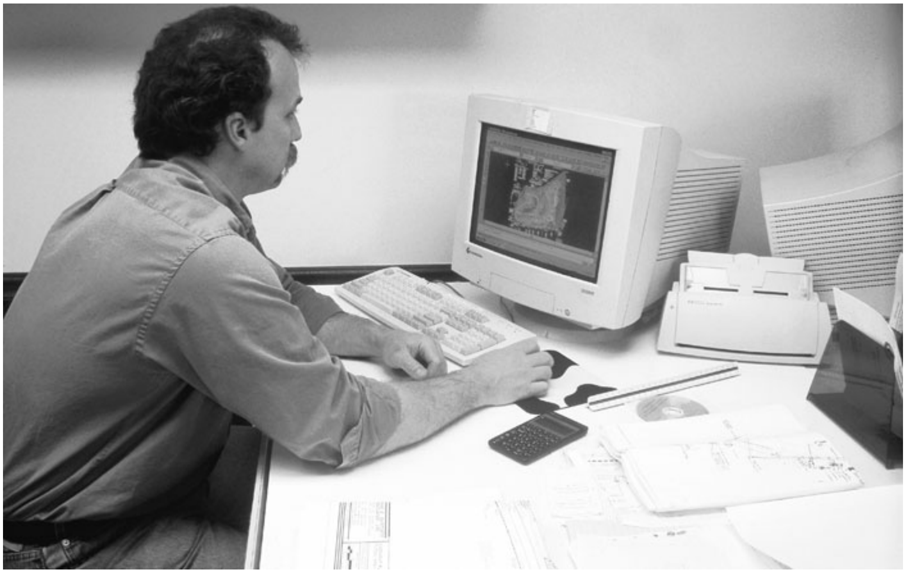


Figure 18.11
Computer-aided drafting and design (CADD) workstation. (Courtesy Tom Pantages.)

designed and drawn using CADD. The condominium plat shown in Figure 21.5 is another example of a product designed and drafted with CADD.

There are numerous advantages derived from using CADD systems in map design and drafting. A major one is increased speed in completing projects. Others include reduction or elimination of errors, increased accuracy, and preparation of a consistently more uniform final product. With completed maps stored in digital form, copies can be quickly reproduced at any time and revisions easily made.

Map data compiled using CADD systems can be stored in a data bank, with different numerical codes for each of the various kinds of features. They can be retrieved later for plotting in total, or in so-called *layers*, or parts, for special-purpose maps. For example, a city engineer may only be interested in a map showing the roads and utilities, while the assessor may want only property boundaries and buildings. This concept of layered maps is fundamental to land and geographic information systems (see Chapter 28).

Another significant advantage of producing maps in digital form is that they can be transmitted electronically from one office to others at remote locations. As an example, the Wisconsin Department of Transportation produces digital maps photogrammetrically for roadway design at its central office in Madison. Using a data modem and/or the Internet, these maps can then be transmitted instantaneously to any of nine district offices located throughout the state, where they are immediately available to engineers there for computer-aided design, or hard copies can be printed. In spite of the many improvements made in automated mapping systems, there is still a possibility that mistakes can occur. For this reason, it is good practice to have the field party chief, who is familiar with the area, review the completed maps. The different AM and CADD systems all have varying individual capabilities. Books and brochures available from the manufacturers provide detailed descriptions.

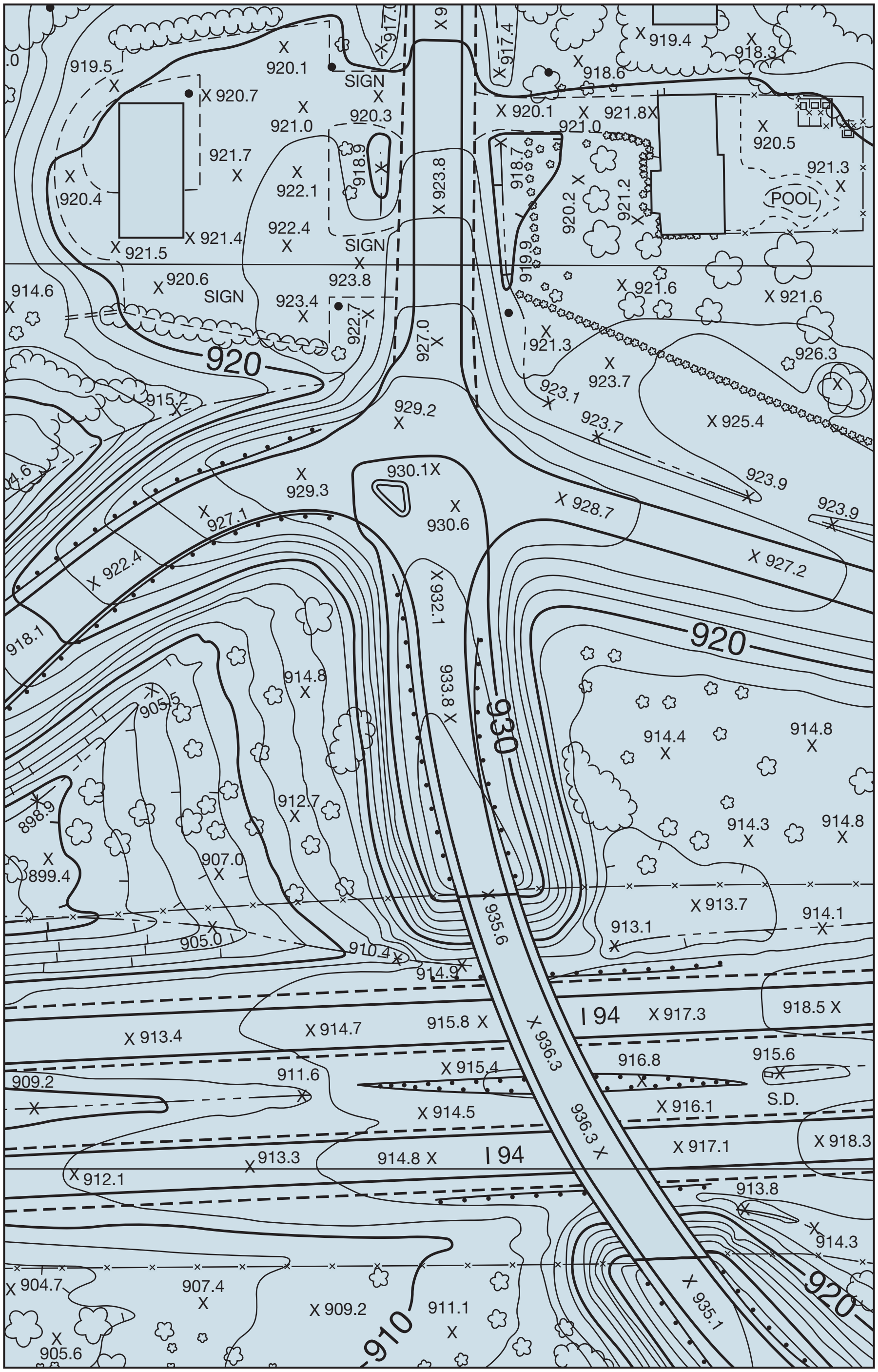
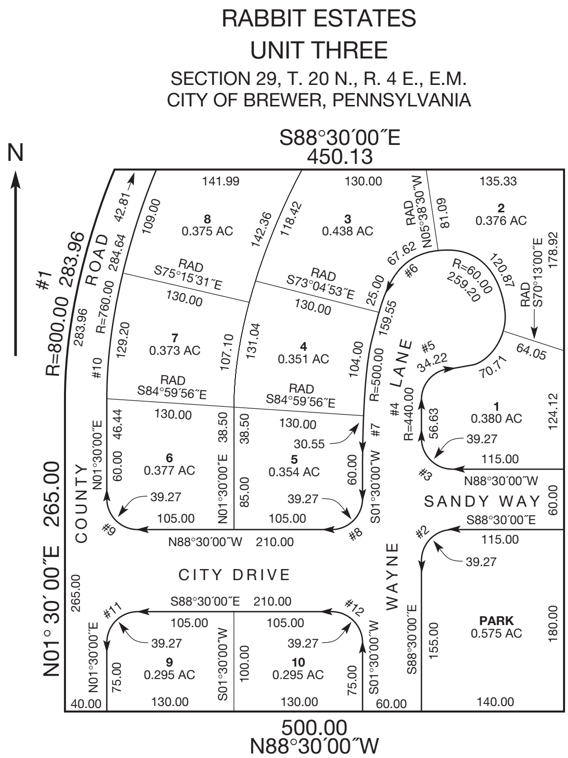


Figure 18.12
Engineering design map prepared using CADD system. (Courtesy Wisconsin Department of Transportation.)



NUMBER	RADIUS	ARC	DELTA	TANGENT	CHORD	CHORD-BEARING
1	800.00	283.96	020° 20′ 14″	143.49	282.47	N11° 40′ 07″E
2	25.00	39.27	090° 00′ 00″	25.00	35.36	N46° 30′ 00″E
3	25.00	39.27	090° 00′ 00″	25.00	35.36	N43° 30′ 00″W
4	440.00	56.63	007° 22′ 26″	28.35	56.59	N05° 11′ 13″E
5	25.00	34.22	078° 25′ 48″	20.40	31.61	N48° 05′ 20″E
6	60.00	259.20	247° 31′ 14″	– – –	99.76	N36° 27′ 23″W
7	500.00	159.55	018° 17′ 00″	80.46	158.88	S10° 38′ 30″W
8	25.00	39.27	090° 00′ 00″	25.00	35.36	S46° 30′ 00″W
9	25.00	39.27	090° 00′ 00″	25.00	35.36	N43° 30′ 00″W
10	760.00	284.64	021° 27′ 32″	144.01	282.98	N12° 13′ 46″E
11	25.00	39.27	090° 00′ 00″	25.00	35.36	N46° 30′ 00″E
12	25.00	39.27	090° 00′ 00″	25.00	35.36	S43° 30′ 00″E
13	630.00	287.96	026° 11′ 19″	146.54	285.46	N14° 35′ 40″E

Figure 18.13 Subdivision map compiled automatically by computer-driven plotter. (Courtesy Technical Advisors, Inc.)

18.15 IMPACTS OF MODERN LAND AND GEOGRAPHIC INFORMATION SYSTEMS ON MAPPING

Land Information Systems (LISs), Geographic Information Systems (GISs), and automated mapping and facilities management (AM/FM) systems all require enormous quantities of position-related land data. From this information, maps and other special-purpose graphic displays can be made and analyzed. For example, a typical LIS or GIS may include attribute data such as political boundaries, land ownership, topography, land use, soil types, natural resources, transportation routes, utilities, and many others. From the stored information, a user can display a map of each attribute category (or *layer*) on a screen, or several layers of data can be merged to produce combination maps. This merging, or *overlaying*, concept, discussed in Chapter 28, greatly facilitates data analysis and aids significantly in management and decision making. If printed maps of any selected layers or combinations are desired, they can be produced rapidly using automated drafting equipment.

The position-related land attribute data needed for LISs and GISs can be collected from a variety of sources and entered in the computer. These sources include field surveys (see Chapter 17), aerial photographs (see Chapter 27), and existing maps. Modern surveying instruments such as the total station instruments, satellite receiver units, and digital photogrammetric plotters can produce huge quantities of digital terrain data in X, Y, Z coordinate form rapidly and economically. *Raster scanners* (see Section 28.7.4) are able to systematically scan existing maps and other printed documents line-by-line, and convert the information to numerical form. The processes of collecting and digitizing data to support LISs and GISs are expected to place a heavy workload on surveyors (geomatics engineers) for many years to come.

18.16 SOURCES OF ERROR IN MAPPING

Sources of error in mapping include:

1. Errors in the data used in plotting.
2. Errors in the scales used for laying out lengths and coordinate values.
3. Errors in laying out grids for plotting by coordinates.
4. Using a soft pencil, or one with a blunt point, for plotting.
5. Variations in the dimensions of map sheets due to temperature and moisture.

18.17 MISTAKES IN MAPPING

Some common mistakes in mapping are:

1. Selecting an inappropriate scale or contour interval for the map.
2. Failing to check grids by measuring diagonals, and not checking points plotted from coordinates by measuring distances between them.
3. Using the wrong edge of an engineer's scale.
4. Making the north arrow too large or too complex.
5. Neglecting to identify the meridian of reference, that is, geodetic, grid, magnetic, etc.
6. Omitting the scale or necessary notes.
7. Failing to balance the sheet by making a preliminary sketch.
8. Drafting the map on a poor-quality medium.
9. Failing to realize that errors are also magnified when maps are enlarged electronically or photographically.
10. Operating AM and CADD systems without sufficient prior training.

PROBLEMS

Asterisks (*) indicate problems that have partial answers given in Appendix G.

18.1 Give the terms to which the acronyms TIN, DTM, and DEM apply.

18.2* On a map drawn to a scale of 1:6000, a point has a plotting error of 1/30 in. What is the equivalent ground error in units of feet?

18.3 What is the scale of the USGS 7 1/2 min quadrangle map sheet?
18.4 What design elements must be considered when laying out a map?
18.5 Why is scale depicted graphically on a map?
18.6 Why should contours be broken where labeling occurs?
18.7 Discuss why a map designed for a planning board hearing may not be the same as a map designed for an engineer?
18.8 What is the content of DTMs?
18.9 List the advantages of compiling maps using field-to-finish software?
18.10* For a 20-ft contour interval, what is the greatest error in elevation expected of any definite point read from a map if it complies with National Map Accuracy Standards?
18.11 An area that varies in elevation from 463 to 634 ft is being mapped. What contour intervals will be drawn if a 20-ft interval is used? Which lines are emphasized?
18.12 Similar to Problem 18.11, except elevations vary from 37 to 165 m and a 10-m interval is used.
18.13 What two questions must be answered before designing a map?
18.14 Describe how clarity can be enhanced through the use of fonts.
18.15* What is the largest acceptable error in position for 90% of the well-defined points on a map with a 1:24,000 scale that meets national map accuracy standards.
18.16 Discuss the balance is achieved on a map.
18.17 Discuss why insets are sometimes used on maps.
18.18* If a map is to have a 1-in. border, what is the largest nominal scale that may be used for a subject area with dimensions of 604 and 980 ft on a paper of dimensions 24 by 36 in?
18.19 Similar to Problem 18.18, except the dimensions of the subject area are 653 and 475 ft.
18.20 Draw 2-ft contours for the data in Plate B.2 of Appendix B.
18.21 If 90% of all elevations on a map must be interpolated to the nearest ±2 ft, what contour interval is necessary according to the National Map Accuracy Standards? Explain.
18.22 If an area having an average slope of 4% is mapped using a scale of 1:1000 and contour interval of 0.5 m, how far apart will contours be on the map?
18.23 Similar to Problem 18.22, except average slope is 6%, map scale is 300 ft/in., and contour interval is 10 ft.
18.24* Similar to Problem 18.22, except average slope is 5%, map scale is 1:500, and contour interval is 0.5 m.
18.25* The three-dimensional (X, Y, Z) coordinates in meters of vertexes A, B, and C in Figure 18.14 are (5412.456, 4480.621, 248.147), (5463.427, 4459.660, 253.121), and (5456.081, 4514.382, 236.193), respectively. What are the coordinates of the intersection of the 248-m contour with side AB? With side BC?

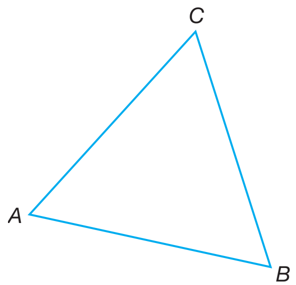


Figure 18. 25 through 18.27

18.26 The three-dimensional (X, Y, Z) coordinates in feet for vertices A, B, and C in Figure 18.14 are (8649.22, 6703.67, 143.86), (8762.04, 6649.77, 165.88), and (8752.64, 6770.20, 146.84), respectively. What are the coordinates of the intersections of the 150-ft contour as it passes through the sides of the triangle?

18.27 Similar to Problem 18.26, except compute the coordinates of the intersection of the 152-ft contour.

18.28 Discuss how contrast can be improved on a map.

The following table gives elevations at the corners of 50-ft coordinate squares, and they apply to Problems 18.29 and 18.31.

84	78	62	55	63	69
78	71	66	61	66	75
76	72	68	62	58	65

18.29 At a horizontal scale of 1 in. = 50 ft, draw 2-ft contours for the area.

18.30 Similar to Problem 18.29, except at the bottom of the table add a fourth line of elevations: 79, 69, 72, 62, 61, and 65 (from left to right).

18.31 In problem 18.30, about which number in the table can a 5-ft closed contour be drawn?

18.32 A rectangular lot running N-S and E-W is 150 by 100 ft. To locate contours it is divided into 50-ft square blocks, and the following horizontal rod readings are taken at the corners successively along the E-W lines, proceeding from west to east, and beginning with the northernmost line: 6.8, 5.6, 3.6; 6.3, 6.9, 5.1; 5.9, 4.7, 2.6; and 3.4, 4.9, 6.3. The HI is 323.9 ft. Sketch 2-ft contours.

18.33 If a map is drawn with 10-ft contour intervals, what contours between 70 and 330 ft are drawn with heavier line weight?

18.34 Find the website at which USGS quadrangle sheets may be purchased.

18.35 Download and prepare a report on Part 3 of the FGDC *Geospatial Positioning Accuracy Standards*.

BIBLIOGRAPHY

American Society for Photogrammetry and Remote Sensing. 1987. "*Large Scale Mapping Guidelines*." Bethesda, MD: American Society for Photogrammetry and Remote Sensing.

______. 1990. "ASPRS Standards for Large-Scale Maps." *Photogrammetric Engineering and Remote Sensing* 56 (No. 7): 1068.

Crawford, K. A. 2009. "Model Behavior: The How-to Guide to Successful Surface Modeling, Part 1." *The American Surveyor* 6 (No. 6): 63.

______. 2009. "Model Behavior: The How-to Guide to Successful Surface Modeling, Part 2." *The American Surveyor* 6 (No. 8): 44.

______. 2009. "Model Behavior: The How-to Guide to Successful Surface Modeling, Part 3." *The American Surveyor* 6 (No. 10): 64.

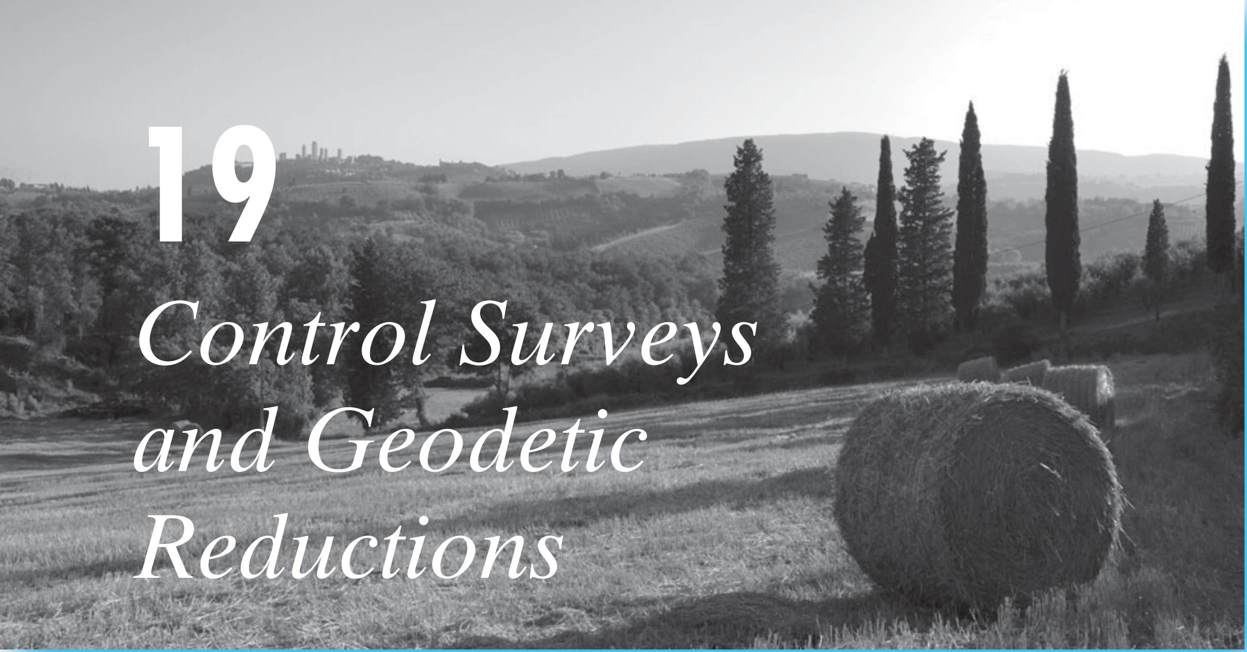
______. 2010. "Model Behavior: The How-to Guide to Successful Surface Modeling, Part 4." *The American Surveyor* 7 (No. 5): 50.

Davis, T. G. 2009. "USGS Quadrangles in Google Earth." *The American Surveyor* 6 (No. 9): 28.

Dronick, G. J. 2007. "Mapping Windmill Farms." *Professional Surveyor* 27 (No. 1): 14.

Ramirez, J. R. 2006. "A New Approach to Relief Representation." *Surveying and Land Information Science* 66 (No. 1): 19.

______. 2006. "Advances in Multimedia Mapping." *Surveying and Land Information Science* 66 (No. 1): 55.



19 Control Surveys and Geodetic Reductions

■ 19.1 INTRODUCTION

Control surveys establish precise horizontal and vertical positions of reference monuments. These serve as the basis for originating or checking subordinate surveys for projects such as topographic and hydrographic mapping; property boundary delineation; and route and construction planning, design, and layout. They are also essential as a reference framework for giving locations of data entered into Land Information Systems (LISs) and Geographic Information Systems (GISs).

Traditionally there have been two general types of control surveys: *horizontal* and *vertical.* Horizontal surveys generally establish *geodetic latitudes* and *geodetic longitudes* (see Section 19.4) of stations over large areas. From these values, plane rectangular coordinates, usually in a state plane or Universal Transverse Mercator (UTM) coordinate system (see Chapter 20) can be computed. On control surveys in smaller areas, plane rectangular coordinates may be determined directly without obtaining geodetic latitudes and longitudes.

Field procedures used in horizontal control surveys have traditionally been the ground methods of *triangulation*, *precise traversing*, *trilateration*, and combinations of these basic approaches (see Section 19.12). In addition, astronomical observations (see Appendix C) were made to determine azimuths, latitudes, and longitudes. Rigorous photogrammetric techniques (see Chapter 27) have also been used to densify the control in areas.

During the 1970s, *inertial surveying systems* (ISSs) were introduced. Their operating principle consisted fundamentally in making measurements of accelerations over time. This was done while the instrument was carried from point to point in a land vehicle or helicopter. The acceleration and time observations were taken independently in three mutually orthogonal planes which were oriented north-south, east-west, and in the direction of gravity. Orientation was achieved by means of gyroscopes. From the acceleration and time data, components of the instrument's movement in each of three reference planes could be computed, and hence relative positions of points determined. Inertial surveying systems were used in a variety of surveying applications, one of the most important being control surveying. Drawbacks of the systems were their high initial cost, equipment that was bulky, and an overall accuracy less than that attainable with GNSS receivers. As a result, ISSs are no longer used for control surveys.

Recently the *global navigation satellite systems*, GNSS (see Chapters 13, 14, and 15), has been employed with increasing frequency, especially on control surveys of larger extent. GNSS surveys are rapidly replacing the other methods because of several advantages including its ease of use, speed, and extremely high accuracy capabilities over long distances. However, in small areas, traditional methods of establishing control are still being used.

Vertical control surveys establish elevations for a network of reference monuments called *benchmarks*. Depending on accuracy requirements, they have traditionally been run by either *differential leveling* or *trigonometric leveling* (see Chapters 4 and 5). GNSS survey can also establish vertical control but are limited by the need for a precise geoid model (see Section 19.2). Thus, the most accurate and widely applied method is still precise differential leveling (see Section 19.13).

This chapter will define elements of geodetic reference systems used for control surveys, describe the *National Spatial Reference System* (NSRS), discuss some of the traditional ground methods used in control surveying, and explain some basic computational methods used in making geodetic reductions on conventional observations.

19.2 THE ELLIPSOID AND GEOID

It was noted in Section 19.1 that horizontal control surveys generally determine geodetic latitudes and geodetic longitudes of points. To explain geodetic latitude and longitude, it is necessary to first define the *geoid* and the *ellipsoid*. The geoid is an equipotential gravitational surface, which is everywhere perpendicular to the direction of gravity. Because of variations in the Earth's mass distribution and the rotation of the Earth, the geoid has an irregular shape.

The ellipsoid is a mathematical surface obtained by revolving an ellipse about the Earth's polar axis. The dimensions of the ellipse are selected to give a good fit of the ellipsoid to the geoid over a large area and are based upon surveys made in the area.

A two-dimensional view, which illustrates conceptually the geoid and ellipsoid, is shown in Figure 19.1. As illustrated, the geoid contains nonuniform undulations (which are exaggerated in the figure for clarity) and is therefore not readily defined mathematically. Ellipsoids, which approximate the geoid and can be defined mathematically, are therefore used to compute positions of widely spaced points that are located through control surveys. The *Clarke Ellipsoid of 1866* approximates the geoid in North America very well and from 1879 until the 1980s it was the ellipsoid used in NAD 27 as a reference surface for specifying

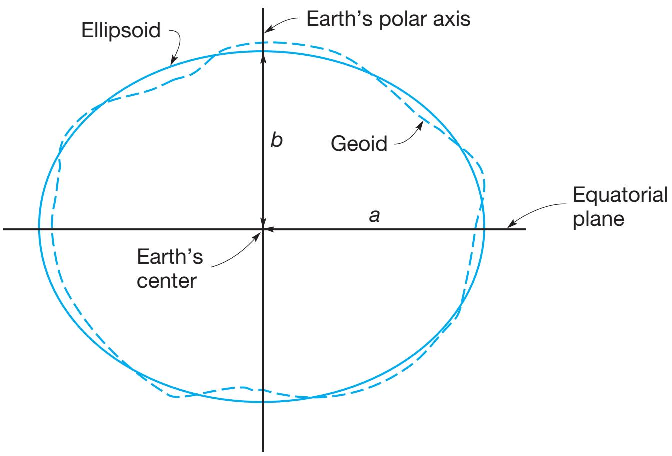


Figure 19.1
Ellipsoid and geoid.

geodetic positions of points in the United States, Canada, and Mexico. Currently, the *Geodetic Reference System of 1980* (GRS80) and *World Geodetic System of 1984* (WGS84) ellipsoids are commonly used in the United States because they provide a good worldwide fit to the geoid. This is important because of the global surveying capabilities of GNSS.

Sizes and shapes of ellipsoids can be defined by two parameters. Table 19.1 lists the parameters for the three ellipsoids noted above. For the Clarke 1866 ellipsoid, the defining parameters were the semiaxes a and b. For GRS80 and WGS84, the defining parameters are the semimajor axis a and flattening f. The relationship between these three parameters is

$$f = 1 - \frac{b}{a} \tag{19.1}$$

Other quantities commonly used in ellipsoidal computations are the *first eccentricity*, e, and the second eccentricity, e', of the ellipse, where

$$e = \frac{\sqrt{a^2 - b^2}}{a} = \sqrt{2f - f^2} \tag{19.2a}$$

$$e' = \frac{\sqrt{a^2 - b^2}}{b} = \frac{e}{\sqrt{1 - e^2}} \tag{19.2b}$$

TABLE 19.1 DEFINING ELLIPSOIDAL PARAMETERS

Ellipsoid	Semiaxis *a* (m)	Semiaxis *b* (m)	Flattening *f*
Clarke, 1866	6,378,206.4*	6,356,583.8*	1/294.978698214
GRS80	6,378,137.0*	6,356,752.3	1/298.257222101*
WGS84	6,378,137.0*	6,356,752.3	1/298.257223563*

*Defining parameters for the ellipsoids.

Often the term eccentricity is understood to mean the first eccentricity and this book will follow that convention. For each ellipsoid, the polar semiaxis is only about 21 km (13 mi) shorter than the equatorial semiaxis *b*. This means the ellipsoid is nearly a sphere; hence, for some calculations involving moderate lengths (usually up to about 50 km) this assumption can be made.[1]

Example 19.1

Using the defining parameters, what are the first eccentricities of the Clarke 1866 and GRS80 ellipsoids?

Solution

For the Clarke 1866 ellipsoid, Equation (19.2a) yields

$$e = \frac{\sqrt{6378206.4^2 - 6356583.8^2}}{6378206.4} = 0.082271854$$

For the GRS80 ellipsoid, Equation (19.2a) yields

$$e = \sqrt{\frac{2}{298.257222101} - \left(\frac{1}{298.257222101}\right)^2} = 0.081819191$$

19.3 THE CONVENTIONAL TERRESTRIAL POLE

As discussed in the preceding section, an ellipsoid is defined on the basis of the size of an ellipse that is rotated about the polar axis of the Earth. In reality, since the principal axis of inertia of the Earth does not coincide with the rotational axis of the Earth, the polar axis at any particular time is not fixed in position. Rather, as illustrated in Figure 19.2, it rotates with respect to the inertial system. This motion is generally divided into two major categories called *precession* and *nutation*. Precession is the greater of the two and is the wander of the polar axis over a long period of time. The pole makes a complete revolution about once every 26,000 years. Additionally, the pole wanders in much smaller radial arcs that are superimposed upon precession. These smaller circles are known as a nutation and are completed about once every 18.6 years. By international convention, the mean rotational axis of the Earth was defined as the "mean" position of the pole between the years of 1900.0 and 1905.0. This position is known as the *Conventional Terrestrial Pole* (CTP).

The CTP defines the *Z*-axis of a three-dimensional global Cartesian coordinate system with the northern portion being positive. The positive *X*-axis lies in

[1]In computations if the ellipsoid is assumed a sphere, its radius is usually taken such that its volume is the same as the reference ellipsoid. It is computed from $r = \sqrt[3]{a^2b}$. For the GRS80 ellipsoid, its rounded value is 6,371,000 m.

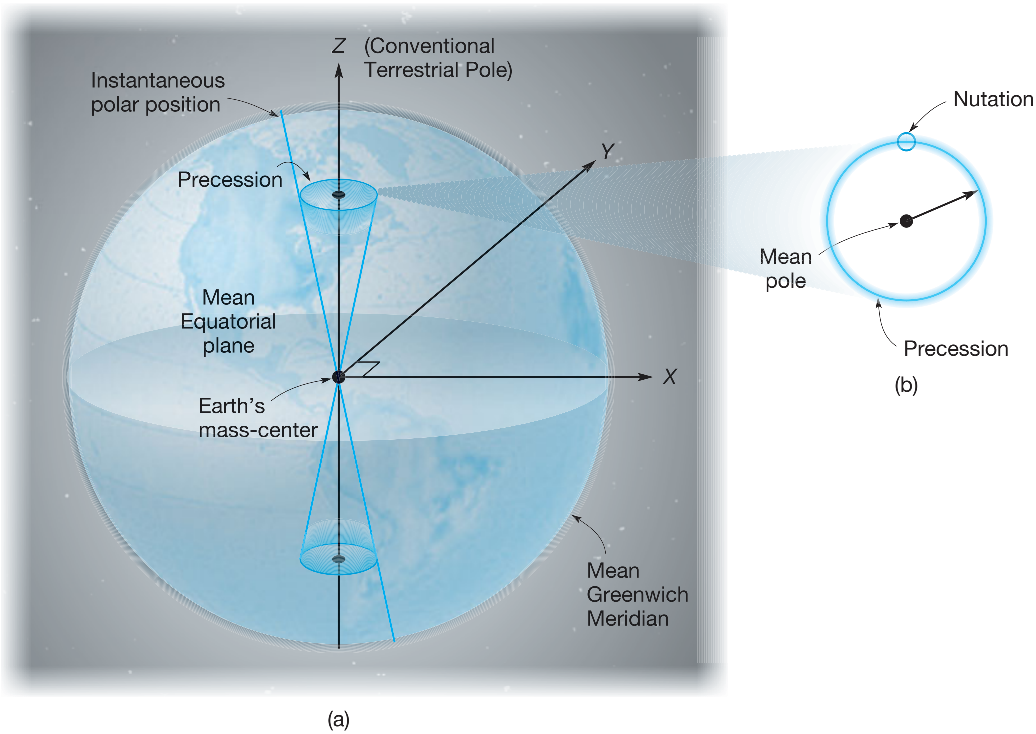


Figure 19.2 Motions of the Earth's polar axis: (a) three-dimensional and (b) plane view.

the mean equatorial plane, begins at the mass-center of the Earth, and passes through the mean Greenwich meridian. Finally, the Y-axis also lies in the mean equatorial plane and creates a right-handed Cartesian coordinate system. This coordinate system is known as the *Conventional Terrestrial System* (CTS). The CTS is shown in Figure 19.2.

Since 1988, the *International Earth Rotation Service* (IERS)[2] has monitored the instantaneous position of the pole with respect to the CTP using observations made by participating organizations employing advanced space methods including *Very Long Baseline Interferometry* (VLBI) and lunar and satellite laser ranging. Consequently, the CTS is now defined by a global set of stations through their instantaneous spatial coordinate positions known as the *International Terrestrial Reference Frame* (ITRF). This system is used in the computation of precise satellite orbits (see Chapters 13, 14, and 15) and is referenced through time-dependent models to other coordinate systems.

The instantaneous position of the pole is given in (x, y) coordinates with respect to the CTP. An application of these positions can be seen in the reduction of astronomical azimuths (see Appendix C) where the observed astronomical

[2]The instantaneous positions of the earth's pole can be found on the IERS website at http://hpiers.obspm.ft/webiers/general/syframes/SY.htm.

azimuth with respect to the instantaneous position of the pole can be related to the CTP by

$$Az_A = Az_{obs} - (x \sin \lambda + y \cos \lambda) \sec \phi \quad \textbf{(19.3)}$$

where Az_A is the astronomic azimuth related to the position of the conventional terrestrial pole, Az_{obs} the observed astronomical azimuth, (x, y) the coordinates of the instantaneous pole, and (ϕ, λ) the geodetic latitude and longitude, respectively, of the observing station. Future references to the polar axis of the Earth will implicitly refer to the CTP.

19.4 GEODETIC POSITION AND ELLIPSOIDAL RADII OF CURVATURE

Figure 19.3 shows a three-dimensional view of the ellipsoid and illustrates a point P on the surface of the Earth (which in this illustration is shown to exist at a distance of h_P above the ellipsoid). Point P' is on the ellipsoid along the *normal* through P. (The normal is defined below.) The geodetic position of point P is given by its *geodetic latitude* ϕ_P, *geodetic longitude* λ_P, and *geodetic height* h_P. To define these three terms, it is necessary to first define *meridians* and *meridian planes*. Meridians are great circles on the circumference of the ellipsoid which pass through the north and south poles. Any plane containing a meridian and the polar axis is a meridian plane. The angle in the plane of the Equator from the

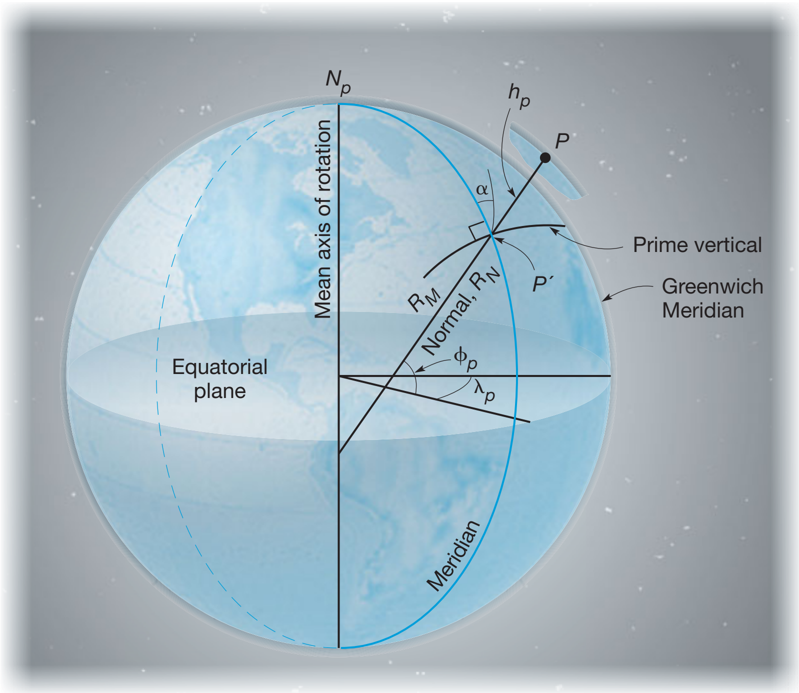


Figure 19.3
Different radii on the ellipsoid.

Greenwich meridian plane to the meridian plane passing through point P defines the geodetic longitude λ_P of the point. The plane defined by the vertical circle that passes through point P, *perpendicular to the meridian plane* on the ellipsoid, is called the plane of the *prime vertical* (also known as the *normal section*). The radius of the prime vertical at point P, R_N, is also called the *normal* since it is perpendicular to a plane that is tangent to the ellipsoid at P. The geodetic latitude ϕ_P is the angle, in the meridian plane containing P, between the equatorial plane and the normal at P.

To uniquely define the location of point P on the surface of the Earth, geodetic height h_P must be included. Geodetic height is the distance measured along the extension of the normal from P' on the ellipsoid to P on the Earth's surface. Geodetic height is not equivalent to elevation determined by differential leveling. (These differences were described in Section 13.4.3 and will be discussed further in Section 19.5.)

Because the Earth is approximated by an ellipsoid and not a sphere, the great circle that defines the prime vertical at P has a radius R_N that is different than the radius in the meridian R_M at P.[3] The lengths of these two radii, which are collinear at any point, are used in many geodetic computations. Figure 19.3 shows R_N. Additionally, the radius R_α of a great circle at any azimuth α to the meridian is different from either R_N or R_M. These three radii are computed as

$$R_N = N = \frac{a}{\sqrt{1 - e^2 \sin^2 \phi}} \tag{19.4}$$

$$R_M = M = \frac{a(1 - e^2)}{(1 - e^2 \sin^2 \phi)^{3/2}} \tag{19.5}$$

$$R_\alpha = \frac{R_N R_M}{R_N \cos^2 \alpha + R_M \sin^2 \alpha} \tag{19.6}$$

where a and e are parameters for the ellipsoid as defined in Section 19.2, and ϕ is the geodetic latitude of the station for which the radii are computed. From an analysis of Equations (19.4) and (19.5), it can readily be shown that R_N equals R_M at the poles where ϕ is 90°. Also, since the quantity $(1 - e^2)$ is less than one, the radius of the prime vertical R_N is greater than the radius of the meridian R_M at every location other than the pole where ϕ is equal to 90°.

Example 19.2

Using the GRS80 ellipsoidal parameters, what are the radii for the meridian and prime vertical for a point of latitude 41°18′15.0132″ N. What is the radius of the great circle that is at an azimuth of 142°14′36″ at this point?

[3]Note that R_N is often referred to as N, and R_M is frequently designated as M.

Solution

From Example (19.1), $e^2 = 0.081819191^2 = 0.00669438$

By Equation (19.4), the radius of the prime vertical is

$$R_N = \frac{6378137.0}{\sqrt{1 - e^2 \sin^2(41°18'15.0132'')}} = 6{,}387{,}458.536 \text{ m}$$

By Equation (19.5), the radius of the meridian is

$$R_M = \frac{6378137(1 - e^2)}{[1 - e^2 \sin^2(41°18'15.0132'')]^{\frac{3}{2}}} = 6{,}363{,}257.346 \text{ m}$$

By Equation (19.6), the radius of the great circle at an azimuth of 142°14′36″ is

$$R_\alpha = \frac{R_N R_M}{R_N \cos^2(142°14'36'') + R_M \sin^2(142°14'36'')} = 6{,}372{,}309.401 \text{ m}$$

■ 19.5 GEOID UNDULATION AND DEFLECTION OF THE VERTICAL

As discussed earlier, the geoid is an equipotential surface defined by gravity. If the Earth was a perfect ellipsoid without internal density variations, the geoid would match the ellipsoid perfectly. However, this is not the case, and thus the geoid can depart from some ellipsoids by as much as 100 m or more in certain locations. Traditional surveying instruments are oriented with respect to gravity and thus observations obtained with them are typically made with respect to the geoid. As can be seen in Figure 19.4, and discussed in Section 13.4.3, the separation between the geoid and the ellipsoid creates a difference between the height of a point above the ellipsoid (*geodetic height*) and its height above the geoid (*orthometric height*, which is commonly known as *elevation*). This difference, known as *geoid height*[4] (also called *geoidal separation*), can often be observed when comparing the geodetic height of a point derived by GNSS surveys, with its elevation as determined by differential leveling. The relationship between the orthometric height H and geodetic height h at any point is

$$h = H + N \tag{19.7}$$

where N is the geoidal height.

[4]The National Geodetic Survey regularly publishes geoid models for the U.S. Its latest version is GEOID09. This model can be obtained from the NGS website at http://www.ngs.noaa.gov/GEOID/GEOID09/.

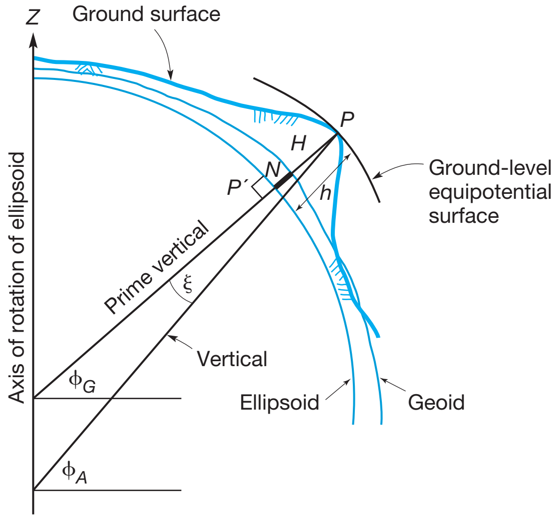


Figure 19.4 Relationships between the ellipsoid and geoid.

The GRS80 and WGS84 ellipsoids were both developed with the intent of providing a good fit to the geoid worldwide. However, they yield a relatively poor fit to the geoid within the continental United States, where the average geoidal height is approximately -30 m (the negative sign means the geoid is below the ellipsoid in the conterminous United States). The Clarke 1866 ellipsoid, on the other hand, provides a very close fit to the geoid in the United States, that is, geoidal heights are typically only a few meters.

In general, equipotential gravitational surfaces are not parallel to either the geoid or ellipsoid. This is partly due to the rotation of the Earth, which causes the surfaces to separate as they approach the equator, and partly due to density anomalies near the surface of the Earth. As a result, certain inconsistencies can occur in different types of field observations and hence corrections must be made. Some of the more significant of these corrections are discussed in Section 19.14.

As illustrated in Figure 19.5, the *deflection of the vertical* (also called *deviation of the vertical*) at any ground point P is the angle between the vertical (direction of gravity) and the normal to the ellipsoid. This angle is generally reported by giving two components: its orthogonal projections onto the meridian and normal planes. In the figure, the zenith of the ground-level equipotential surface is called the astronomical zenith Z_A since it corresponds to the direction of gravity (zenith) of a leveled instrument during astronomical observations. Also in Figure 19.5, Z_G is the normal at point P. The projected components of the total deviation of the vertical onto the meridian and normal planes are called ξ(xi) and η(eta), respectively.[5] The relationships between the astronomic latitude (ϕ_A),

[5]As with geoidal undulations, values for ξ and η can also be obtained through the National Geodetic Survey's deflection of vertical models. The latest version, DEFLEC09, can be obtained from the NGS website at http://www.ngs.noaa.gov/GEOID/DEFLEC09/.

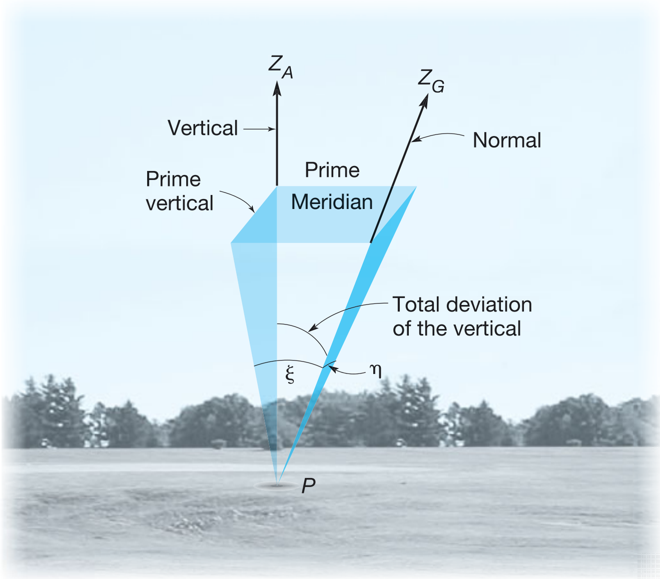


Figure 19.5
ξ and η components of deflection of the vertical.

astronomic longitude (λ_A), and astronomic azimuth (Az_A); the geodetic latitude (ϕ_G), geodetic longitude (λ_G), and geodetic azimuth (Az_G); and ξ and η are

$$\xi = \phi_A - \phi_G \tag{19.8}$$

$$\eta = (\lambda_A - \lambda_G) \cos \phi = (Az_A - Az_G) \cot \phi \tag{19.9}$$

In Equation (19.9), ϕ can be either the astronomic or geodetic latitude. From this equation, the so-called *Laplace equation* can be derived as

$$Az_G = Az_A - (\lambda_A - \lambda_G) \sin \phi = Az_A - \eta \tan \phi \tag{19.10}$$

Stations at which the necessary parameters are known, such that Equation (19.10) can be formed are called *Laplace stations.* Note that in Equation (19.9), for points near the equator, latitude ϕ approaches 0°, and the astronomic and geodetic azimuths become essentially the same. As will be discussed in Section 19.14.3, additional corrections are needed to properly reduce an observed azimuth to its geodetic equivalent on the ellipsoid.

■ 19.6 U.S. REFERENCE FRAMES

Horizontal and vertical datums consist of a network of control monuments and benchmarks whose horizontal positions and/or elevations have been determined by precise geodetic control surveys. These monuments serve as reference points for originating subordinate surveys of all types and as such are known as *reference frames*. The horizontal and vertical reference systems used in the immediate past and at present in the United States are described in the following subsections.

19.6.1 North American Horizontal Datum of 1927 (NAD27)

In 1927, a least-squares adjustment was performed, which incorporated all horizontal geodetic surveys that had been completed up to that date. This network of monumented points included in the adjustment, together with their adjusted geodetic latitudes and longitudes, was referred to as the *North American Datum of 1927* (NAD27). The adjustment utilized the Clarke ellipsoid of 1866 and held fixed the latitude and longitude of an "initial point," station *Meades Ranch* in Kansas, along with the azimuth to nearby station *Waldo.* The project yielded adjusted latitudes and longitudes for some 25,000 monuments existing at that time. Until the advent of the current datum (NAD83), positions of stations established after 1927 were adjusted in processes that held NAD27 monuments fixed.

19.6.2 North American Horizontal Datum of 1983 (NAD83)

The National Geodetic Survey (NGS) began a new program in 1974 to perform another general adjustment of the North American horizontal datum. The adjustment was deemed necessary because of the multitude of post-1927 geodetic observations that existed and because many inconsistencies had been discovered in the NAD27 network. The project was originally scheduled for completion in 1983, hence its name *North American Datum of 1983* (NAD83), but it was not actually finished until 1986. The adjustment was a huge undertaking, incorporating approximately 270,000 stations and all geodetic surveying observations on record—nearly 2 million of them! About 350 person-years of effort were required to accomplish the task.

The initial point in the new adjustment is not a single station such as Meades Ranch in Kansas; rather, the Earth's mass-center and numerous other points whose latitudes and longitudes had been precisely established using radio astronomy and satellite observations were used. The GRS80 ellipsoid was employed since, as noted earlier, it fits the Earth, in a global sense, more accurately than the Clarke ellipsoid of 1866.

Within the United States, the adjusted latitudes and longitudes of all monuments in NAD83 differ from their NAD27 values. These differences result primarily because of the different ellipsoids and origins used, but part of the change is also due to the addition of the many post-NAD27 observations in the NAD83 adjustment. The approximate magnitudes of these changes in the conterminous United States, expressed in meters, are illustrated in Figure 19.6. Of course, these differences have a significant impact on all existing control points and map products that were based on NAD27. Various mathematical models were developed to transform NAD27 values to their NAD83 positions (see Section 20.11).

19.6.3 Later Versions of NAD83

The internal consistency of first-order points adjusted in NAD83 was specified to be at least 1:100,000, but tests have verified that on average it is probably 1:200,000 or better. However, there are some areas where relative accuracies fall

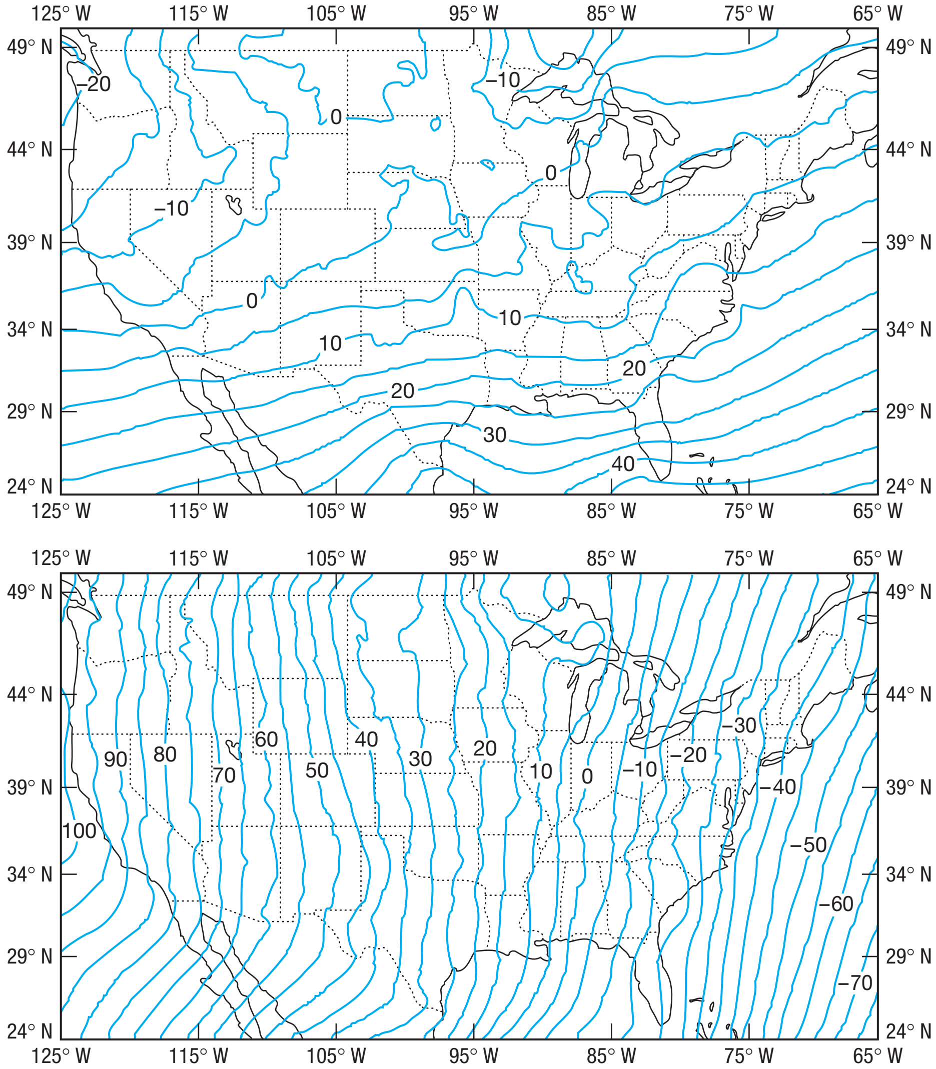


Figure 19.6 Approximate changes in latitude and longitude (in meters) in the conterminous United States from NAD27 to NAD83. Upper figure: Latitude. Lower figure: Longitude. (Courtesy National Geodetic Survey.)

below 1:100,000. Since GNSS survey accuracies are often better than 1:100,000, there was concern among GNSS users about how to fit their observations to a network of reference points whose inherent accuracies were less.

State governments and the National Geodetic Survey cooperatively sought to solve this problem by adding *High-Accuracy Reference Networks* (HARNs) to the National Spatial Reference System (see Section 19.8). From 1987 to 1997, HARN networks were created in every state. When each state's HARN was completely observed, an adjustment of these new stations was performed. This created an interim reference frame that was available to surveyors using GPS. This second version of NAD83 is known as NAD83 (HARN). For these regional reference systems, NGS retained the location of the mass-center of the Earth and orientation of the Cartesian coordinate axes,[6] but introduced a new scale that was consistent with the International Terrestrial Reference Frame of 1989 (ITRF89).

With the introduction of the CORS network in 1994 (see Section 14.3.5), the NGS was again faced with the problem of having stations of higher accuracy that were being adjusted using a different reference frame. Thus, a third realization of NAD83 was obtained using the ITRF93 reference frame. This transformation created another datum involving only the CORS stations and was known as NAD83 (CORS93).

In the spring of 1996, the NGS computed the positional coordinates for all existing CORS stations using ITRF94. This created the fourth realization of NAD83 and was known as NAD83 (CORS94). In 1998, the NGS computed positional coordinates for all existing CORS sites using ITRF96 as their reference frame. This version of NAD83 is known as NAD83 (CORS96). Each datum differs slightly from the previous definition of NAD83. For instance, the positions of sites in NAD83 (CORS96) differ by a maximum of about 2 cm in horizontal and 4 cm in vertical from their equivalent NAD83 (CORS94) values. Furthermore, the difference between any NAD83 CORS adjustment and NAD83 (HARN) is under 10 cm in horizontal and 20 cm in vertical.

With completion of the last statewide HARN in 1997, the NGS had two spatial reference systems NAD83 (HARN) and NAD83 (CORS96). Since GPS technology and related accuracies had improved over the time of the HARN creation, the NGS decided in 1998 to reobserve all HARN stations. This process known as the *Federal Base Network* (FBN) survey was initiated in 1999. In 2007 a simultaneous readjustment of all HARN and CORS observations was completed. This created a new definition of NAD83 known as the NAD83 (2007). This system is connected to the International Terrestrial Frame (ITRF) using the ITRF coordinates of the CORS sites. This datum removes problems of having two different reference frames available for use in a GNSS survey.

19.6.4 National Geodetic Vertical Datum of 1929 (NGVD29)

Vertical datums for referencing benchmark elevations are based on a single equipotential surface. Prior to the NAVD88 readjustment (see Section 19.6.5), the vertical datum used in the United States was the *National Geodetic Vertical Datum of 1929* (NGVD29). The NGVD29 was obtained from a best fit of mean sea level observations taken at 26 tidal gauge stations in the United States and Canada, and thus is often referred to as "mean sea level (MSL)." Unfortunately, the use of the term "mean sea level" is still used today when expressing elevations of benchmarks. As will be discovered in the next subsection, the use of "mean sea level" to define the elevation of a station is incorrect since the current datum was arbitrarily defined using a single benchmark.

[6]Using high-precision surveys, it was determined that the NAD83 (1986) Cartesian coordinate axes were misaligned by 0.03″, and its scale differed by 0.0871 ppm from the true definition of the meter.

19.6.5 North American Vertical Datum of 1988 (NAVD88)

Between 1929 and 1988 more than 625,000 km of additional control leveling lines had been run. Furthermore, crustal movements and subsidence had changed the elevations of many benchmarks. To incorporate the additional leveling, and correct elevations of erroneous benchmarks, a general vertical adjustment was performed. This adjustment included the new observational data, as well as an additional 81,500 km of releveled lines, and leveling observations from both Canada and Mexico. It was originally scheduled for completion in 1988 and named the *North American Vertical Datum of 1988* (NAVD88), but it was not actually released to the public until 1991. This adjustment shifted the position of the reference equipotential surface from the mean of the 26 tidal gauge stations used in NGVD29 to a single tidal gage benchmark known as *Father Point*, which is in Rimouski on the Saint Lawrence Seaway in Quebec, Canada. As a result of these changes, published elevations of benchmarks in NAVD88 have shifted from their NGVD29 values. The magnitudes of these changes in the conterminous United States, expressed in millimeters, are shown in Figure 19.7. Note that the changes are largest in the western half of the country, with shifts of more than 1.5 m occurring in the Rocky Mountain region.[7]

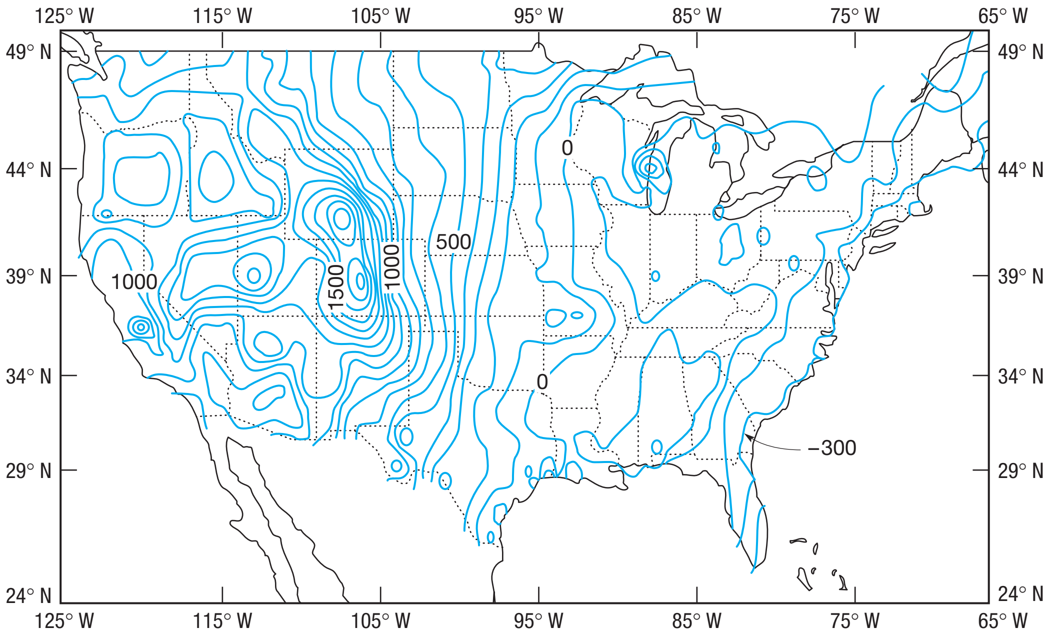


Figure 19.7 Approximate shift in vertical datum (in millimeters). Values shown are NAVD88 minus NGVD29. (Courtesy National Geodetic Survey.)

[7]Those wishing to convert NGVD29 benchmark elevations to NAVD88 values can use the software VERTCON available from the NGS on their website at http://www.ngs.noaa.gov/TOOLS/Vertcon/vertcon.html.

19.6.6 Transforming Coordinates Between Reference Frames

Historically, a goal of geodesy has been to obtain one common reference frame for coordinates. However, realistically, each country or region has often developed its reference frame independently. Today, we often need to transform station coordinates from those derived using GNSS surveys and those developed in some local reference frame such as NAD83. In Section 15.9, this process was introduced as *localization*. To do this, stations with known geodetic coordinates in both reference frames are required. If sufficient common stations are known, a three-dimensional coordinate transformation (see Section 19.17) can be used to convert the coordinates of stations from one reference frame into another. Since most reference frames have nearly aligned coordinate axes, the three-dimensional coordinate transformation can be simplified to the so-called *Helmert transformation*. The Helmert transformation is mathematically expressed as

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_2 = (1 + \Delta S)\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_1 + \begin{bmatrix} 0 & R_Z & -R_Y \\ -R_Z & 0 & R_X \\ R_Y & -R_X & 0 \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_1 + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix} \qquad (19.11)$$

where $[X\ \ Y\ \ Z]_1^T$ and $[X\ Y\ Z]_2^T$ are the geocentric coordinates of the common stations in the two reference frames derived using Equation (13.1); ΔS is the scale factor change between reference frames; R_X, R_Y, and R_Z the rotations in radian units for the X, Y, and Z axes, respectively; T_X, T_Y, and T_Z the translations between the two reference frames. A minimum of two stations known in horizontal position and three stations in elevation are required to perform this transformation. When insufficient common stations are not available, it is possible to perform this transformation using only the translations. However, this produces considerably lower-quality results.

Besides different reference frames, the crustal plates of the Earth are constantly in motion. For example, some parts of California are moving at a rate of 4 cm per year. Thus, the coordinates of points in any reference frame must be tied to a specific moment in time, or epoch. The National Geodetic Survey has combined the Helmert transformation with the velocity vectors of the crustal plates to produce a coordinate transformation software package known as *Horizontal Time-Dependent Positioning* (HTDP) software.[8] This software allows users to transform coordinates across time and between reference frames. Computations for the Helmert transformation as used in the HTDP software are demonstrated in a Mathcad® worksheet on the companion website for this book at http://www.pearsonhighered.com/ghilani.

It is also possible to perform the transformation in two separate transformations (horizontal and vertical). This is especially useful when the design coordinates are in a local reference frame that is arbitrarily assigned. In this case, the geodetic coordinates derived from GNSS survey are converted to coordinates in a map projection system (see Chapter 20). Since only horizontal positioning is

[8]The HTDP software can be found at http://www.ngs.noaa.gov/TOOLS/Htdp/Htdp.html.

involved, the transformation between the reference frames is further simplified using a modified version of the two-dimensional conformal coordinate transformation shown in Equation (11.37). To perform this transformation, the centroids of the coordinates for the common stations are computed and the coordinates in both frames translated by these values. This places the origins of both reference frames at the centroid of the common points thus removing the translations from Equation (11.37). Using the translated coordinates, the remaining scale and rotation of Equation (11.37) is computed. Following these procedures, any remaining GNSS-derived coordinates can be transformed into the local reference frame. This process is demonstrated in the following example.

Example 19.3

A surveyor establishes a control network of points using an arbitrary coordinate system. In preparation for staking out the project using kinematic survey (see Chapter 15), the surveyor reoccupies each station with a receiver. The resulting GNSS-derived coordinates are transformed into a two-dimensional map projection coordinate system with the common station coordinate values in both systems listed in Table 19.2. Determine the rotation and scale between the arbitrary and worldwide reference frames.

Solution

An oblique stereographic map projection (see Section 20.13.1) was used to transform the observed geodetic coordinates derived by the application software to a coordinate system with the centroid of the project at its origin. The average values of the coordinates (centroid) in local reference system are

$$X_0 = \frac{5000.00 + 1978.54 + 6328.46 + 6058.04}{4} = 4841.26$$

$$Y_0 = \frac{5000.00 + 6075.88 + 5983.64 + 5000.00}{4} = 5514.88$$

These values are then used to translate the arbitrary coordinates to a common origin resulting in the following set of coordinates. The local orthometric height and GNSS-derived orthometric height are also shown.

TABLE 19.2 COORDINATE VALUES OF COMMON STATIONS

	Arbitrary Reference Frame		GPS Reference Frame	
Station	***X* (ft)**	***Y* (ft)**	***E* (m)**	***N* (m)**
A	5000.00	5000.00	635797.076	464685.605
B	1978.54	6075.88	625530.377	462379.464
C	6328.46	5983.64	637760.165	469740.901
D	6058.04	5000.00	638732.517	466538.417

	Arbitrary			GNSS		
Station	$x' = X - X_0$ **(ft)**	$y' = Y - Y_0$ **(ft)**	H **(m)**	$e' = E - E_0$ **(m)**	$n' = N - N_0$ **(m)**	H**(m)**
A	158.74	−514.88	282.486	45.212	−157.888	282.476
B	−2862.72	561.00	296.577	−869.005	−188.611	296.571
C	1487.20	468.76	313.819	456.132	133.694	313.814
D	1216.78	−514.88	304.191	367.660	−164.417	304.205

Since the two coordinate systems share a common origin, Equation (11.37) is modified as

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} e' \\ n' \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} v_x \\ v_y \end{bmatrix} \tag{19.12}$$

Substituting the above coordinates into this equation results in

$$A = \begin{bmatrix} 45.212 & 157.888 \\ -157.888 & 45.212 \\ -869.005 & -188.611 \\ 188.611 & -869.005 \\ 456.132 & -133.694 \\ 133.694 & 456.132 \\ 367.660 & 164.417 \\ -164.417 & 367.660 \end{bmatrix} \quad X = \begin{bmatrix} a \\ b \end{bmatrix} \quad L = \begin{bmatrix} 158.74 \\ -514.88 \\ -2862.72 \\ 561.00 \\ 1487.20 \\ 468.76 \\ 1216.78 \\ -514.88 \end{bmatrix}$$

Using Equation (16.6), the solution of this system ($AX = L + V$) results in $a = 3.27987$ and $b = 0.066308$. Recognizing that $\tan(\theta) = \dfrac{b}{a}$, the scale and rotation between the two systems of coordinates are $s = 3.28054$ and $\theta = 1°09'29.4''$, respectively. These values are now used in conjunction with Equation (19.12) to transform the GNSS-derived map projection coordinates into the arbitrary local coordinate system for any additional points.

Note that the scale factor is approximately equal to the conversion factor going from meters to feet of 3.28083. Note that the computed residuals for the observations are 0.020, 0.027, −0.011, −0.003, −0.009, −0.016, −0.001, and −0.007, respectively, which are within the observational precisions of kinematic survey for horizontal work.

The GNSS-derived heights can also be transformed onto a local level plane. This process must account for the translation between the two systems and the obliquity of the two planes. The obliquity of the two level surfaces is corrected applying two rotations in the cardinal north and east directions to bring the GNSS-derived level surface parallel to the local horizontal plane. This is performed as

$$T_0 + r_e N_{GNSS} + r_n E_{GNSS} = H_L - H_G + \nu \tag{19.13}$$

where T_0 is the translation between the two level planes; N_{GNSS} and E_{GNSS} the northing and easting of map projection coordinates derived from the GNSS application software (see Example 19.3), respectively; r_e and r_n the rotations about the x and y axes, respectively; H_L the local height of the control points used in the project design; and H_G is either the geodetic height of the point as derived by GNSS or the orthometric height of the point as derived by the combination of GNSS-derived geodetic heights and a geoid model.

As can be seen, Equation (19.13) involves three unknown parameters, T_0, r_e, and r_n. Thus, a minimum of three benchmarks with local heights must be known. However, it is wise to always have a fourth for the purposes of redundancy and a check. Again when less than three benchmarks are known, the translation can be computed from a single station. However, the accuracy of this transformation will decrease significantly.

Example 19.4

Using the data given in Example 19.3, determine the three transformation parameters of Equation (19.13).

Solution

Using the northing (N), easting (E), and heights from Example 19.3 in concert with Equation (19.13) yields the following observation equations

$$A = \begin{bmatrix} 1 & -157.888 & 45.212 \\ 1 & 188.611 & -869.005 \\ 1 & 133.694 & 456.132 \\ 1 & -164.417 & 367.660 \end{bmatrix} \quad X = \begin{bmatrix} T_0 \\ r_e \\ r_n \end{bmatrix} \quad L = \begin{bmatrix} 0.025 \\ 0.021 \\ 0.018 \\ 0.001 \end{bmatrix}$$

Solving the system of equations, $AX = L + V$, using Equation (16.6) yields $T_0 = 0.016$, $r_e = 2.2''$, and $r_n = -1.3''$. With these transformation parameters and the GNSS-derived map projection coordinates, a geodetic height can be transformed into a local height. The resulting residuals for the observations are −0.011, 0.003, −0.003, and 0.011, respectively. Again these values are well within the vertical accuracy of GNSS-observed heights.

In order to have the distances obtained from GNSS surveys match the equivalent ground distances, the map projection plane is brought to the surface using an appropriate scaling factor. As discussed in Section 20.13.1, the oblique stereographic map projection has a defining scale factor of k_0 at its origin, which is the centroid of the projection. This makes the oblique stereographic map projection the preferred projection for this process. If this value is set to an appropriate scale, the plane surface of the map projection will be coincident with the elevation of the centroid. As an example, a scale factor of

$$k_0 = 1 + \frac{H_{centroid}}{R_e} \tag{19.14}$$

This scale factor is often used as one of the defining parameters for the oblique stereographic map projection system (see Section 20.13.1). The lengths of the distances are further refined with scaling factor as derived from the two-dimensional conformal coordinate transformation as defined in Example 19.3. The combination of these scales applied to the observed and transformed geodetic coordinates will result in distances that closely match their equivalent ground values. Computations for this problem are demonstrated in the Mathcad® worksheets *Helmert.xmcd* and *C19-6.xmcd*, which are on the companion website for this book.

■ 19.7 ACCURACY STANDARDS AND SPECIFICATIONS FOR CONTROL SURVEYS

The required accuracy for a control survey depends primarily on its purpose. Some major factors that affect accuracy are type and condition of equipment used, field procedures adopted, and the experience and capabilities of personnel employed. In 1984 and again in 1998, the Federal Geodetic Control Subcommittee (FGCS) published different sets of detailed standards of accuracy and specifications for geodetic surveys.[9] The rationale for both sets of standards is twofold: (1) to provide a uniform set of standards specifying minimum acceptable accuracies of control surveys for various purposes and (2) to establish specifications for instrumentation, field procedures, and misclosure checks to ensure that the intended level of accuracy is achieved.

Table 19.3 lists the 1998 FGCS accuracy standards for control points. These standards are independent of the method of survey and based on a 95% confidence level (see Section 16.12). In order to meet these standards, control points in the survey must be consistent with all other points in the geodetic control network and not merely those within that particular survey. In Table 19.3, for horizontal surveys the accuracy standard specifies the radius of a circle within which the true or theoretical location of the survey point falls 95% of the time. The vertical accuracy standard specifies a linear value (plus or minus) within which the true or theoretical location of the point falls 95% of the time. Procedures leading to classification according to these standards involve four steps:

1. The survey observations, field records, sketches, and other documentation are examined to insure their compliance with specifications for the intended accuracy of the survey.

[9]The 1998 standards are titled **GEOSPATIAL POSITIONING ACCURACY STANDARDS,** Part 2: **STANDARDS FOR GEODETIC NETWORKS.** They can be downloaded at the following website: http://www.fgdc.gov. The 1984 standards, published in a booklet entitled **STANDARDS AND SPECIFICATIONS FOR GEODETIC CONTROL NETWORKS,** are available from the National Geodetic Information Center, NOAA, National Geodetic Survey, N/CG17, SSMC3 Station 09535, 1315 East-West Highway, Silver Spring, MD 20910.

TABLE 19.3 1998 FGCS ACCURACY STANDARDS: HORIZONTAL HEIGHT, ELLIPSOID HEIGHT, AND ORTHOMETRIC HEIGHT

Accuracy Classifications	95% Confidence Less Than or Equal to
1 millimeters	0.001 meters
2 millimeters	0.002 meters
5 millimeters	0.005 meters
1 centimeters	0.010 meters
2 centimeters	0.020 meters
5 centimeters	0.050 meters
1 decimeters	0.100 meters
2 decimeters	0.200 meters
5 decimeters	0.500 meters
1 meters	1.000 meters
2 meters	2.000 meters
5 meters	5.000 meters
10 meters	10.000 meters

2. A minimally constrained least-squares adjustment of the survey observations is analyzed to guarantee that the observations are free from blunders and have been correctly weighted.
3. The accuracy of control points in the local existing network to which the survey is tied is computed by random error propagation and weighted accordingly in the least-squares adjustment of the survey network.
4. The survey accuracy is checked at the 95% confidence level by comparing minimally constrained adjustment results against established control. The comparison takes into account the network accuracy of the existing control as well as systematic effects such as crustal motion or datum distortion.

Because many existing products, including control datasheets in the NAD83 datum, refer to the 1984 standards, these will also be described. This earlier set of standards established three distinct *orders of accuracy* to govern traditional control surveys, given in descending order: *first-order*, *second-order*, and *third-order.* For horizontal control surveys, second-order and third-order each have two separate accuracy categories, *class I* and *class II.* For vertical surveys, first-order and second-order each have class I and class II accuracy divisions. In 1985, three new orders of accuracy were defined for GNSS surveys (see Section 14.5.1). These were orders AA, A, and B. Another lower order of accuracy for GNSS surveys, identified as Order C, was also specified in these standards. It overlaps the three orders of accuracy applied to traditional horizontal surveys (see Tables 19.4 and 14.5).

TABLE 19.4 1984 AND 1985 FGCS HORIZONTAL CONTROL SURVEY ACCURACY STANDARDS

GPS Order*	Traditional Surveys Order and Class**	Relative Accuracy Required Between Points
Order AA		1 part in 100,000,000
Order A		1 part in 10,000,000
Order B		1 part in 1,000,000
Order C-1	First Order	1 part in 100,000
	Second Order	
Order C-2-1	Class I	1 part in 50,000
Order C-2-II	Class II	1 part in 20,000
	Third Order	
Order C-3	Class I	1 part in 10,000
	Class II	1 part in 5000

*Published in 1985.
**Published in 1984.

Triangulation, trilateration, and traverse surveys are included in the 1984 horizontal control standards and specifications, and differential leveling is covered in the vertical control section.

Tables 19.4 and 19.5 give the 1984 FGCS accuracy standards required for the various orders and classes of horizontal and vertical control surveys, respectively. Values in Table 19.4 are ratios of allowable relative positional errors of a pair of horizontal control points to the horizontal distance separating them. Thus, two first-order stations located 100 km (60 mi) apart are expected to be correctly located with respect to each other to within ± 1 m.

Table 19.5 gives maximum relative elevation errors allowable between two benchmarks, as determined by a weighted least-squares adjustment (see Chapter 16).

TABLE 19.5 1984 FGCS VERTICAL CONTROL SURVEY ACCURACY STANDARDS

Order and Class	Relative Accuracy Required Between Benchmarks*
First Order	
Class I	$0.5 \text{ mm} \times \sqrt{K}$
Class II	$0.7 \text{ mm} \times \sqrt{K}$
Second Order	
Class I	$1.0 \text{ mm} \times \sqrt{K}$
Class II	$1.3 \text{ mm} \times \sqrt{K}$
Third Order	$2.0 \text{ mm} \times \sqrt{K}$

*K is distance between benchmarks, in kilometers.

Thus, elevations for two benchmarks 25 km apart, established by second-order class I standards, should be correct to within $\pm 1.0\sqrt{25} = \pm 5$ mm. These standards are not the same as the maximum allowable loop misclosures for the five classes of leveling given in Section 5.5. Table 19.5 values specify relative accuracies of benchmarks after adjustment, whereas loop misclosures enable assessment of results in differential leveling prior to adjustment.

The ultimate success of any engineering or mapping project depends on appropriate survey control. The higher the order of accuracy demanded, the more time and expense are required. It is therefore important to select the proper order of accuracy for a given project and carefully follow the specifications. Note that no matter how accurately a control survey is conducted, errors will exist in the computed positions of its stations, but a higher order of accuracy presumes smaller errors.

■ 19.8 THE NATIONAL SPATIAL REFERENCE SYSTEM

To meet the various local needs of surveyors, engineers, and scientists, the federal government has established a National Spatial Reference System (NSRS) consisting of more than 270,000 horizontal control monuments and approximately 600,000 benchmarks throughout the United States. The National Geodetic Survey (NGS) (which began control surveying operations as the *Survey of the Coast* in 1807, changed to *Coast Survey* in 1836, to *Coast and Geodetic Survey* in 1878, and to a division of the *National Ocean Survey* (NOS) in 1970) has primary responsibility for the NSRS. It continues to assist with, and to coordinate, geodetic control surveying activities with other agencies and with all states to establish new NSRS control stations and upgrade and maintain existing ones. It also disseminates a variety of publications and software related to geodetic surveying.

The NSRS is split into horizontal and vertical divisions. All control within each part is classified in a ranking scheme based on purpose and order of accuracy. These are described in the following two sections.

■ 19.9 HIERARCHY OF THE NATIONAL HORIZONTAL CONTROL NETWORK

The hierarchy of control stations within the NSRS Horizontal Control Network, from highest to lowest order, and their primary uses, are as follows:

Global-regional geodynamics consist of GPS-surveyed points that meet the Order AA accuracy requirements. These are primarily used for international deformation studies.

Primary control consists of GPS-surveyed points that meet the Order A accuracy requirements. These points are used for regional-local geodynamic and deformation studies.

Secondary control densifies the network within areas surrounded by primary control, especially in high-value land areas and for high-precision engineering surveys. Secondary control surveys are executed to GPS Order B standards.

Terrestrial-based control is used for dependent control surveys to meet mapping, land information systems, property survey, and engineering needs. This network consists principally of stations set by traverse and triangulation to first- and second-order standards, and stations set by GPS to Order C standards.

Local control provides reference points for local construction projects and small-scale topographic mapping. These surveys are referenced to higher-order control monuments and, depending on accuracy requirements, may be third-order class I, or third-order class II.

19.10 HIERARCHY OF THE NATIONAL VERTICAL CONTROL NETWORK

The scheme of benchmarks within the National Vertical Control Network may be classified as follows:

Basic framework is a uniformly distributed nationwide network of benchmarks whose elevations are determined to the highest order of accuracy. It consists of nets *A* and *B*. In net *A*, adjacent level lines are ideally spaced at an average of about 100 to 300 km apart using first-order class I standards; in net *B* the average separation is ideally about 50 to 100 km, and first-order class II standards are specified. Benchmarks are placed intermittently along the level lines at convenient locations.

Secondary network densifies the basic framework, especially in metropolitan areas and for large engineering projects. It is established to second-order class I standards.

General area control consists of vertical control for local engineering, surveying, and mapping projects. It is established to second-order class II standards.

Local control provides vertical references for minor engineering projects and small-scale topographic mapping. Benchmarks in this category satisfy third-order standards.

19.11 CONTROL POINT DESCRIPTIONS

To obtain maximum benefit from control surveys, horizontal stations and benchmarks are placed in locations favorable to their subsequent use. The points should be permanently monumented and adequately described to ensure recovery by future potential users. Reference monuments placed by the NGS are marked by bronze disks about 3.5 in. in diameter set in concrete or bedrock. Figure 19.8 shows two types of these disks.

Procedures for establishing permanent monuments vary with the type of soil or rock, climatic conditions, and intended use for the monument. In cases where soil can be excavated, monuments are commonly set in concrete that goes a foot or more below the local maximum frost depth. The bottom of the excavation is generally wider than the top to maximize monument stability during periods of freeze and thaw. Another option commonly used today is to drive a stainless steel rod to refusal using powered tools. Driving depths of 10 ft or more are common when using this technique. In bedrock, holes are often drilled into the rock and

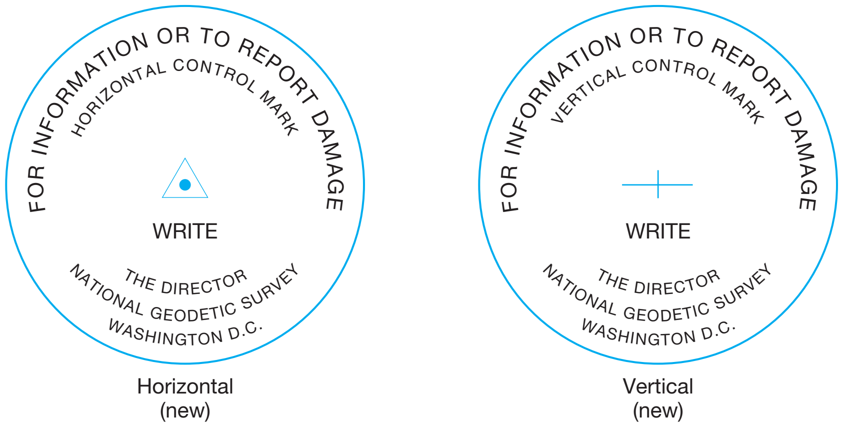


Figure 19.8 Bronze disks used by the National Geodetic Survey to mark horizontal and vertical control stations.

the monument is simply cemented into the hole. Other variations for monumenting can be used, as long as the resulting objects will remain stable in their positions.

The NGS ranks their monuments in their database by their stability. Quality code A monuments are the most reliable and are expected to hold a precise elevation. These monuments are typically rock outcrops, bedrock, and similar features plus massive structures with deep foundations; large structures with foundations on bedrock; or sleeved deep settings (10 ft or more) with galvanized steel pipe or galvanized steel, stainless steel, or aluminum rods. Quality code B monuments are those that will probably hold a precise elevation. Examples include unsleeved deep settings (10 ft or more) with galvanized steel pipe or galvanized steel, stainless steel, or aluminum rods; massive structures other than those listed under Quality code A, massive retaining walls, abutments and piers of large bridges or tunnels, unspecified rods or pipe in a sleeve less than 10 ft, or sleeved copper-clad steel rods. Quality code C monuments are those that may hold a precise elevation but are subject to ground movement. Examples of these monuments include metal rods with base plates less than 10 ft deep, concrete posts (3 ft or more deep), unspecified rods or pipe more than 10 ft deep, large boulders, retaining walls for culverts or small bridges, footings or foundation walls of small to medium-size structures, or foundations such as landings, platforms, or steps. Quality code D monuments are those of questionable stability. Examples include objects of unknown character, shallow set rods or pipe (less than 10 ft), light structures, pavements such as street, curbs, or aprons, piles and poles such as spikes in utility poles, masses of concrete, or concrete posts less than 3 ft deep.

The NGS makes complete descriptions of all its control stations available to surveyors. As an example, a partial listing from an actual NGS horizontal control station description is given in Figure 19.9. These descriptions give each station's general placement in relation to nearby towns, instructions on how to reach the station following named or numbered roads in the area, and the monument's precise location by means of distances and directions to several nearby objects. The station's specific description, such as, "a triangulation disk set in drill hole in rock outcrop," is given, along with a record of recovery history. Data supplied with

```
       National Geodetic Survey,   Retrieval Date = OCTOBER 18, 1999
 LZ1878 ***********************************************************************
 LZ1878  DESIGNATION -  HAYFIELD NE 1974
 LZ1878  PID         -  LZ1878
 LZ1878  STATE/COUNTY-  PA/LUZERNE
 LZ1878  USGS QUAD   -  HARVEYS LAKE (1979)
 LZ1878
 LZ1878                         *CURRENT SURVEY CONTROL
 LZ1878  ___________________________________________________________________
 LZ1878* NAD 83(1986)-  41 18 20.25410(N)    076 00 57.00239(W)     ADJUSTED
 LZ1878* NAVD 88     -       398.7    (meters)    1308.     (feet)  VERTCON
 LZ1878  ___________________________________________________________________
 LZ1878  LAPLACE CORR-           0.27  (seconds)                    DEFLEC99
 LZ1878  GEOID HEIGHT-         -31.73  (meters)                     GEOID99
 LZ1878
 LZ1878  HORZ ORDER  -  SECOND
 LZ1878
 LZ1878.The horizontal coordinates were established by classical geodetic methods
 LZ1878.and adjusted by the National Geodetic Survey in July 1986.
 LZ1878.No horizontal observational check was made to the station.
 LZ1878
 LZ1878.The NAVD 88 height was computed by applying the VERTCON shift value to
 LZ1878.the NGVD 29 height (displayed under SUPERSEDED SURVEY CONTROL.)
 LZ1878
 LZ1878.The Laplace correction was computed from DEFLEC99 derived deflections.
 LZ1878
 LZ1878.The geoid height was determined by GEOID99.
 LZ1878
 LZ1878;                   North         East     Units   Scale      Converg.
 LZ1878;SPC PA N     -   127,939.400   745,212.637   MT  0.99995873 +1 08 50.0
 LZ1878;UTM  18      - 4,573,182.989   414,962.076   MT  0.99968900 -0 40 14.0
 LZ1878
 LZ1878                          SUPERSEDED SURVEY CONTROL
 LZ1878
 LZ1878  NAD 27      -  41 18 19.96597(N)    076 00 58.27835(W) AD(       ) 2
 LZ1878  NGVD 29     -       398.9    (m)         1309.     (f) VERT ANG
 LZ1878
 LZ1878.Superseded values are not recommended for survey control.
 LZ1878.NGS no longer adjusts projects to the NAD 27 or NGVD 29 datums.
 LZ1878.See file dsdata.txt to determine how the superseded data were derived.
 LZ1878
 LZ1878_STABILITY: C = MAY HOLD, BUT OF TYPE COMMONLY SUBJECT TO
 LZ1878+STABILITY: SURFACE MOTION
 LZ1878
 LZ1878  HISTORY     - Date     Condition        Recov. By
 LZ1878  HISTORY     - 1975     MONUMENTED       NGS
 LZ1878
 LZ1878                          STATION DESCRIPTION
 LZ1878
 LZ1878'DESCRIBED BY NATIONAL GEODETIC SURVEY 1975 (CLN)
 LZ1878'THE STATION IS LOCATED ABOUT 3/4 MILE SOUTHEAST OF LEHMAN AND ON THE
 LZ1878'GROUNDS OF THE PENNSYLVANIA STATE UNIVERSITY (WILKES-BARRE
 LZ1878'CAMPUS).
(continues...)
```

Figure 19.9 Partial listing of station data sheet in the National Spatial Reference System for horizontal control station Hayfield NE. (Courtesy National Geodetic Survey.)

horizontal control-point descriptions include the datum(s) used and the station's geodetic latitude and longitude. Also given are the state plane coordinates, convergence angle and scale factor, UTM coordinates (see Chapter 20), and approximate elevation and geoidal height (in meters).

```
NB0293 *********************************************************************
NB0293  DESIGNATION -  F 137
NB0293  PID         -  NB0293
NB0293  STATE/COUNTY-  NY/TIOGA
NB0293  USGS QUAD   -  ENDICOTT (1978)
NB0293
NB0293                         *CURRENT SURVEY CONTROL
NB0293  ___________________________________________________________________
NB0293* NAD 83(1986)-  42 04 10.     (N)    076 07 04.     (W)     SCALED
NB0293* NAVD 88     -       252.471  (meters)     828.32   (feet)  ADJUSTED
NB0293  ___________________________________________________________________
NB0293  GEOID HEIGHT-         -32.70  (meters)                     GEOID99
NB0293  DYNAMIC HT  -         252.373 (meters)      827.99  (feet)  COMP
NB0293  MODELED GRAV-     980,231.5   (mgal)                        NAVD 88
NB0293
NB0293  VERT ORDER  -  FIRST     CLASS II
```

Figure 19.10 An excerpt from a NGS data sheet for benchmark F 137. (Courtesy National Geodetic Survey.)

Some station descriptions give geodetic and grid azimuths (see Chapter 20) to a nearby station or stations. Geodetic and grid azimuths differ by an amount equal to the convergence angle, and therefore the appropriate azimuth must be selected for the particular surveying methods used.

As shown in Figure 19.10, published benchmark data include approximate station locations and adjusted elevations in both meters and feet. Again the relevant datum is identified.

Besides control within the national network set by the NGS, additional marks have been placed in various parts of the United States by other federal agencies such as the USGS, Corps of Engineers, and Tennessee Valley Authority. State, county, and municipal organizations may have also added control. When this work is coordinated through the NGS, the descriptions of the stations involved are distributed by the NGS.

As noted earlier, complete descriptions for all points in the National Spatial Reference System can be obtained from the NGS.[10] This includes horizontal, vertical, and GNSS control points. The descriptions can be obtained in hard copy form, or in computer format on diskettes or compact disks. Only five compact disks are needed to store all NSRS data for the entire United States!

■ 19.12 FIELD PROCEDURES FOR TRADITIONAL HORIZONTAL CONTROL SURVEYS[11]

As noted earlier, in spite of the increasing prominence of GNSS surveys, horizontal control surveys over limited areas are still being accomplished by the traditional methods of triangulation, trilateration, precise traverse, or a combination

[10]It is possible to obtain *data sheets* directly from the NGS Internet site at http://www.ngs.noaa.gov/datasheet.html. This website allows the user to search for data sheets of control points based on the station's name, *permanent identifier* (PID), and perform radial and rectangular searches from a location, or from a clickable image map.

[11]Traditional as used here implies non-GNSS ground-surveying methods.

of these techniques. These methods are described briefly in the subsections that follow.

Traditional methods in horizontal control surveys require observations of horizontal distances, angles, and observations of astronomic azimuths. Basic theory, equipment, and procedures for making these observations have been covered in earlier chapters. The following sections concentrate on procedures specific to control surveys, and on matters related to obtaining the higher orders of accuracy generally required for these types of surveys. Readers interested in performing traditional geodetic surveys should refer to the FGCS manuals.

19.12.1 Triangulation

Prior to the emergence of electronic distance-measuring equipment, triangulation was the preferred and principal method for horizontal control surveys, especially if extensive areas were to be covered. Angles could be more easily observed compared with distances, particularly where long lines over rugged and forested terrain were involved, by erecting towers to elevate the operators and their instruments. Triangulation possesses a large number of inherent checks and closure conditions that help detect blunders and errors in field data, and increase the possibility of meeting a high standard of accuracy.

As implied by its name, triangulation utilizes geometric figures composed of triangles. Horizontal angles and a limited number of sides called *baselines* are observed for length. By using the angles and baseline lengths, triangles are solved trigonometrically and positions of stations (vertices) calculated.

Different geometric figures are employed for control extension by triangulation, but chains of quadrilaterals called arcs (Figure 19.11) have been most common. They are the simplest geometric figures that permit rigorous closure checks and adjustments of field observational errors, and they enable point positions to be calculated by two independent routes for computational checks. More complicated figures like that illustrated in Figure 19.12 have frequently been used to establish horizontal control by triangulation in metropolitan areas.

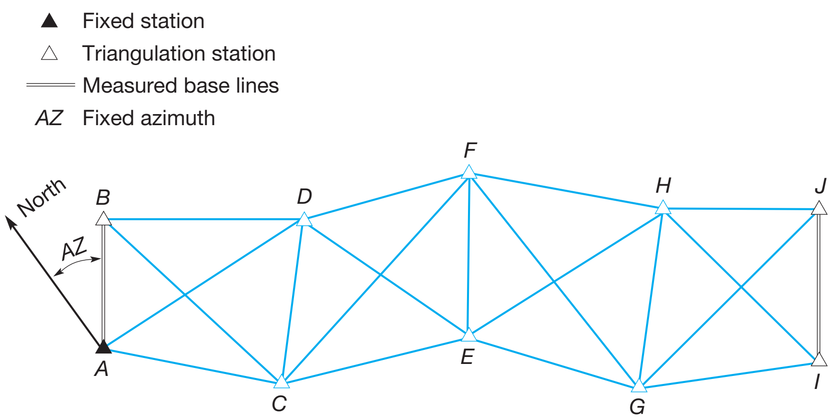


Figure 19.11 Chain of quadrilaterals.

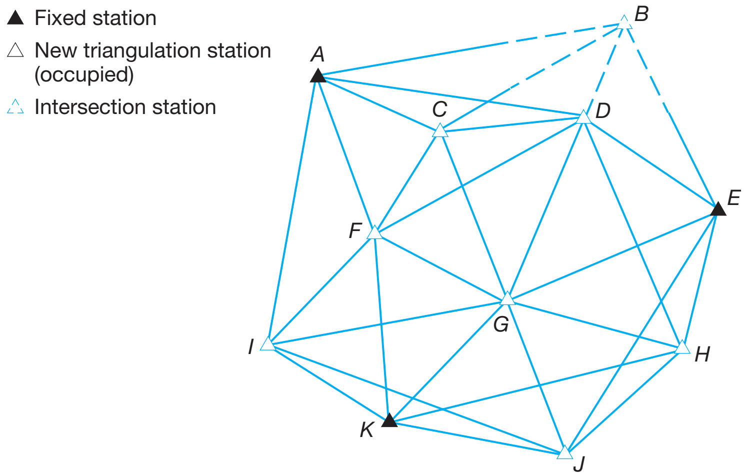


Figure 19.12 Triangulation network for a metropolitan control survey.

In executing triangulation surveys, *intersection stations* can be located as part of the project. In this process, angles are observed from as many occupied points as possible to tall prominent objects in the area such as church spires, smokestacks, or water towers. The intersection stations are not occupied, but their positions are calculated; thus, they become available as local reference points. An example is station *B* in Figure 19.12.

To compensate for the errors that occur in the observations, triangulation networks must be adjusted. The most rigorous method utilizes least squares (see Chapter 16). In that procedure all angle, distance, and azimuth observations are simultaneously included in the adjustment and given appropriate relative weights based on their precisions. The least-squares method not only yields the most probable adjusted station coordinates for a given set of data and weights, but it also gives their precisions.

19.12.2 Precise Traverse

Precise traversing is common among local surveyors for horizontal control extension, especially for small projects. Fieldwork consists of two basic parts: observing horizontal angles at the traverse hubs and observing distances between stations. With total station instruments, these observations can be observed simultaneously. Precise traverses always begin and end on stations established by equal- or higher-order surveys.

Unlike triangulation, in which stations are normally widely separated and placed on the highest ridges and peaks in an area, traverse routes generally follow the cleared rights-of-way of highways and railroads, with stations located relatively close together. Besides easing fieldwork, this provides a secondary benefit in accessibility to the stations. Traverses lack the automatic checks inherent in triangulation, and extreme observational caution must therefore be exercised to avoid blunders. Also, since traverses usually run along single lines, they are generally not as good as triangulation for establishing control over large areas.

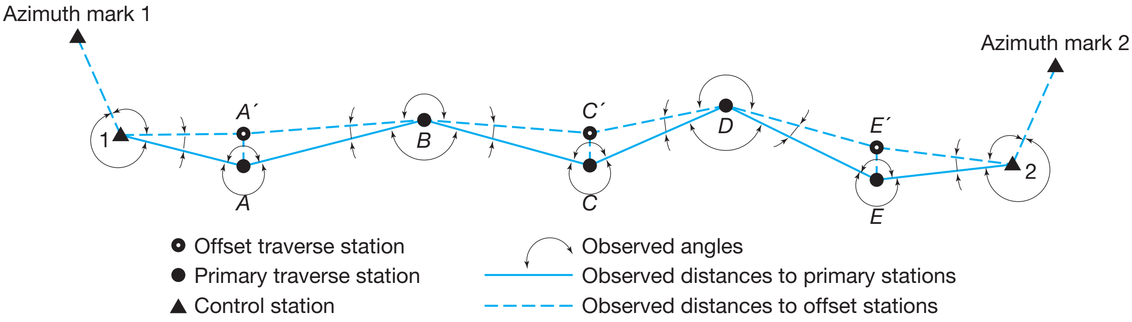


Figure 19.13 Control traverse with offset stations.

Control traverses can be strengthened to provide additional checks in the data by establishing "offset stations" such as A', C', and E' of Figure 19.13. An offset station is set near every-other primary traverse station. In performing the field observations, instrument setups are made only at the primary traverse stations. All possible angles are observed with horizon closures at each station; thus, four angles are determined at interior single primary stations and two angles are observed at primary stations with nearby offset stations. This observation scheme is shown in Figure 19.13. Additionally, all distances are observed, that is, at station 1 distances $1A$ and $1A'$ are observed, at station A, lengths $A1$, AA', and AB are observed, and so on. When the field observations have been completed, the network can be adjusted using all observations in a least-squares adjustment, thereby providing geometric checks for all angle and distance observations in the traverse. Additional geometric strength in the figure could be obtained by also observing angles at the offset stations.

In traversing to gain overall project efficiency and improve angle accuracy, it is always preferable to have long sight distances. Also to avoid mistakes it is advisable to avoid having nearly "flat" angles (values near 180°) whenever possible. To accomplish this, presurvey reconnaissance is recommended. An oft-made mistake is to construct the traverse while collecting the observations. This technique works in low-order surveys but frequently results in poorly designed control traverses.

For long traverses, checks on the observed horizontal angles can be obtained by making periodic astronomical azimuth observations (see Appendix C). These should agree with the values computed from the direction of the starting line and the observed horizontal angles. However, if a traverse extends an appreciable east-west direction, as illustrated in Figure 19.14, meridian convergence will cause the two azimuths to disagree. For example, in Figure 19.14, azimuth FG obtained from direction AB and the observed horizontal angles should equal astronomic azimuth $FG + \theta$, where θ is the meridian convergence. A good approximation for meridian convergence between two points on a traverse is

$$\theta'' = \frac{\rho d \tan \phi}{R_e} \tag{19.15}$$

where θ'' is meridian convergence, in seconds; d the east-west distance between the two points in meters; R_e the mean radius of the Earth (6,371,000 m); ϕ the mean

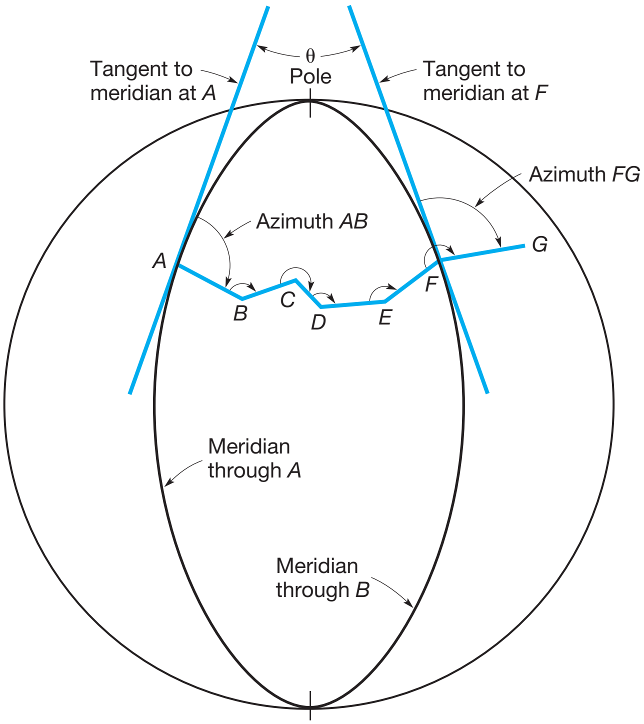


Figure 19.14 Meridian convergence on long east-west traverses.

latitude of the two points; and ρ the number of seconds per radian (206,265). Because of meridian convergence, forward and back azimuths of long east-west lines do not differ by exactly 180°, but rather by $180° \pm \theta$. (If the traverse proceeds in an easterly direction the sign of θ is positive, if it goes westerly θ is negative. A sketch will clarify the situation.) From Equation (19.15) an east-west traverse of 1 mi length at latitude 30° produces a convergence angle of approximately 30″. At latitude 45°, convergence is approximately 51″/mi east-west. These calculations illustrate that the magnitude of convergence can be appreciable and must be considered when astronomic observations are made in connection with plane surveys that assume the y-axis parallel throughout the project area.

Procedures for precise traverse computation vary depending on whether a geodetic or a plane reference coordinate system is used. In either case, it is necessary first to eliminate mistakes and compensate for systematic errors. In the adjustment, closure conditions are enforced for (1) azimuths or angles, (2) departures, and (3) latitudes. The most rigorous process, the least-squares method (see Chapter 16), should be used because it simultaneously satisfies all three conditions and gives residuals having the highest probability.

19.12.3 Trilateration

Trilateration, a method for horizontal control surveys based exclusively on observed horizontal distances, has gained acceptance because of electronic distance measuring capability. Both triangulation and traversing require time-consuming horizontal angle measurement. Hence, trilateration surveys often can be executed faster and produce equally acceptable accuracies.

The geometric figures used in trilateration, although not as standardized, are similar to those employed in triangulation. Stations should be intervisible and are therefore placed in elevated locations, sometimes using towers to elevate instruments and observers if necessary.

Because of intervisibility requirements and the desirability of having essentially square networks, trilateration is ideally suited to densify control in metropolitan areas and on large engineering projects. In special situations where topography or other conditions require elongated narrow figures, reading some horizontal angles can strengthen the network. Also, for long trilateration arcs, astronomic azimuth observations can help prevent the network from deforming in direction.

As in triangulation, surveys by trilateration can be extended from one or more monuments of known position. If only a single station is fixed, at least one azimuth must be known or observed.

Trilateration computations consist of reducing observed slope distances to horizontal lengths, then to the ellipsoid and finally to grid lengths if the calculations are being done in state plane coordinate systems (see Chapter 20). Observational errors in trilateration networks must be adjusted, preferably by the least-squares method.

19.12.4 Combined Networks

With the ability to easily observe both distances and angles in the field, networks similar to that shown in Figure 19.12 are becoming increasingly popular. In a combined network, many or all angles and distances are observed. These surveys provide the greatest geometric strength and the highest coordinate accuracies for traditional survey techniques. As described in Section 19.14, all observations must be corrected to the ellipsoid or a mapping grid (see Chapter 20). The least-squares method as described in Chapter 16 is used to adjust the observations.

■ 19.13 FIELD PROCEDURES FOR VERTICAL CONTROL SURVEYS

Vertical control surveys are generally run by either direct differential leveling or trigonometric leveling. The method selected will depend primarily on the accuracy required, although the type of terrain over which the leveling will be done is also a factor. Differential leveling, described in Section 5.4 produces the highest order of accuracy typically. The GNSS surveys can be used for lower-order vertical control surveys, but to get accurate elevations using this method, *geoidal heights* in the area must be known and applied (see Section 19.5).

Although trigonometric leveling produces a somewhat lower order of accuracy than differential leveling, the method is still suitable for many projects such as establishing vertical control for topographic mapping or for lower-order construction stakeout. It is particularly convenient in hilly or mountainous terrain where large differences in elevation are encountered. Field procedures for trigonometric leveling and methods for reducing the data are discussed in Section 4.5.4.

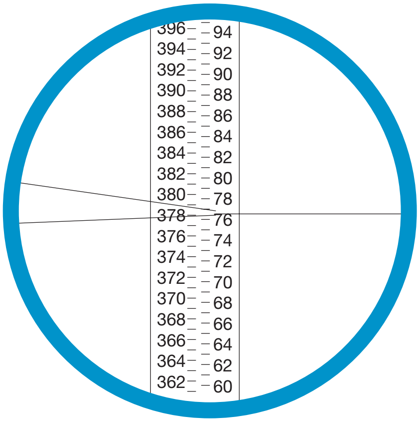


Figure 19.15
Reticle of a precise level shown with dual metric-scale precise leveling rod.

Differential leveling can produce varying levels of accuracy, depending on the precautions taken. In this section only *precise differential leveling*, which produces the highest quality results, is considered.

As noted in Section 19.7 and Table 19.5, the FGCS-established accuracy standards and specifications for various orders of differential leveling. To achieve the higher orders, special care must be exercised to minimize errors, but the same basic principles apply.

Special level rods are needed for precise work. They have scales graduated on Invar strips, which are only slightly affected by temperature variations. Precise level rods are equipped with rod levels to facilitate plumbing and special braces aid in holding the rod steady. They usually have two separate graduated scales. One type of rod is divided in centimeters on an Invar strip on the rod's front side, with a scale in feet painted on the back for checking readings and minimizing blunders. A second type of rod, shown in Figure 19.15, has two sets of centimeter graduations on the Invar strip with the right one precisely offset from the left by a constant, thereby giving checks on readings.

Cloudy weather is preferable for precise leveling, but an umbrella can be used on sunny days to shade the instrument and prevent uneven heating, which causes the bubble to run. (One design encases the vial in a Styrofoam shield.) Automatic levels are not as susceptible to errors caused by differential heating. Precise work should not be attempted on windy days. For best results, short and approximately equal backsight and foresight distances are recommended. Table 19.6 lists the maximum sight distances and allowable differences between backsight and foresight lengths for first-, second-, and third-order leveling. Rodpersons can pace or count rail lengths or highway slab joints to set sight distances, which are then checked for accuracy by three-wire stadia methods (see Section 5.8). Precise leveling demands good-quality turning points. Lines of sight should not pass closer than about 0.5 m from any surface, for example, the ground, to minimize refraction. Readings at any setup must be completed in rapid succession; otherwise changes in atmospheric conditions might significantly alter refraction characteristics between them.

Three-wire leveling has been employed for much of the precise work in the United States. In this procedure, as described in Section 5.8, rod readings at the

TABLE 19.6 RECOMMENDED FIELD CONDITIONS FOR PRECISE LEVELING

Order Class	First I	First II	Second I	Second II	Third
Maximum sight length (m)	50	60	60	70	90
Difference between foresight and backsight lengths never to exceed					
per setup (m)	2	5	5	10	10
per section (m)	4	10	10	10	10

upper, middle, and lower cross wires are taken and recorded for each backsight and foresight. The difference between the upper and middle readings is compared with that between the middle and lower values for a check, and the average of the three readings is used. A sample set of field notes for three-wire leveling is illustrated in Figure 5.9. When using a digital level, the elevation difference along with the backsight or foresight length can be digitally recorded for each sight. A second technique in precise leveling employs the parallel-plate micrometer attached to a precise leveling instrument and a pair of precise rods like those described earlier.

It is generally advisable to design large level networks so that several smaller circuits are interconnected. This enables making checks that isolate blunders or large errors. For example, in Figure 19.16, it is required to determine the elevations of points X, Y, and Z by commencing from BM A and closing on BM B. As a minimum, running level lines 1 through 4 could do this, but if an unacceptable misclosure were obtained at BM B, it would be impossible to discover which line the blunder occurred. If additional lines 5, 6, and 7 are run, calculating differences in elevation by other routes through the network should isolate the blunder. Furthermore, by including the supplemental observations, precisions of the resulting elevations at X, Y, and Z are increased.

For long lines, one procedure used to help isolate mistakes and minimize field time is to run small loops with approximately five setups between temporary benchmarks. In this procedure as each loop is completed, it is checked for acceptable closure before proceeding forward to the next loop. This procedure increases the number of observations, but helps minimize the amount of time that

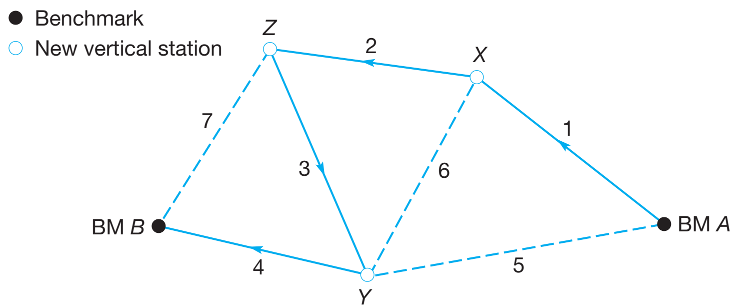


Figure 19.16 Interconnecting level network.

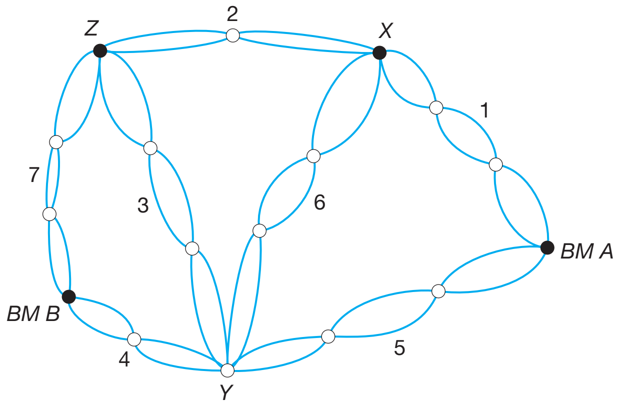


Figure 19.17 Leveling of network performed in small closing loops.

is required to uncover mistakes. Each smaller loop is connected to subsequent loops until the entire network is observed. Figure 19.17 depicts this procedure.

In precise differential leveling, frequent calibration of the leveling instrument is necessary to determine its collimation error. A collimation error exists if, after leveling the instrument, its line of sight is inclined or depressed from horizontal. This causes errors in determining elevations when backsight and foresight distances are not equal. But they can be eliminated if the magnitude of the collimation error is known.

A method that originated at the National Geodetic Survey can be used to determine the collimation error. It requires a baseline approximately 300 ft (90 m) long. Stakes are set at each end of the line and at two intermediate stations located approximately 20 ft (6 m) and 40 ft (12 m) from the two ends. Figure 19.18 shows an example set of field notes for determining the collimation error and includes a sketch illustrating the baseline layout. With the instrument at station 1, middle wire r_1 and R_1 are observed on stations A and B, respectively. If there were no collimation error, the true elevation difference ΔH from these observations would be $r_1 - R_1$. However, if a collimation error is present, each observation must be corrected by adding an amount proportional to the horizontal distance from the level to the rod. The horizontal distance is measured by the stadia interval (see Section 16.9.2). Introducing collimation corrections, the true elevation difference ΔH is

$$\Delta H = [r_1 + C(i_1)] - [R_1 + C(I_1)] \qquad \text{(a)}$$

In Equation (a), i_1 and I_1 are the stadia intervals (differences between top and bottom cross-wire values) for the rod readings on stations A and B, respectively, and C is the collimation factor (in feet per foot, or meters per meter, of stadia interval). A similar equation for the true elevation difference can be written for rod readings R_2 and r_2 taken on stations A and B, respectively, from station 2, or

$$\Delta H = [R_2 + C(I_2)] - [r_2 + C(i_2)] \qquad \text{(b)}$$

LEVEL CALIBRATION

STA	A	INT	B	INT
	4.516		5.567	
1	4.414	0.102	4.644	0.923
	4.312	0.102	3.721	0.923
	r_1	i_1	R_1	I_1
	4.414	0.204	4.644	1.846
	4.699		4.182	
2	3.825	0.874	4.042	0.140
	2.950	0.875	3.901	0.141
	R_2	I_2	r_2	i_2
	3.825	1.749	4.042	0.281

$$C = \frac{(4.414 + 4.042) - (4.644 + 3.825)}{(1.846 + 1.749) - (0.204 + 0.281)}$$

$$C = \frac{-0.013}{3.110} = -0.0042\ \mathrm{m/m}$$

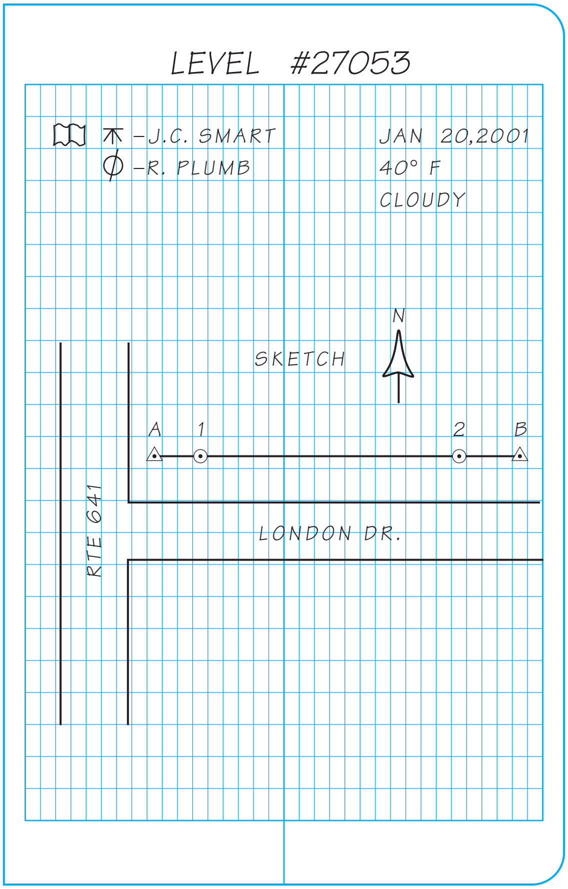


Figure 19.18 Field notes for determining collimation factor.

Note that in Equations (a) and (b), uppercase R and I apply to the longer sights, and lowercase r and i are for the shorter sights. Equating the right sides of Equations (a) and (b), and reducing, yields

$$C = \frac{(r_1 + r_2) - (R_1 + R_2)}{(I_1 + I_2) - (i_1 + i_2)} \tag{19.16}$$

As previously noted, the units of the collimation factor calculated by Equation (19.16) are either in feet per feet, or meters per meter, of stadia interval. Computation of the factor is illustrated in Figure 19.18 for the data of the field notes (given in feet).

Because the collimation correction increases linearly with distance, it is unnecessary to apply it to each backsight and foresight. Rather the corrected elevation difference $\Delta H'$ for any loop or section leveled is computed as

$$\Delta H' = \Sigma BS - \Sigma FS + C(\Sigma I_{BS} - \Sigma I_{FS}) \tag{19.17}$$

where ΣBS is the sum of the middle cross-wire readings of backsights in the loop or section; ΣFS is the middle cross-wire total of foresights; and ΣI_{BS} and ΣI_{FS} are the sums of the stadia intervals for the backsights and foresights, respectively.

Example 19.5

The section from BM A to BM 3 is leveled using the instrument whose collimation factor of −0.00012 m/m of interval was determined in the field notes in Figure 19.18. The sum of the backsights is 125.590 m, and the sum of the foresights is 88.330 m. Backsight stadia intervals total 351.52 m, while the sum of foresight intervals is 548.40 m. Find the corrected elevation difference.

Solution

By Equation (19.17)

$$\begin{aligned}\Delta H' &= (125.590 - 88.330) + (-0.00012)(351.52 - 548.40)\\ &= 37.260 + 0.024 = 37.284 \text{ m}\end{aligned}$$

Regardless of precautions taken in field observations, errors accumulate in leveling and must be adjusted to provide perfect mathematical closure. For simple level loops, adjustment procedures presented in Section 5.6 can be followed; for interconnected level networks such as those of Figures 19.16 and 19.17, the method of least squares is preferable. An example least-squares adjustment of an interconnected network is given in Section 16.6.

19.14 REDUCTION OF FIELD OBSERVATIONS TO THEIR GEODETIC VALUES

Traditional surveying instruments, such as levels and total stations, are oriented with respect to the local gravity surface. In geodetic work, since horizontal surveys are referenced to an ellipsoid and vertical surveys to the geoid, corrections must be made to field observations to obtain their equivalent geodetic values. The following subsections discuss some of these corrections. Many of these computations are demonstrated in the Mathcad® worksheet *geodobs.xmcd*, which is available on the Prentice Hall companion website for this book. Additionally, WOLFPACK, which is available from the same website, has options to perform these computations.

19.14.1 Reduction of Distance Observations Using Elevations

In geodetic control survey computations, observed slope distances (sometimes called *slant* distances) must first be reduced to the surface of the ellipsoid. Observed

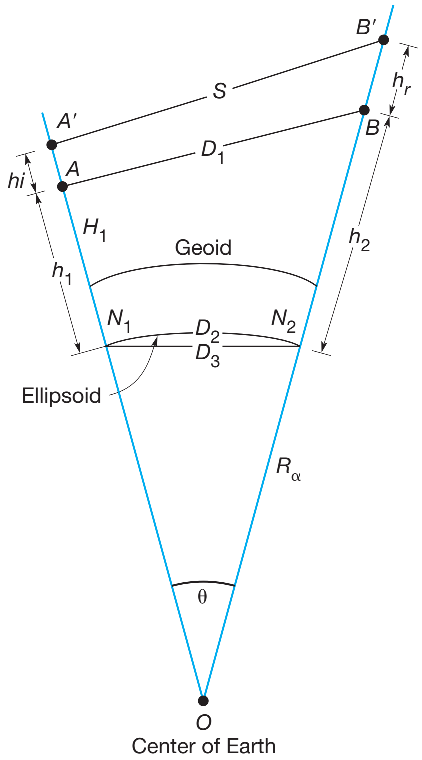


Figure 19.19 Reduction of long lengths to the ellipsoid based on elevations.

distances in geodetic surveys are often long and thus the short-line reduction techniques given in Section 6.13 do not provide satisfactory accuracy. This is especially true for long lines that are steeply inclined.

A procedure for reducing long slope distances to their ellipsoid lengths is discussed here. The method is based on elevation differences between the end points of the sloping line. In Figure 19.19, an EDM instrument is at A, a reflector is at B, and S is the observed slope distance from A to B. (Assume that S has been corrected for meteorological conditions.) Length D_1 is the "mark-to-mark" distance between stations A and B. Mark-to-mark distances apply for EDM calibration lines, as well as for GNSS baselines. Length D_2 is the arc distance on the ellipsoid, which is also known as the *geodetic distance*. It is the length required for most geodetic computations. Distance D_3 is the ellipsoidal chord length between stations A and B.

In Figure 19.19, let $h_1' = h_1 + hi$ and $h_2' = h_2 + hr$, where hi and hr are instrument and reflector heights, respectively, above stations A and B, and h_1 and h_2 are the geodetic heights at A and B, respectively. Expressing the relationship of the three sides of triangle ABO using the law of cosines [see Equation (11.2)] gives

$$S^2 = (R_\alpha + h_1')^2 + (R_\alpha + h_2')^2 - 2(R_\alpha + h_1')(R_\alpha + h_2')\cos\theta \quad \textbf{(19.18)}$$

where R_α is the radius of the earth in the azimuth of the distance from point A as defined by Equation (19.6), and θ the angle subtended by the verticals from points

A and B. Substituting the trigonometric identity of $\cos\theta = 1 - 2\sin^2(\theta/2)$ into Equation (19.18) and expanding yields

$$S^2 = (h_2' - h_1')^2 + 4R_\alpha^2\left(1 + \frac{h_1'}{R_\alpha}\right)\left(1 + \frac{h_2'}{R_\alpha}\right)\sin^2\left(\frac{\theta}{2}\right) \tag{19.19}$$

Substituting $\Delta h' = h_2' - h_1'$ and $D_3 = 2R_\alpha \sin(\theta/2)$ into Equation (19.19), the expression reduces to

$$S^2 = \Delta h'^2 + \left(1 + \frac{h_1'}{R_\alpha}\right)\left(1 + \frac{h_2'}{R_\alpha}\right)D_3^2 \tag{19.20}$$

Solving Equation (19.20) for D_3 yields the following expression for the ellipsoidal chord length:

$$D_3 = \sqrt{\frac{S^2 - \Delta h'^2}{\left(1 + \frac{h_1'}{R_\alpha}\right)\left(1 + \frac{h_2'}{R_\alpha}\right)}} \tag{19.21}$$

The arc length on the ellipsoid (*geodetic distance* or *geodetic length*) can be computed from this chord distance as

$$D_2 = 2R_\alpha \sin^{-1}\left(\frac{D_3}{2R_\alpha}\right) \tag{19.22}$$

Equations (19.21) and (19.22) can be used to compute the distance on any level surface by simply modifying the heights of the endpoints as appropriate. It is important to realize that the unit of the arcsine in Equation (19.22) is radians. To compute the chord distance between two points at different elevations, for example, D_1 in Figure 19.19, the following equation is used

$$D_1 = \sqrt{D_3^2\left(1 + \frac{h_1}{R_\alpha}\right)\left(1 + \frac{h_2}{R_\alpha}\right) + (h_2 - h_1)^2} \tag{19.23}$$

Example 19.6

A slope distance of 5000.000 m is observed between two points A and B whose orthometric heights are 451.200 and 221.750 m, respectively. The geoidal undulation at point A is −29.7 m and is −29.5 m at point B. The height of the instrument at the time of the observation was 1.500 m and the height of the reflector was 1.250 m. What are the geodetic and mark-to-mark distances for this observation? (Use a value of 6,386,152.318 m for R_α in the direction AB.)

Solution

By Equation (19.7), the geodetic heights at points A and B are

$$h_A = 451.200 - 29.7 = 421.500 \text{ m}$$
$$h_B = 221.750 - 29.5 = 192.250 \text{ m}$$

Thus, $h'_A = 421.500 + 1.500 = 423.000$ m, $h'_B = 192.250 + 1.250 = 193.500$, $\Delta h' = 193.500 - 423.000 = -229.500$ m, and by Equation (19.20), the ellipsoidal chord distance D_3 is

$$D_3 = \sqrt{\frac{5000^2 - 229.5^2}{\left(1 + \frac{423.0}{R_\alpha}\right)\left(1 + \frac{193.5}{R_\alpha}\right)}} = 4994.489 \text{ m}$$

By Equation (19.22), the reduced ellipsoidal arc, or geodetic length, for the line AB is

$$D_2 = 2R_\alpha \sin^{-1}\left(\frac{4994.489}{2R_\alpha}\right) = 4994.489 \text{ m}$$

Finally by Equation (19.23), the mark-to-mark distance is

$$D_1 = \sqrt{4994.489^2\left(1 + \frac{421.500}{R_\alpha}\right)\left(1 + \frac{192.250}{R_\alpha}\right) + (192.25 - 421.5)^2}$$
$$= 4999.987 \text{ m}$$

Note that the ellipsoid arc and chord lengths are the same to the nearest millimeter. As lines become longer, however, this will not necessarily be the case. Nevertheless for most geodetic observations, these arc and chord values will generally be nearly the same. Note also that the observed slope distance differs from the mark-to-mark distance by 13 mm. Finally, if the short-line reduction procedure of Section 6.13 had been used, an error of more than 0.2 m would have resulted.

19.14.2 Reduction of Distance Observations Using Vertical Angles

Figure 19.20 illustrates a slope distance S observed from A to B. Points A and B represent an EDM instrument and a reflector, respectively, O is the Earth's center, and R_α its radius in the direction of the azimuth as defined by Equation (19.6). Vertical angles α and β were observed at A and B, respectively. Arc AB_2, which is closely approximated by its chord, is the required horizontal distance. If the short-line reduction procedures given in Section 6.13 were used, horizontal distance AB_1 would result, which would be in error by B_1B_2. Arc $A'B'$ is the required ellipsoid distance.

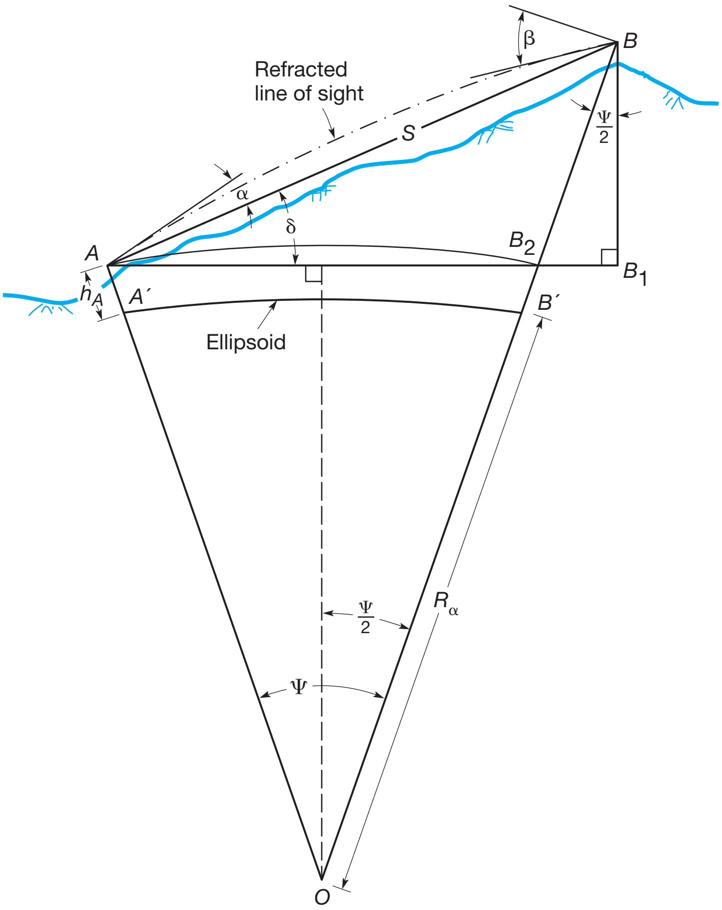


Figure 19.20
Reduction of long lengths to the ellipsoid based on vertical angles.

From Figure 19.20 the following equations can be written to compute required horizontal (chord) distance AB_2:

$$\delta = \frac{\alpha - \beta}{2} \qquad \textbf{(19.24)}$$

$$AB_1 = S \cos \delta \qquad \textbf{(19.25)}$$

$$BB_1 = S \sin \delta \qquad \textbf{(19.26)}$$

$$\psi = \frac{AB_1}{R_\alpha + h_A} \times \frac{180^\circ}{\pi} \text{ (approx.)} \qquad \textbf{(19.27)}$$

$$B_1B_2 = BB_1 \tan\left(\frac{\psi}{2}\right) \qquad \textbf{(19.28)}$$

$$AB_2 = AB_1 - B_1B_2 \qquad \textbf{(19.29)}$$

Finally, ellipsoidal (chord) length $A'B'$ can be computed from

$$A'B' = AB_2\left(\frac{R_\alpha}{R_\alpha + h_A}\right) \tag{19.30}$$

where h_A is the ellipsoidal height.

Example 19.7

In Figure 19.20, slope distance L and vertical angles α and β were observed as 14,250.590 m, 4°32′18″, and −4°38′52″, respectively. If the geodetic height at A is 438.4 m, what is distance $A'B'$ reduced to the ellipsoid? (Use the mean radius of 6,371,000 m for R_α.)

Solution

Solving Equations (19.24) through (19.30) in sequence

$$\delta = \frac{4°32'18'' - (-4°38'52'')}{2} = 4°35'35''$$

$$AB_1 = 14{,}250.590 \cos 4°35'35'' = 14{,}204.826 \text{ m}$$

$$BB_1 = 14{,}250.590 \sin 4°35'35'' = 1141.160 \text{ m}$$

$$\psi = \frac{14{,}204.826}{6{,}371{,}000 + 438.4} \times \frac{180°}{\pi} = 0.127738° = 0°07'40''$$

$$B_1B_2 = 1141.160 \tan\left(\frac{0°07'40''}{2}\right) = 1.272 \text{ m}$$

$$AB_2 = 14{,}204.826 - 1.272 = 14{,}203.554 \text{ m}$$

$$A'B' = 14{,}203.554\frac{6{,}371{,}000}{6{,}371{,}000 + 438.4} = 14{,}202.577 \text{ m}$$

(Note that if the short-line reduction procedure of Section 6.13 had been used, an error equal to B_1B_2 or 1.272 m, would have resulted.)

In the foregoing computations, the effects of refraction were eliminated by averaging reciprocal vertical angles α and β in Equation (19.24). This procedure yields the best results. If the vertical angle were observed at only one end of the line, as angle α at A, then because refraction is approximately 1/7 curvature, or $\psi/7$, a correction for its effect can be applied. In that case instead of using Equation (19.24), angle δ is computed as

$$\delta = \alpha + \frac{\psi}{2} - \frac{\psi}{7} = \alpha + \frac{5\psi}{14} \tag{19.31}$$

where

$$\psi = \frac{S \cos \alpha}{R_e + h_A} \times \frac{180°}{\pi}$$

Then Equations (19.25), (19.26), and (19.28) through (19.30) are solved as before.

Example 19.8

Compute the geodetic (ellipsoidal) length of the line in Example 19.7 using only the observed vertical angle at A.

Solution

$$\psi = \frac{14{,}250.590 \ \cos 4°32'18''}{6{,}371{,}000 + 438.4} \times \frac{180°}{\pi} = 0.127748° = 0°07'40''$$

By Equation (19.31)

$$\delta = 4°32'18'' + \frac{5}{14}(0°07'40'') = 4°35'02''$$

Then Equations (19.25), (19.26), and (19.28) through (19.30) are solved in order

$$AB_1 = 14{,}250.590 \ \cos 4°35'02'' = 14{,}205.008 \text{ m}$$

$$BB_1 = 14{,}250.590 \ \sin 4°35'02'' = 1138.888 \text{ m}$$

$$B_1B_2 = 1138.888 \tan \ = 1.270 \text{ m}$$

$$AB_2 = 14{,}205.008 - 1.270 = 14{,}203.738 \text{ m}$$

$$A'B' = 14{,}203.738 \frac{6{,}371{,}000}{6{,}371{,}000 + 438.4} = 14{,}202.761 \text{ m}$$

Note that this answer differs by 0.184 m from the one obtained in Example 19.7. This can be expected, because refraction varies with atmospheric conditions, and the correction $\psi/7$ only approximates its effects. Thus, as stated earlier, it is best to measure the vertical angles at both ends of the line if possible.

19.14.3 Reduction of Directions and Angles

As was discussed in Section 19.5, deflection of the vertical varies at different locations on the surface of the Earth. Because of this, during the angle observing process both the horizontal and vertical circles of a total station instrument are, in general, misaligned with the horizontal and vertical. Thus, for geodetic calculations, the observed directions must be corrected according to Equation (19.11).

Additionally, because of the sphericity of the Earth, the normals at the observing and target stations are skewed with respect to each other, and hence two additional corrections may be necessary. First, if the height of the target above the ellipsoid is substantial enough, this may necessitate a correction. Second, if the latitudes of the occupied and sighted stations differ significantly, this may also require a correction. However, for the target heights and relatively small latitude differences that apply in most surveys, corrections to azimuths for these conditions are often smaller than the observational errors. Thus, except for very precise geodetic work they are not made. Finally, a correction that stems from deflection of the vertical is often significant and the procedures for making it are given below.

Correction for deflection of the vertical Since the instrument is set up with respect to the local vertical, angular measurements in the vertical and horizontal will both be affected by deflection of the vertical. The correction C''_{defl} in an observed horizontal direction in units of arc seconds for deflection of the vertical is

$$C''_{defl} = Az - \alpha = -\eta \tan\phi - (\xi \sin Az - \eta \cos Az) \cot z \qquad \textbf{(19.32)}$$

where Az is the astronomic azimuth of the observed direction, α is the geodetic azimuth, and z is the zenith angle.

Adding the correction determined in Equation (19.32) yields the corrected geodetic azimuth of an observed direction as

$$\alpha = Az_A + C''_{defl} \qquad \textbf{(19.33)}$$

For angles, the backsight and foresight directions that compose the angle are each corrected according to Equation (19.33) and subtracted according to Equation (11.11). This is demonstrated in Example 19.9.

■ Example 19.9

As shown in Figure 19.21, a horizontal angle BIF of 57°44′06″ is observed at instrument station I, whose latitude is 41°13′24.67″ as scaled from a topographic map. Orthometric heights were also scaled from a topographic map and estimated as 613.8 m at backsight station B and 853.9 m at foresight station F. The geoidal separation is estimated using the NGS software GEOID09 as −29.45 m at station B and −29.84 m at station F. The reflector height hr at both target stations is set to 1.650 m. The azimuth from the observing station to station B is 23°16′24″. The components of the deflection of the vertical at the observing station are estimated using the NGS software DEFLEC09 as η = 4.82″ and ξ = 0.29″. The geodetic distance IB, as defined by Equation (19.21) is 975.548 m, and for IF it is 883.49 m. The zenith angle to station B is 71°30′56″ and to station F is 57°21′46″. What is the corrected geodetic angle BIF?

Solution

In this solution, the individual correction components are independently computed for each sight of the angle and subtracted to determine the corrected angle.

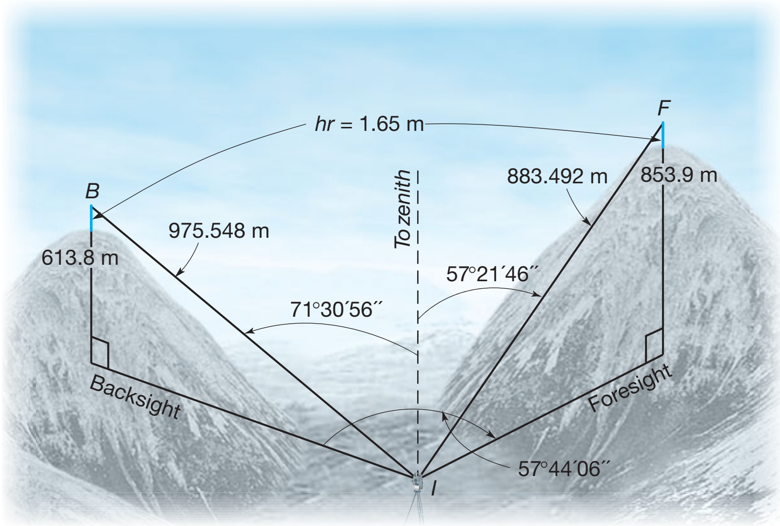


Figure 19.21
Figure for Example 19.9.

Geodetic heights: using Equation (19.7), the geodetic heights of h_B at the backsight station and h_F are

$$h_B = 613.8 - 29.45 + 1.65 = 586.00 \text{ m}$$
$$h_F = 853.9 - 29.84 + 1.65 = 825.71 \text{ m}$$

Correction for deflection of the vertical: by Equation (19.32)

Backsight:

$$C''_{defl} = -4.82'' \tan(\phi) - [0.29'' \sin(Az_B) - 4.82'' \cos(Az_B)] \cot(71°30'56'') = -2.78''$$

where ϕ and Az_B are 41°13′24.67″ and 23°16′24″, respectively.

Foresight:

$$Az_F = 23°16'24'' + 57°44'06'' = 81°00'30''$$

$$C''_{defl} = -4.82'' \tan(\phi) - [0.29'' \sin(Az_F) - 4.82'' \cos(Az_F)] \cot(57°21'46'') = -3.92''$$

where ϕ again is 41°13′24.67″ and Az_F is 81°00′30″.

Corrected azimuths:

$$\text{Backsight: } 23°16'24'' - 2.78'' = 23°16'21.22''$$
$$\text{Foresight: } 81°00'30'' - 3.92'' = 81°00'26.08''$$

By Equation (11.11), the corrected angle is: 81°00′26.08″ − 23°16′21.22″ = 57°44′04.9″

Note that the correction to the angle over these short distances is $-1.1''$. Also note in Equation (19.32) that $\eta \tan \varphi$ is the same for both the backsight and foresight directions of an angle. Thus, it did not have to be included in correction of directions for angles. The correction for deflection of vertical in angles could be rewritten as

$$C''_{\angle} = (\xi \sin Az_{FS} - \eta \cos Az_{FS}) \cot z_{FS} - (\xi \sin Az_{BS} - \eta \cos Az_{BS}) \cot z_{BS}$$

where FS represents the foresight values and BS represents the backsight values.

For precise geodetic control surveys, a correction due to deflection of the vertical must also be made to observed vertical angles. For zenith angles, the following equation applies

$$z_C = z_{\text{obs}} + \xi \cos Az + \eta \sin Az \quad \textbf{(19.34)}$$

where z_C is the corrected zenith angle, z_{obs} the observed zenith angle, and Az the azimuth of the line of sight when the zenith angle is observed.

19.14.4 Leveling and Orthometric Heights

Distances observed along plumb lines (elevation differences) between equipotential gravitational surfaces provide the basis for specifying orthometric heights. One of these equipotential surfaces, the geoid, is defined as the datum for observing these heights. For example, in Figure 19.22, the orthometric heights of points A and B are H_A and H_B, respectively.

The Earth spins about its rotational axis in approximately a 24-h period. From physics it is known that this spin results in a centrifugal acceleration that acts on all

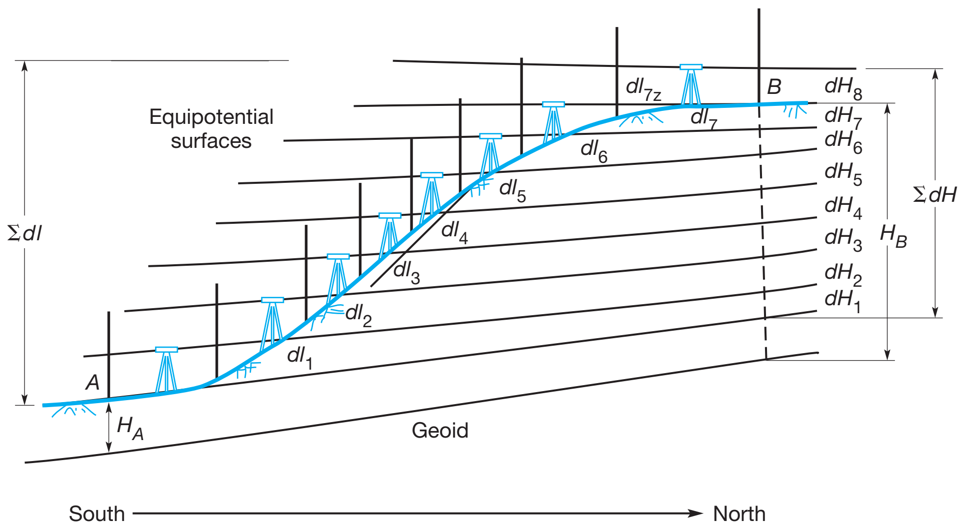


Figure 19.22 Leveled height versus orthometric height.

bodies that take part in the spin. The *gravitational force*[12] is a combination of the attractive force of the Earth's mass and centrifugal force that acts on a body on the surface of the Earth. Since these two forces are in different directions, their combined effect results in a gravitational force that is less than the attractive force and is directed radially toward the mass-center of the Earth. Thus, points on the equator, which experience the greatest rotational velocity, have the weakest gravitational force. Conversely, points at the poles that are not subject to rotational velocity do not experience any centrifugal acceleration component and thus experience the maximum gravitational force. It should be pointed out here that the difference between gravitational force at the pole and at the equator is only about 5 gals.[13]

As stated earlier, an equipotential surface has the same gravitational potential throughout its extent. From physics, potential can be defined as

$$W = mar \tag{19.35}$$

where m is the mass of the body, a the acceleration of the body, and r the radial distance from the mass-center of the Earth. In the case of orthometric elevation differences, the acceleration a in Equation (19.35) can be replaced by the gravitational acceleration $-g$ acting on the point,[14] or

$$W = -mgr \tag{19.36}$$

From Equation (19.36) and the previous discussion concerning changes in the values of gravitational force on the Earth, it can be stated that as gravitational force increases, the radial distance must decrease in order that the potential of points on a given equipotential surface remain equal. Thus, equipotential surfaces are not concentric, and instead they converge as they approach the poles, as shown in Figure 19.22.

The difference in potential is defined as

$$\delta W = -g\,\delta l \tag{19.37}$$

where δW is the change in potential and δl the separation between the two equipotential surfaces. A proper expression for the potential of a point on an equipotential surface is the *geopotential number* C of a point, where C is defined as a negative difference in the potential of a point P and the potential of the geoid, and is mathematically defined as

$$C = -(W - W_0) = \int_{P_0}^{P} g\,dl \tag{19.38}$$

where W_0 is the potential of a point P_0 on the geoid. The units of the geopotential number are kgal-meters (kgal-m) where 1 kgal-m is called a *geopotential unit*

[12] Force is defined from physics as mass times acceleration.

[13] A gal is a unit of acceleration where 1 gal = 1 cm/sec^2; 1 kgal = 1000 gal; 1 mgal = 0.001 gal.

[14] A negative sign is introduced in Equation (19.35) to account for the fact that gravitational attraction points radially inward while increases in elevation occur radially outward.

(GPU). Since neither l nor g are known as a continuous function, in practice the difference in geopotential number is approximated as

$$\Delta C_{ij} = \sum_{k=i}^{j} \overline{g}_k \delta l_k \tag{19.39}$$

where $\overline{g}_k$ is the average gravitational attraction between two adjacent benchmarks i and j and δl_k the observed level difference between the two adjacent benchmarks. Because it is impractical to observe g at every benchmark, and since the units of geopotential numbers are unfamiliar to surveyors, *dynamic heights* were introduced. Dynamic heights are the geopotential number divided by a *reference gravity*, or

$$H_i^D = \frac{C_i}{g_R} \tag{19.40}$$

where H_i^D is the dynamic height of a point, C_i the geopotential number of the point, and g_R the appropriate gravitational constant. In the United States, the NGS has adopted a reference gravity value of 980.6199 gal. The reader should note that the units of dynamic height are traditional distance units. In Figure 19.10, the dynamic height of point F 137 is 252.373 m where the reference gravity is 0.9806199 kgal. Thus, by rearranging Equation (19.40), the geopotential number of point F 137 is 0.9806199(252.373) = 247.482 kgal-m.

Without the inclusion of gravity observations, height differences determined by leveling do not produce true orthometric height differences, and thus a correction must be applied. As can be seen in Figure 19.22, as the leveling process proceeds from station A to B, the equipotential surfaces converge. Letting dl_i represent the leveled height differences and dH_i represent the orthometric height differences between incremental equipotential surfaces, it is apparent that the sum of dl_1, dl_2, dl_3, dl_4, dl_5, dl_6, dl_7, and dl_8 (Σdl_i) will not equal the sum of dH_1, dH_2, dH_3, dH_4, dH_5, dH_6, dH_7, and dH_8 (ΣdH_i) because of convergence. The difference between the leveled height difference and the orthometric height difference is called the *orthometric correction*. The orthometric correction O_c is added to the leveled height to obtain the orthometric height, or

$$dH = dl + O_c \tag{19.41}$$

where dH represents the orthometric height difference between two points and dl represents the leveled height difference between the two points.

Because gravitational surfaces converge near the Earth's poles, and orthometric corrections are a function of gravity values observed along the leveling lines, the largest corrections occur along lines that are run in the north-south direction. When running a line of levels between two NSRS benchmarks, it is possible to estimate the orthometric correction from data that is published on the data sheets. Refer again to Figure 19.10, which shows an excerpt from the data sheet for benchmark F 137. Note that the NAVD88 (orthometric) height is 252.471 m, the modeled gravity value is 980,231.5 mgals, and the modeled geoidal undulation is −32.70 m. To compute the leveled height difference, the potential heights for two control benchmark stations must be computed as

$$H_C(A) = H_A\left(g_A + \frac{0.0424}{1{,}000{,}000}H_A\right) \tag{19.42}$$

where $H_C(A)$ is the potential height of station A in units of kgal-meters, g_A the modeled gravity value at station A in units of kgals, and H_A the orthometric height of the station. Following this, the difference in the potential heights is computed and divided by the average gravity value for the two benchmarks, or

$$dl = \frac{2[H_C(B) - H_C(A)]}{g_A + g_B} \tag{19.43}$$

Example 19.10

Given the following information from the control data sheets for Stations F 137 and J 231, what should be the leveled height difference between stations?

Station	Orthometric Height (m)	Gravity (mgal)
F 137	252.471	980,231.5
J 231	294.548	980,143.5

Solution

By Equation (19.42)

$$H_C(F137) = 252.471\left[0.9802315 + \frac{0.0424}{1{,}000{,}000}252.471\right] = 247.4827 \text{ GPU}$$

$$H_C(J237) = 294.548\left[0.9801435 + \frac{0.0424}{1{,}000{,}000}294.548\right] = 288.7030 \text{ GPU}$$

By Equation (19.43)

$$dl(F, J) = \frac{2(288.7030 - 247.4827)}{0.9802315 + 0.9801435} = 42.053 \text{ m}$$

Note that in Example 19.8, the difference in orthometric heights is 294.548 − 252.471 = 42.077 m, but the leveled height difference is 42.053 m yielding a difference of 2.4 cm. This difference represents the orthometric correction for the leveled line and would be seen as part of the misclosure of the line if this computation is not considered. In this example, Stations F 137 and J 231 are approximately 120 km apart in the north-south direction. As can be seen by example, the convergence of the equipotential surfaces is very modest over long distances. Thus, it is only considered in high-precision surveys involving long north-south extents.

After applying the orthometric correction, the resultant misclosure in the leveling circuit can be adjusted using least squares. However, for the most precise surveys, the gravity values at the intermediate benchmarks must also be considered. Readers, who wish to learn more on this topic, should consult the references at the end of this chapter.

19.15 GEODETIC POSITION COMPUTATIONS

Geodetic position computations involve two basic types of calculations, the *direct* and the *inverse* problems. In the direct problem, given the latitude and longitude of station A and the geodetic length and azimuth of line AB, the latitude and longitude of station B are computed. In the inverse problem, the geodetic length of AB and its forward and back azimuths are calculated, given the latitudes and longitudes of stations A and B.

For long lines it is necessary to account for the ellipsoidal shape of the Earth in these calculations to maintain suitable accuracy. Many formulas are available for making direct and inverse calculations, some of which are simplified approximations that only apply for shorter lines. This book will present those developed by Vincenty (1975). The procedures presented in the following subsections have a stated accuracy of a few centimeters for lines up to 20,000 km in length. These computations are demonstrated in the Excel® spreadsheet *vincenty.xls*, which is available on the companion website for this book at http://www.pearsonhighered.com/ghilani.

19.15.1 Direct Geodetic Problem

In the direct problem, ϕ_1 and λ_1 represent the latitude and longitude, respectively, s the geodetic length from station 1 to station 2, and α_1 the forward azimuth from station 1 to station 2. The variables a, b, and f are the defining parameters of the ellipsoid as presented in Section 19.2. The unknowns in the problem are ϕ_2 and λ_2, the geodetic latitude and longitude of the sighted station, and α_2 the azimuth of the line from station 2 to station 1. Note that the observations used in this computation must be corrected to the ellipsoid using procedures outlined in Section 19.14.

The computational steps are as follows:[15]

1. $\tan U_1 = (1 - f) \tan \phi_1$
2. $\tan \sigma_1 = \tan U_1 / \cos \alpha_1$

[15]For the derivation of this formulation, and that of the inverse problem which follows, consult the publication by T. Vincenty cited in this chapter's bibliography.

3. $u = e' \cos \alpha$
4. $\sin \alpha = \cos U_1 \sin \alpha_1$
5. $A = 1 + \dfrac{u^2}{16{,}384}\{4096 + u^2[-768 + u^2(320 - 175u^2)]\}$
6. $B = \dfrac{u^2}{1024}\{256 + u^2[-128 + u^2(74 - 47u^2)]\}$
7. $2\sigma_m = 2\sigma_1 + \sigma$; where the first iteration uses $\sigma = \dfrac{s}{bA}$
8. $$\Delta\sigma = B \sin \sigma \left\{\cos(2\sigma_m) + \frac{B}{4}\left[\cos \sigma\,(-1 + 2\cos^2 \sigma_m) - \frac{B}{6}\cos(2\sigma_m)(-3 + 4\sin^2 \sigma)(-3 + 4\cos^2 \sigma_m)\right]\right\}$$
9. $\sigma = \dfrac{s}{bA} + \Delta\sigma$
10. Repeat steps 7, 8, and 9 until the $\Delta\sigma$ becomes negligible.
11. $\tan \phi_2 = \dfrac{\sin U_1 \cos \sigma + \cos U_1 \sin \sigma \cos \alpha_1}{(1 - f)\sqrt{\sin^2\alpha + (\sin U_1 \sin \sigma - \cos U_1 \cos \sigma \cos \alpha_1)^2}}$
12. $\tan \lambda = \dfrac{\sin \sigma \sin \alpha_1}{\cos U_1 \cos \sigma - \sin U_1 \sin \sigma \cos \alpha_1}$
13. $C = \dfrac{f}{16}\cos^2\alpha[4 + f(4 - 3\cos^2 \alpha)]$
14. $L = \lambda - (1 - C)f \sin \alpha \{\sigma + C \sin \sigma [\cos 2\sigma_m + C \cos \sigma(-1 + 2\cos^2 2\sigma_m)]\}$
15. $\lambda_2 = \lambda_1 + L$
16. $\tan \alpha_2 = \dfrac{\sin \alpha}{-\sin U_1 \sin \sigma + \cos U_1 \cos \sigma \cos \alpha_1}$

19.15.2 Inverse Geodetic Problem

In the inverse geodetic problem, ϕ_1, λ_1, ϕ_2, and λ_2 represent the latitude and longitude of the first and second stations, respectively. In the solution, the geodetic length, s, between the two points, and the forward and back azimuths of the line, α_1 and α_2, respectively, are determined. The variables a, b, and f again are the defining parameters of the ellipsoid as presented in Section 19.2.

Steps:

1. $L = \lambda = \lambda_2 - \lambda_1$
2. $\tan U_1 = (1 - f) \tan \phi_1$
3. $\tan U_2 = (1 - f) \tan \phi_2$
4. $\sin^2 \sigma = (\cos U_2 \sin \lambda)^2 + (\cos U_1 \sin U_2 - \sin U_1 \cos U_2 \cos \lambda)^2$

5. $\cos\sigma = \sin U_1 \sin U_2 + \cos U_1 \cos U_2 \cos\lambda$
6. $\sin\alpha = \cos U_1 \cos U_2 \sin\lambda/\sin\sigma$
7. $\cos 2\sigma_m = \cos\sigma - 2\sin U_1 \sin U_2/\cos^2\alpha$
8. $C = \dfrac{f}{16}\cos^2\alpha[4 + f(4 - 3\cos^2\alpha)]$
9. $\lambda = L - (1 - C)f\sin\alpha\{\sigma + C\sin\sigma[\cos 2\sigma_m$
   $+ C\cos\sigma(-1 + 2\cos^2 2\sigma_m)]\}$
10. Repeat steps 8 and 9 until changes in λ become negligible.
11. $s = bA(\sigma - \Delta\sigma)$ where $\Delta\sigma$ comes from the steps 12 to 15 below.
12. $u = e'\cos\alpha$
13. $A = 1 + \dfrac{u^2}{16{,}384}\{4096 + u^2[-768 + u^2(320 - 175u^2)]\}$
14. $B = \dfrac{u^2}{1024}\{256 + u^2[-128 + u^2(74 - 47u^2)]\}$
15. $\Delta\sigma = B\sin\sigma\left\{\cos(2\sigma_m) + \dfrac{1}{4}B\left[\cos\sigma(-1 + 2\cos^2\sigma_m)\right.\right.$
    $\left.\left. - \dfrac{1}{6}B\cos(2\sigma_m)(-3 + 4\sin^2\sigma)(-3 + 4\cos^2\sigma_m)\right]\right\}$
16. $\tan\alpha_1 = \dfrac{\cos U_2 \sin\lambda}{\cos U_1 \sin U_2 - \sin U_1 \cos U_2 \cos\lambda}$
17. $\tan\sigma_2 = \dfrac{\cos U_1 \sin\lambda}{-\sin U_1 \cos U_2 + \cos U_1 \sin U_2 \cos\lambda}$

The software WOLFPACK on the companion website for this book at http://www.pearsonhighered.com/ghilani can be used to do both of these computations. Figure 19.23 shows the data entry screen for the direct geodetic problem. A similar

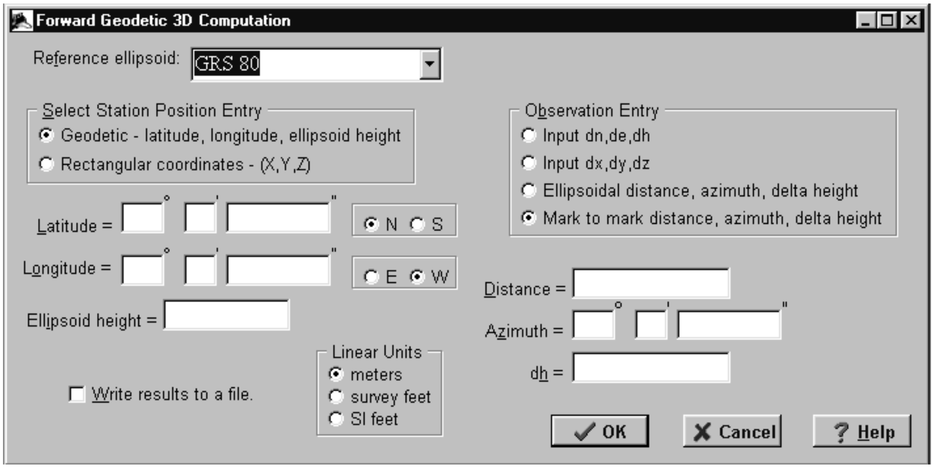


Figure 19.23 Data entry screen for forward computation from WOLFPACK.

data screen is available in WOLFPACK to compute the inverse geodetic problem. An Excel spreadsheet and Mathcad® worksheet that demonstrate these computations are also provided on the companion website for this book.

To simplify position calculations for long lines, and yet maintain geodetic accuracy, state plane coordinate systems have been developed. These are described in Chapter 20.

■ 19.16 THE LOCAL GEODETIC COORDINATE SYSTEM

GNSS surveys (see Chapters 13, 14, and 15) yield three-dimensional baseline components (ΔX, ΔY, ΔZ). These vector components are in the geocentric coordinate system (see Section 13.4.3). It is a common practice to transform these geocentric coordinate vector components into a *local geodetic system of easting* (Δe), *northing* (Δn), and *local up* (Δu). The two coordinate systems are illustrated in Figure 19.24, where XYZ represents the geocentric system and e, n, u is the local geodetic system. The local geodetic system is user oriented in that the e and n axes are in a local horizontal plane (u is coincident with the normal at the origin of the local coordinate system) and n is in the direction of local north.

To perform a transformation from the geocentric coordinate system to local geodetic, a set of three-dimensional rotation matrices must be employed. These rotation matrices are similar to their two-dimensional counterparts as shown in Equation (11.36). In the transformation process, rotations occur about each of the three coordinate axes. Letting the rotation angle around the X-axis be θ_1, the rotation angle around the Y-axis be θ_2, and the rotation angle around the Z-axis be θ_3, the three-dimensional rotation matrices are

$$R_X(\theta_1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & \sin\theta_1 \\ 0 & -\sin\theta_1 & \cos\theta_1 \end{bmatrix} \tag{19.44}$$

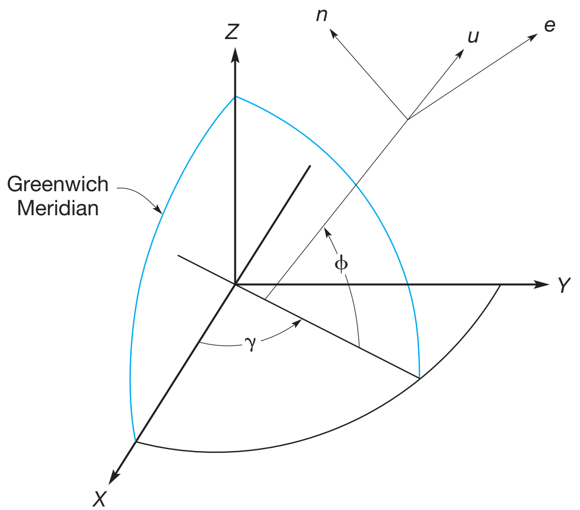


Figure 19.24 The relationship between the geocentric coordinate system and the local geodetic coordinate system.

$$R_Y(\theta_2) = \begin{bmatrix} \cos\theta_2 & 0 & -\sin\theta_2 \\ 0 & 1 & 0 \\ \sin\theta_2 & 0 & \cos\theta_2 \end{bmatrix} \tag{19.45}$$

$$R_Z(\theta_3) = \begin{bmatrix} \cos\theta_3 & \sin\theta_3 & 0 \\ -\sin\theta_3 & \cos\theta_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{19.46}$$

As shown in Figure 19.24, a rotation about the Z-axis by an amount of $\lambda - 270°$ must occur to align the X-axis with the local e-axis. Following this rotation, the once-rotated Z-axis is brought into coincidence with the u-axis by a rotation of $90° - \phi$ about the once-rotated X-axis.[16] The resultant expression is

$$\begin{bmatrix} \Delta e \\ \Delta n \\ \Delta u \end{bmatrix} = R_X(90° - \phi)R_Z(\lambda - 270°)\begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} \tag{19.47}$$

Performing the proper trigonometric substitutions, and rearranging the equations to place them in the standard order of (n, e, u), the final transformation equations are

$$\begin{aligned}
\begin{bmatrix} \Delta n \\ \Delta e \\ \Delta u \end{bmatrix} &= \begin{bmatrix} -\sin\phi & 0 & \cos\phi \\ 0 & 1 & 0 \\ \cos\phi & 0 & \sin\phi \end{bmatrix}\begin{bmatrix} \cos\lambda & \sin\lambda & 0 \\ -\sin\lambda & \cos\lambda & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} \\
&= \begin{bmatrix} -\sin\phi\cos\lambda & -\sin\phi\sin\lambda & \cos\phi \\ -\sin\lambda & \cos\lambda & 0 \\ \cos\phi\cos\lambda & \cos\phi\sin\lambda & \sin\phi \end{bmatrix}\begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} \\
&= R(\phi, \lambda)\begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix}
\end{aligned} \tag{19.48}$$

■ 19.17 THREE-DIMENSIONAL COORDINATE COMPUTATIONS

Sometimes it is advantageous to compute three-dimensional changes in a local geodetic coordinate system from reduced field observations. In Figure 19.25, the reduced observations of azimuth Az, ellipsoid distance s, and altitude

[16] A three-dimensional animation viewable in a web browser having the appropriate plug-in demonstrates the rotation of the local geodetic coordinate system axes into a geocentric coordinate system. The animation *LOC2GEO.wrl* is available on the companion website for this book at http://www.pearsonhighered.com/ghilani.

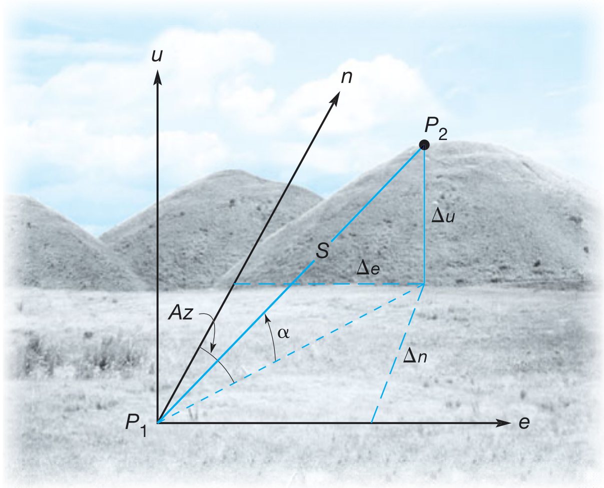


Figure 19.25 Reduction of observations in a local geodetic coordinate system.

angle α can be used to derive the changes in the local geodetic coordinate system as

$$\begin{aligned}\Delta n &= s\cos(\alpha)\cos(Az)\\ \Delta e &= s\cos(\alpha)\sin(Az)\\ \Delta u &= s\sin(\alpha)\end{aligned}\qquad(19.49)$$

Equations (19.49) can be modified by Equations (8.2) to incorporate a zenith angle z as

$$\begin{aligned}\Delta n &= s\sin(z)\cos(Az)\\ \Delta e &= s\sin(z)\sin(Az)\\ \Delta u &= s\cos(z)\end{aligned}\qquad(19.50)$$

From Figure 19.25, the following inverse relationships can be developed

$$\begin{aligned}s &= \sqrt{\Delta n^2 + \Delta e^2 + \Delta u^2}\\ Az &= \tan^{-1}\left(\frac{\Delta e}{\Delta n}\right)\\ \alpha &= \sin^{-1}\left(\frac{\Delta u}{\Delta s}\right) \text{ or } z = \cos^{-1}\left(\frac{\Delta u}{\Delta s}\right)\end{aligned}\qquad(19.51)$$

By combining Equation (19.48) with Equations (19.51), the reduced observations can be obtained directly from changes in geocentric coordinates as

$$\begin{aligned}s &= \sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2}\\ Az_1 &= \tan^{-1}\left(\frac{-\Delta X\sin\lambda_1 + \Delta Y\cos\lambda_1}{-\Delta X\sin\phi_1\cos\lambda_1 - \Delta Y\sin\phi_1\sin\lambda_1 + \Delta Z\cos\phi_1}\right)\\ z_1 &= \cos^{-1}\left(\frac{\Delta X\cos\phi_1\cos\lambda_1 + \Delta Y\cos\phi_1\sin\lambda_1 + \Delta Z\sin\phi_1}{s}\right)\end{aligned}\qquad(19.52)$$

It is important to note that Equation (19.52) uses the latitude and longitude of the observation station P_1 in Figure 19.25. These values can be computed from the geocentric coordinate values of P_1 based on Equation (13.3) through (13.7).

Example 19.11

The mark-to-mark distance from station Bill to station Red is 568.138 m, and the zenith angle and azimuth of this course are 92°14′25″ and 40°36′23″, respectively. If the geodetic coordinates of station Bill are 61°10′42.1058″ N latitude and 149°11′12.1033″ W longitude, what are the changes in the geocentric coordinate system?

Solution

Using Equation (19.50), the changes in the local coordinate system are

$$\begin{aligned}
\Delta n &= 568.138 \sin(92°14'25'') \cos(40°36'23'') = 431.000 \text{ m} \\
\Delta e &= 568.138 \sin(92°14'25'') \sin(40°36'23'') = 369.495 \text{ m} \\
\Delta u &= 568.138 \cos(92°14'25'') = -22.209 \text{ m}
\end{aligned}$$

To solve for the changes in geocentric coordinates in Equation (19.48), the inverse of $R(\phi, \lambda)$ must be determined. *Since $R(\phi, \lambda)$ is an orthogonal rotation matrix, its inverse is $R^T(\phi, \lambda)$.* Thus, the changes in geocentric coordinates for these observations are

$$\begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} = \begin{bmatrix} -\sin\phi\cos\lambda & -\sin\lambda & \cos\phi\cos\lambda \\ -\sin\phi\sin\lambda & \cos\lambda & \cos\phi\sin\lambda \\ \cos\phi & 0 & \sin\phi \end{bmatrix} \begin{bmatrix} 431.0000 \\ 369.4950 \\ -22.2087 \end{bmatrix}$$

$$= \begin{bmatrix} 522.773 \\ -118.425 \\ 188.321 \end{bmatrix}$$

Note that the changes in geodetic coordinates computed in Example 19.11 should closely agree with GNSS baseline vector observations that would be obtained if the baseline between Bill and Red was observed. Equations (19.52) can be used to determine the mark-to-mark distance, geodetic azimuth, and zenith angle of the observations that would be obtained if a survey were conducted at station Bill. An advantage of local coordinate systems is that unlike plane coordinate systems as presented in Chapter 20, all three dimensions are simultaneously computed using standard observations collected with a total station. Additionally, these local systems can be integrated with GNSS observations for simultaneous GNSS and terrestrial least-squares adjustments. However, it should always be remembered that the computed values would differ from the field-observed values if the appropriate corrections discussed in Section 19.14 were not applied.

19.18 SOFTWARE

In this chapter, an introductory discussion of geodetic computations was presented, datums were briefly discussed, and transformations between surveyed, geocentric and geodetic observations presented. WOLFPACK, which is available on the companion website for this book at http://www.pearsonhighered.com/ghilani, contains options to perform many of the computations presented in this chapter. Figure 19.23 shows the entry screen in WOLFPACK for the forward geodetic problem. Readers wishing to see these reductions in a higher-level programming language should explore the Mathcad® worksheets, which are on the companion website for this book also. Readers who wish to explore these topics in more detail should refer to the bibliography at the end of this chapter.

PROBLEMS

Asterisks (*) indicate problems that have partial answers given in Appendix G.

19.1 Define the *geoid* and *ellipsoid*.

19.2 What are the possible monumentation types for a control station with a quality code of *A*?

19.3 What is *nutation*?

19.4* What is the difference between the equatorial circumference of the Clarke 1866 ellipsoid and that of the GRS80 ellipsoid?

19.5 Determine the first and second eccentricities for the GRS80 ellipsoid.

19.6 Why shouldn't points defined in different versions of the NAD83 datum be mixed in a survey?

19.7 What are the radii in the meridian and prime vertical for a station with latitude 29°06′58.29740″ using the GRS80 ellipsoid?

19.8 For the station listed in Problem 19.7, what is the radius of the great circle at the station that is at an azimuth of 302°49′21″ using the GRS80 ellipsoid?

19.9* What are the radii in the meridian and prime vertical for a station with latitude 42°37′26.34584″ using the GRS80 ellipsoid?

19.10 For the station listed in Problem 19.9, what is the radius of the great circle at the station that is at an azimuth of 153°29′32″ using the GRS80 ellipsoid?

19.11* The orthometric height at Station Y 927 is 304.517 m and the geoidal undulation at that station is −31.893 m. What is its geodetic height?

19.12 The geodetic height at Station Z104 is 102.054 m. Its geoidal undulation is −22.101 m. What is its orthometric height?

19.13 The orthometric height of a particular benchmark is 87.95 ft. The geoidal height at the station is −30.66 m. Is the station above or below the ellipsoid? Draw a sketch depicting the geoid, ellipsoid, and benchmark.

19.14 The instantaneous position of the pole at the time of an azimuth observation is $x = -0.55''$ and $y = -0.83''$. The position of the station is (42°37′23.0823″ N, 128°56′01.0089″ W) and the observed azimuth of a line is 18°52′37″. What is the astronomic azimuth of the line corrected for polar motion?

19.15* The deflection of the vertical components ξ and η are $-2.85''$ and $-5.94''$, respectively. The observed zenith angle is $42°36'58.8''$. What is the geodetic zenith angle for the observations in Problem 19.14?

19.16 To within what tolerance should the elevations of two benchmarks 15 km apart be established if first-order class II standards were used to set them? What should it be if second-order class I standards were used?

19.17 Name the orders and classes of accuracy of both horizontal and vertical control surveys, and give their relative accuracy requirements.

19.18* Given the following information for stations *JG00050* and *KG0089*, what should be the leveled height difference between them?

Station	Height (m)	Gravity (mgal)
JG0050	474.442	979,911.9
KG0089	440.552	979,936.2

19.19 Similar to Problem 19.18 except that the station data for EY5664 and EY1587 is

Station	Height (m)	Gravity (mgal)
EY5664	2.960	979,811.6
EY1587	3.679	979,791.2

19.20 Similar to Problem 19.18 except that the station data for CV0178 and DQ0080 is

Station	Height (m)	Gravity (mgal)
CV0178	1028.652	979,279.4
DQ0080	1013.487	979,269.9

19.21* A slope distance of 2458.663 m is observed between stations *Gregg* and *Brian*, whose orthometric heights are 458.966 m and 566.302 m, respectively. The geoidal undulations are -25.66 and -25.06 m at *Gregg* and *Brian*, respectively. The height of the instrument at station *Gregg* at the time of the observation was 1.525 m and the height of the reflector at station *Brian* was 1.603 m. What are the geodetic and mark-to-mark distances for this observation? (Use an average radius for the Earth of 6,371,000 m for R_α.)

19.22 If the latitude of station *Gregg* in Problem 19.21 was $45°22'58.6430''$ and the azimuth of the line was $110°33'03.8''$, what are the geodetic and mark-to-mark distances for this observation? (Use the GRS80 ellipsoid.)

19.23 A slope distance of 6704.511 m is observed between two stations *A* and *B* whose geodetic heights are 916.963 and 928.578 m respectively. The height of the instrument at the time of the observation was 1.500 m and the height of the reflector was 1.825 m. The latitude of Station *A* is $33°08'36.2947''$, and the azimuth of *AB* is $202°28'21.9''$. What are the geodetic and mark-to-mark distances for this observation?

19.24 Discuss the advantages of combined networks for establishing control.

19.25* Compute the back azimuth of a line 5863 m long in the east-west direction at a mean latitude of $45°01'32.0654''$, whose forward azimuth is $88°16'33.2''$ from north. (Use an average radius for the Earth of 6,371,000 m.)

19.26 Compute the back azimuth of a line 6505 m long in the east-west direction at mean latitude of 38°52′02″, whose forward azimuth is 83°24′37.5″ from north. (Use an average radius for the Earth of 6,371,000 m.)

19.27 In Figure 19.14, azimuth of AB is 102°36′20″ and the angles to the right observed at B, C, D, E, and F are 132°01′054, 241°45′12″, 141°15′01″, 162°09′24″, and 202°33′19″, respectively. An astronomic observation yielded an azimuth of 82°24′03″ for line FG. The mean latitude of the traverse is 42°16′00″, and the total departure between points A and F was 24,986.26 ft. Compute the angular misclosure and the adjusted angles. (Assume the angles and distances have already been corrected to the ellipsoid.)

19.28 In Figure 19.20, slope distance S and vertical angles α and β were observed as 76,953.82 ft, +4°18′42″, and −4°26′28″, respectively. Ellipsoid height of point A is 1672.21 ft. What is length $A'B'$ on the ellipsoid? (Use an average radius for the Earth of 6,371,000 m.)

19.29 In Figure 19.19, slope distance S was observed as 19,875.28 m. The orthometric elevations of points A and B were 657.73 and 1805.54 m, respectively, and the geoidal undulation at both stations was −30.5 m. The instrument and reflector heights were both set at 1.50 m. Calculate geodetic distance $A'B'$. (Use an average radius for the Earth of 6,371,000 m.)

19.30 In Figure 19.20, slope distance S and zenith angle α at station A were observed as 2072.33 m and 82°17′18″, respectively, to station A. If the elevation of station A is 435.967 m and the geoidal undulations at stations A and B are both −28.04 m , what is ellipsoid length $A'B'$? (Use an average radius for the Earth of 6,371,000 m.)

19.31* Components of deflection of the vertical at an observing station of latitude 43°15′47.5864″ are $\xi = -6.87''$ and $\eta = -3.24''$. If the observed zenith angle on a course with an astronomic azimuth of 204°32′44″ is 85°56′07″, what are the azimuth and zenith angles corrected for deviation of the vertical?

19.32 At the same observation station as for Problem 19.31, the observed zenith angle on a course with an azimuth of 12°08′07″ is 80°15′54″, what are the azimuth and zenith angles corrected for deviation of the vertical?

19.33 Using the reduced azimuths of Problems 19.31 and 19.32, what is the reduced geodetic angle that is less than 180°?

19.34 What is a dynamic height of a point?

19.35 Compute the collimation correction factor C for the following field data, taken in accordance with the example and sketch in the field notes of Figure 19.18. With the instrument at station 1, high, middle, and low crosshair readings were 5.512, 5.401, and 5.290 ft on station A and 4.978, 3.728, and 2.476 ft on station B. With the instrument at station 2, high, middle, and low readings were 7.211, 6.053, and 4.894 ft on station A and 4.561, 4.358, and 4.155 ft on station B.

19.36 A leveling instrument having a collimation factor of −0.0007 m/m of interval was used to run a section of three-wire differential levels from BM A to BM B. Sums of backsights and foresights for the section were 1320.892 m and 1333.695 m, respectively. Backsight stadia intervals totaled 1557.48, while the sum of foresight intervals was 805.67. What is the corrected elevation difference from BM A to BM B?

19.37 The relative error of the difference in elevation between two benchmarks directly connected in a level circuit and located 90 km apart is ±0.009 m. What order and class of leveling does this represent?

19.38 Similar to Problem 19.37, except the relative error is ±0.025 ft for benchmarks located 35 km apart.

19.39 The baseline components of a GPS baseline vector observed at a station A in meters are (1204.869, 798.046, −666.157). The geodetic coordinates of the first base station are 44°27′36.0894″ N latitude and 74°44′09.4895″ W longitude. What are the changes in the local geodetic coordinate system of (Δn, Δe, Δu)?

19.40 In Problem 19.39, what are the slant distance, zenith angle, and azimuth for the baseline vector?

19.41 If the slant distance between two stations is 843.273 m, the zenith angle between them is 85°58′44″ and the azimuth of the line is 312°23′59″, what are the changes in the local geodetic coordinates?

19.42 Create a computational program to solve Problem 19.40.

19.43 Create a computational program to solve Problem 19.41.

BIBLIOGRAPHY

Carlson, E., D. Doyle, and D. Smith. 2009. "Development of Comprehensive Geodetic Vertical Datums for the United States Pacific Territories of American Somoa, Guam, and the Northern Marianas." *Surveying and Land Information Science* 69 (No. 1): 5.

Diemirkesen, A. C. and N. W. J. Hazelton. 2009. "Fundamental Principles of Deformation Measurement." *Surveying and Land Information Science* 69 (No. 2): 89.

Ghilani, C. D. 2007. "Animating Three-D Concepts in Geodesy." *Surveying Land Information Science* 67 (No. 4): 205.

Heiskanen, W.A. and H. Moritz. 1967. *Physical Geodesy*. San Francisco, CA: W. H. Freeman and Co.

Meyer, T., et. al. 2006. "What Does Height Really Mean? Part III: Height Systems." *Surveying and Land Information Science* 66 (No. 2): 149.

______ 2006. "What Does Height Really Mean? Part IV: GPS Heighting." *Surveying and Land Information Science* 66 (No. 3): 165.

Reilly, J. P. 2007. "Gravity Anomalies." *Point of Beginning* 33 (No. 3): 68.

______. 2007. "Physical Geodesy 101." *Point of Beginning* 32 (No. 11): 54.

______. 2007. "Physical Geodesy 201." *Point of Beginning* 33 (No. 1): 60.

Smith, D. A. and D. R. Doyle. 2006. "The Future Role of Geodetic Datums in Control Surveying in the United States." *Surveying and Land Information Science* 66 (No. 2): 101.

This page intentionally left blank

20 State Plane Coordinates and Other Map Projections

■ 20.1 INTRODUCTION

Most surveys of small areas are based on the assumption that the Earth's surface is a plane. However, as explained in Chapter 19, for large-area surveys it is necessary to consider Earth curvature. Unfortunately, the calculations necessary to determine geodetic positions from survey observations and get distances and azimuths from them are lengthy, and practicing surveyors often are not familiar with these procedures. Clearly, a system for specifying positions of geodetic stations using plane rectangular coordinates is desirable, since it allows computations to be made using simple coordinate geometry formulas, such as those presented in Chapter 11. The National Geodetic Survey (NGS) met this need by developing a state plane coordinate system for each state.

A state plane coordinate system provides a common datum of reference for horizontal control of all surveys in a large area, just as the geoid furnishes a single datum for vertical control. It eliminates having individual surveys based on different assumed coordinates, unrelated to those used in other adjacent work. State plane coordinates are available for all control points in the National Spatial Reference System and for many other control points as well. They are widely used as the reference points for initiating surveys of all types, including those for highway construction projects, property boundary delineation, and photogrammetric mapping.

There are many examples illustrating the value of state plane coordinates. They make it possible for extensive surveys on highway projects to begin on one control station and close on another that is tied to the same coordinate system.

On boundary surveys, if a parcel's corners are referenced to the state plane coordinate system, their locations are basically indestructible. The iron pipes, posts, or other monuments marking their positions may disappear, but their original locations can be restored from surveys initiated at other nearby monuments referenced to the state plane coordinate system. For this reason, some states require that state plane coordinates be included on all new subdivisions. State plane coordinates are highly recommended as the reference frame for entering maps and other data into Land and Geographic Information Systems. This allows all data to be referenced to a common system, and thus can be accurately registered and overlaid for analysis purposes.

20.2 PROJECTIONS USED IN STATE PLANE COORDINATE SYSTEMS

To convert geodetic positions of a portion of the Earth's surface to plane rectangular coordinates, points are projected mathematically from the ellipsoid to some imaginary *developable surface*—a surface that can conceptually be developed or "unrolled and laid out flat" without distortion of shape or size. A rectangular grid can be superimposed on the *developed* plane surface and the positions of points in the plane specified with respect to X and Y grid axes. A plane grid developed using this mathematical process is called a *map projection*.

There are several types of map projections, with the oldest known projections dating back to the times of the ancient Greeks. Today, two of the most commonly used mapping projections are the *Lambert conformal conic* and the *Transverse Mercator* projections. Johann Heinrich Lambert initially developed both of these projections. The Transverse Mercator projection was further developed and redefined by Carl Friedrich Gauss and L. Krüger, and thus is also known as the *Gauss-Krüger* projection. These two projections are used in state plane coordinate systems. The Lambert conformal conic projection utilizes an imaginary cone as its developable surface and the Transverse Mercator employs a fictitious cylinder. These are shown in Figure 20.1(a) and (b). The cone and cylinder are secant to the ellipsoid in the state plane coordinate systems, that is, they intersect the ellipsoid along two arcs AB and CD as shown. With this placement, the conical and cylindrical surfaces conform better to the ellipsoid over larger areas than they would if placed tangent.

Figure 20.1(c) and (d) illustrate plane surfaces "developed" from the cone and cylinder. Here, points are projected mathematically from the ellipsoid to the surface of the imaginary cone or cylinder based on their geodetic latitudes and longitudes. For discussion purposes, this may be considered a radial projection from the Earth's center. Figure 20.2 illustrates this process diagrammatically and displays the relationship between the length of a line on the ellipsoid and its extent when projected onto the surface of either a cone or a cylinder. Note that distance $a'b'$ on the projection surface is greater than ab on the ellipsoid, and similarly $g'h'$ is longer than gh. From this observation it is clear that map projection scale is larger than true ellipsoid scale where the cone or cylinder is outside the ellipsoid. Conversely distance $d'e'$ on the projection is shorter than de on the ellipsoid, and thus map scale is smaller than true ellipsoid scale when the projection surface is

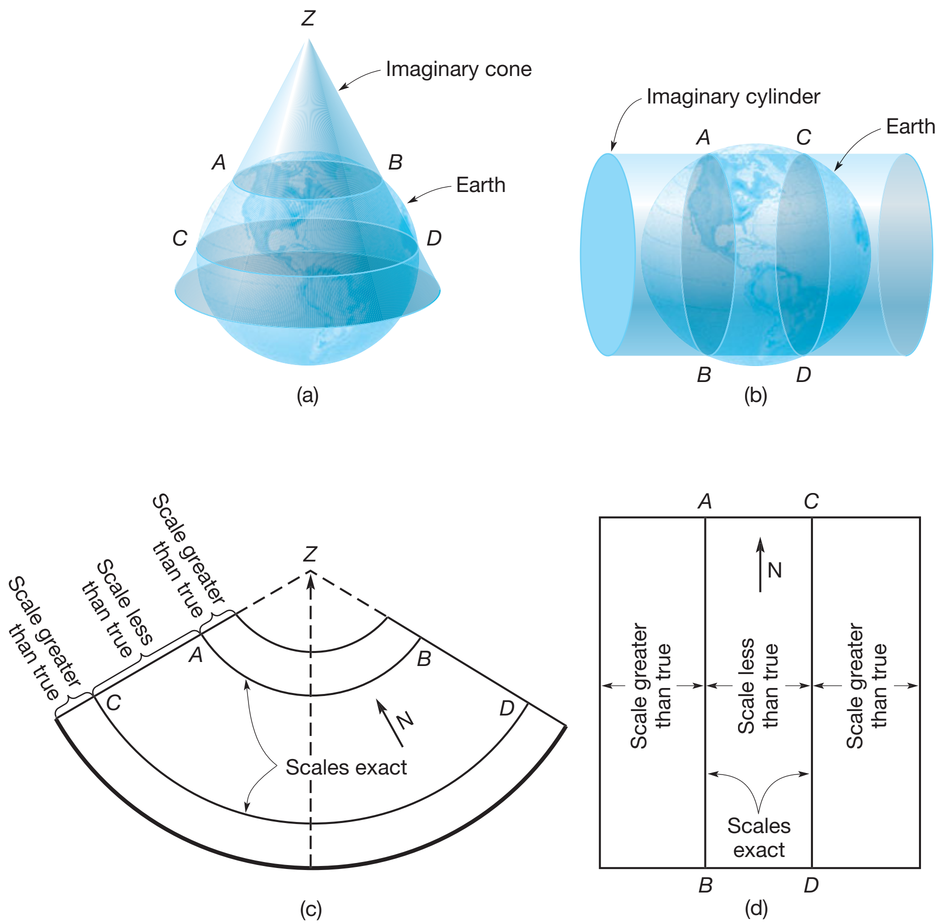


Figure 20.1 Surfaces used in state plane coordinate systems.

inside the ellipsoid. Points *c* and *f* occur at the intersection of the projection and ellipsoid surfaces, and therefore map scale equals true ellipsoid scale along the lines of intersection. These relationships of map scale to true ellipsoid scale for various positions on the two projections are indicated in Figure 20.1(c) and (d). As will be discussed later, these length differences are accounted for by means of a *scale factor*.

From the foregoing discussion it should be clear that points couldn't be projected from the ellipsoid to developable surfaces without introducing distortions in the lengths of lines or the shapes of areas. However, these distortions are held to a minimum by selected placement of the cone or cylinder secant to the ellipsoid, by choosing a *conformal projection* (one that preserves true angular relationships around points in a small region), and also by limiting the zone size or extent of coverage on the Earth's surface for any particular projection. If the width of zones is held to a maximum of 158 mi (254 km), and if two thirds of this zone width is between the secant lines, distortions (differences in line lengths on the two surfaces) are kept to 1 part in 10,000 or less. The NGS intended this accuracy

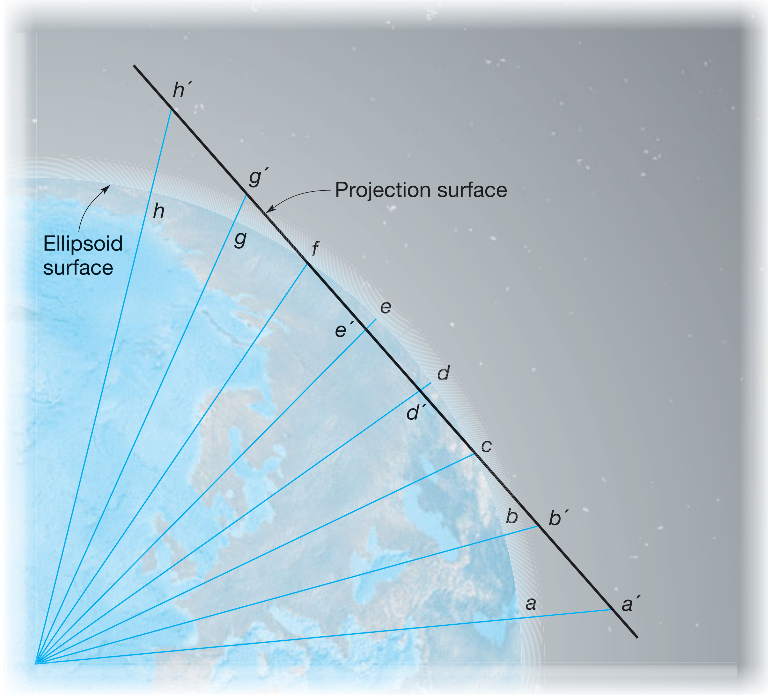


Figure 20.2 Method of projection.

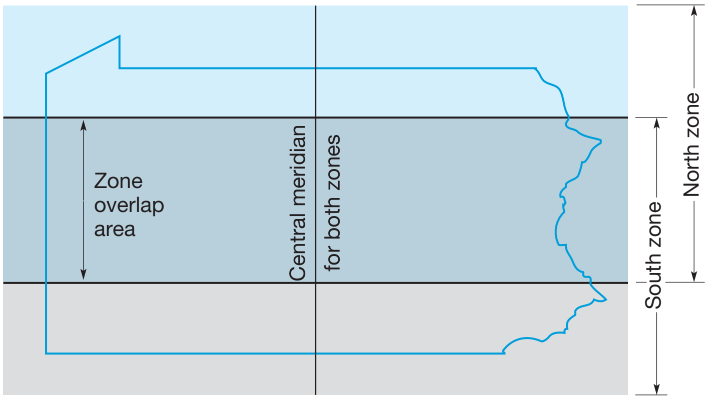


Figure 20.3 Coverages and overlap of zones in Pennsylvania's Lambert conformal conic state plane coordinate system.

in its development of the state plane coordinate systems. For small states such as Connecticut and Delaware, one state plane coordinate zone is sufficient to cover the entire state. Larger states require several zones to encompass them; for example, Alaska has 10, California 6, and Texas 5. Where multiple zones are needed to cover a state, adjacent zones overlap each other. As explained in Section 20.10, this is important when lengthy surveys extend from one zone to another. Figures 20.3 and 20.4 show the coverage of zones in Pennsylvania and Indiana, respectively. Both states have two zones, with Pennsylvania using the Lambert conformal conic projection and Indiana the Transverse Mercator projection.

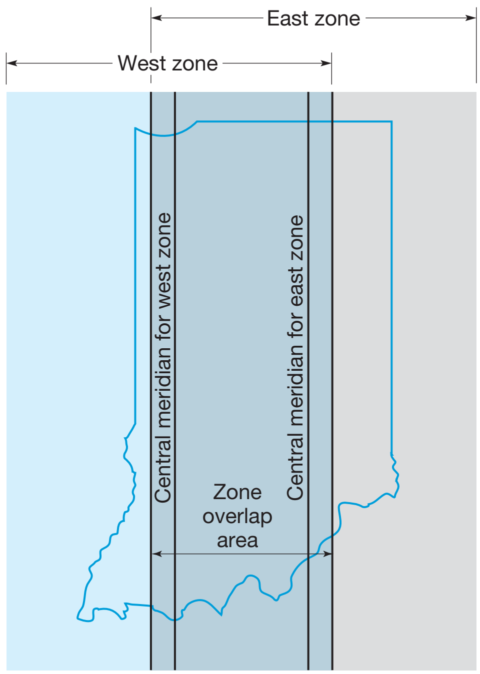


Figure 20.4 Coverages and overlap of zones in Indiana's Transverse Mercator state plane coordinate system.

■ 20.3 LAMBERT CONFORMAL CONIC PROJECTION

The Lambert conformal conic projection, as its name implies, is a projection onto the surface of an imaginary cone. The term *conformal*, as noted earlier, means that true angular relationships are retained around all points in small regions. Scale on a Lambert projection varies from north to south but not from east to west, as shown in Figure 20.1(c). Zone widths in the projection are therefore limited north-south, but not east-west. The Lambert projection is thus ideal for mapping states that are narrow north-south, but which extend long distances in an east-west direction—for example, Kentucky, Montana, Pennsylvania, and Tennessee.

Figure 20.5 shows the portion of the developed cone of a Lambert projection covering an area of interest. In the Lambert projection, the cone intersects the ellipsoid along two parallels of latitude, called *standard parallels*, at one sixth of the zone width from the north and south zone limits. All meridians are straight lines converging at Z, the apex of the cone. An example is ZM, which is the *central meridian*. All parallels of latitude are the arcs of concentric circles having centers at the apex. The projection is located in a zone in an east-west direction by assigning the central meridian a longitudinal value that is near the middle of the area to be covered. The direction of the central meridian on the projection establishes grid north. All lines parallel with the central meridian point in the direction of grid north. Therefore, except at the central meridian, directions of "true" and "grid" north do not coincide because true meridians converge. As shown in Figure 20.5, the *easting* (E) and *northing* (N) coordinates of points are

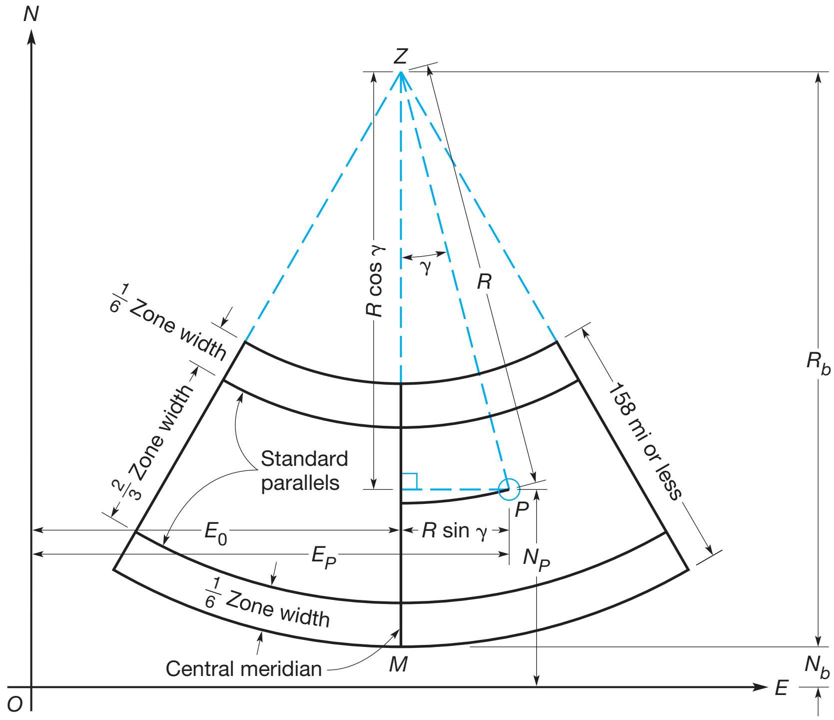


Figure 20.5 The Lambert conformal conic projection (SPCS83 symbology).

measured perpendicular and parallel to the central meridian, respectively, from a reference *E-N* axis system that is offset to the west and south.

20.4 TRANSVERSE MERCATOR PROJECTION

The Transverse Mercator projection is also a conformal projection, but is based on an imaginary secant cylinder as its developable surface. As illustrated in Figure 20.1(d), scale in the Transverse Mercator projection varies from east to west, but not from north to south. Thus this projection is used for states like Illinois, Indiana, and New Jersey, which are narrow east-west and longer north-south.

In developing the Transverse Mercator projection, the axis of the imaginary cylinder is placed in the plane of the Earth's equator. The cylinder cuts the spheroid along two small circles equidistant from the central meridian. On the developed plane surface (Figure 20.6) all parallels of latitude, and all meridians except for the central meridian are curves (shown in light broken lines). As with the Lambert conformal conic projection, the central meridian establishes the direction of grid north, and the zone is centered over an area of interest by assigning the central meridian a longitude value that applies approximately at the center of the region to be mapped. Also, the *E* and *N* coordinates of points are measured perpendicular to and parallel with the central meridian, respectively, from an *E-N* axis system offset to the west and south.

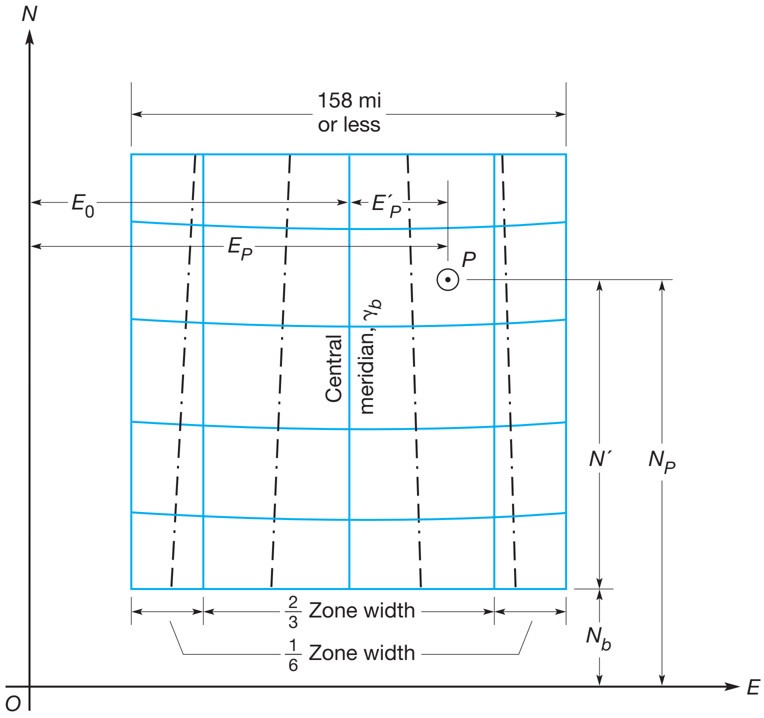


Figure 20.6
The Transverse Mercator projection.

20.5 STATE PLANE COORDINATES IN NAD27 AND NAD83

The first state plane coordinate system was developed by the NGS in 1933 for the State of North Carolina. Systems for all other states followed shortly thereafter. As noted earlier, state plane coordinates of points are computed from geodetic latitudes and geodetic longitudes. Section 19.4 and Figure 19.3 define and illustrate these two terms and explain that geodetic latitudes and longitudes are defined with respect to a reference ellipsoid and associated datum. From 1927 until the inception of NAD83, the reference datum used in the United States was NAD27, based on the Clarke 1866 ellipsoid. All original state plane coordinates were therefore developed by the NGS in accordance with that ellipsoid and datum. This system is referred to as the *State Plane Coordinate System of 1927* (SPCS27).

As noted in Chapter 19, NAD83 employs a different set of defining parameters than NAD27 and it uses a reference surface of different dimensions, the GRS80 ellipsoid. Thus, the latitudes and longitudes of points in NAD83 are somewhat different from their values in NAD27. Because of these changes, the constants and variables that define the state plane coordinate systems also changed. Thus following completion of NAD83, it was necessary to develop a new state plane coordinate system. This has been completed by the NGS and is called the *State Plane Coordinate System of 1983* (SPCS83).

In the SPCS83 most states retained the same projections they had used in SPCS27, with basically the same central meridian positioning. There were changes, however; some major ones being (1) Montana reduced its number of zones from three to one; and (2) Nebraska and South Carolina reduced their number of zones from two to one. In SPCS83, 29 states employ the Lambert conformal conic projection; 18 use the Transverse Mercator; and Alaska, Florida, and New York utilize both. In addition to using eight Transverse Mercator zones for its mainland and a Lambert conformal conic for the Aleutian Islands, Alaska also employs an *oblique Mercator* projection for the southeast part of the state (see Section 20.13).

Although the same fundamental approach was used to develop both the SPCS27 and the SPCS83, as noted above, the parameters that define the two systems are different. Accordingly, different symbols are used, and the equations for computing coordinates in SPCS83 have been changed. Thus, slightly different procedures are used in making computations in the two systems. For those wishing to review the procedures used in SPCS27, please refer to an earlier edition of this book.

20.6 COMPUTING SPCS83 COORDINATES IN THE LAMBERT CONFORMAL CONIC SYSTEM

A procedure for computing E and N coordinates of points from their geodetic latitudes and longitudes in the Lambert conformal conic system are illustrated in Figure 20.5. This process of converting from geodetic to plane coordinates is called the *direct problem*. The fundamentals described here apply to both SPCS27 and SPCS83, although the symbology used in Figure 20.5 and this section is applicable only to SPCS83. Computation in the reverse manner can also be performed, that is, calculating geodetic latitude and longitude from a given set of E, N state plane coordinates. This reverse form of calculation is called the *inverse problem*.

Most state plane coordinate computations are performed using software. However, the use of tables is often simpler when handheld calculators are used. For these reasons both methods of computations will be presented.

20.6.1 Zone Constants

A zone is defined in the Lambert conformal conic map projection by the selection of four sets of parameters. They are the defining ellipsoidal parameters a and f, grid origin (ϕ_0, λ_0), latitudes of the northern and southern standard parallels ϕ_N and ϕ_S, and false easting and northing (E_0, N_b). From these defining parameters a set of zone constants are mathematically defined that are used in both the direct and inverse problems. Common functions used in the definition of the map projection are

$$W(\phi) = \sqrt{1 - e^2 \sin^2 \phi} \tag{20.1}$$

$$M(\phi) = \frac{\cos \phi}{W(\phi)} \tag{20.2}$$

$$T(\phi) = \sqrt{\left(\frac{1 - \sin \phi}{1 + \sin \phi}\right)\left(\frac{1 + e \sin \phi}{1 - e \sin \phi}\right)^e} \tag{20.3}$$

where e is the eccentricity of the ellipse as defined by Equation (19.2c). The defining zone constants for a Lambert conformal conic map projection are

$$w_1 = W(\phi_S) \quad (20.4)$$

$$w_2 = W(\phi_N) \quad (20.5)$$

$$m_1 = M(\phi_S) \quad (20.6)$$

$$m_2 = M(\phi_N) \quad (20.7)$$

$$t_0 = T(\phi_0) \quad (20.8)$$

$$t_1 = T(\phi_S) \quad (20.9)$$

$$t_2 = T(\phi_N) \quad (20.10)$$

$$n = \sin \phi_0 = \frac{\ln(m_1) - \ln(m_2)}{\ln(t_1) - \ln(t_2)} \quad (20.11)$$

$$F = \frac{m_1}{n \cdot t_1^n} \quad (20.12)$$

$$R_b = aFt_0^n \quad (20.13)$$

20.6.2 The Direct Problem

In Figure 20.5, line ZM is the central meridian of the projection, the N axis (line ON) is parallel to ZM, and point O is the origin of the rectangular coordinate system. A constant E_0 is adopted to offset the N grid axis from the central meridian and make E coordinates of all points positive. Similarly, a constant N_b can be adopted to offset the E grid axis from the southern edge of the projection. The N coordinate of the cone's apex is the constant $(R_b + N_b)$, the numerical values of these terms being such that all N coordinates are positive.

The coordinates E and N of point P, whose geodetic latitude ϕ_P, and geodetic longitude λ_P, are known, are to be determined. Line ZP represents a portion of the meridian through point P with its length designated as R. Angle γ between the central meridian and meridian ZP represents the amount of convergence between these two meridians. In SPCS83 it is termed the *convergence angle*. In SPCS27 it was called the *mapping angle*.

From Figure 20.5, the following equations of the direct problem can be solved for the *easting* (E) and *northing* (N) coordinates of point P:

$$\begin{aligned} E_p &= R \sin \gamma + E_0 \\ N_P &= R_b - R \cos \gamma + N_b \end{aligned} \quad (20.14)$$

To solve Equations (20.14), values for E_0, N_b, R_b, R, and γ must be known. The quantities E_0 and N_b are constants for any zone. In many states, the SPCS83 value for E_0 has been arbitrarily assigned a value of 600,000 m, and N_b assigned a value of 0.000 m; R_b is a computed zone constant defined by Equation (20.13). Values for R and γ also depend on the ellipsoid used and vary with changing

TABLE 20.1 EXCERPT FROM THE PENNSYLVANIA NORTH ZONE TABLES

Zone constants: N_b = 0.000 m E_0 = 600,000.000 m l_{CM} = 778,459
R_b = 7,379,348.3668 m sin ϕ_0 = 0.661539733812

Latitude (ϕ)	R (m)	Tab. Diff.	k
41°10′	7268294.836	30.84819	0.99996637
41°11′	7266443.945	30.84824	0.99996514
41°12′	7264593.050	30.84830	0.99996400
41°13′	7262742.152	30.84836	0.99996295
41°14′	7260891.251	30.84842	0.99996198
41°15′	7259040.346	30.84848	0.99996109
41°16′	7257189.437	30.84855	0.99996029
41°17′	7255338.524	30.84862	0.99995957
41°18′	**7253487.607**	**30.84869**	**0.99995893**
41°19′	7251636.685	30.84876	0.99995838

locations of points in the zone; R changes with latitude and γ with longitude. The convergence angle γ can be computed as

$$\gamma = (\lambda_0 - \lambda)n \tag{20.15}$$

$$t = T(\phi_P) \tag{20.16}$$

$$R = aFt^n \tag{20.17}$$

where Equation (20.15) is adjusted for western longitudes. On the companion website for this book at http://www.pearsonhighered.com/ghilani is an Excel® spreadsheet *map_projections.xls* and Mathcad® worksheet *Lamber.xmcd* that demonstrate these computations.

To aid in hand solutions of Equations (20.14), the NGS has computed and published individual SPCS83 booklets of projection tables for each state. These give the constants for each zone and tabulate R and scale factor k values versus latitude. Thus given the geodetic latitude of any point, R for that point can be interpolated from the tables for use in Equations (20.14). Table 20.1 shows an excerpt from the north zone of the Pennsylvania tables. The column labeled *Tab. Diff.* provides the change in radius R per second of latitude. The use of this table is demonstrated in Example 20.1. The equations are solved using a calculator and the computational procedure is called the *tabular method*.

Example 20.1

Using values in Table 20.1, what are the coordinates E (easting) and N (northing) for station "Hayfield NE" that lies in Pennsylvania's north Lambert conformal conic zone? The station's geodetic latitude is 41°18′20.25410″ N, and its geodetic longitude is 76°00′57.00239″ W.

Solution

Step 1: Determine the radius to Hayfield NE.
The tabular difference (Tab. Diff. column) listed in Table 20.1 for latitude of 41°18′ is 30.84869. Thus, the change in the radius ΔR from 41°18′ to 41°18′20.25410″ is

$$\Delta R = 20.25410'' \; (30.84869) = 624.8125 \text{ m}$$

As latitude increases, the radius decreases, and thus ΔR must be subtracted from the tabulated R of 7,253,487.607 m. Thus the radius to Hayfield NE is

$$R = 7{,}253{,}487.607 \; - \; 624.8125 = 7{,}252{,}862.794 \text{ m}$$

The equivalent linear interpolation formula for the radius is

$$\begin{aligned} R &= 7{,}253{,}487.607 + (7{,}251{,}636.685 - 7{,}253{,}487.607)\; 20.25410''/60'' \\ &= 7{,}252{,}862.794 \text{ m} \end{aligned}$$

Step 2: Compute the convergence angle, γ, using Equation (20.15) as

$$\gamma = (77°45' - 76°00'57.00239'') \times 0.661539733812 = 1°08'49.991''$$

Step 3: Solve Equations (20.14) as

$$\begin{aligned} E &= 7{,}252{,}862.795 \sin 1°08'49.991'' + 600{,}000.000 \\ &= 745{,}212.637 \text{ m} \\ N &= 7{,}379{,}348.3668 - 7{,}252{,}862.795 \cos 1°08'49.991'' + 0.000 \\ &= 127{,}939.400 \text{ m} \end{aligned}$$

20.6.3 The Inverse Problem

The inverse problem in state plane coordinate computations is the determination of the geodetic latitude and geodetic longitude of a station based on its state plane coordinates and zone. The inverse equations also utilize the zone constants computed in Section 20.6.1 or given in Table 20.1. The remaining equations for accomplishing this can be derived from Equations (20.14) as

$$\begin{aligned} N' &= R_b - (N - N_b) \\ E' &= E - E_0 \\ \gamma &= \tan^{-1}\left[\frac{E'}{N'}\right] \\ R &= \frac{N'}{\cos \gamma} = \sqrt{E'^2 + N'^2} \\ \lambda &= \lambda_0 + \frac{\gamma}{n} \end{aligned} \tag{20.18}$$

The solution for the latitude of the station is iterative. This process is

$$t = \left(\frac{R}{aF}\right)^{\frac{1}{n}} \tag{20.19}$$

$$\chi = \frac{\pi}{2} - 2\tan^{-1}(t) \tag{20.20}$$

$$\phi_P = \frac{\pi}{2} - 2\tan^{-1}\left[t\left(\frac{1 - e\sin\phi_p}{1 + e\sin\phi_p}\right)^{e/2}\right] \tag{20.21}$$

Iterate Equation (20.21) starting with ϕ_P equal to χ on the first iteration and continuing until changes in ϕ_P are negligible. This procedure is demonstrated in the previously mentioned Excel® spreadsheet and Mathcad® worksheet on the companion website for this book.

When using the tables, the latitude ϕ_P of the station can be interpolated, versus R, from the tables. The procedure is demonstrated in Example 20.2.

Example 20.2

What are the geodetic latitude and geodetic longitude for station Hayfield NE given the following SPCS83 coordinates? (The station lies in the north zone of Pennsylvania's Lambert conformal conic projection.)

$$\text{Easting} = 745{,}212.637\text{ m}$$
$$\text{Northing} = 127{,}939.400\text{ m}$$

Solution

Using Table 20.1 and Equations (20.18)

$$E' = 745{,}212.637 - 600{,}000.000 = 145{,}212.637\text{ m}$$

$$N' = 7{,}379{,}348.3668 - (127{,}939.400 - 0.000) = 7{,}251{,}408.9668\text{ m}$$

$$\gamma = \tan^{-1}\left(\frac{145{,}212.637}{7{,}251{,}408.9668}\right) = 1°08'49.99''$$

$$R = \sqrt{145{,}212.637^2 + 7{,}251{,}408.9668^2} = 7{,}252{,}862.7943$$

$$\lambda = 77°45' - \frac{1°08'49.99''}{0.661539733812} = 76°00'57.00239''$$

Now the latitude of the station can be interpolated from the values in Table 20.1. As can be seen in the table, the computed radius R is between 41°18′ and 41°19′. To determine the number of arc seconds to be added to 41°18′, the difference between the tabulated radius for 41°18′ of 7,253,487.607 and the computed radius R for the station is evaluated and divided by the tabulated difference of 30.84869. That is

$$\Delta\phi'' = \frac{7{,}253{,}487.607 - 7{,}252{,}862.7943}{30.84869} = 20.25411''$$

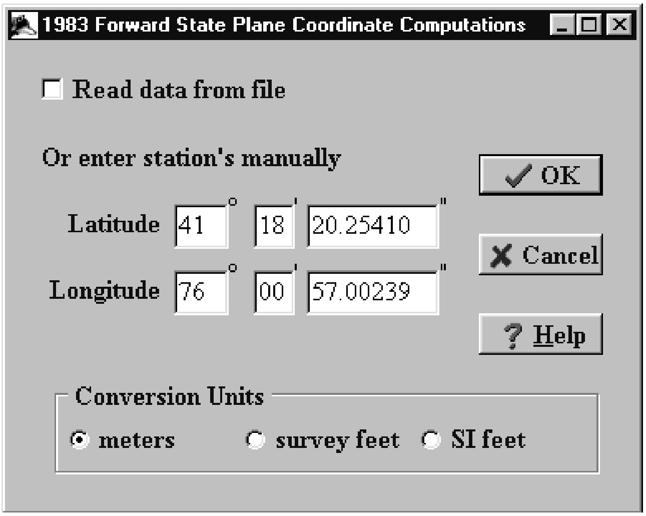


Figure 20.7 Data entry screen for direct SPCS83 computations of Example 20.1.

Thus, the latitude and longitude of the station are computed as 41°18′20.25411″ and 76°00′57.00239″, respectively. Rounding errors in both the forward and inverse problems caused the slight difference of 0.00001″ in the computed and given latitude from Example 20.1.

To aid in the solution of state plane coordinate computations, the NGS has published a booklet[1] entitled "State Plane Coordinate System of 1983" (Stem, 1989) that gives the zone constants and formulas for every zone in the United States and its territories. These constants are repeated in the Excel® spreadsheet on the CD that accompanies this book. The computer program WOLFPACK on the companion website for this book, which is at http://www.pearsonhighered.com/ghilani, contains a state plane coordinate option under its coordinate computation menu. The forward data entry screen from WOLFPACK for Example 20.1 is shown in Figure 20.7. Of course the state, and zone within that state, must also be specified.

It should be emphasized that computation of coordinates in the SPCS83 should only be done using points whose geodetic positions are given in NAD83.

20.7 COMPUTING SPCS83 COORDINATES IN THE TRANSVERSE MERCATOR SYSTEM

The NGS has also published SPCS83 Transverse Mercator state plane coordinate zones for those states that use this projection. All necessary constants and variables are computed and tabulated, and instructions together with sample problems are given to illustrate the computational procedures.

[1]Individual booklets for each state for making SPCS83 tabular solutions and NOAA Manual NOS NGS 5 can be obtained from the National Geodetic Information Center, NOAA, National Geodetic Survey, N/CG174, SSMC3 Station 09535, 1315 East West Highway, Silver Spring, MD 20910, telephone (301) 713-3242. Similar products are also available for SPCS27.

20.7.1 Zone Constants

A zone is defined in the Transverse Mercator map projection by the selection of four sets of parameters. They are the defining ellipsoidal parameters a and f, grid origin (ϕ_0, λ_0), scale factor at the central meridian λ_0, and false easting and northing (E_0, N_b). From these defining parameters a set of zone constants are defined mathematically, which are used in both the direct and inverse problems. Common functions used in the definition of the map projection are

$$C(\phi) = e'^2 \cos^2\phi \qquad T(\phi) = \tan^2\phi$$

$$M(\phi) = a\left[\begin{aligned}&\left(1 - \frac{e^2}{4} - \frac{3e^4}{64} - \frac{5e^6}{256}\right)\phi - \left(\frac{3e^2}{8} + \frac{3e^4}{32} + \frac{45e^6}{1024}\right)\sin 2\phi \\ &+ \left(\frac{15e^4}{256} + \frac{45e^6}{1024}\right)\sin 4\phi - \left(\frac{35e^6}{3072}\right)\sin 6\phi\end{aligned}\right] \tag{20.22}$$

where e and e' are the first and second eccentricities of the ellipse as defined by Equations (19.2a) and (19.2b), respectively.

20.7.2 The Direct Problem

The computations for the direct problem using the map projection formulas are

$$m_0 = \mathrm{M}(\phi_0) \tag{20.23}$$

$$m = \mathrm{M}(\phi) \tag{20.24}$$

$$t = T(\phi) \tag{20.25}$$

$$c = \mathrm{C}(\phi) \tag{20.26}$$

$$A = (\lambda_0 - \lambda)\cos\phi \tag{20.27}$$

where Equation (20.27) has been adjusted for western longitudes.

$$E = k_0 R_N\left[A + (1 - t + c)\frac{A^3}{6} + (5 - 18t + t^2 + 72c - 58e'^2)\frac{A^5}{120}\right] + E_0 \tag{20.28}$$

$$N = k_0\left\{m - M_0 + R_N \tan\phi\left[\begin{aligned}&\frac{A^2}{2} + (5 - t + 9c + 4c^2)\frac{A^4}{24} \\ &+ (61 - 58t + t^2 + 600c - 330e'^2)\frac{A^6}{720}\end{aligned}\right]\right\} + N_b \tag{20.29}$$

where R_N is the radius in the prime vertical for the latitude ϕ as defined by Equation (19.4). The convergence angle γ is computed as

$$\begin{gathered} c_2 = \frac{1 + 3c + 2c^2}{3} \qquad c_3 = \frac{2 - \tan^2\phi}{15} \\ \gamma = A\tan\phi[1 + A^2(c_2 + c_3A^2)] \end{gathered} \tag{20.30}$$

TABLE 20.2 EXCERPT FROM THE TRANSVERSE MERCATOR PROJECTION TABLES FOR NEW JERSEY FOR THE DIRECT PROBLEM

Zone constants: $E_0 = 150{,}000$ m $\phi_b = 38°50'$ $\lambda_0 = 74°30'$ $N_b = 0.000$ m $k_0 = 0.9999$

Latitude	(I) (IV)	Difference 1″	(II) (V)	Difference 1″	(III) (VI)
39°00′	18500.4650	30.834594	3670.4645	0.007613	1.902
	240604.8369	−0.910942	19.8284	−0.000975	−0.026
39°01′	20350.5407	30.834682	3670.9213	0.007592	1.901
	240548.3803	−0.941283	19.7699	−0.000975	−0.026
39°02′	**22200.6216**	**30.834771**	**3671.3768**	**0.007572**	**1.900**
	240491.9034	**−0.941623**	**19.7114**	**−0.000974**	**−0.026**
39°03′	24050.7079	30.834859	3671.8311	0.007551	1.899
	240435.4060	−0.941964	19.6529	−0.000974	−0.026
39°04′	25900.7994	30.834947	3672.2842	0.007530	1.898
	240378.8882	−0.942304	19.5945	−0.000974	−0.026

Second-difference corrections

	00″	10″	20″	30″		00″	10″	20″	30″
	60″	50″	40″	30″		60″	50″	40″	30″
(I)	0.0000	−0.0004	−0.0006	−0.0007	(IV)	0.0000	0.0014	0.0023	0.0025

Computations utilizing the Equations (20.23) to (20.30) are demonstrated in an Excel® spreadsheet *map_projections.xls* and Mathcad® worksheet *TM.xmcd*, which are on the companion website for this book at http://www.pearsonhighered.com/ghilani.

With reference to Figure 20.6 and appropriate Transverse Mercator projection tables such as those shown in Table 20.2, the following SPCS83 equations for hand computations yield the solution of the *direct problem*, that is, obtaining the E_P and N_P coordinates of any point P from its geodetic coordinates:

$$\begin{aligned}
\Delta\lambda &= (\lambda_0 - \lambda)3600''/° \\
p &= 10^{-4}\Delta\lambda'' \\
N_P &= (I) + (II)p^2 + (III)p^4 + N_b \\
E_P &= (IV)p + (V)p^3 + (VI)p^5 + E_0
\end{aligned} \tag{20.31}$$

In Equations (20.32), $\Delta\lambda''$ is the difference in longitude between the central meridian and the point in seconds, and N and E are the state plane coordinates of the point in meters. The values for λ_0 and E_0 are zone constants, and are supplied with the table as shown in Table 20.2. The values for Roman numerals (I), (II), (III), (IV), (V), and (VI) are interpolated from the Transverse Mercator tables using the appropriate value given in the *"Difference"* column. For the first (I) and fourth (IV) column values, a small second-difference correction must also be interpolated using the numbers given at the bottom of the Table. Example 20.3 demonstrates the use of Equations (20.31) and Table 20.2.

Example 20.3

The geodetic latitude and geodetic longitude of station Stone Harbor in the state of New Jersey, which uses the Transverse Mercator projection, are 39°02′21.63632″ and 74°46′08.80133″, respectively. What are the station's SPCS83 coordinates?

Solution

The determination of the second differences and (III) and (VI) column values involve a linear interpolation. For example, the second difference for column (I) is determined as

$$-0.0006 + [-0.0007 - (-0.0006)]\ 1.63632''/10'' = -0.00062$$

where 1.63632″ comes from the second's portion of the station's latitude. The three values (III), (IV), and (VI) are interpolated in similar fashion and shown in the computations that follow. Using Table 20.2, the appropriate column values are

$$\begin{aligned}
\text{I} &= 22{,}200.6216 + 30.834771(21.63632) + (-0.00062) &&= 22{,}867.77195\\
\text{II} &= 3671.3768 + 0.007572\,(21.63632) &&= 3671.54063\\
\text{III} &= 1.900 + (1.899 - 1.900)(21.63632/60) &&= 1.89964\\
\text{IV} &= 240{,}491.9034 + (-0.941623)(21.63632) + 0.00233 &&= 240{,}471.53247\\
\text{V} &= 19.7114 + (-0.000974)(21.63632) &&= 19.690326\\
\text{VI} &= -0.026 + (-0.026 + 0.026)(21.63632/60) &&= -0.026
\end{aligned}$$

Substituting these values into Equations (20.31) yields

$$\begin{aligned}
\Delta\lambda'' &= (74°30' - 74°46'08.80133'')3600''/° &&= -968.80133''\\
p &= -968.80133 \times 10^{-4} &&= -0.096880133\\
N &= 22{,}867.77197 + 3671.54063p^2 + 1.89964p^4 + 0 &&= 22{,}902.2323 \text{ m}\\
E &= 240{,}471.5324p + 19.690326p^3 - 0.026p^5 + E_0 &&= 126{,}703.0680 \text{ m}
\end{aligned}$$

20.7.2 The Inverse Problem

The Transverse Mercator inverse problem is solved as

$$E' = E - E_0 \tag{20.32}$$

$$N' = N - N_b \tag{20.33}$$

$$e_1 = \frac{1 - \sqrt{1 - e^2}}{1 + \sqrt{1 - e^2}} \tag{20.34}$$

$$m = m_0 + \frac{N'}{k_0} \tag{20.35}$$

$$\chi = \frac{m}{a\left(1 - \frac{e^2}{4} - \frac{3e^4}{64} - \frac{5e^6}{256}\right)} \tag{20.36}$$

The foot point latitude ϕ_f is

$$\begin{aligned}\phi_f = \chi &+ \left(\frac{3e_1}{2} - \frac{27e_1^3}{32}\right)\sin 2\chi + \left(\frac{21e_1^2}{16} - \frac{55e_1^4}{32}\right)\sin 4\chi \\ &+ \left(\frac{151e_1^3}{96}\right)\sin 6\chi + \left(\frac{1097e_1^4}{512}\right)\sin 8\chi\end{aligned} \tag{20.37}$$

$$c_1 = C(\phi_f) \tag{20.38}$$

$$t_1 = T(\phi_f) \tag{20.39}$$

$$N_1 = R_N \text{ evaluated using Equation (19.4) with } \phi_f \tag{20.40}$$

$$M_1 = R_M \text{ evaluated using Equation (19.5) with } \phi_f \tag{20.41}$$

$$D = \frac{E'}{N_1 k_0} \tag{20.42}$$

$$\begin{aligned}B = \frac{D^2}{2} &- (5 + 3t_1 + 10c_1 - 4c_1^2 - 9e')\frac{D^4}{24} \\ &+ (61 + 90t_1 + 298c_1 + 45t_1^2 - 252e' - 3c_1^2)\frac{D^6}{720}\end{aligned} \tag{20.43}$$

$$\phi = \phi_f - \left(\frac{N_1 \tan \phi_f}{R_1}\right)B \tag{20.44}$$

Computations using Equations (20.37) to (20.44) are demonstrated in an Excel® spreadsheet *map_productions.xls* and Mathcad® worksheet *NJ_Table.xmcd* on the companion website for this book at http://pearsonhighered.com/ghilani. Similar to Table 20.2, Table 20.3 contains the necessary parameters and second differences

TABLE 20.3 EXCERPT FROM THE TRANSVERSE MERCATOR PROJECTION TABLES FOR NEW JERSEY FOR THE INVERSE PROBLEM

Zone constants: E_0 = 150,000 m ϕ_b = 38°50′ λ_0 = 74°30′ N_b = 0.000 m k_0 = 0.9999

Latitude	(I) (IX)	Difference 1″	(VII) (X)	Difference 1″	(VIII) (XI)
39°00′	18500.4650	30.834594	2056.2443	0.020257	29.191
	41561.9242	0.162576	393.3224	0.005941	7.024
39°01′	20350.5407	30.834682	2057.4597	0.020267	29.218
	41571.6788	0.162711	393.6788	0.005949	7.035
39°02′	**22200.6216**	**30.834771**	**2058.6757**	**0.020276**	**29.245**
	41581.4415	**0.163846**	**394.0357**	**0.005957**	**7.047**
39°03′	24050.7079	30.834859	2059.8923	0.020286	29.272
	41591.2122	0.162982	394.3931	0.005965	7.058
39°04′	25900.7994	30.834947	2061.1094	0.020295	29.299
	41600.9911	0.163117	394.7510	0.005973	7.069

Second-difference corrections

	00″	10″	20″	30″		00″	10″	20″	30″
	60″	50″	40″	30″		60″	50″	40″	30″
(I)	0.0000	−0.0004	−0.0006	−0.0007	(IX)	−0.0000	−0.0006	−0.0009	−0.0010

to compute the inverse problem. The formulas necessary to perform this calculation are

$$\begin{aligned}
E' &= E - E_0 \\
N' &= N - N_b \\
q &= E' \times 10^{-6} \\
\phi_f &= \phi_{\text{table}} + \Delta\phi'' \text{ (interpolated from table)} \\
\Delta\phi_f'' &= -(\text{VII})q^2 + (\text{VIII})q^4 \\
\Delta\lambda'' &= -(\text{IX})q + (\text{X})q^3 + (\text{XI})q^5 \\
\phi &= \phi_f + \Delta\phi_f'' \\
\lambda &= \lambda_b + \Delta\lambda''
\end{aligned} \tag{20.45}$$

In Equations (20.45) E' and N' are in meters, $\Delta\phi_f''$ and $\Delta\lambda''$ are in seconds, and ϕ_f, ϕ, and λ are in degrees, minutes, and seconds. The value of ϕ_f is interpolated from column (I) in Table 20.3, and is known as the *foot point latitude*. Notice that in this table, the first (I) and the ninth (IX) values have second-difference corrections. Example 20.4 demonstrates the use of these tables in the inverse problem.

Example 20.4

What are the geodetic coordinates of station Stone Harbor if its SPCS83 easting and northing coordinates are 126,703.0681 m and 22,902.2323 m, respectively? (The station lies in New Jersey's Transverse Mercator zone.)

Solution

Step 1: Calculate E', N', and q as

$$\begin{aligned}
E' &= 126{,}703.0680 - 150{,}000.000 &&= -23{,}296.9320 \text{ m}\\
N' &= 22{,}902.2323 - 0.000 &&= 22{,}902.2323 \text{ m}\\
q &= -23{,}296.9320 \times 10^{-6} &&= -0.023296932
\end{aligned}$$

Step 2: Looking at Table 20.3, it can be seen that the northing coordinate lies between the (I) values of 39°02′ and 39°03′. Thus $\Delta\phi_f''$ can be interpolated from column (I) as

$$\Delta\phi_f'' = (22{,}902.2323 - 22{,}200.6216)/30.834771 = 22.75388068''$$

Hence the foot point latitude is

$$\phi_f = 39°02' + 22.75388068'' = 39°02'22.75388068''$$

Step 3: Using $\Delta\phi_f''$, evaluate the tabular values for VII, VIII, IX, X, and XI. Note in this procedure, the values for VIII and XI must be linearly interpolated using the tabular values for 39°02′ and 39°03′.

$$\begin{aligned}
\text{VII} &= 2058.6757 + 0.020276(22.75388068) &&= 2059.13706\\
\text{VIII} &= 29.245 + (29.272 - 29.245)(22.75388068/60) &&= 29.25524\\
\text{IX} &= 41{,}581.4415 + 0.162846(22.75388068) - 0.0009 &&= 41{,}585.14688\\
\text{X} &= 394.0357 + 0.005957(22.75388068) &&= 394.17124\\
\text{XI} &= 7.047 + (7.058 - 7.047)(22.75388068/60) &&= 7.05117
\end{aligned}$$

Step 4: Applying the computed tabular values of step 3, compute the latitude and longitude of station Stone Harbor using Equations (20.45) as

$$\begin{aligned}
\Delta\phi'' &= -2059.13706q^2 + 29.25524q^4 &&= -1.11758''\\
\Delta\lambda'' &= -41{,}585.14688q + 394.17124q^3 + 7.05117q^5 &&= 968.80135''\\
\phi &= 39°02'22.75388068'' - 1.11758'' &&= 39°02'21.6363''\\
\lambda &= 74°30' + 968.80135'' &&= 74°46'08.8013''
\end{aligned}$$

Again, except for small rounding errors from both the forward and inverse computations, the solution produces the geodetic latitude and longitude of Stone Harbor that were given for Example 20.3.

As with the Lambert conformal conic map conversions, the solutions of the direct and inverse problems are typically performed with computers. Several programs have been written that allow for the easy conversion from geodetic to state plane coordinates and vice versa. WOLFPACK, which is available on the companion website for this book at http://www.pearsonhighered.com/ghilani contains this option under its coordinate computation menu.

20.8 REDUCTION OF DISTANCES AND ANGLES TO STATE PLANE COORDINATE GRIDS

Ground-surveyed distances and angles must undergo corrections prior to using them in making computations in state plane coordinate systems. As shown in Figure 20.8, distances must first be reduced from their ground-surveyed lengths to their ellipsoidal equivalents. Furthermore as shown in Figure 20.2, these ellipsoidal distances must then be reduced to the developable surface of the state plane system being used. Also, whenever angles or azimuths are used in these computations, they can be reduced to their grid equivalents. Once these reductions have been completed, traverse computations (see Chapter 10), and adjustment procedures given in both Chapters 10 and 16 may be performed. This section describes the processes of reducing these observations to state plane grids.

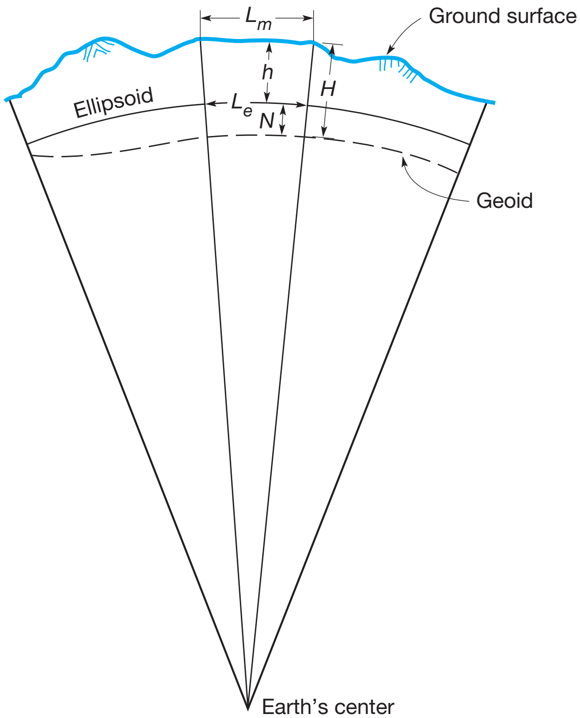


Figure 20.8 Reduction of lengths from surface observations to the ellipsoid.

20.8.1 Grid Reduction of Distances

The reduction of distances is normally done in two steps: (1) reduce the observations from their ground lengths to ellipsoid lengths (geodetic distance) and (2) reduce the ellipsoid lengths to their grid equivalents. The most precise formulas for reduction of slope distances to the ellipsoid were given in Section 19.14. However in most local surveys, the lengths of the distances are short and simpler methods of reduction can be used since the horizontal distance L_m closely matches the arc distance at the surface of the Earth. With this simplification, the relationship between the ground-surveyed length and the ellipsoid length L_e is

$$L_e = L_m\left(\frac{R_\alpha}{R_\alpha + H + N}\right) \tag{20.46}$$

where R_α is the radius of the Earth in the azimuth of the line as given by Equation (19.6), H the average orthometric height of the observed line above the geoid, and N the geoidal separation. The ratio $R_\alpha/(R_\alpha + H + N)$ is commonly called the *elevation factor*. For all but the most rigorous surveys, acceptable results can be obtained from Equation (20.46) by substituting the mean radius of the Earth (20,902,000 ft or 6,371,000 m) for R_α.

After a distance has been reduced to its ellipsoidal equivalent, it must then be scaled to its grid equivalent. This is accomplished by multiplying the ellipsoidal length of the line by an appropriate *scale factor*. For Lambert conformal conic mapping projections, the scale factor for any latitude ϕ can be computed as

$$m = M(\phi) \tag{20.47}$$

$$k = \frac{Rn}{am} \tag{20.48}$$

where M, R, and n are previously defined in this chapter. The scale factor at a point can be also interpolated from tables. This procedure is demonstrated in Example 20.5 using Table 20.1.

■ Example 20.5

What is the scale factor for Hayfield NE of Example 20.1? (This station is in the north zone of Pennsylvania's Lambert conformal conic projection.)

Solution

From Example 20.1, the geodetic latitude of station Hayfield NE is 41°18′20.25410″. Using the tabulated scale factors for latitudes of 41°18′ and 41°19′ from Table 20.1, the interpolated scale factor is

$$k = 0.99995893 + (0.99995838 - 0.99995893)(20.25410''/60'') = 0.999958744$$

For a Transverse Mercator map projection, the scale factor k for any point is computed as

$$k = k_0\left[1 + (1 + c)\frac{A^2}{2} + (5 - 4t + 42c + 13c^2 - 28e'^2)\frac{A^4}{24} + (61 - 148t + 16t^2)\frac{A^6}{720}\right] \tag{20.49}$$

where k_0 is a zone constant and c, A, and t are previously defined. To compute the scale factor k using tables, the following formula is used:

$$k = k_0[1 + (\text{XVI})q^2 + 0.00003q^4] \tag{20.50}$$

where q is defined in Equation (20.45), and (XVI) comes from Table 20.4.

TABLE 20.4 EXCERPT FROM THE TRANSVERSE MERCATOR PROJECTION TABLES FOR NEW JERSEY FOR COMPUTATION OF THE SCALE FACTOR. (NOTE: COLUMNS ARE REARRANGED FOR PUBLICATION PURPOSES ONLY.)

Zone constants: E_0 = 150,000 m ϕ_b = 38°50′ λ_b = 74°30′ N_b = 0.000 m k_0 = 0.9999

Latitude	(I) (XIV)	Difference 1″	(XII) (XV)	Difference 1″	(XIII) (XVI)
39°00′	18500.4650	30.834594	6293.2039	0.037673	3.014
	26155.7664	0.258924	353.110		0.012311
39°01′	20350.5407	30.834682	6295.4643	0.037664	3.014
	26171.3018	0.259046	353.486		0.012311
39°02′	**22200.6216**	**30.834771**	**6297.7241**	**0.037655**	**3.014**
	26186.8446	**0.259167**	**353.863**		**0.012311**
39°03′	24050.7079	30.834859	6299.9834	0.037646	3.013
	26202.3946	0.259289	354.240		0.012310
39°04′	25900.7994	30.834947	6302.2422	0.037637	3.013
	26217.9520	0.259412	354.618		0.012310

Second-difference corrections

	00″	10″	20″	30″		00″	10″	20″	30″
	60″	50″	40″	30″		60″	50″	40″	30″
(I)	0.0000	−0.0004	−0.0006	−0.0007	(XIV)	−0.0000	−0.0005	−0.0008	−0.0009

Example 20.6

Compute the scale factor for station Stone Harbor of Example 20.3. (This station lies in New Jersey's Transverse Mercator projection.)

Solution

From Example 20.3, the geodetic latitude of station Stone Harbor is 39°02′21.63632″. Using Equation (20.50), the scale factor for this station is

$$k = 0.9999[1 + 0.0123106q^2 + 0.00003q^4] = 0.99990668$$

where q is −0.023296932 as determined in Example 20.4, and the value for (XVI) is interpolated from the values at the bounding latitudes of 39°02′ and 39°03′.

A line consists of many points and thus there are several approaches to computing the scale factor k of a line. The easiest and least precise involves determining an average scale factor for an entire survey project and applying this single value to all reduced distances. This approach is adequate for low-accuracy surveys and for surveys that cover small areas. To achieve a higher level of accuracy, an "average scale factor" can be applied to each individual line. In this method the values are obtained by averaging the scale factors of the end points of the lines. This method works well for moderately long distances. However, for the most precise surveys, an additional scale factor at the midpoint of the line should be computed. Then an improved scale factor k_{12} is determined as

$$k_{12} = \frac{k_1 + 4k_m + k_2}{6} \tag{20.51}$$

where k_1 and k_2 are scale factors for the end points of the line, and k_m is the scale factor for the midpoint of the line. Obviously, this method requires that coordinates and scale factor k_m for the midpoint of the line be computed.

The product of the elevation factor and the scale factor is the so-called *combined factor* and mathematically expressed as

$$\text{combined factor} = \text{elevation factor} \times \text{scale factor} \tag{20.52}$$

In NAD27, the combined factor was known as the *grid factor*.

It is common in lower-order surveys and surveys that cover small areas, to use a single combined factor for the entire survey. Often data collectors allow for the entry of this value, so that coordinates are computed directly on the grid. With surveys that cover larger areas, or require more rigorous procedures, the scale factor and elevation factors for each line should be computed. The reduced grid distance is

$$\text{grid distance} = \text{ground distance} \times \text{combined factor} \tag{20.53}$$

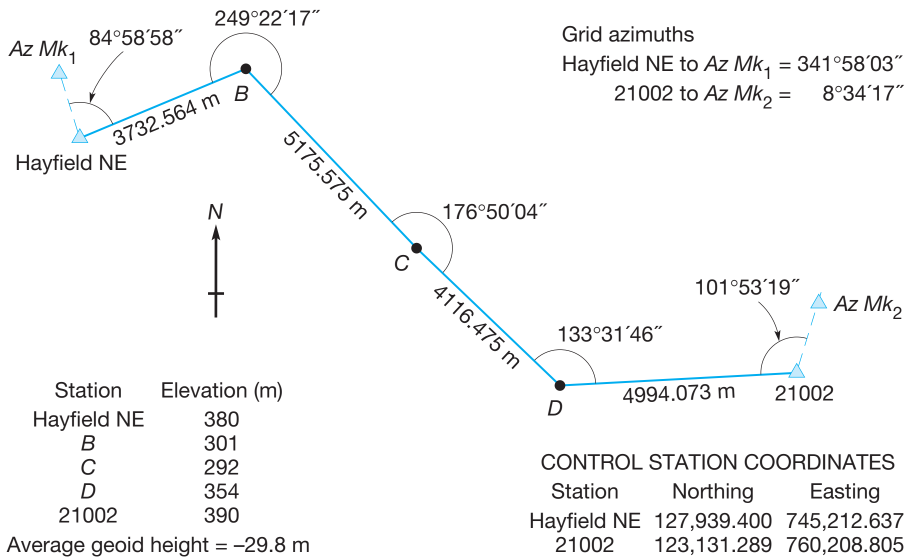


Figure 20.9 Field observed traverse.

Example 20.7

What are the grid lengths for the observed distances in Figure 20.9? The scale factor of station Hayfield NE in the figure was determined in Example 20.5.

Solution

For this example, elevation factors were computed using Equation (20.6) and employing the given average geoidal separation of −29.8 m and the mean radius of the Earth. Using computational procedures as given in Chapter 12, approximate coordinate values for each station were computed and then used to determine the scale factor at each station (see Example 20.5). For this survey, it was appropriate to use scale factors obtained by averaging the end point values. Elevation factors were obtained by using average elevations for each line in Equation (20.46). Grid factors for each line were then determined by multiplying the elevation factor by the average scale factor. Finally, grid distances were computed by multiplying each observed distance by its corresponding combined factor. The results of these calculations are listed in Table 20.5.

20.8.2 Grid Reduction of Azimuths and Angles

As shown in Figure 20.10, all grid meridians are parallel while all geodetic meridians converge to a single point. The primary difference between these directions is the convergence angle γ. Computation of the convergence angle for the Lambert conformal conic was demonstrated in both Examples 20.1 and 20.2.

TABLE 20.5 REDUCED GRID DISTANCES FOR FIGURE 20.9

Station	Elev. (m)	Distance (m)	Elev. Factor	k	k_{avg}	Combined Factor	Grid Dist. (m)
Hayfield NE	380			0.99995874			
		3732.564	0.99995123		0.99995854	0.99990978	3732.2272
B	301			0.99995834			
		5175.575	0.99995814		0.99995895	0.99991709	5175.1459
C	292			0.99995956			
		4116.475	0.99995398		0.99996014	0.99991412	4116.1215
D	354			0.99996072			
		4994.073	0.99994629		0.99996068	0.99990698	4993.6084
21002	390			0.99996065			

The convergence angle (in seconds) for the Transverse Mercator projection is computed as

$$c_2 = \frac{1 + 3c + 2c^2}{3}$$
$$c_3 = \frac{2 - \tan^2\phi}{15} \tag{20.54}$$
$$\gamma = A \tan \phi[1 + A^2(c_2 + c_3A^2)]$$

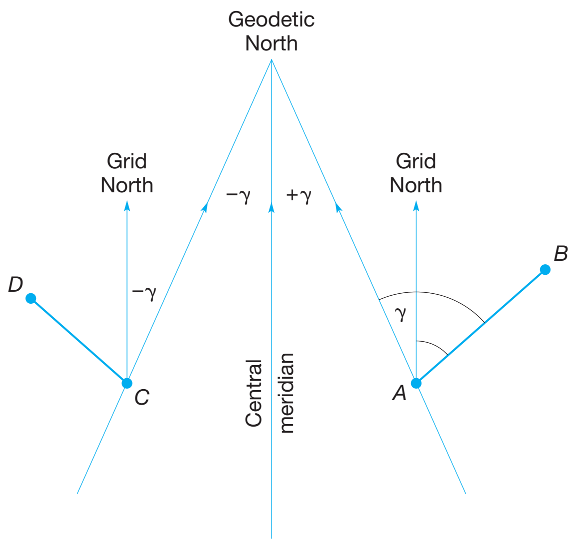


Figure 20.10 Relationship between geodetic azimuth, grid azimuth, and convergence angle γ.

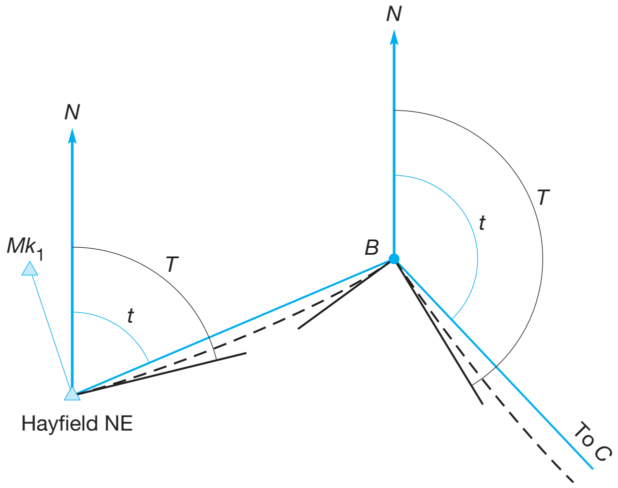


Figure 20.11 Part of Figure 20.9 showing arc-to-chord ($t - T$) correction at stations *Hayfield NE* and *B*.

where A and c are previously defined in this chapter. It can be computed using tables as

$$\gamma'' = (\text{XII})p + (\text{XIII})p^3 = (\text{XIV})q - (\text{XV})q^3 \qquad \textbf{(20.55)}$$

Using the value for q determined in Example 20.4 and Equation (20.55), the convergence angle at station Stone Harbor in New Jersey is

$$\gamma'' = (26{,}192.4512)q - (353.9989482)q^3 = -610.2'' = -0°10'10.2''$$

where (XIV) and (XV) were interpolated from Table 20.4.

Another factor that affects the reduction of azimuths is the projection of the geodetic azimuth onto a developable mapping surface. As can be seen in Figure 20.11, the projection of geodetic azimuths onto a flat surface results in an arc between the occupied and sighted stations. Letting the grid azimuth be t and the geodetic azimuth be T, the difference in these values is known as the *arc-to-chord* correction, also called the *second-term correction*, and is designated as δ.

The sign of the arc-to-chord correction is given by the location of the line and the type of map projection. For the Lambert conformal conic, the projected geodetic arc is always concave toward the central parallel of the zone. For the Transverse Mercator, the projected geodetic arc is always concave toward the central meridian for the zone. In Lambert conformal conic projections, the value for N_0 (the north-south center of the zone) can be determined from the derived zone constant sin ϕ_0, as

$$N_0 = R_b + N_b - R_0 \qquad \textbf{(20.56)}$$

where R_b and N_b are zone constants, and R_0 can be interpolated from the table using the latitude of ϕ_0. Generally, the precise numerical value for N_0 is not necessary for computations because the line's location in relation to the zone's north-south center, and the latitude ϕ_0 is all that is needed to determine the concavity

TABLE 20.6 The Sign of the δ Correction

Map Projection		Azimuth of the Line from North	
Lambert	Sign of $N - N_0$	0° to 180°	180° to 360°
	Positive	+	−
	Negative	−	+
Transverse Mercator	Sign of $E - E_0$, or E_3	270° to 90°	90° to 270°
	Positive	−	+
	Negative	+	−

of the projected geodetic arc. For example, in the north zone of Pennsylvania, the sin ϕ_0 is given as 0.661539733812 (see Table 20.1), which yields a central parallel of approximately 41°25′. This value is sufficient to determine the concavity of the projected geodetic azimuth. Table 20.6 shows the sign of δ'' based on these criteria.

The arc-to-chord correction is a function of the positions of the end points of the line and the type of map projection. For a Lambert conformal conic map projection δ is given as

$$\delta_{12} = 0.5(\sin \phi_3 - \sin \phi_0)(\lambda_1 - \lambda_2) \tag{20.57}$$

where (ϕ_1, λ_1) and (ϕ_2, λ_2) are the geodetic positions of the end points of the line using positive values for western longitudes, sin ϕ_0 is a zone constant, and $\phi_3 = (2\phi_1 + \phi_2)/3$. For the Transverse Mercator projection, δ can be computed from the northing and easting coordinates for the end points of the line as

$$E' = E - E_0 \tag{20.58a}$$

$$\Delta N = N_2 - N_1 \tag{20.58b}$$

$$\eta_f^2 = e'^2 \cos^2 \phi_f \tag{20.58c}$$

$$F = (1 - e^2 \sin^2 \phi_f)(1 + \eta_f^2)/(k_0 a)^2 \tag{20.58d}$$

$$E_3 = 2E'_1 + E'_2 \tag{20.58e}$$

$$\delta_{12} = -\frac{1}{6}\Delta N E_3 F\left(1 - \frac{1}{27}E_3^2 F\right) \tag{20.58f}$$

An approximate value for δ_{12} can be computed as $-25.4\ \Delta N(E_3/3)10^{-10}$ seconds where the coordinate values are in meters.

Combining the two aforementioned corrections to the grid azimuth t, the geodetic azimuth T can be computed as

$$T = t + \gamma - \delta \tag{20.59}$$

In Example 20.8, the arc-to-chord correction is computed for the lines in Figure 20.9 to demonstrate the procedure. The arc-to-chord correction is generally ignored for short lines since it is typically below the error of the angle or azimuth observation. The NGS has recommended that this correction only be applied for lines longer than 8 km. However, each surveyor should determine the maximum arc-to-chord correction acceptable for his or her survey.

Example 20.8

Evaluate the arc-to-chord correction for the lines in Figure 20.9.

Solution

Using the observed angles and computational procedures as demonstrated in Chapter 10, approximate coordinates were computed for each station. From these approximate coordinate values, the magnitudes of the second-term corrections were calculated using Equation (20.58). Note as can be shown from Table 20.6, the sign of all the corrections is negative since $N - N_0$ is negative for all lines in this example. The results of these computations are shown in Table 20.7.

The values in Table 20.7 are used to correct geodetic directions between the stations. Since Figure 20.9 has two grid azimuths, these azimuths do not need corrections. However when the lines are over 8 km or 5 mi in length, or a rigorous reduction is desired, the observed angles should be corrected. The arc-to-chord correction to observed angles can be found by taking the difference in the forward and back azimuths. The corrections for the backsight azimuths have the same magnitude, but opposite signs of the computed foresight corrections in Table 20.7. For example, the correction for the azimuth from station C to B is 1.42″. Since the sight distances to both the azimuth marks was short, the

TABLE 20.7 COMPUTATION OF THE ARC-TO-CHORD CORRECTION FOR FIGURE 20.9

Station	Approx. Northing	Approx. Easting	Approx. ϕ	Approx. λ	ϕ_3	δ
Hayfield NE	127,939.400	745,212.637	41°18′20.2541″	76°00′57.0024″		
					41°18′35.2908″	−0.10″
B	129,400.865	748,646.890	41°19′05.3642″	75°58′28.1096″		
					41°18′24.1269″	−0.11″
C	125,657.841	752,220.832	41°17′01.6524″	75°55′57.8495″		
					41°16′30.5490″	−0.12″
D	122,842.331	755,223.511	41°15′28.3421″	75°53′51.4269″		
					41°15′30.2932″	−0.22″
21002	123,131.289	760,208.805	41°15′34.1953″	75°50′17.0477″		

TABLE 20.8 ARC-TO-CHORD CORRECTION FOR ANGLES IN FIGURE 20.9

Station	Obs. Angle	Backsight δ	Foresight δ	Total δ	Corr. Angle
Hayfield NE	84°58′58″	0.00″	−0.10″	−0.10″	84°58′57.9″
B	249°22′17″	0.10″	−0.11″	−0.21″	249°22′16.8″
C	176°50′04″	0.11″	−0.12″	−0.23″	176°50′03.8″
D	133°31′46″	0.12″	−0.22″	−0.34″	133°31′45.7″
21002	101°53′19″	0.22″	0.00″	−0.22″	101°53′18.8″

corrections were assumed to be zero. The reduction of the observed angles is shown in Table 20.8.

As can be seen in Table 20.8, the corrections are small. Often the arc-to-chord correction is ignored for traverses involving lines under 8 km, and for lower-order surveys. However, the reduction of observed distances to the mapping grid is generally significant for most traverse surveys. Failure to account for these corrections will result in incorrect misclosures and subsequent incorrect adjustments and coordinate values. When these corrections are properly performed, the resulting adjustments will yield results similar to those achieved with geodetic computations. If adjusted ground distances are needed after an adjustment, a rearranged form of Equation (20.10) can be used to determine their values.

■ 20.9 COMPUTING STATE PLANE COORDINATES OF TRAVERSE STATIONS

Determining state plane coordinates of new traverse stations is a problem routinely solved by local surveyors. Normally it requires only that traverses (or triangulation or trilateration surveys) start and end on existing stations having known state plane coordinates, and from which known grid azimuths have been established. Generally these data are available for immediate use, but if not, they can be calculated as indicated in Sections (20.6) and (20.7) when geodetic latitude and geodetic longitude are known. State plane coordinates and the grid azimuths to a nearby azimuth mark are published by the NGS for most stations in the national horizontal network. In most areas, many other stations set by local surveyors exist that also have state plane coordinates and reference grid azimuths.

It is important to note that if a survey begins with a given grid azimuth and ties into another, directions of all intermediate lines are automatically grid azimuths. Thus, corrections for convergence of meridians are not necessary when the state plane coordinate system is used throughout the survey. However, as demonstrated in Section 20.8.2, the arc-to-chord correction should be considered and applied to observed angles when appropriate. Assuming that starting stations meeting the above described conditions are available, then there isn't any difference between making traverse computations in state plane coordinates and the procedures given for plane surveys in Chapter 10.

TABLE 20.9 REDUCED HORIZONTAL DISTANCES, ANGLES TO THE RIGHT, ANGLE MISCLOSURE, AND ADJUSTED AZIMUTHS FOR EXAMPLE 20.9

Station	Reduced Horizontal Distance (m)	Corrected Angle to the Right	Preliminary Azimuth	Adjusted Azimuth
AZ Mk$_1$				
			161°58′03.0″ (fixed)	161°58′03.0″ (fixed)
HAYFIELD NE		84°58′57.9″		
	3732.227		66°57′00.9″	66°56′59.1″
B		249°22′16.8″		
	5175.146		136°19′17.7″	136°19′14.1″
C		176°50′03.8″		
	4116.122		133°09′21.5″	133°09′16.1″
D		133°31′45.7″		
	4993.608		86°41′07.2″	86°41′00.0″
21002		101°53′18.8″		
			8°34′26.0″	8°34′17.0″ (fixed)
AZ Mk$_2$				

Angular misclosure = 8°34′26.0″ − 8°34′17″ = +9″
Correction per angle = −9.0″/5 = −1.8″

To illustrate the procedure of computing a traverse in the SPCS83, the following example is solved step-by-step.

Example 20.9

The traverse illustrated in Figure 20.9 originates from station Hayfield NE and closes on station 21002, both in the north zone of Pennsylvania. The reduction of both the distances and angles to the SPCS83 grid were demonstrated in Section 20.8. Using these values, compute and adjust the traverse, and determine the state plane coordinates of all traverse stations.

Solution

1. The computed azimuth of the line station 21002 to Az Mk$_2$ is compared with its fixed control value. The difference (+9.0″) represents the traverse angular misclosure. This misclosure is divided by the number of angles (five) to get the correction per angle (−1.8″). [This calculation is shown at the bottom of Table 20.9.] It should be interesting to note that if the observed angles had been used in the computations, the angular misclosure would have been +10″, or 1″ greater than is appropriate. Also note that for lines

Course	Length	Azimuth	Unbalanced Dep	Unbalanced Lat
A-B	3,732.227	66°56'59.1"	3434.2528	1461.3098
B-C	5,175.146	136°19'14.1"	3574.0730	-3742.7447
C-D	4,116.122	133°09'16.1"	3002.7618	-2815.2943
D-E	4,993.608	86°41'00.0"	4985.2439	288.9023
Sum =	18,017.103		14996.3515	-4807.8269

Misclosure in Departure = 14,996.3315 - 14,996.1680 = 0.1635
Misclosure in Latitude = -4,807.8269 - -4,808.1110 = 0.2841

Balanced Dep	Balanced Lat	Point	Coordinates X	Coordinates Y
3434.2189	1461.2509	A	745,212.637	127,939.400
3574.0261	-3742.8264	B	748,646.856	129,400.651
3002.7245	-2815.3592	C	752,220.882	125,657.825
4985.1986	288.8236	D	755,223.606	122,842.465
		E	760,208.805	123,131.289

Linear misclosure = 0.3278
Relative Precision = 1 in 55,000

Figure 20.12 Modified excerpt of compass rule adjustment done using WOLFPACK.

this short the arc-to-chord corrections are minimal and could easily have been avoided without appreciably affecting the final solution.

2. Traverse computations are performed using the same steps as described in Chapter 10. The procedure, shown in Figure 20.12, includes (a) calculating departures and latitudes [columns (1) and (2)], (b) adjusting the departures and latitudes [columns (3) and (4)], and (c) determining the station coordinates [columns (5) and (6)]. Adjustment of departures and latitudes in this example has been done by compass rule, but any method could be used including least squares. In the adjustment, the differences in eastings (X) and northings (Y) between control points were computed and checked against their fixed values to obtain the misclosures in departure (+0.164 m) and latitude (+0.284 m). An adjustment was then made to correct these computed differences to the required totals. The relative precision of the traverse was 1:55,000. Had the original distance and angle observations been used in the computations instead of their reduced equivalents, the relative precision of the traverse would have been only 1:10,000. This demonstrates the importance of making proper observational reductions before attempting an adjustment.

In summary, the following steps are necessary for performing traverse computations in state plane coordinates:

1. Obtain a starting and closing azimuth, and, if necessary, reduce them to grid azimuths.
2. Analyze the scale factor for the project. A mean of the published scale factors may be adequate for the project. This can be done by analyzing the number of significant figures in the longest measured length versus the

number of common digits in the scale factors. To avoid rounding errors there should be one more common digit in the scale factors than there is in the longest observed distance.

3. Analyze the elevation factor for the project. A mean factor may be adequate in terrain with small relief. Again the elevation factor for the highest and lowest station elevations in the project should have one more common digit than the number of significant figures in the longest observed length to avoid rounding errors.
4. If a project scale factor and elevation factor can be used, compute a combined factor for the project.
5. Reduce all horizontal distances to their grid equivalents.

For lines under 8 km and lower-order surveys, steps 6 through 8 are typically skipped.

6. Using preliminary azimuths derived from unreduced angles and grid distances, compute approximate coordinates.
7. Analyze the magnitude of the arc-to-chord correction for each line using the approximate coordinates.
8. Apply the arc-to-chord corrections to the observed angles.
9. Compute and adjust the traverse.
10. Compute the final adjusted SPCS83 coordinates for the new stations. If adjusted ground distances are required, apply the inverse of the combined factor to each line.

Procedures for computing traverses in SPCS27 follow the same steps.

20.10 SURVEYS EXTENDING FROM ONE ZONE TO ANOTHER

Surveys in border areas often cross into different zones or even abutting states. This does not represent a problem, however, because adjacent zones overlap by appreciable distances, as shown in Figures 20.3 and 20.4.

The general procedure for extending surveys from one zone to another requires that the survey proceed from the first zone into the overlap area with the second. Then the geodetic latitudes and longitudes are computed for two intervisible stations using their grid coordinates in the first zone. (Recall that this conversion is called the *inverse problem*.) Using the geodetic positions of the two points, their state plane coordinates in the new zone are then computed. (This is the *direct problem*.) Finally the grid azimuth for the line in the new zone can be obtained from the new coordinates of the two points.

Suppose, for example, that a survey being computed in SPCS83 originates in southern Wisconsin, which uses the Lambert conformal conic projection, and extends into the western zone in northern Illinois, which uses the Transverse Mercator grid. With Wisconsin south zone SPCS83 coordinates of two intervisible points in the overlap area of the two zones known, Equations (20.3) are solved for the geodetic latitudes and longitudes of the points.

With the geodetic latitudes and longitudes of the two points known, Equations (20.4) are solved using constants for the appropriate Illinois zone entered to obtain their E and N coordinates in that zone. From these coordinates, the grid azimuth of the line joining the intervisible points can be calculated by using Equation (11.5) and the survey can continue into Illinois.

If immediate coordinate values in the new zone are not required, the entire survey can be computed in one zone. The inverse problem can then be used to compute geodetic coordinate values of the points followed by the direct problem to compute grid coordinates in the second zone. For example, the survey in the previous paragraph could be computed entirely in the Wisconsin South Zone. Following this, the inverse problem can be used to compute the geodetic coordinates for the stations in northern Illinois. These geodetic coordinates can then be used to determine their grid coordinates equivalents in the western Illinois zone. However to do this, the coordinates values of the control station(s) in Illinois would need to be converted to their geodetic and Wisconsin southern zone equivalents. Additionally, the control azimuth in the Illinois zone would need to be converted to its geodetic equivalent and then to its Wisconsin south zone grid value. That is all control used in the survey must be converted into common, single zone values.

It should be remembered that the zone limits do not mark the end of the map projection, but simply the extents of the zone where 1:10,000 precisions are maintained between grid and ellipsoidal lengths. If the proper reductions as presented in Section 20.8 are performed to the observed distances, the use of a single zone for computational purposes can be extended well into neighboring zones without loss of accuracy to the survey. For example, the zones in Pennsylvania can be used to perform traverse computations in neighboring New Jersey, Ohio, Maryland, New York, and so on. Once the Pennsylvania grid coordinates for the points are determined, their geodetic equivalents can be determined using inverse computations and converted to the appropriate state zone with direct computations.

Extending surveys from zone to zone in SPCS27 follows the same procedure. Solving the direct and inverse problems that are necessary in this procedure is most conveniently handled using the computer programs described previously.

■ 20.11 CONVERSIONS BETWEEN SPCS27 AND SPCS83

Many stations that have coordinates known in SPCS27 were not included as part of the adjustment of NAD83 and thus they do not have SPCS83 coordinates. Coordinates in SPCS83 can be determined for these stations in different ways, depending on conditions and on the level of accuracy required in the conversion process. The most accurate procedure is to readjust the original survey data to control points whose positions were included in the NAD83 general adjustment, and thus known in SPCS83. This process is only possible if the original survey was tied to one or more of these control points. If such points do not exist, they can be established using GNSS receivers or traversing as described earlier in this chapter.

A second method involves knowing coordinate values for stations in both NAD27 and NAD83. These stations should encompass other stations that are to be converted. As discussed in Section 11.8, a two-dimensional conformal coordinate transformation can be used to transform all the coordinates known in NAD27 into NAD83.

Another method, which yields an intermediate level of accuracy, computes *expected* changes for all points in a large area based on known changes for a selected network of points in the region. Changes of the selected points, which should be uniformly distributed throughout the area, are obtained by computing differences between their SPCS27 and SPCS83 coordinates. Mathematical functions (polynomials) are then used to predict changes of all other points in the area based on the pattern of changes for the known points. This procedure gives reasonably good accuracy, suitable for many needs. The NGS has two software packages available in their website[2] that will perform this type of conversion: the software package NADCON, which was developed by the NGS; and CORPSCON, which is an enhanced version of NADCON from the U.S. Corps of Engineers.

A final lower-order conversion procedure uses either linear interpolation of changes in E and N coordinates between points of known change, or an average change for a given area. It produces results suitable for certain purposes; for example, correcting coordinate grids of small-scale maps and nautical charts.

20.12 THE UNIVERSAL TRANSVERSE MERCATOR PROJECTION

The Universal Transverse Mercator (UTM) system is another important map projection that has worldwide use. Originally developed by the Department of Defense primarily for artillery use, it provides worldwide coverage from 80°S latitude to 80°N latitude. Each zone has a 6° longitudinal width; thus 60 zones are required to encircle the globe. The UTM system is a Transverse Mercator map projection and thus uses the equations presented in this chapter. It has recently taken on added importance for surveyors, since UTM coordinates in metric units are now being included along with state plane and geodetic coordinates for all published NAD83 station descriptions. UTM grids are also being included on all maps in the national mapping program, and UTM coordinates are being used more frequently for referencing positions of data entered into Land and Geographic Information Systems.

UTM zones are numbered easterly from 1 through 60, beginning at longitude 180°W. The United States is covered from zone 10 (west coast) through zone 20 (east coast). The central meridian for each zone is assigned a false easting E_0 of 500,000 m. A false northing N_b of zero is applied for the northern hemisphere of each zone, and 10,000,000 m is assigned for the southern hemisphere to avoid negative Y coordinates. To specify the position of any point in the UTM system, the zone number must be given as well as its northing and easting.

In the UTM system, each zone overlaps adjacent ones by 0°30′. Because zone widths of 6° are considerably larger than those used in state plane systems, lower accuracies result, and 1 part in 2500 ($k_0 = 0.9996$) applies at the center and edges of zones. Equations for calculating X and Y coordinates in the UTM system are the same as those for the Transverse Mercator projection. As with the

---

[2]Both NADCON and CORPSCON are available for downloading at the NGS website http://www.ngs.noaa.gov/PC_PROD/pc_prod.shtml

state plane systems, tables giving formulas and constants for the system are available. Also, like the state plane coordinate system, the datum of reference must be specified. Because UTM coordinates are available for all points within the NAD83, calculations between very widely spaced points can readily be made. This is convenient and entirely consistent with current capabilities for conducting surveys of global extent with new devices such as GNSS receivers. The Universal Transverse Mercator map projection with a graphic showing zone boundaries is included in the Excel® spreadsheet *map_projections.xls*, which is on the companion website for this book.

20.13 OTHER MAP PROJECTIONS

The Lambert conformal conic and Transverse Mercator map projections are designed to cover areas extensive in east-west and north-south directions, respectively. However, these systems do not conveniently cover circular areas or long strips of the Earth that are skewed to the meridians. Two other systems, the *oblique stereographic* and the *Oblique Mercator* projections, satisfy these problems.[3]

20.13.1 Oblique Stereographic Map Projection

The oblique stereographic projection can be divided into two classes: *tangent plane* and *secant plane*. In either case, as illustrated in Figure 20.13, the projection

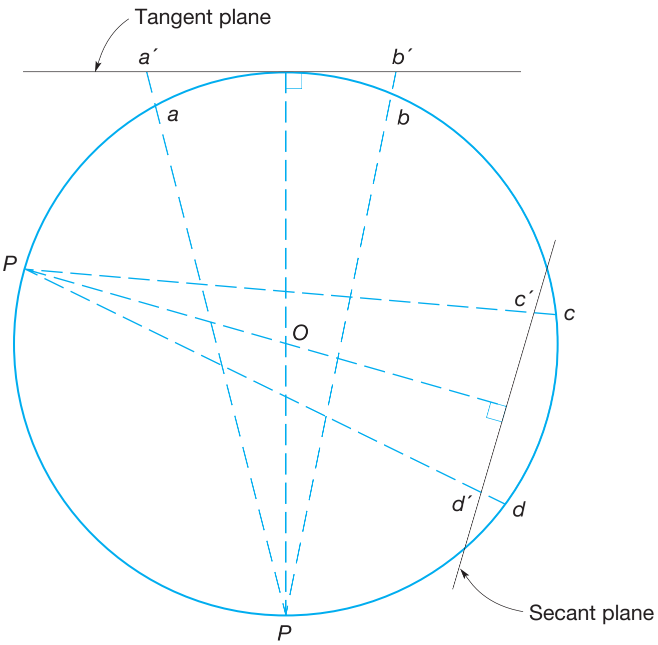


Figure 20.13 Tangent plane and secant plane horizon stereographic map projections.

[3]The oblique Mercator projection is also called the Hotine skew orthomorphic projection, named after the English geodesist Martin Hotine.

point P (the origin) is on the ellipsoid where a line perpendicular to the map plane and passing through center point O intersects the ellipsoid. In the tangent plane system, ellipsoid points a and b are projected outward to a' and b', respectively, on the map plane. For the secant plane system, ellipsoid points c and d are projected inward to c' and d' on the map plane. (If they were outside of the secant points, projection would be outward.) The oblique stereographic map projection is conformal and thus preserves the shapes of objects.

Oblique stereographic projections are not employed in the United States typically, but are used in Canada and other parts of the world. As discussed in Section 19.6.6, these projections are also used to convert geodetic coordinates determined by GNSS surveys into map projection coordinates for use in the localization process. If point P is the North or South Pole, the projection is called *polar stereographic*; if it is on the equator, *equatorial stereographic*.

The defining parameters of this projection are the latitude and longitude of the grid origin (φ_0, λ_0) and the scale factor at the grid origin k_0. It uses common functions

$$\chi(\varphi) = 2\tan^{-1}\left[\tan\left(\frac{\pi}{4} + \frac{\varphi}{2}\right)\left(\frac{1 - e\sin\varphi}{1 + e\sin\varphi}\right)^{e/2}\right] - \frac{\pi}{2} \tag{20.60}$$

$$M(\varphi) = \frac{\cos\varphi}{\sqrt{1 - e^2\sin^2\varphi}} \tag{20.61}$$

where e is the eccentricity of the ellipsoid as defined by Equation (19.2a) or (19.2b). Using Equation (20.6), the zone constants for the projection are

$$\chi_0 = \chi(\varphi_0) \tag{20.62}$$

$$m_0 = M(\chi_0) \tag{20.63}$$

Using the latitude φ and longitude λ of the point, and the semimajor axis a for the ellipsoid, the equations for the direct problem are

$$\chi = \chi(\varphi) \tag{20.64}$$

$$m = M(\varphi) \tag{20.65}$$

$$A = \frac{2ak_0m_0}{\cos\chi_0[1 + \sin\chi_0\sin\chi + \cos\chi_0\cos\chi\cos(\lambda - \lambda_0)]} \tag{20.66}$$

$$E = A\cos\chi\sin(\lambda - \lambda_0) \tag{20.67}$$

$$N = A[\cos\chi_0\sin\chi - \sin\chi_0\cos\chi\cos(\lambda - \lambda_0)] \tag{20.68}$$

$$k = \frac{A\cos\chi}{am} \tag{20.69}$$

Example 20.10

The following geodetic coordinates are observed using GNSS methods. What are the oblique stereographic map projection coordinates for Station A using a grid origin of (41°18′15″N,76°00′00″W) and $k_0 = 1$? (Use WGS84 ellipsoidal parameters.)

Station	Latitude	Longitude	Height (m)
A	41°18′09.88223″N	75°59′58.05637″W	282.476
B	41°18′21.11176″N	76°00′37.35445″W	296.571
C	41°18′19.33293″N	75°59′40.39279″W	313.814
D	41°18′09.67030″W	75°59′44.19645″W	304.205

Solution

The zone constants are

By Equation (20.62):

$$\chi_0 = 2\tan^{-1}\left[\tan\left(\frac{\pi}{4} + \frac{41°18'15''}{2}\right)\left(\frac{1 - e\sin 41°18'15''}{1 + e\sin 41°18'15''}\right)^{e/2}\right] - \frac{\pi}{2} = 41°06'48.66298''$$

By Equation (20.63):

$$m_0 = \frac{\cos 41°18'15''}{\sqrt{1 - e^2\sin^2 41°18'15''}} = 0.752314$$

By Equation (20.64):

$$\chi = 2\tan^{-1}\left[\tan\left(\frac{\pi}{4} + \frac{41°18'09.88223''}{2}\right)\left(\frac{1 - e\sin 41°18'09.88223''}{1 + e\sin 41°18'09.88223''}\right)^{e/2}\right]$$
$$= 41°06'43.54972''$$

By Equation (20.65):

$$m = \frac{\cos 41°18'09.88223''}{\sqrt{1 - e^2\sin^2 41°18'09.88223''}} = 0.752330$$

By Equation (20.66):

$$A = \frac{2(6378137)(1)(0.752314)}{\cos\chi_0[1 + \sin\chi_0\sin\chi + \cos\chi_0\cos\chi\cos(-75°59'58.05637'' + 76°)]}$$
$$= 6{,}368{,}873.344 \text{ m}$$

By Equation (20.67):

$$E = A\cos 41°06'43.54972''\sin(-75°59'58.05637'' + 76°) = 45.218 \text{ m}$$

By Equation (20.68):

$$N = A[\cos\chi_0\sin\chi - \sin\chi_0\cos\chi\cos(-75°59'58.05637'' + 76°)] = -157.869 \text{ m}$$

By Equation (20.69):

$$k = \frac{6{,}368{,}873.344 \cos 41°06'43.54972''}{6{,}378{,}137(0.752330)} = 1.00000$$

Note in the example that the direct problem was not defined with false easting or northings, and thus one of the resultant coordinates has a negative value. For localization, as discussed in Section 19.6.6, this is not a problem since the user never views the coordinates. However, a false easting and northing could be added to Equations (20.67) and (20.68) to provide positive coordinates values in a limited region.

The inverse problem converts the map projection coordinates of (N,E) to geodetic latitude and longitude. The equations used in the inverse problem for the oblique stereographic map projection are

$$\rho = \sqrt{E^2 + N^2} \tag{20.70}$$

$$c = 2\tan^{-1}\left(\frac{\rho \cos \chi_0}{2ak_0m_0}\right) \tag{20.71}$$

$$\chi = \sin^{-1}\left(\cos c \sin \chi_0 + \frac{N \sin c \cos \chi_0}{\rho}\right) \tag{20.72}$$

$$\lambda = \lambda_0 + \tan^{-1}\left(\frac{E \sin c}{\rho \cos \chi_0 \cos c - N \sin \chi_0 \sin c}\right) \tag{20.73}$$

$$\phi = 2\tan^{-1}\left[\tan\left(\frac{\pi}{4} + \frac{\chi}{2}\right)\left(\frac{1 + e\sin\varphi}{1 - e\sin\varphi}\right)^{e/2}\right] - \frac{\pi}{2} \tag{20.74}$$

Using $\varphi = \chi$ in the first iteration, Equation (20.74) is iterated until the change in φ becomes negligible. The scale factor is computed using Equation (20.69). Computations for the oblique stereographic map projection are demonstrated in a Mathcad® worksheet *oblique.xmcd* on the companion website for this book.

20.13.2 Oblique Mercator Map Projection

The oblique Mercator projection is designed for areas whose major extent runs obliquely to a meridian, such as northwest to southeast. The projection is conformal and is developed by projecting points from the ellipsoid to an imaginary cylinder that is oriented with its axis skewed to the equator. It is used as the state plane coordinate projection for the southeast portion of Alaska. Computations for the aforementioned map projections are demonstrated in the Excel® spreadsheet and Mathcad® worksheet on the companion website for this book.

Also on the companion website are an Excel® spreadsheet, *map_projections.xls*, which contains all the map projections presented in this chapter.

■ 20.14 MAP PROJECTION SOFTWARE

The computations presented in this chapter are often difficult to perform correctly using a calculator due to their length and the magnitude of the values involved. Thus, programs have been developed that allow users to easily and conveniently perform these computations. On the companion website for this book are three such software packages. As discussed earlier, WOLFPACK has the ability to perform direct and inverse computations for any 1983 state plane coordinate system zone, universal Transverse Mercator zone, or a user-specified oblique stereographic map projection. The software accepts manually entered coordinate values or can read a file of coordinates. Figure 20.14 shows the format for a file of geodetic coordinates as presented in Table 20.7 in the upper-half of the window. After performing direct problem in the Pennsylvania north zone, a portion of the resulting output file is shown in the lower-half of the screen. A complete explanation of the input file is discussed in the help file that accompanies the software. WOLFPACK can also perform reductions on the distance and azimuth observations. That is, it can scale an observed horizontal distance to its *grid* equivalent and reduce a geodetic azimuth to its equivalent grid value. These options, shown in Figure 20.15, are under the *map projections* submenu in the *programs* menu.

Also on the companion website at http://www.pearsonhighered.com/ghilani is the Excel® spreadsheet, *map_projections.xls*. This spreadsheet demonstrates

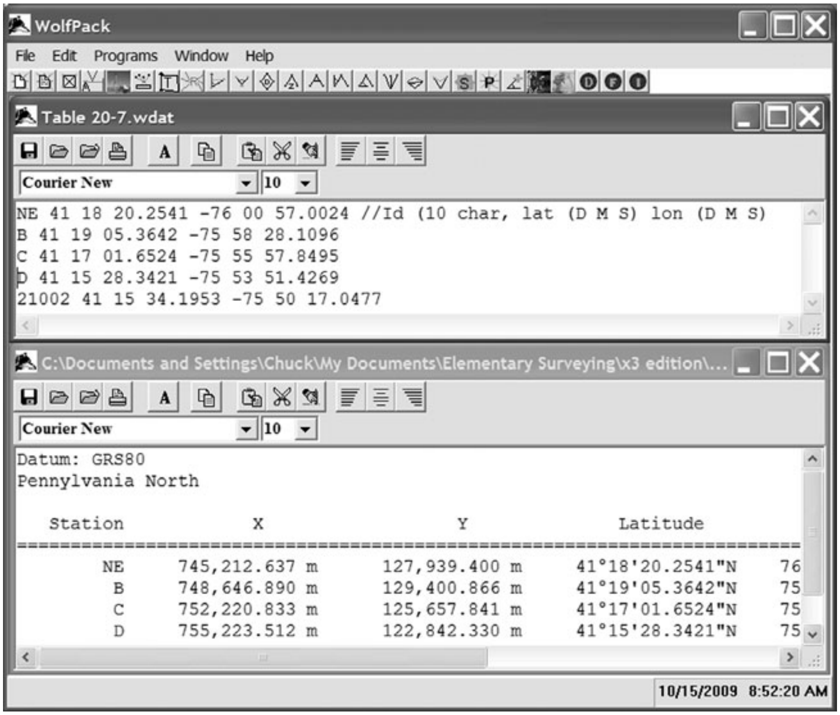


Figure 20.14 WOLFPACK with file for direct problem computations in Pennsylvania north zone shown in upper-half of screen and the resulting computed file shown in the lower-half.

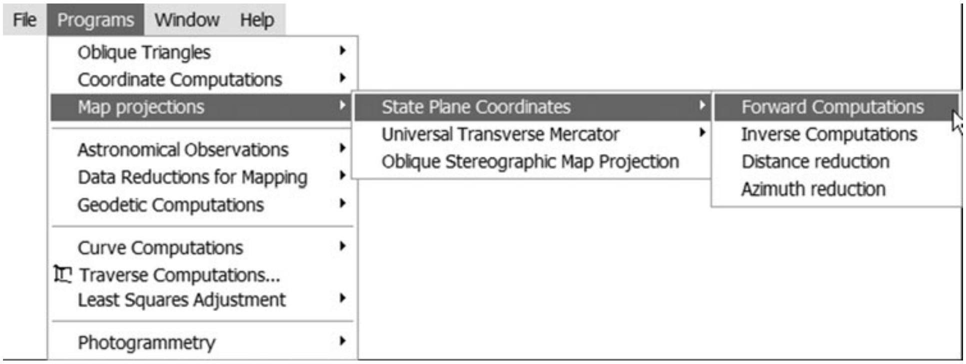


Figure 20.15 WOLFPACK menu items for map projection computations.

the computations discussed in this chapter. For those interested in programming these computations using a higher-level language, there are several Mathcad® worksheets on the companion website that can be explored. The file *Lambert.xmcd* demonstrates the direct and inverse computations using the Lambert conformal conic map projection. The file *TM.xmcd* contains the direct and inverse computations for the Transverse Mercator map projection and file *oblique.xmcd* contains the direct and inverse computations for the oblique Mercator map projection used in Alaska. Finally, the file *ostereo.xmcd* contains the programming for the direct and inverse computations using an oblique stereographic map projection. For those wishing to compare their intermediate computations using tables with software-derived values, the files *PA_Table.xmcd* and *NJ_Table.xmcd* provide this feature. Finally, the file *GridObs.xmcd* demonstrates the reduction of observations as presented in this chapter. These additional features are provided as learning aids so that the user can check their values when exploring the example problems presented in this chapter.

PROBLEMS

Asterisks (*) indicate problems that have partial answers given in Appendix G.

20.1 Name three developable surfaces that are used in map projections.

20.2 Name the two basic projections used in state plane coordinate systems. What are their fundamental differences? Which one is preferred for states whose long dimensions are north-south? East-west?

20.3 What does it mean to be a conformal projection?

20.4 How is a geodetic distance reduced to the map projection surface in state plane coordinate system map projections?

20.5 What corrections must be made to measured slope distances prior to computing state plane coordinates?

20.6 What is the primary difference between grid and geodetic azimuths?

20.7 Develop a table of SPCS83 elevation factors for geodetic heights ranging from 0 to 3000 ft. Use increments of 500 ft and an average radius for the Earth of 6,371,000 m.

20.8 Similar to Problem 20.7, except for geodetic heights from 0 to 1200 m using 100 m increments.

20.9 Explain how surveys can be extended from one state plane coordinate zone to another or from one state to another.

20.10 Develop a table similar to Table 20.1 for a range of latitudes from 40°30′ N to 40°39′ N in the Pennsylvania north zone with standard parallels of 40°53′ N and 41°57′ N, and a grid origin at (40°10′ N, 77°45′ W).

20.11* The Pennsylvania north zone SPCS83 state plane coordinates of points *A* and *B* are as follows:

Point	*E* (m)	*N* (m)
A	541,983.399	115,702.804
B	541,457.526	115,430.257

Calculate the grid length and grid azimuth of line *AB*.

20.12 Similar to Problem 20.11, except points *A* and *B* have the following New Jersey SPCS83 state plane coordinates:

Point	*E* (m)	*N* (m)
A	135,995.711	19,969.603
B	136,459.204	20,302.524

20.13 What are the SPCS83 coordinates (in ft) and convergence angle for a station in the north zone of Pennsylvania with geodetic coordinates of 41°12′33.0745″ N and 76°23′48.9765″ W?

20.14* Similar to Problem 20.13 except that the station's geodetic coordinates are 41°14′20.03582″ N and 80°58′46.28764″ W. Give coordinates in meters.

20.15 What is the scale factor for the station in Problem 20.13?

20.16 What is the scale factor for the station in Problem 20.14?

20.17* What are the SPCS83 coordinates for a station in New Jersey with geodetic coordinates of 40°50′23.2038″ N and 74°15′36.4908″ W?

20.18 Similar to Problem 20.17 except that the geodetic coordinates of the station are 41°11′25.0486″ N and 74°29′36.9641″ W.

20.19 What are the convergence angle and scale factor at the station in Problem 20.17?

20.20 What are the convergence angle and scale factor at the station in Problem 20.18?

20.21* What are the geodetic coordinates for a point *A* in Problem 20.11?

20.22 Similar to Problem 20.21 except for point *B* in Problem 20.11?

20.23* What are the geodetic coordinates for a point *A* in Problem 20.12?

20.24 Similar to Problem 20.23 except for point *B* in Problem 20.12.

20.25 In computing state plane coordinates for a project area whose mean orthometric height is 305 m, an average scale factor of 0.99992381 was used. The average geoidal separation for the area is −23.83 m. The given distances between points in this project area were computed from SPCS83 state plane coordinates. What horizontal length would have to be observed to lay off these lines on the ground? (Use 6,371,000 m for an average radius for the Earth.)

(a)* 2834.79 ft
(b) 230.204 m
(c) 823.208 ft

20.26 Similar to Problem 20.25, except that the mean project area elevation was 38 m, the geoidal separation −31.37 m, the scale factor 0.99997053, and the computed lengths of lines from SPCS83 were

(a) 132.028 m
(b) 502.39 ft
(c) 910.028 m

20.27 The horizontal ground lengths of a three-sided closed-polygon traverse were measured in feet as follows: $AB = 501.92$, $BC = 336.03$, and $CA = 317.88$ ft. If the average orthometric height of the area is 4156.08 ft and the average geoid separation is −23.05 m, calculate ellipsoid lengths of the lines suitable for use in computing SPCS83 coordinates. (Use 6,371,000 m for an average radius for the Earth.)

20.28 Assuming a scale factor for the traverse of Problem 20.27 to be 0.99996294, calculate grid lengths for the traverse lines.

20.29 For the traverse of Problem 20.27, the grid azimuth of a line from A to a nearby azimuth mark was 10°07′59″ and the clockwise angle measured at A from the azimuth mark to B, 213°32′06″. The measured interior angles were $A = 41°12'26''$, $B = 38°32'50''$, and $C = 100°14'53''$. Balance the angles and compute grid azimuths for the traverse lines. (*Note*: Line BC bears easterly.)

20.30 Using grid lengths of Problem 20.28 and grid azimuths from Problem 20.29, calculate departures and latitudes, linear misclosure, and relative precision for the traverse.

20.31 If station A has SPCS83 state plane coordinates $E = 634{,}728.082$ m and $N = 384{,}245.908$ m, balance the departures and latitudes computed in Problem 20.30 using the compass rule, and determine SPCS83 coordinates of stations B and C.

20.32* What is the combined factor for the traverse of Problems 20.27 and 20.28?

20.33 The horizontal ground lengths of a four-sided closed-polygon traverse were measured as follows: $AB = 479.549$ m, $BC = 830.616$ m, $CD = 685.983$ m, and $DA = 859.689$ m. If the average orthometric height of the area is 1250 m, the geoidal separation is −30.0 m, and the scale factor for the traverse is 0.9999574, calculate grid lengths of the lines for use in computing SPCS83 coordinates. (Use 6,371,000 m for an average radius of the Earth.)

20.34 For the traverse of Problem 20.33, the grid bearing of line BC is N57°39′48″ W. Interior angles were measured as follows: $A = 120°26'28''$, $B = 73°48'56''$, $C = 101°27'00''$, and $D = 64°17'26''$. Balance the angles and compute grid bearings for the traverse lines. (*Note*: Line CD bears southerly.)

20.35 Using grid lengths from Problem 20.33 and grid bearings from Problem 20.34, calculate departures and latitudes, linear misclosure, and relative precision for the traverse. Balance the departures and latitudes by the compass rule. If the SPCS83 state plane coordinates of point B are $E = 255{,}096.288$ m and $N = 280{,}654.342$ m, calculate SPCS83 coordinates for points C, D, and A.

20.36 The traverse in Problems 10.9 through 10.11 was performed in the Pennsylvania north zone of SPCS83. The average elevation for the area was 505.87 m and the average geoidal separation was −28.25 m. Using the data in Table 20.1 and a mean radius for the Earth, reduce the observations to grid and adjust the traverse. Assume that the azimuths given in Chapter 10 are grid azimuths. Compare this solution with that obtained in Chapter 10. (Use 6,371,000 m for an average radius of the Earth.)

20.37 The traverse in Problems 10.12 through 10.14 was performed in the New Jersey zone of SPCS83. The average elevation for the area was 234.93 m and the average geoidal separation was −32.86 m. Using the data in Tables 20.3 and 20.4 and a mean radius for the Earth, reduce the observations to grid and adjust the traverse. Assume that the azimuths given in Chapter 10 are grid azimuths. Compare this solution with that obtained in Chapter 10.

20.38 The traverse in Problem 10.22 was performed in the New Jersey SPCS83. The average elevation of the area was 67.2 m and the average geoidal separation was −28.5 m. Using 6,371,000 m for the mean radius of the Earth, reduce the observations to grid and adjust the traverse using the compass rule. Assume that the azimuths given in Problem 10.22 are grid azimuths. Compare this solution with that obtained in Problem 10.22.

20.39 The traverse in Problem 10.21 was performed in the Pennsylvania north zone of SPCS83. The average elevation for the area was 367.89 m and the average geoidal separation was −30.23 m. Using the mean radius of the Earth of 6,371,000 m, reduce the observations to grid and adjust the traverse using the compass rule. Assume that the azimuths given in Problem 10.21 are grid azimuths. Compare this solution with that obtained in Problem 10.21.

20.40* If the geodetic azimuth of a line is 205°06′36.2″, the convergence angle is −0°42′26.1″ and the arc-to-chord correction is +0.8″, what is the equivalent grid azimuth for the line?

20.41 If the geodetic azimuth of a line is 18°47′20.1″, the convergence angle is −1°08′06.8″ and the arc-to-chord correction is −1.5″, what is the equivalent grid azimuth for the line?

20.42 Using the values given in Problems 20.40 and 20.41, what is the obtuse grid angle between the two azimuths?

20.43 The grid azimuth of a line is 102°37′08″. If the convergence angle at the end point of the azimuth is 2°05′52.9″ and the arc-to-chord correction is 0.7″, what is the geodetic azimuth of the line?

20.44 Similar to Problem 20.43, except the convergence angle is −1°02′20.7″ and the arc-to-chord correction is −0.6″.

20.45 What is the arc-to-chord correction for the line from *A* to *B* in Problems 20.21 and 20.22?

20.46 What is the geodetic azimuth of the line from *A* to *B* in Problem 20.45?

20.47 Using the defining parameters given in Example 20.10, compute oblique stereographic map projection coordinates for Station B.

20.48 Similar to Problem 20.47 except for Station C.

20.49 Similar to Problem 20.47 except for Station D.

20.50 Create a computational program that reduces distances from the ground to a mapping grid.

BIBLIOGRAPHY

Bunch, B. W. 2002. "A New Projection: Developing and Adopting a Single Zone State Plane Coordinate System for Kentucky, Part 1." *Professional Surveyor* 22 (No. 4): 26.

______. 2002. "A New Projection: Developing and Adopting a Single Zone State Plane Coordinate System for Kentucky, Part 2." *Professional Surveyor* 22 (No. 5): 34.

GIA. 2006. "How Things Work: Scale, Elevation, Grid, and Combined Factors Used in Instrumentation." *Professional Surveyor* 26 (No. 2): 47.

Hartzell, P., L. Strunk and C. Ghilani. 2002. "Pennsylvania State Plane Coordinate System: Converting to a Single Zone." *Surveying and Land Information Science* 62 (No. 2): 95.

Snay, R. A. 1999. "Using the HTDP Software to Transform Spatial Coordinates Across Time and Reference Frames." *Surveying and Land Information Systems* 59 (No. 1): 15.

Snyder, J. P. 1987. *Map Projections—A Working Manual*. Washington, DC: U.S. Government Printing Office.

Stachurski, R. 2002. "History of American Projections: The American Projection, Part 1." *Professional Surveyor* 22 (No. 4): 16.

______. 2002. "History of American Projections: The American Projection, Part 2." *Professional Surveyor* 22 (No. 5): 32.

Stem, J. E. 1989. "State Plane Coordinate System of 1983." *NOAA Manual NOS NGS 5*. Rockville, MD: National Geodetic Information Center.

21 Boundary Surveys

■ 21.1 INTRODUCTION

The oldest types of surveys in recorded history are boundary surveys, which date back to about 1400 B.C. when plots of ground were subdivided in Egypt for taxation purposes. Boundary surveys still are one of the main areas of surveying practice. From Biblical times[1] when the death penalty was assessed for destroying corners, to the colonial days of George Washington[2] who was licensed as a land surveyor by William and Mary College of Virginia, and through the years to the present, natural objects (i.e., trees, rivers, rock outcrops, etc.) and man-made objects (i.e., fences, wooden posts, iron, steel or concrete markers, etc.) have been used to identify land parcel boundaries.

As property increased in value and owners disputed rights to land, the importance of more accurate surveys, monumentation of the boundaries, and written records became obvious. When Texas became a state in 1845, its public domain amounted to about 172,700,000 acres, which the U.S. government could have acquired by payment of the approximately $13,000,000 in debts accumulated by the Republic of Texas. However, Congress allowed the Texans to retain their land and pay their own debts—a good bargain at roughly 7.6 cents/acre!

The term *land tenure system* applies to the manner in which rights to land are held in any given country. Such a system, as a minimum, must provide (1) a means for transferring or changing the title and rights to the land, (2) permanently monumented or marked boundaries that enable parcels to be found on the ground, (3) officially retained records defining who possesses what rights to the land, and (4) an official legal description of each parcel. In the United States

[1]"Cursed be he that removeth his neighbor's landmark. And all the people shall say Amen." Deut. 27:17.

[2]"Mark well the land, it is our most valuable asset." George Washington.

a two-tier land tenure system exists. At the federal level, records of surveys and rights to federal land are maintained by the U.S. *Bureau of Land Management* (*BLM*). At the state and local levels, official records concerning land tenure are held in county courthouses.

Land titles in the United States are now transferred by written documents called *deeds* (*warranty*, *quitclaim*, or *agreement*), which contain a description of the property. Property descriptions are prepared as the result of a land survey. The various methods of description include (1) metes and bounds, (2) block-and-lot number, (3) coordinate values for each corner, and (4) township, section, and smaller subdivisions of the U.S. *Public Land Survey System* (*PLSS*) commonly referred to as the *aliquot part*. Often a property description will combine two or more of these methods. The first three methods are discussed briefly in this chapter; the PLSS is covered in Chapter 22.

■ 21.2 CATEGORIES OF LAND SURVEYS

Activities involved in the practice of land surveying can be classified into three categories: (1) *original surveys* to subdivide the remaining unsurveyed U.S. public lands, most of which are in Alaska, (2) *retracement surveys* to recover and monument or mark boundary lines that were previously surveyed, and (3) *subdivision surveys* to establish new smaller parcels of land within lands already surveyed. The last two categories are described in this chapter; the first is discussed in Chapter 22.

In establishing new property lines, and especially in retracing old ones, surveyors must exercise acute judgment based on education, practical experience, and knowledge of land laws. They must also be accurate and articulate in making observations. This background must be bolstered by tenacity in searching the records of all adjacent property as well as studying descriptions of the land in question. In fieldwork, surveyors must be untiring in their efforts to find points called for by the deed. Often it is necessary to obtain *parol evidence*, that is, testimony from people who have knowledge of accepted land lines and the location of corners, reference points, fences, and other information about the correct lines.

Modern-day land surveyors are confronted with a multitude of problems created over the past two centuries under different technology and legal systems that now require professional solutions. These include defective compass and chain surveys; incompatible descriptions and plats of common lines for adjacent tracts; lost or obliterated corners and reference marks; discordant testimony by local residents; questions of riparian rights; and a tremendous number of legal decisions on cases involving property boundaries.

The responsibility of a professional surveyor is to weigh all evidence and try to establish the originally intended boundary between the parties involved in any property-line dispute, although without legal authority to force a compromise or settlement. Fixing title boundaries must be done by agreement of adjacent owners or court action. Surveyors are often called upon to serve as expert witnesses in proceedings to establish boundaries, but to do so they should be registered.

Because of the complicated technical judgment decisions that must be made, the increasing cost of land surveyors' professional liability insurance for "errors and omissions" has become a major part of operating expenses. Some states demand that a surveyor have it for the protection of all parties.

■ 21.3 HISTORICAL PERSPECTIVES

In the eastern part of the United States, individuals acquired the first land titles by gifts or purchase from the English Crown. Surveys and maps were completely lacking or inadequate, and descriptions could be given in only general terms. The remaining land in the 13 colonies was transferred to the states at the close of the Revolutionary War. Later this land was parceled out to individuals, generally in irregular tracts. Boundary lines were described by *metes and bounds* (see Section 21.4).

Many original transfers and subsequent ownerships and subdivisions were not recorded. Those that were usually had scanty or defective descriptions, since land was cheap and abundant. Trees, rocks, and natural landmarks defining the corners, as in the first example metes-and-bounds description (see Section 21.4), were soon disturbed. The intersection of two property lines might be described only as "the place where John killed a bear" or "the bend in a footpath from Jones's cabin to the river."

Numerous problems in land surveying stem from the confusion engendered by early property titles, descriptions, and compass surveys. The locations of thousands of corners have been established by compromise after resurveys or by court interpretation of all available evidence pertinent to their original or intended positions. *Squatters' rights*, *adverse possession*, and *riparian changes* have fixed other corners. Many boundaries are still in doubt, particularly in areas having marginal land where the cost of a good retracement survey exceeds the property's value.

The fact that four corners of a field can be found and the distances between them agree with the "calls" in a description does not necessarily mean that they are in the proper place. Title or ownership is complete only when the land covered by a deed is positively identified and located on the ground.

Land law from the time of the Constitution has been held as a state's right, subject to interpretation by state court systems. Many millions of land parcels have been created in the United States over the past four centuries under different technology and legal systems. Some of the countless problems passed on to today's professional surveyors, equipped with immensely improved equipment, are discussed in this chapter and in Chapter 22.

Land surveying measurements and analysis follow basic plane surveying principles. But a land surveyor needs years of experience in a given state to become familiar with local conditions, basic reference points, and legal interpretations of complicated boundary problems. Methods used in one state for prorating differences between recorded and measured distances may not be acceptable in another. Rules on when and how fences determine property lines are not the same in all states or even in adjacent ones.

The term *practical location* is used by the legal profession to describe an agreement, either explicit or implied, in which two adjoining property owners mark out an ambiguous boundary, or settle a boundary dispute. Fixed principles enter the process and the boundary established, called an *agreed-upon boundary*, can become permanent.

Different interpretations are given locally to (1) the superiority or definiteness of one distance over another associated with it, (2) the position of boundaries shown by occupancy, (3) the value of corners in place in a tract and its subdivisions, and (4) many other factors. Registration of land surveyors is therefore required in all states to protect the public interest.

21.4 PROPERTY DESCRIPTION BY METES AND BOUNDS

As noted earlier, metes and bounds is one of the methods commonly used in preparing legal descriptions of property. Descriptions by metes (to measure, or assign by measure) and bounds (boundary lines or property limits) have a *point of commencement* (POC) such as a nearby existing corner of the PLSS. Commencing at this point, successive lengths and directions of lines are given that lead to the *point of beginning* (POB). The POB is usually a fence post, iron or steel rod, or some natural feature, which marks one corner of the property. Lengths and directions (bearings or azimuths) of successive lines from the point of beginning that enclose or bound the property are then given. Early distance units of chains, poles, and rods are now replaced by feet and decimals and sometimes by metric units. Bearings or azimuths may be geodetic, astronomic, magnetic, or grid. Care must be exercised to indicate clearly which of these is the basis of directions so that no confusion arises. In the past, *assumed* bearings or azimuths have sometimes been used, but many states no longer allow them because they are not readily reproducible. In some states, survey regulations call for exterior lines of new subdivisions to be based on the true meridian.

Surveyors write metes-and-bounds property descriptions, and they are included in the legal documents that accompany the transfer of title to property. In preparing the descriptions, extreme care must be exercised. A single mistake in transcribing a numerical value, or one incorrect or misplaced word or punctuation mark, may result in litigation for more than a generation, since the intentions of the *grantor* (person selling property) and the *grantee* (person buying property) may then be unclear. If numbers are both spelled out and given as figures, words control in the case of conflicts unless other proof is available. There is a greater likelihood of transposing than misspelling—and lawyers prefer words!

The importance of permanent monuments to mark property is evident. In fact, some states require pipes, iron pins, and/or concrete markers set deep enough to reach below the frost line at all property corners before surveys will be accepted for recording. Actually, almost any suitable marker could be called for as a monument. A map attached to the description will contain a legend, which identifies all monuments. By scaling from the map, a rough check on the distances and directions in the description can be obtained.

To increase precision in property surveys and to simplify the process of preparing property descriptions and relocating the corners of established parcels, large cities and some states have established a network of control monuments to supplement stations of the National Spatial Reference System (NSRS). These points are then available as control for initiating topographic and construction surveys and can also serve as POCs for commencing property descriptions. Some states have embarked on statewide programs in which a relatively dense grid of control monuments is being set on a county-by-county basis. For example, the State of Florida established a 6-mi grid. State plane coordinates are determined for the monumented positions, and they are tied to corners of the PLSS.

Description of land by metes and bounds in a deed should always contain the following information in addition to the recital:

1. *Point of commencement (POC)*. This is an established reference point such as a corner of the PLSS or NSRS monument to which the property description is tied or referenced. It serves as the starting point for the description.
2. *Point of beginning (POB)*. This point must be identifiable, permanent, well referenced, and one of the property corners. Coordinates, preferably state plane, should be given if known or computable. Note that a POB is no more important than others and a called-for monument in place at the next corner establishes its position, even though bearing and distance calls to it may not agree.
3. *Definite corners*. Such corners are clearly defined points with coordinates if possible.
4. *Lengths and directions of the property sides*. All lengths in feet and decimals (or metric units), and directions by angles, true bearings, or azimuths must be stated to permit computation of any misclosure error. Omitting the length or direction of a closing line to the POB and substituting a phrase "and thence to the point of beginning" is not acceptable. The survey date is required and particularly important if directions are referred to magnetic north.
5. *Names of adjoining property owners*. These are helpful to show the intent of a deed in case an error in the description leaves a gap or creates an overlap. However, called-for monuments in place will control title over calls for adjoiners unless precluded by *senior rights* (see Section 21.7).
6. *Areas*. The included area is normally given as an aid in the valuation and identification of a piece of property. Areas of rural land are given in acres or hectares and those of city lots in square feet or square meters. Because of differences in measurements and depending on the adjustment method used for a traverse (compass rule, least squares, etc.), one surveyor's calculated area, directions, and distances may differ slightly from another's. The expression "more or less," which may follow a computed area, allows for minor errors, and avoids nuisance suits for insignificant variations. Additionally as discussed in Section 12.5, random errors propagate from the surveying observations to the computed area. As a rule of thumb, the uncertainty in the area of a parcel can be computed using Equation (12.9). The number of significant figures in the area of a parcel should not be expressed with too many significant figures.

Note that items 4, 5, and 6 above establish the *shape*, *intent*, and *size*, respectively, of the parcel. Without item 5, the description would be strictly "metes" and would lack intent. Also if items 4 and 6 were excluded, the description would be purely "bounds" and would lack the important measurements needed to establish

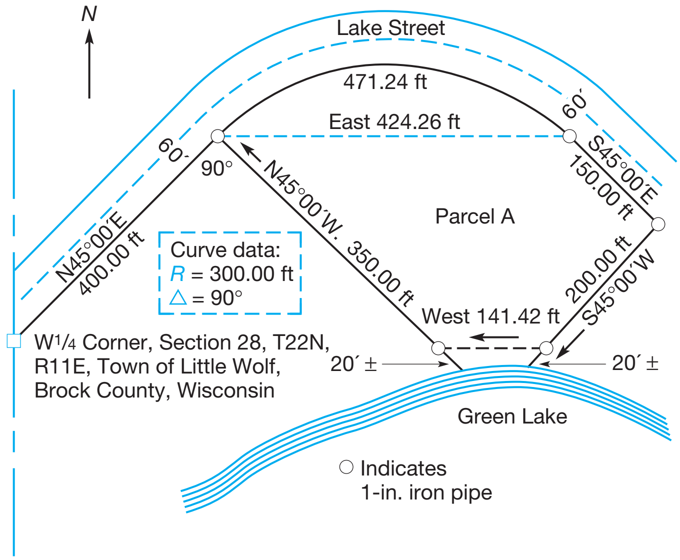


Figure 21.1 Metes-and-bounds tract.

the shape and size of the parcel. A partial metes-and-bounds description for the tract shown in Figure 21.1 is given as an example.

> That part of the SW 1/4 of the NW 1/4 of Section 28, T 22 N, R 11 E, 4'th P.M., Town of Little Wolf, Brock County, Wisconsin, described as follows: Starting at the point of commencement, which is a stone monument at the W 1/4 corner of said Section 28; thence N 45 degrees 00 minutes E, four hundred (400.00) feet along the Southeasterly R/W line of Lake Street to a 1 inch iron pipe at the point of beginning of this description, said point also being the point of curvature of a tangent curve to the right having a central angle of 90 degrees 00 minutes and radius of three hundred (300.00) feet; thence Easterly, four hundred seventy one and twenty four hundredths (471.24) feet along the arc of the curve, the long chord of which bears East, four hundred twenty four and twenty six hundredths (424.26) feet, to a 1 inch iron pipe at the point of tangency thereof, said arc also being the aforesaid Southwesterly R/W line of Lake Street; thence continuing along the Southerly R/W line of Lake Street, S 45 degrees 00 minutes E, one hundred fifty (150.00) feet to a 1 inch iron pipe; thence S 45 degrees 00 minutes W, two hundred (200.00) feet to a 1 inch iron pipe located N 45 degrees 00 minutes E, twenty (20) feet, more or less, from the water's edge of Green Lake, and is the beginning of the meander line along the lake; thence West one hundred forty one and forty two hundredths (141.42) feet along the said meander line to a 1 inch iron pipe at the end of the meander line; said pipe being located N 45 degrees 00 minutes W, twenty (20) feet, more or less from the said water's edge; thence N 45 degrees 00 minutes W, three hundred fifty (350.00) feet to a 1 inch iron pipe at the point of beginning . . . including all lands lying between the meander line herein described and the Northerly shore of Green Lake, which lie between true extensions of the Southeasterly and Southwesterly boundary lines of the parcel herein described, said parcel containing 2.54 acres, more or less. Bearings are based on astronomic north.

Many metes-and-bounds descriptions have been prepared, which because of a variety of shortcomings have created later problems. To illustrate, portions of two *old* metes-and-bounds descriptions from the eastern United States follow. The first, part of an early deed registered in Maine, is

> Beginning at an apple tree at about 5 minutes walk from Trefethen's Landing, thence easterly to an apple tree, thence southerly to a rock, thence westerly to an apple tree, thence northerly to the point of beginning.

With numerous apple trees and an abundance of rocks in the area, the dilemma of a surveyor trying to retrace the boundaries many years later is obvious.

The second, a more typical old description of a city lot showing lack of comparable precision in angles and distances, follows:

> Beginning at a point on the west side of Beech Street marked by a brass plug set in a concrete monument located one hundred twelve and five tenths (112.5) feet southerly from a city monument No. 27 at the intersection of Beech Street and West Avenue; thence along the west line of Beech Street S 15 degrees 14 minutes 30 seconds E fifty (50) feet to a brass plug in a concrete monument; thence at right angles to Beech Street S 74 degrees 45 minutes 30 seconds W one hundred fifty (150) feet to an iron pin; thence at right angles N 15 degrees 14 minutes 30 seconds W parallel to Beech Street fifty (50) feet to an iron pin; thence at right angles N 74 degrees 45 minutes 30 seconds E one hundred fifty (150) feet to place of beginning; bounded on the north by Norton, on the east by Beech Street, on the south by Stearns, and on the west by Weston.

■ 21.5 PROPERTY DESCRIPTION BY BLOCK-AND-LOT SYSTEM

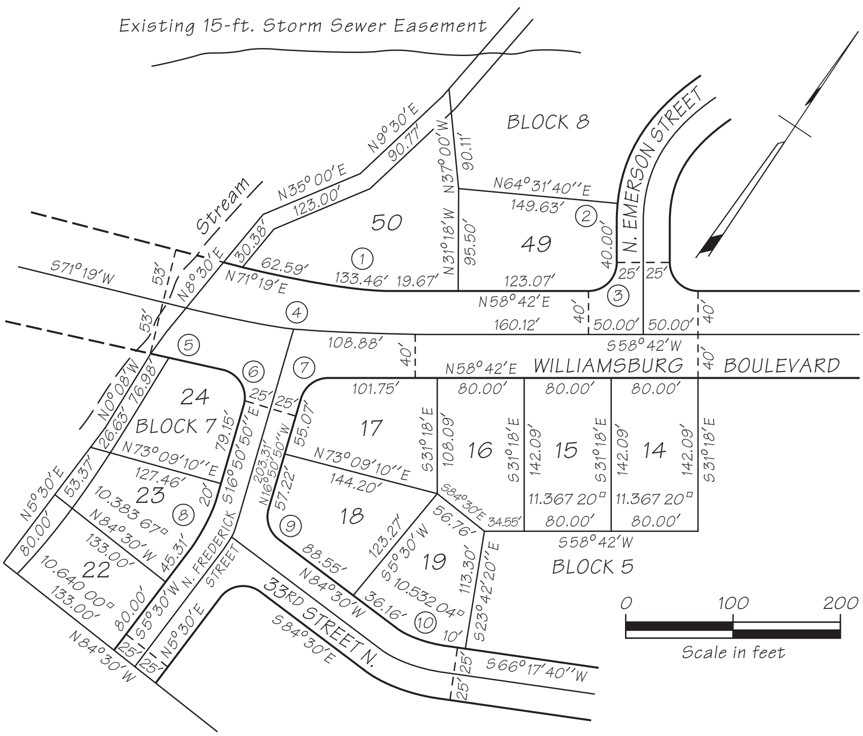
As urban areas grow, adjoining parcels of land are subdivided to create streets, blocks, and lots according to an orderly and specific plan. Each new subdivided parcel, called a *subdivision*, is assigned a name and annexed by the city. Block and lot number, tract and lot number, or subdivision name and lot number identifies the individual lots within the subdivided areas. Examples are

> Lot 34 of Tract 12314 as per map recorded in book 232, pages 23 and 24 of maps, in the office of the county recorder of Los Angeles County.
>
> Lot 9 except the North twelve (12) feet thereof, and the East twenty-six (26) feet of Lot 10, Broderick's Addition to Minneapolis. [Parts of two lots are included in the parcel described.]
>
> That portion of Lot 306 of Tract 4178 in the City of Los Angeles, as per Map recorded in Book 75, pages 30 to 32 inclusive of maps in the office of the County Recorder of said County, lying Southeasterly of a line extending Southwesterly at right angles from the Northeasterly line of said lot, from a point in said Northeasterly line distant Southeasterly twenty-three and seventy-five hundredths (23.75) feet from the most Northerly corner of said Lot.

The block-and-lot system is a short and unique method of describing property for transfer of title. Standard practice calls for a map or plat of each subdivision to be filed with the proper office. The plat must show the types and locations of monuments, dimensions of all blocks and lots, and other pertinent information such as the locations and dimensions of streets and easements, if any. These subdivision plats are usually kept in map books in the city or county recorder's office.

Lot-and-block descriptions typically are created simultaneously and thus are not subject to *junior* or *senior rights* (described in Section 21.7). In performing resurveys to find or reestablish lot corners, therefore, any excess or deficiency found in the measurements is prorated equally among the lots.

Figure 21.2 is an example of a small hand-drafted block-and-lot subdivision. Figure 18.13 shows a portion of a subdivision map of blocks and lots produced using a CADD system.



NO	DELTA (I)	TAN	ARC	RADIUS	CHORD	CHORD BEAR
1	12°37′	67.00′	133.46′	606.07′	133.19′	N65°00′30″E
2	5°49′40″	7.67′	15.33′	150.72′	15.32′	N28°23′10″W
3	90°00′	25.00′	39.27′	25.00′	35.36′	N13°42′E
4	12°37′	105.57′	210.28′	954.93′	209.85′	S65°00′30″W
5	5°50′15″	33.59′	67.11′	658.72′	67.08′	N68°23′30″E
6	97°40′27″	28.60′	42.62′	25.00′	37.64′	S65°41′05″E
7	75°32′50″	19.37′	32.96′	25.00′	30.63′	N20°55′35″E
8	22°20′50″	19.75′	39.00′	100.00′	38.76′	S5°40′25″E
9	67°39′10″	16.75′	29.52′	25.00′	27.83′	N50°40′25″W
10	29°12′20″	36.38′	71.16′	139.62′	70.41′	S80°53′50″W

Figure 21.2 Small subdivision plat.

■ 21.6 PROPERTY DESCRIPTION BY COORDINATES

State plane coordinate systems provide a common reference system for surveys in large regions, even entire states (see Chapter 20). Several advantages result from using them on property surveys. One of the most significant is that they greatly facilitate the relocation of lost and obliterated corners. Every monument, which has known state plane coordinates, becomes a "witness" to other corner markers whose positions are given in the same system. State plane coordinates also enable the evaluation of adjoiners with less fieldwork. As cities and counties develop *land information systems* (LISs) and *geographic information systems* (GISs) (see Chapter 28), descriptions by coordinates are becoming the norm. Many local and state governments currently require state plane coordinates on the corners of recorded subdivision plat boundaries, and on the exteriors of large boundary surveys. It is only a matter of time before state plane coordinates are required on all descriptions of land.

A coordinate description of corners may be used alone, but is usually prepared in conjunction with an alternative method. An example of a description by coordinates of a parcel in California follows.

> A parcel of tide and submerged land, in the state-owned bed of Seven Mile Slough, Sacramento County, California, in projected Section 10, T 3 N, R 3 E, Mt. Diablo Meridian, more particularly described as follows:
>
> BEGINNING at a point on the southerly bank of said Seven Mile Slough which bears S62°37′ E, 860 feet from a California State Lands Commission brass cap set in concrete stamped "JACK 1969," said point having coordinates of X = 2,106,973.68 and Y = 164,301.93 as shown on Record of Survey of Owl Island, filed October 6, 1969, in Book 27 of Surveys, Page 9, Sacramento County Records, thence to a point having coordinates of X = 2,107,196.04 and Y = 164,285.08; thence to a point having coordinates of X = 2,107,205.56, Y = 164,410.72; thence to a point having coordinates of X = 2,106,983.20, Y = 164,427.57; thence to the point of beginning. Coordinates, bearings, and distances in the above description are based on the California Coordinate System, Zone II.

When the preceding description was prepared, the writer could not have anticipated that an ambiguity would later arise concerning the datum of reference. From the dates given in the description, of course it can be concluded that the coordinates are referred to NAD27. In preparing coordinate descriptions nowadays, the datum upon which the coordinates are based should be identified as being either in NAD27 or NAD83, as appropriate, to avoid any confusion.

Earthquakes in Alaska, California, and Hawaii and the subsidence caused by the withdrawal of oil and groundwater in many states have caused ground shifts that move corner monuments and thereby change their coordinates. If undisturbed relative to their surroundings, the monuments, rather than the coordinates, then have greater weight in ownership rights.

■ 21.7 RETRACEMENT SURVEYS

Retracement surveys are run for the purpose of relocating or reestablishing previously surveyed boundary lines. *They are perhaps the most challenging of all types of surveys*. The rules used in retracement surveys are guided by case law and as such can vary from state to state and with time. However, the fundamental precept governing retracement surveys is that the monuments as originally placed and agreed to by the grantee and grantor constitutes the correct boundary location. *The objective of resurveys therefore is to restore boundary markers to their original locations, not to correct them,* and this should guide all of the surveyor's actions.

In making a retracement survey, written evidence of title for the parcel involved should first be obtained. This will normally be in the form of a deed, but could also be obtained from an abstract or title policy. Even if the deed is available, it is a good practice to trace it back to its creation in order to ensure that no transcription errors have been made and to check for possible modifiers (such as the term "surface measure"). Deeds of all adjoining properties should also always be obtained and matched to (1) determine if any gaps or overlaps exist and (2) understand any junior–senior rights (defined below) that might possibly apply. The possibility that the written title could be supplanted by an *unwritten conveyance* must also be investigated. In the absence of any alterations of the written title by unwritten means, an evaluation of all evidence related to the written conveyance should be made in order to properly establish the property boundaries.

Various types of evidence are considered and used when retracing boundary surveys. When conflicts exist between the different types, the order of importance, or weight, generally assigned in evaluating that evidence, is as follows:

1. *Senior rights*. When parcels of land are conveyed in sequence, the one created first (senior) receives all that was specified in the written documents, and in case of any overlap of descriptions, the second (junior) receives the remainder. In case of overlaps in other subsequent conveyances, the elder receives the benefits.
2. *Intent of the parties*. The intent of the grantee and grantor at the time of conveyancing must be considered in resurveys. Usually the best evidence of intent is contained in the written documents themselves.
3. *Call for a survey*. If the written documents describe a survey, an attempt should be made to locate the stakes or monuments placed as a result of that survey.
4. *Monuments*. If the written documents describe original monuments that were set to mark the boundaries, these must be searched out. When there are conflicts in monuments, natural ones such as trees or streams receive more weight than artificial ones such as stakes and iron pipes.
5. *Measurements*. Courts have consistently ruled that measurements called for in a description merely describe the positions of the corners. Consequently they generally receive the least weight in interpreting a conveyance. When measurements are evaluated as evidence, the order of importance that generally applies to them is (1) distance, (2) direction, (3) area, and (4) coordinates. However, in some states, the order of distances and directions is reversed.

It should be noted that the foregoing are "general rules" for evaluating conflicting evidence, but it is possible, in certain circumstances, for a supposedly inferior element to control a superior one.

Good judgment is especially important when old lines are being restored and the original corners are lost. Then all possible evidence that applies to the original location must be found, and a decision made as to what evidence is good and what can or should be discarded. All found evidence should be recorded in the resurvey field notes and reasons for using or rejecting any of it noted. Then if the survey is questioned in the courts, all actions taken can be justified. Familiarity with state and local laws, and past court decisions affecting surveys in an area, is valuable when making evaluations of evidence.

A corner that has been preserved is the best evidence of the original location of a line in question. The original notes are used as guides in locating these monuments, but a marker's actual position is the governing factor. Thus, if a monument is found that is and has been accepted for years as the location of a particular corner, the surveyor should also accept it.

A monument should not be assumed lost unless every possible source of information has been exhausted and no trace of it can be found. Even then the surveyor should hesitate before disturbing settled possessions. It may be possible, for example, that where one or more corners are lost, all concerned parties have acquiesced in lines or corners based upon some other corner or landmark. These acquiesced corners may not be the original ones, but it would be unreasonable to discredit them when the people concerned do not question them. In a legal controversy, the law as well as common sense will normally declare that a boundary line, long acquiesced in, is better evidence of where the real line should be than one established by a survey done long after the original monuments have disappeared.

In retracement surveys, a determination must be made as to whether directions cited in the documents are astronomic or magnetic. If they are magnetic, the declination at the time of the original survey must be determined so that astronomic bearings or azimuths can be determined and lines retraced following them. A good assumption to adopt in retracing lines is that the boundaries are where the description says they are. If they aren't, some indications of their actual locations will probably be found when the distances and directions in the description are followed.

Testimony of persons who remember boundary locations is always valuable, but not always reliable. Therefore, when such testimony is taken, a careful search must be made to find some corroborative evidence. Old fences or decayed wood at the location where a stake was purportedly originally set are examples of extremely valuable corroborative evidence. When two possibilities exist for establishing a boundary, evidence for rejecting one is often as important as evidence for accepting the other.

In retracement surveys, measurements made with a total station instrument between found monuments may not agree with distances on record. This situation provides a real test for surveyors. Perhaps the original chain or tape had an uncorrected systematic error, or a mistake was made. Alternatively, the markers may have been disturbed, or are the wrong ones. If differences exist between distances of record and measured values for found monuments, it may be helpful to determine a "scale factor,"[3] which relate original recorded distances to actual

[3]A "scale factor" can be obtained by carefully measuring the distance between two found monuments whose positions appear to be undisturbed and reliable, and dividing the distance between these monuments as recorded on the original survey by the measured distance. As an example, if the recorded distance between two found monuments was 750.00 ft, and the measured distance was 748.62 ft, then the scale factor would be $748.62/750.00 = 0.99816$.

measurements. In searching for monuments and evidence, this scale factor can then be applied to lay off the actual distances noted by the original surveyor.

Measurements may indicate that bearings and angles between adjacent sides also do not fit those called for in the writings. These discrepancies could be due to a faulty original compass survey, incorrect corner marks, or other causes. If, while conducting a resurvey, discrepancies are found between adjacent descriptions, which cannot be resolved, they should be called to the attention of the owners so that steps can be taken to harmonize the land actually owned with that contained in the titles. Once the location of the original boundary is decided upon, it should be carefully marked and witnessed so that it can be easily found in the future.

■ 21.8 SUBDIVISION SURVEYS

Subdivision surveys consist in establishing new smaller parcels of land within larger previously surveyed tracts. In these types of surveys, one or only a few new parcels may be created, in which case they may be described using the metes-and-bounds system. Conversely, in areas where new housing is planned, a block-and-lot subdivision survey can be conducted, thus creating many small lots simultaneously. Laws governing subdivision surveys vary from state to state and surveyors must follow them carefully in performing these types of surveys.

Whether a single new lot is created or a block-and-lot survey performed, generally a closed traverse is first run around the larger tract, with all corners being occupied if possible. Fences, trees, shrubbery, hedges, common (party) walls, and other obstacles may necessitate running the traverse either inside or outside the property. If corners cannot be occupied, stub (side shot) measurements can be made to them and their coordinates computed, from which side lengths and bearings can be computed (see Section 10.12). All measurements should be made with a precision suited to the specifications and land values.

Before traversing around the larger tract, it is necessary to first establish a reference direction for one of its lines. Also, if state plane coordinates are to be used in the new parcel's description, starting coordinates must be determined for one of the parcel's corners. A reference direction can be established by making an astronomic observation. Alternatively, as illustrated in Figure 21.3, both a reference direction and coordinates can be transferred to the parcel by including one of its lines in a separate traverse initiated at existing nearby control monuments.

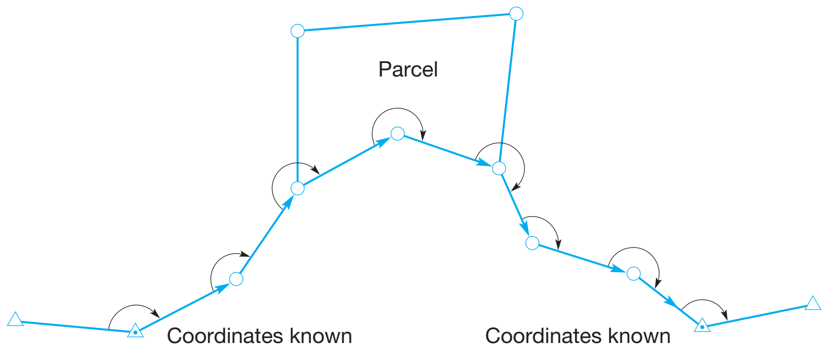


Figure 21.3 Transfer of direction and coordinates to a parcel by traverse.

A check is secured on the traverse by closing back on the starting station, or on a second nearby monument, as indicated in the figure.

After the traverse around the larger exterior parcel has been completed and adjusted, the new parcels can be surveyed. Where a single new parcel is being created, its corners are established according to the owner's specifications within requirements of the statutes. From the survey, the parcel description is prepared, certified, and recorded.

Block-and-lot subdivisions must not only conform to state statutes, but in addition many municipalities have laws covering these types of surveys. Regulations may specify minimum lot size, allowable misclosures for surveys, types of corner marks to be used, minimum width of streets and the procedure for dedicating them, rules for registry of plats, procedures for review, and other matters. Often several jurisdictions and agencies, each with its own laws and regulations, may have authority over the subdivision of land. In these cases, if standards conflict, the most stringent usually applies. The mismatched street and highway layouts of today could have largely been eliminated by suitable subdivision regulations and by a thorough review of them in past years.

A portion of a small block-and-lot subdivision is shown in Figure 21.2. (Some lot areas have been deleted so that their calculations can become end-of-chapter problems.) Computers with appropriate software such as coordinate geometry (see Chapter 11) and CADD greatly reduce the labor of computing subdivisions. They are especially valuable for large plats and designs with curved streets. Automatic plotters using stored computer files make drafting the final plat map accurate, simple, and fast.

Critical subdivision design and layout considerations include creating good building sites, an efficient street and utility layout, and assured drainage. Furthermore, subdivision rules and regulations must be followed and the developer's desires met as nearly as possible. A subdivision project involves a survey of the *exterior* boundaries of the tract to be divided, followed by a topographic survey, design of the subdivision, and layout of the interior of the tract. Following is a brief outline of the steps to be accomplished in these procedures.

1. Exterior survey
   (a) Obtain recorded deed descriptions of the parent tract of land to be subdivided, and of all adjoiners, from the Register of Deeds office. Note any discrepancies between the parent tract and its adjoiners.
   (b) Search for monuments marking corners of the parent tract, those of its adjoiner's where necessary, and, where appropriate, for U.S. Public Land Survey monuments to which the survey may be referred or tied. Resolve any discrepancies with adjoiners.
   (c) Make a closed survey of the parent tract and adequately reference it to existing monuments.
   (d) Compute departures, latitudes, and misclosures to see whether the survey meets requirements. Balance the survey if the misclosure is within allowable limits.
   (e) Resolve, if possible, any encroachments on the property, or differences between occupation lines and title lines, so there will be no problems later with the final subdivision boundary.

2. Interior survey, design, and layout
   (a) Perform a topographic survey of the area within the tract. Include existing utility infrastructure, which may extend into the subdivision. From the survey data, prepare a topographic map for evaluating drainage, and for use in preliminary street layout and lot design.
   (b) Develop a preliminary plat showing the streets and the blocks and lots into which the tract is to be subdivided. Compute departures and latitudes on every block and lot to ensure their mathematical closures.
   (c) Obtain the necessary approvals of the preliminary plat.
   (d) Prepare a final plat map that will conform to state and city platting regulations; and include whatever certificates may be required.
   (e) Set block-and-lot corners of the tract according to the approved preliminary plat. Block corners should be located first, and lot corners set by measuring along lines between block ends. (In some cases final stakeout may be delayed until utilities are placed so corners are not destroyed in this process.)
   (f) Have the certificates executed and witnessed, and the final plat approved and recorded.

■ 21.9 PARTITIONING LAND

A common problem in property surveys is partitioning land into two or more parcels for sale or distribution to family members, heirs, and so on. Prior to partitioning, a boundary survey of the parent tract is run, departures and latitudes computed, the traverse balanced, and the total enclosed area calculated. Computational procedures involved in partitioning land vary depending on conditions. Some parent tract shapes and partitioned parcel requirements permit formula solutions, often by using analytical geometry. Others require trial-and-error methods. These procedures were discussed in Chapter 11 and Section 12.8 where examples were computed to illustrate the procedures.

Typical partitioning problems consist of cutting a certain number of acres from a parent parcel, or dividing the parent parcel into halves, thirds, and so on. Required *cutoff lines* to separate a certain area from the parcel may have (1) a specific starting point (distance from one corner of the tract polygon and run to the midpoint, or any other location on the opposite side); or (2) a required direction (parallel with, perpendicular to, or on a designated bearing angle from a selected line). These cases can often be handled by trial-and-error solutions involving an initial assumption such as the cutoff line direction or the starting point. Certain problems are amenable to solution using coordinate formulas for the intersection of two lines (see Chapter 11).

Other partitioning problems may call for dividing a parcel into an easterly and westerly half, or a northerly and southerly half. But such descriptions could be ambiguous as illustrated in Figure 21.4. This figure shows that a statement "the southerly half" of tract *ABCD* can have a number of meanings—the most southerly half, half the frontage, or half the actual acreage. An important consideration in land partitioning is the final shape of each lot. Connecting midpoints *G* and *H* leaves the southerly "half" smaller than the northerly "half" but provides

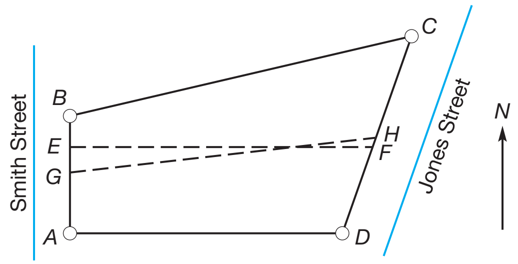


Figure 21.4
"The southerly half."

equal frontage for both parts on the two streets. Course *EF* parallel with *AD* produces one trapezoidal lot but a poorly shaped northerly parcel with meager frontage on Smith Street. In either case, the intent of the deed must be clearly stated. This is best accomplished by describing the dividing line by metes and bounds, which clearly depicts the intent.

■ 21.10 REGISTRATION OF TITLE

To remedy difficulties arising from inaccurate descriptions and disputed boundary claims, some states provide for registration of property titles under rigid rules. The usual requirements include marking each corner with standardized monuments referenced to established points, recording a plat drawn to scale, and containing specified items. The court under certain conditions then guarantees titles.

A number of states have followed Massachusetts' example and maintain separate land courts dealing exclusively with *land titles*. As the practice spreads, the accuracy of property surveys will be increased and transfer of property simplified.

Title insurance companies search, assemble, and interpret official records, laws, and court decisions affecting ownership of land, and then insure purchasers against loss regarding title defects and recorded liens, encumbrances, restrictions, assessments, and easements. Defense in lawsuits is provided by the company against threats to a clear title from claims shown in public records and not exempted in the policy. The locations of corners and lines are not guaranteed; hence it is necessary to establish, on the ground, the exact boundaries called for by the deed and title policy. Close cooperation between surveyors and title insurance companies is necessary to prevent later problems for their clients.

Many technical and legal problems are considered before title insurance is granted. In some states, title companies refuse to issue a policy covering a lot if fences in place are not on the property line and exclude from the contract "all items that would be disclosed by a property survey."

To guide surveyors in their conduct of land title survey the American Land Title Association (ALTA) and the National Society of Professional Surveyors (NSPS) have established a set of standards. Named the "ALTA/ACSM Land Title Surveys,"[4] they set forth concise guidelines on what must be included in property surveys for title insurance purposes and also state the following regarding point positional tolerance: *Relative Positional Accuracy* may be tested by (1) comparing

[4] A copy of the current ALTA/ACSM Land Title Surveys is available at http://www.acsm.net/alta.html.

the relative location of points in a survey as measured by an independent survey of higher accuracy or (2) the results of a minimally constrained, correctly weighted least-squares adjustment of the survey. *Allowable Relative Positional Accuracy for Measurements Controlling Land Boundaries on ALTA/ACSM Land Title Surveys* is $\pm[0.07 \text{ ft } (20 \text{ mm}) + 50 \text{ ppm}]$ at the 95% confidence level. Other accuracy criteria specified relate to the required precisions of instruments used and acceptable field procedures. Benefits derived from these guidelines are clarification of the exact requirements of land title surveys so that uniformly high-quality results are obtained.

■ 21.11 ADVERSE POSSESSION AND EASEMENTS

Adverse rights can generally be applied to gain ownership of property by occupying a parcel of land for a period of years specified by state law and performing certain acts.[5] To claim land or rights to it by *adverse possession*, its occupation or use must be (1) actual, (2) exclusive, (3) open and notorious, (4) hostile, and (5) continuous. It may also be necessary for the property to be held under *color of title* (a claim to a parcel of real property based on some written instrument, though a defective one). In some states all taxes must be paid. The time required to establish a claim of adverse possession varies from a minimum of 5 years in California to a maximum of 60 years for urban property in New York. The customary period is 20 years.

The occupation and use of land belonging to a neighbor but outside his or her apparent boundary line as defined by a fence may lead to a claim of adverse possession. Continuous use of a street, driveway, or footpath by an individual or the general public for a specified number of years results in the establishment of a right-of-way privilege, which cannot be withheld by the original owner.

An *easement* is a right, by grant or agreement, which allows a person or persons to use the land of another for a specific purpose. It always implies an interest in the land on which it is imposed. *Black's Law Dictionary* lists and defines 18 types of easements; hence, the exact purpose of an easement should be clearly stated. The discussion of property surveys has necessarily been condensed in this text, but it provides helpful information to readers while deterring inexperienced people from attempting to run boundary lines. For more extensive coverage, references are listed in the Bibliography.

■ 21.12 CONDOMINIUM SURVEYS

The word condominium is derived from the prefix *con* meaning "together" and from classical Roman law, *dominium* meaning "ownership." In the United States, the term *condominium* refers to a type of property ownership, where individual units within a multiple-unit building are owned separately. Every unit owner receives a deed describing their property and is able to buy, mortgage, or sell their unit independent of the other owners. Thus, legal descriptions based on surveys are required. The condominium concept of ownership is relatively new in the United

---

[5]Except in a few special circumstances, adverse rights cannot be claimed against public lands.

States, as compared with other countries. They have been present in Europe since the Middle Ages and appeared in this country in the later part of the 19th century. The number of condominiums in the United States has been growing rapidly as more families discover the many benefits this type of living offers. Condominium ownership has tax advantages, investment benefits, and most of all, eliminates rent increases. This form of ownership can be an economical solution to rising land values, building costs, and maintenance expenses. It can also provide shared recreational facilities and other amenities that might otherwise be unaffordable.

A *condominium association* is the entity responsible for the operation of a condominium and the unit owners are members of the association. The document creating the association is called the *Articles of Incorporation*, which describes the purpose, powers, and duties of the association. The *By-laws* provide for the administration of the association, including meetings, quorums, voting, and other rules. The document that establishes a condominium is known as the *Condominium Declaration*, and once it is filed in the public records, the condominium is legally created. The Declaration contains important information including a legal description of the property, descriptions of the units, designation of *common elements* (those jointly owned and used by all units such as sidewalks, stairways, swimming pool, tennis courts, etc.), and identification of *limited common elements* (those reserved for the exclusive use of a particular unit such as a designated parking space). It also describes any *covenants* or restrictions on the use of the units, common elements, and limited common elements.

Although condominium ownership often applies to multistory residential buildings, it is also used in commercial and industrial situations. The condominium concept has been applied to mobile home lots, travel trailer and camper sites, boat slips and docks, horse stables, shopping centers, and other types of properties. Special types of condominiums include: *timeshare*, in which the owner purchases an interest in a unit for a specified time period each year; *mixed-use*, which includes both residential and commercial units; and *multiple condominium community*, which is a development containing several separate condominiums that share a single common recreation area.

Condominium surveys differ from ordinary land surveys in several ways. They also bear many similarities with some property surveys, especially those for creating subdivision plats. Each state adopts statutory laws and promulgates rules that govern the procedures and requirements for creating condominiums. In many states, the statute is known as the *Condominium Act*. Preparation of the required materials for a condominium project is a joint effort, which typically includes an architect, engineer, attorney, and surveyor. The architect prepares the building plans and specifications; the engineer designs the construction plans; and the attorney creates the legal documentation for the condominium and the association. The surveyor assembles necessary information; prepares the required condominium plat, graphic plans, and descriptions; and performs the surveys needed for describing the parcel boundary and for locating the "as-built" improvements. It is important for the information shown on the graphic plans to agree with the provisions described in the Declaration.

Figures 21.5 through 21.8 illustrate a four-sheet sample set of plans for a proposed multistory residential condominium. The main purpose of these graphic

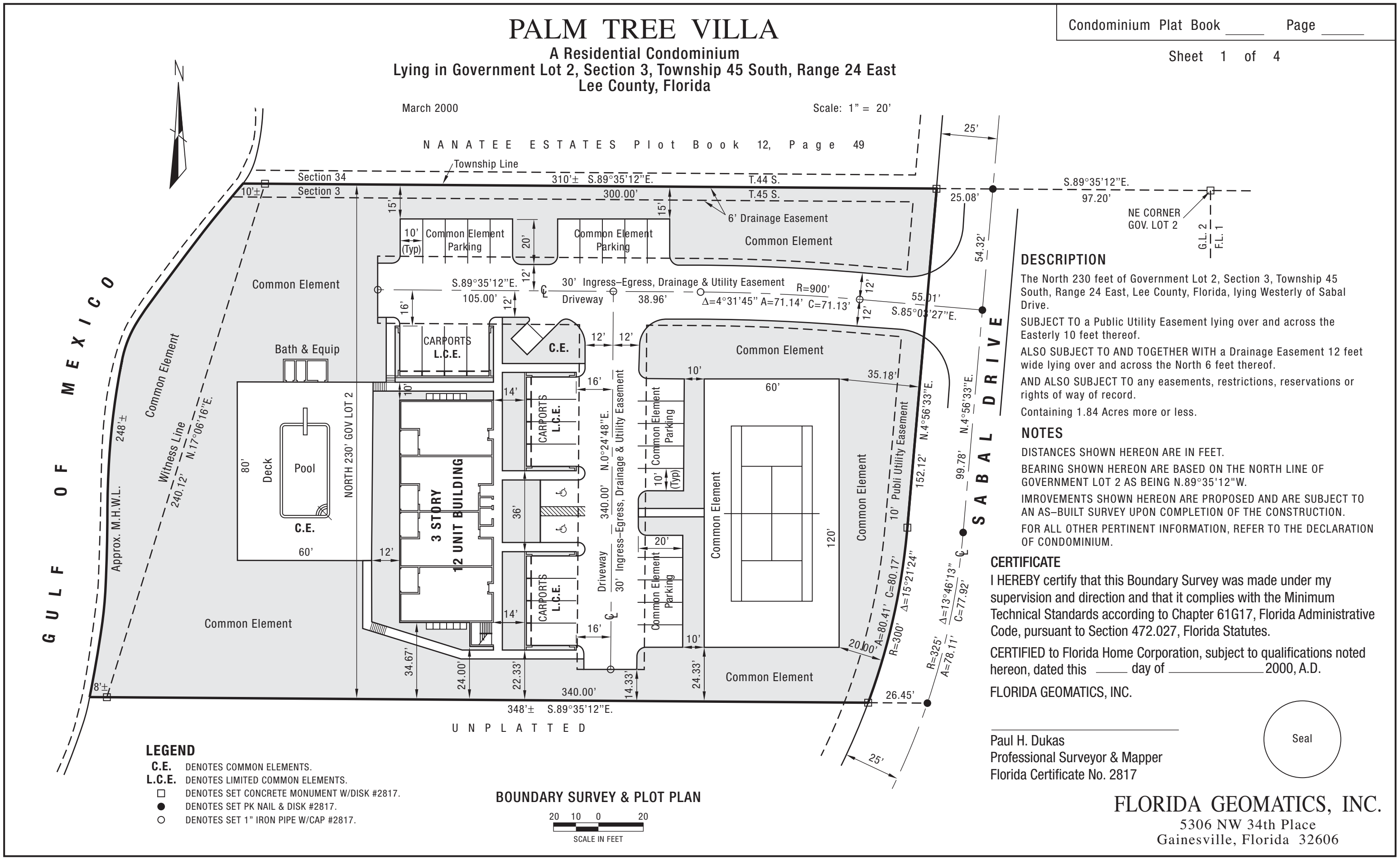


Figure 21.5 Condominium boundary survey and proposed plot plan. (Reprinted with permission from Paul Dukas, Professional Surveyor and Mapper, Geomatics Consultant.)

plans is to clearly and accurately represent the locations of the units, common areas, and limited common areas of the condominium parcel. Figure 21.5 is the *Boundary Survey and Plot Plan*. It shows the exterior boundary survey of the condominium parcel, gives a description for the parcel, and locates the proposed improvements with dimensions given from parcel boundary lines. Also included on this diagram are some general notes and the Surveyor's *Certificate* for the boundary survey. The Boundary Survey and Plot Plan is typically a small-scale drawing and is therefore generally unsuitable for showing sufficient detail and dimensions for all improvements. Thus, it is necessary to attach additional sheets at larger scales.

Figure 21.6 is the *Building and Carport Plot Plan*. It is a larger-scale graphic that not only depicts the residential building and carport areas, but also delineates the common and limited common elements and labels the individual units with identification numbers. Again, some general notes are included for clarification purposes. Although the scale of this drawing is larger than the Boundary Survey and Plot Plan, it is still too small to effectively show necessary details and dimensions for the individual unit areas. Thus, *Typical Unit Plans* are prepared at a still larger scale, as shown in Figure 21.7. These show the interior floor plans of the units and their *perimetrical boundaries* (the perimeter or horizontal dimensions encompassing the vertical planes of the interior surfaces of the walls bounding the unit). In addition, this drawing includes a *Typical Wall Section* showing the elevation of the ground floor and heights of the units and building. Typically, the elevation of the ground floor of the building is referenced to a well-established vertical datum. Also included on the plan is a description of the boundaries for the units, and some general notes.

Figure 21.8 details a *Typical Carport and Storage Plan*, with dimensions showing the sizes of the storage and parking areas assigned to individual units. A tabular list of the *Undivided Share of Common Elements* for each unit is also shown on this sample plan, but alternatively it could be included only in the declaration. Computation of the undivided shares is sometimes prorated on the basis of a unit's area to the total area of all units. Another method uses the number of bedrooms in the unit relative to the total number of bedrooms for all units. Figure 21.8 also shows an *As-Built Dimension Table* used to record the actual field measurements of selected portions of each unit or limited common element. An ideal time to measure as-built dimensions is during construction just after the exterior unit walls are completed, but before the interior room partitions are added. Not all as-built dimensions are tabularized. Some measurements, such as the building ties from the boundary lines, may be revised directly on the appropriate plan sheets to reflect the as-built location of the building. If the difference between a measured distance and its corresponding proposed distance is within the construction tolerance, then the dimension need not be revised. For example, if the measured width of a driveway pavement was within ± 0.1 ft of the proposed width, then the proposed dimension generally would not be revised. Should the size or location of an improvement differ substantially from that proposed, then it would be changed on the graphic to reflect the as-built condition.

The last item on Figure 21.8 is the sample *Surveyor's Certificate*. This is executed only upon "substantial completion" of construction of the proposed improvements. Definitions for "substantial completion" vary. As a general rule, if a local building department issues a certificate of occupancy or other similar permit,

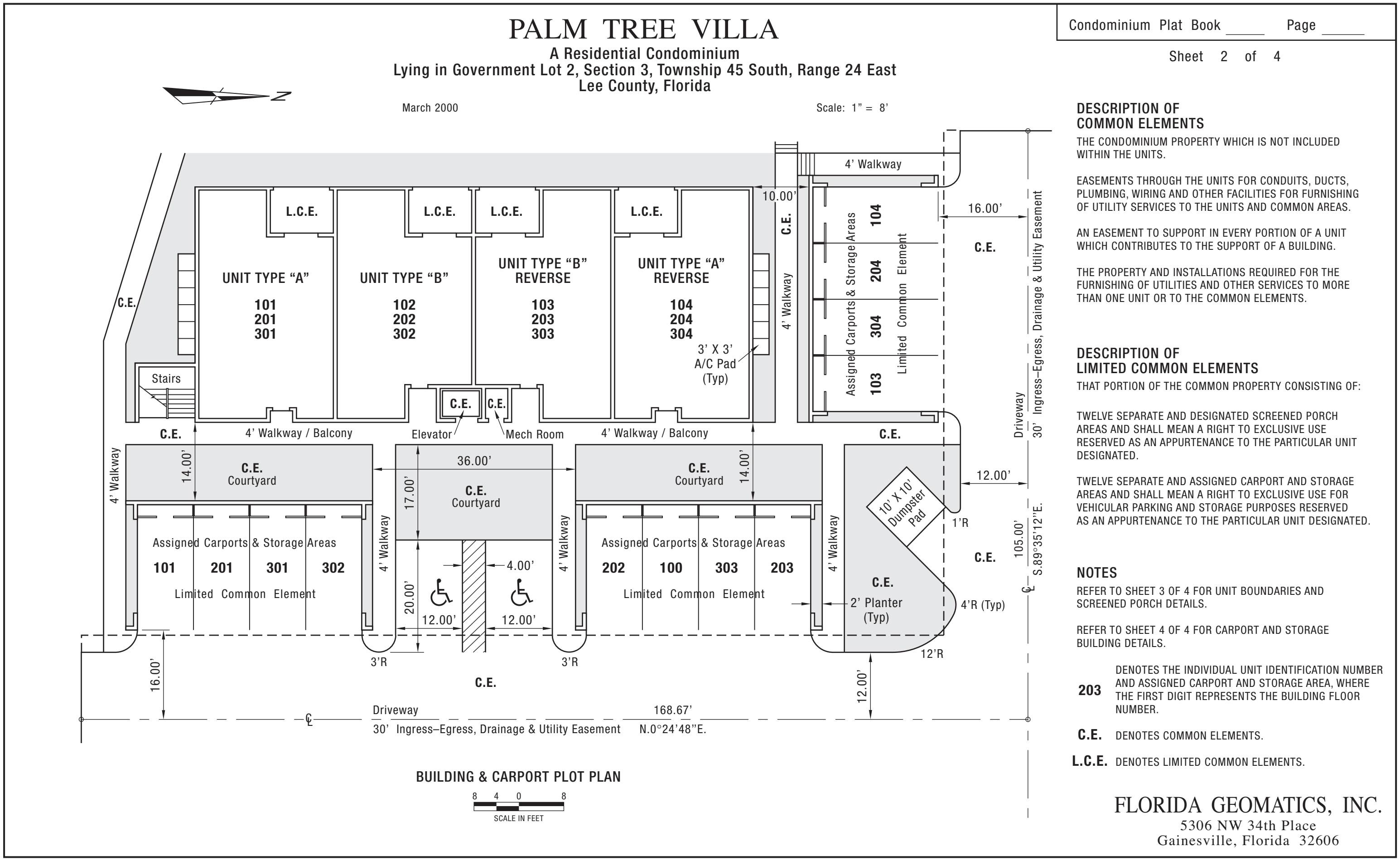


Figure 21.6 Condominium building and carport plan. (Reprinted with permission from Paul Dukas, Professional Surveyor and Mapper, Geomatics Consultant.)

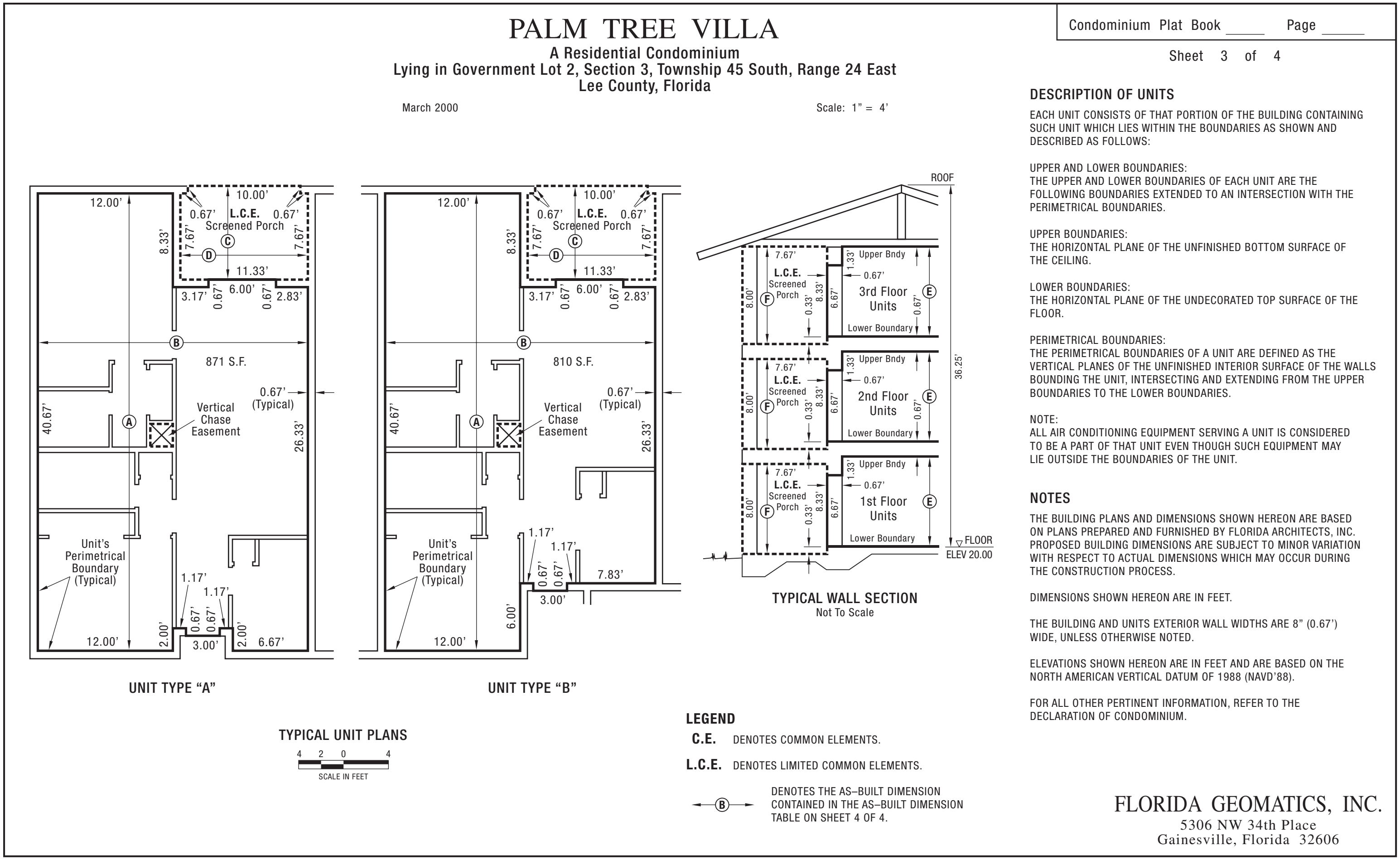


Figure 21.7 Typical unit plans and wall section. (Reprinted with permission from Paul Dukas, Professional Surveyor and Mapper, Geomatics Consultant.)

PALM TREE VILLA

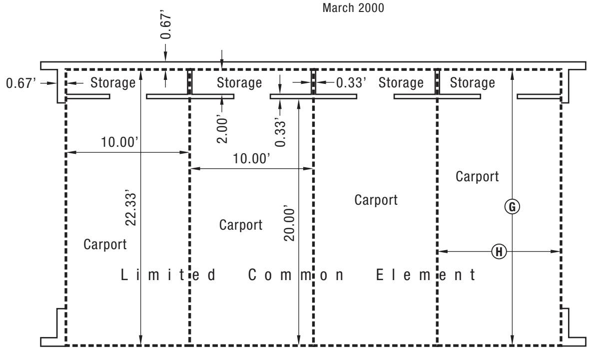
A Residential Condominium
Lying in Government Lot 2, Section 3, Township 45 South, Range 24 East
Lee County, Florida

March 2000

Scale: 1" = 4'

Condominium Plat Book _____ Page _____

Sheet 4 of 4



TYPICAL CARPORT & STORAGE PLAN

4 2 0 4
SCALE IN FEET

UNDIVIDED SHARE OF COMMON ELEMENTS

UNIT NO.	FRACTION	PERCENT
101	1/12	8 1/3%
102	1/12	8 1/3%
103	1/12	8 1/3%
104	1/12	8 1/3%
201	1/12	8 1/3%
202	1/12	8 1/3%
203	1/12	8 1/3%
204	1/12	8 1/3%
301	1/12	8 1/3%
302	1/12	8 1/3%
303	1/12	8 1/3%
304	1/12	8 1/3%
Total	12/12	100%

NOTES

PALM TREE VILLA IS A PROPOSED RESIDENTIAL CONDOMINIUM AND THE CONSTRUCTION OF THE IMPROVEMENTS IS NOT SUBSTANTIALLY COMPLETE. UPON SUBSTANTIAL COMPLETION OF THE CONSTRUCTION OF THE IMPROVEMENTS, THE DEVELOPER OR THE ASSOCIATION SHALL AMEND THE DECLARATION TO INCLUDE THE SURVEYOR'S CERTIFICATE SHOWN HEREON.

REFER TO SHEET 1 OF 4 FOR BOUNDARY SURVEY AND PLOT PLAN INFROMATION.

REFER TO SHEET 2 OF 4 FOR BUILDING AND CARPORT PLOT PLAN.

REFER TO SHEET 3 OF 4 FOR UNIT BOUNDARIES AND SCREENED PORCH DETAILS.

PROPOSED DIMENSIONS ARE SHOWN TO THE NEAREST HUNDREDTH OF A FOOT.

AS–BUILT DIMENSIONS ARE SHOWN TO THE NEAREST TENTH OF A FOOT.

THE BUILDING AND UNITS EXTERIOR WALL WIDTHS ARE 8" (0.67') WIDE, UNLESS OTHERWISE NOTED.

FOR ALL OTHER PERTINENT INFORMATION, REFER TO THE DECLARATION OF CONDOMINIUM.

C.E. DENOTES COMMON ELEMENTS.

L.C.E. DENOTES LIMITED COMMON ELEMENTS.

←B→ DENOTES THE AS–BUILT DIMENSION CONTAINED IN THE AS–BUILT TABLE SHOWN HEREON.

AS–BUILT DIMENSIONS TABLE

FLOOR	UNIT NO.	Horizontal UNIT A	Horizontal UNIT B	Horizontal PORCH C	Horizontal PORCH D	Vertical UNIT E	Vertical PORCH F	Horizontal CARPORT G	Horizontal CARPORT H	UNIT Floor Elev
1st	101									
	102									
	103									
	104									
2nd	201									
	202									
	203									
	204									
3rd	301									
	302									
	303									
	304									

SURVEYOR'S CERTIFICATE

The undersigned, a surveyor duly authorized to practice under the laws of the State of Florida, hereby certifies that the construction of the improvements is substantially complete so that the material, together with the provisions of the declaration describing the condominium property, relating to matters of survey, is an accurate representation of the location and dimensions of the common elements and of each unit, and where applicable, the limited common elements.

CERTIFIED to Florida Home Corporation, subject to qualifications noted hereon, dated this _____ day of September, 2000 A.D.

FLORIDA GEOMATICS, INC.

Seal

By: ______________________

Paul H. Dukas
Professional Surveyor & Mapper
Florida Certificate No. 2817

FLORIDA GEOMATICS, INC.
5306 NW 34th Place
Gainesville, Florida 32606

Figure 21.8 Typical carport plan, shares of common element, as-built data and certificate. (Reprinted with permission from Paul Dukas, Professional Surveyor and Mapper, Geomatics Consultant.)

then this signifies the construction of improvements as being substantially complete. However, if the certificate of occupancy is issued only for the building and a proposed improvement, for example, the pool and deck, is not substantially complete, then the surveyor's certificate should exclude those improvements from the certificate and label those amenities as "under construction" or "proposed."

Because this sample condominium is in the proposed stage, a statement of that fact is necessary. Note that this statement appears in the "Notes" of both Figures 21.5 and 21.8. For convenience many declarations include a reduction of the full-sized condominium plat and graphic plans. In this case, careful consideration of the graphic scales, text sizes, and line weights should be exercised when preparing the original full-sized drawings.

■ 21.13 GEOGRAPHIC AND LAND INFORMATION SYSTEMS

As noted in Chapter 28, surveyors are playing a major role in the development and implementation of modern *Geographic Information Systems* (GISs) and *Land Information Systems* (LISs), and this activity will continue in the future. These systems include computerized data banks that contain descriptive information about the land such as its shape, size, location, topography, ownership, easements, zoning, flood plain extent if any, land use, soil types, existence of mineral and water resources, and much more. The information is available for rapid retrieval and is invaluable to surveyors, government officials, lawyers, developers, planners, environmentalists, and others.

Boundary surveys are fundamental to GISs and LISs since knowledge about the land is meaningless unless its position on Earth is specified. In most modern systems being developed, positional data are established by associating attributes about the land with individual lots or tracts of ownership. Perhaps the most fundamental information associated with each individual parcel is its legal description, which gives the unique location on the Earth of the parcel and thus provides the positional information needed to support the system. As described in previous sections, legal descriptions are written instruments, based on measurements and prepared to exacting standards and specifications. Thus, the land surveyor's role in modern LISs and GISs is an important one.

■ 21.14 SOURCES OF ERROR IN BOUNDARY SURVEYS

Some sources of error in boundary surveys follow:

1. Errors in measured distances and directions.
2. Corners not defined by unique monuments.
3. Judgment errors in evaluating evidence.

■ 21.15 MISTAKES

Some typical mistakes in connection with boundary surveys are:

1. Failure to perform closed traverse surveys around parcels or not closing on a control station.
2. Not properly adjusting errors of closure.
3. Use of the wrong corner marks.

4. Failure to check deeds of adjacent property as well as the description of the parcel in question.
5. Failure to describe "intent" in deed descriptions, or preparing ambiguous deed descriptions.
6. Omission of the length or direction of the closing line in parcel descriptions.
7. Magnetic bearings not properly corrected to the date of the new survey.

PROBLEMS

Asterisks (*) indicate problems that have partial answers given in Appendix E.

21.1 Define the following terms:
(a) Easement
(b) Subdivision surveys
(c) Color of title
(d) Adverse possession

21.2 What is the responsibility of a surveyor in a retracement survey?

21.3 Visit your county courthouse and obtain a copy of a metes-and-bounds property description. Write a critique of the description, with suggestions on how the description could have been improved.

21.4 In a description by metes and bounds, what purpose may be served by the phrase "more or less" following the acreage?

21.5 Write a metes-and-bounds description for the exterior boundary of lot 16 in Figure 21.2.

21.6 Write a metes-and-bounds description for the house and lot where you live. Draw a map of the property.

21.7 What are the advantages of the block-and-lot system of describing property?

21.8 From the property description of the parcel of tide and submerged land described in Section 21.6, compute the parcel's area.

21.9 What is the difference between the point of commencement and point of beginning in a property description?

21.10 What major advantage does the coordinate method of property description have over other methods?

21.11 What are the advantages of the block-and-lot system of describing property?

21.12* List in their order of importance the following types of evidence when conducting retracement surveys: (a) measurements, (b) call for a survey, (c) intent of the parties, (d) monuments, and (e) senior rights.

21.13 In performing retracement surveys, list in their order of importance, the four different types of measurements called for in a description for your state.

21.14 Why do rules governing retracement surveys vary by state?

21.15 Identify all types of pertinent information or data that should appear on the plat of a completed property survey.

21.16 What are the first steps in performing a subdivision survey?

21.17 Two disputing neighbors employ a surveyor to check their boundary line. Discuss the surveyor's authority if (a) the line established is agreeable to both clients and (b) the line is not accepted by one or both of them.

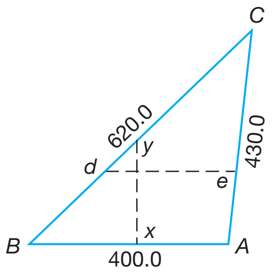
21.18 What is required to adversely possess land?

21.19 Compute the misclosures of lots 16 and 17 in Figure 21.2. On the basis of your findings, would this plat be acceptable for recording? Explain.

21.20* Compute the areas of lots 16 and 17 of Figure 21.2.

21.21 Determine the misclosure of lot 50 of Figure 21.2, and compute its area.

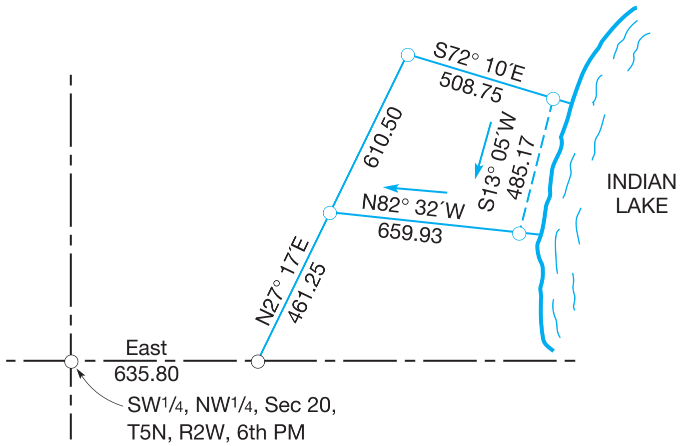
21.22 For the accompanying figure; using a line perpendicular to *AB* through *x*, divide the parcel into two equal parts, and determine lengths *xy* and *By*.



Problem 21.22

21.23 For the figure of Problem 21.22, calculate the length of line *de*, parallel to *BA*, which will divide the tract into two equal parts. Give lengths *Bd*, *de*, and *eA*.

21.24 Prepare a metes-and-bounds description for the parcel shown in the accompanying figure. Assume all corners are marked with 1-in.-diameter steel rods, and a 20-ft meander line setback from Indian Lake.

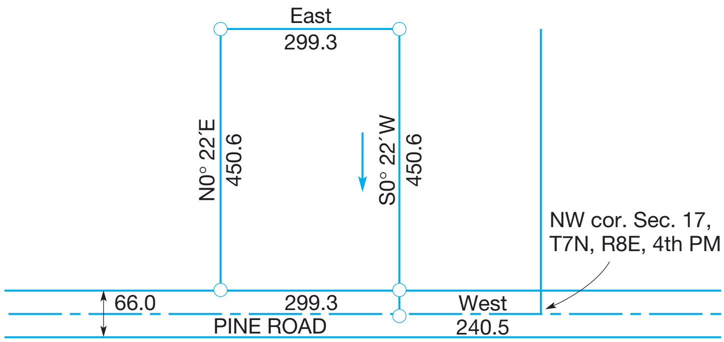


Problem 21.24

21.25 Draw a plat map of the parcel in Problem 21.24 at a convenient scale. Label all monuments and the lengths and directions of each boundary line on the drawing. Include a title, scale, north arrow, and legend.

21.26 Prepare a metes-and-bounds description for the property shown in the accompanying figure. Assume all corners are marked with 2-in.-diameter iron pipes.

21.27 Create a 1-acre tract on the westerly side of the parcel in Problem 21.26 with a line parallel to the westerly property line. Give the lengths and bearings of all lines for both new parcels.



Problem 21.26

21.28 Discuss the ownership limits of a condominium unit.

21.29 Define *common elements* and *limited common elements* in relation to condominiums. Give examples of each.

21.30 What types of measurements are typically made by surveyors in performing work for condominium developments?

BIBLIOGRAPHY

Cliff, C., et al. 2008. "Major Andrew Ellicott's Survey of the First Southern Boundary of the United States." *Surveying and Land Information Science* 68 (No. 4): 251.

Deakin, A. K. 2007. "Debating the Boundary between Geospatial Technology and Licensed Land Surveying." *Surveying and Land Information Science* 68 (No. 1): 5.

Edwards, W. D. 2009. "Oklahoma v. Texas Court Case and Texas Land Surveying." *Surveying and Land Information Science* 69 (No. 2): 129.

Gletne, J. 2008. "Changes in Riparian Boundary Location Due to Accretion, Avulsion, and Erosion." *Surveying and Land Information Science* 68 (No. 1): 47.

Kellie, A. C. 2004. "Accretion, Avulsion, and Riparian Boundaries." *Surveying and Land Information Science* 64 (No. 1): 5.

Liuzzo, T. 2007. "Encroachments: To State or Not to State." *Professional Surveyor* 4 (No. 1): 32.

Marsico, S. A. 2009. "Deeds: Types, Formalities, and Warranties." *Surveying and Land Information Science* 69 (No. 3): 121.

Miller, C. 2007. "Monuments vs. Distance and Direction." *Surveying and Land Information Science* 67 (No. 2): 101.

van der Molen, P. 2007. "Corruption and Land Administration." *Surveying and Land Information Science* 67 (No. 1): 5.

Ovans, N., et al. 2008. "The Michigan–Indiana Border Survey." *Surveying and Land Information Science* 68 (No. 4): 209.

Schultz, R. 2006. "Education in Surveying: Fundamentals of Surveying Exam." *Professional Surveyor* 26 (No. 3): 38.

Coalter, J. S. 2004. "The Fabric of Surveying in America: Surveying Texas." *American Surveyor* 1 (No. 3): 20.

U.S. Department of the Interior, Bureau of Land Management. 1973. *Manual of Surveying Instructions* 1973. Washington, DC: U.S. Government Printing Office.

Wilson, D. A. 2005. "Rules of Evidence I: Judicial Notice." *Professional Surveyor* 25 (No. 3): 51.

_____ 2005. "Rules of Evidence II: Presumptions." *Professional Surveyor* 25 (No. 5): 50.

_____ 2005. "Rules of Evidence II: Exceptions to the Hearsay Rule." *Professional Surveyor* 25 (No. 9): 48.

_____ 2005. "Rules of the Game: Rules for Investigation." *Professional Surveyor* 25 (No. 11): 46.

_____ 2006. "Second Thoughts: Undoing a Survey." *Professional Surveyor* 26 (No. 1): 43.

22 Surveys of the Public Lands

■ 22.1 INTRODUCTION

The term *public lands* is applied broadly to the areas that have been subject to administration, survey, and transfer of title to private owners under the public lands laws of the United States since 1785. These lands include those turned over to the federal government by the colonial states and the larger areas acquired by purchase from (or treaty with) the Native Americans or foreign powers that had previously exercised sovereignty.

Thirty states, including Alaska, constitute the public land survey states that have been, or are being, subdivided into rectangular tracts (see Figure 22.1). The area of these states represents approximately 72% of the United States. Title to the vacant lands, and therefore direction over the surveys within their own boundaries, was retained by the colonial states, the other New England and Atlantic coast states (except Florida), and later by the states of West Virginia, Kentucky, Tennessee, Texas, and Hawaii. In these areas the U.S. public land laws have not been applicable.

The beds of navigable bodies of water are not public domain, and are not subject to survey and disposal by the United States. Sovereignty is in the individual states.

The survey and disposition of the public lands were governed originally by two factors:

1. A recognition of the value of grid-system subdivision based on experience in the colonies and another large-scale systematic boundary survey—the 1656 Down Survey in Ireland.
2. The need of the colonies for revenue from the sale of public land. Monetary returns from their disposal were disappointing, but the planners' farsighted vision of a grid system of subdivision deserves commendation.

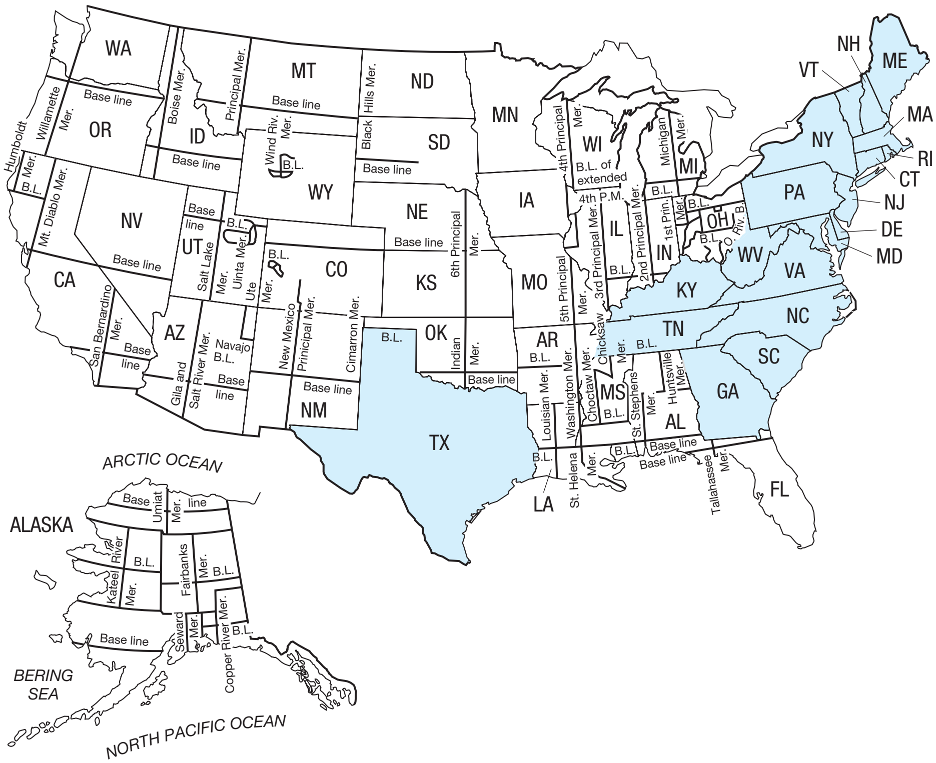


Figure 22.1 Areas covered by the public land surveys, with principal meridian shown. Areas excluded are shaded. (Hawaii, although not shown on this map, would also be shaded. Texas has a rectangular system similar to the U.S. public land system.)

Although nearly a billion acres of public land has either been sold or granted since 1785, approximately one third of the area of the country is still federally owned. The U.S. *Bureau of Land Management* (BLM), within the Department of the Interior, was created in 1946 as a merger of the U.S. *Grazing Service* and the U.S. *General Land Office*, and is responsible for surveying and managing a significant portion of these federal lands.

22.2 INSTRUCTIONS FOR SURVEYS OF THE PUBLIC LANDS

The U.S. *Public Land Surveys System* (PLSS) was inaugurated in 1784 and the territory immediately northwest of the Ohio River in what is now eastern Ohio served as a test area. Sets of instructions for the surveys were issued in 1785 and 1796. Manuals of instructions were later issued in 1855, 1881, 1890, 1894, 1902, 1930, 1947, 1973, and 2009.

In 1796 General Rufus Putnam was appointed as the first U.S. *Surveyor General*, and at that time the numbering of sections changed to the system now in use (see the excerpt from the 1973 *Manual of Surveying Instructions* below, and Figure 22.8). Supplementary rules were promulgated by each succeeding surveyor general "according to the dictates of his own judgment" until 1836, when the General Land Office (GLO) was reorganized. Copies of changes and instructions for local use were not always preserved and sent to the GLO in Washington.

As a result, no office in the United States has a complete set of instructions under which the original surveys were supposed to have been made.

Many of the later public land surveys have been run by the procedures to be described in this chapter, or variations of them. The task of present-day surveyors consists primarily of retracing the original lines set by earlier PLSS surveyors and/or further subdividing sections. To accomplish these tasks, they must be thoroughly familiar with the rules, laws, equipment, and conditions that governed their predecessors in a given area.

Basically, the rules of survey stated in the 1973 *Manual of Surveying Instructions* are as follows:

> The public lands shall be divided by north and south lines run according to the true[1] meridian, and by others crossing them at right angles, so as to form townships six miles square.
>
> The corners of the townships must be marked with progressive numbers from the beginning; each distance of a mile between such corners must be also distinctly marked with marks different from those of the corners.
>
> The township shall be subdivided into sections, containing as nearly as may be, six hundred and forty acres each, by running parallel lines through the same from east to west and from south to north at the distance of one mile from each other (originally at the end of every two miles but amended in 1800), and marking corners at the distance of each half mile. The sections shall be numbered, respectively, beginning with the number one in the northeast section, and proceeding west and east alternately through the township with progressive numbers until the thirty-six be completed.

Additional rules of survey covering field books, subdivision of sections, adjustment for excess and deficiency, and other matters are given in the manuals and special instructions. Private surveyors, on a contract basis, were paid $2/mi of line run until 1796 and $3/mi thereafter. Sometimes the amount was adjusted in accordance with the importance of a line, the terrain, location, and other factors. From this meager fee surveyors had to pay and feed a party of at least four while on the job, and in transit to and from distant points. They had to brush out and *blaze* (mark trees by scarring the bark) the line, set corners and other marks, and provide satisfactory notes and one or more copies of completed plats. The contract system was completely discarded in 1910. Public land surveyors are now appointed.

Since meridians converge, it is evident that the requirements that "lines shall conform to the true meridians and townships shall be 6 mi square" are mathematically impossible. An elaborate system of subdivision was therefore worked out as a practical solution.

Two principles furnished the legal background for stabilizing lines on the land:

1. Boundaries of public lands established by duly appointed surveyors are unchangeable.
2. Original township and section corners established by surveyors must stand as the true corners they were intended to represent, whether or not in the place shown by the field notes.

[1] As previously noted, the true meridian may be taken as either astronomic or geodetic.

Expressed differently, the original surveyors had an official plan with detailed instructions for its layout and presumably set corners to the best of their ability. After title passed from the United States, their established corners (monuments), regardless of errors, became the lawful ones. Therefore, if monuments have disappeared, *the purpose of resurveys is to determine where they were, not where they should have been*. Correcting mistakes or errors now would disrupt too many accepted property lines and result in an unmanageable number of lawsuits.

In general, the procedure in surveying the public lands provides for the following subdivisions:

1. Division into quadrangles (tracts) approximately 24 mi on a side (after about 1840)
2. Division of quadrangles into townships (16), approximately 6 mi on a side
3. Division of townships into sections (36), approximately 1 mi square
4. Subdivision of sections (usually by local surveyors)

It will be helpful to keep in mind that the purpose of the grid system was to obtain sections 1 mi on a side. To this end, surveys proceeded from south to north and east to west, and all discrepancies were thrown into the sections bordering the north and west township boundaries to get as many regular sections as possible.

Although the general method of subdivision outlined above was normally followed, detailed procedures were altered in surveys made at different times in various areas of the country. As examples, instructions for New Mexico said only township lines were to be run where the land was deemed unfit for cultivation, and in Wisconsin the first four correction lines north of the baseline were 60 mi apart rather than 24.

Another current example relates to the surveys in Alaska, where the area's sheer vastness requires changes. When Alaska gained statehood in 1958, only 2% of its 375 million acres had been surveyed. Priorities were set for conducting the remaining surveys and plans developed that called for subdividing some 155 million acres to be transferred to that state and native Alaskans. To accelerate the project, the 18,651 townships in Alaska were first established on protraction diagrams and latitude and longitude determined for each corner. In executing the surveys, markers are being set at 2-mi intervals in most areas, and GPS is being utilized extensively. But even with this modern technology and relaxed procedures, with so much area involved, it will take many years to complete the job.

Distances given in the instructions were in *chains* and miles. The particular chain referred to is the *Gunter's chain*, which was 66 ft long. It was selected for two reasons: (1) it was the best measuring device available to surveyors in the United States at the inception of the PLSS and (2) it had a convenient relationship to the rod, mile, and acre, that is, 1 chain (ch) $=$ 4 rods, 80 ch $=$ 1 mile, and 10 square chains (ch^2) $=$ 1 acre. The Gunter's chain was introduced in this text as a unit of length in Section 2.2.

Figure 22.2 illustrates a Gunter's chain. It had 100 links, each link equal to 0.66 ft or 7.92 in. The links were made of heavy wire, had a loop at each end, and were joined together by three rings. The outside ends of the handles fastened to the end links were the 0 and 66 ft marks. Tags with one, two, three, and four teeth were fastened to the 10th, 20th, 30th, and 40th links, respectively, from both ends of

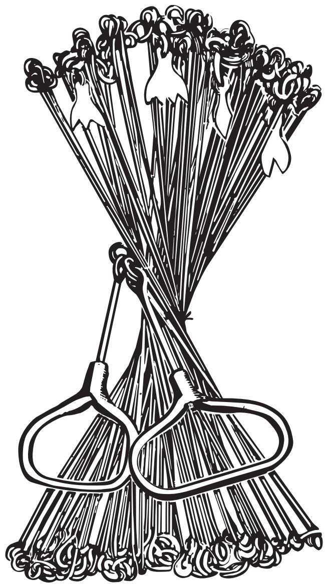


Figure 22.2
Gunter's chain.

the chain. The 50th link was marked with a round tag. These tags saved time when measuring partial chain lengths. With its many connecting link and ring surfaces subject to frictional wear, hard use elongated the chain, and its length had to be adjusted by means of bolts in the handles. Distances measured with Gunter's chains were recorded either in chains and *links* or in chains and decimals of chains—for example, 7 ch 94.5 lk or 7.945 ch. Decimal parts of links were estimated.

Gunter's chains are no longer manufactured and are seldom, if ever, used today. Nevertheless, the many chain surveys on record oblige the modern practitioner to understand the limits of accuracy possible with this equipment, and the conversion of distances recorded in chains and links to feet or meters. Descriptions of field procedures for conducting PLSS surveys given in the following sections of this chapter are taken from the *1973 Manual of Surveying Instructions*.[2] Again, because lengths are frequently given in chains, familiarity with this unit of measure is essential to understanding the material presented. Of course surveyors involved in PLSS work today would likely use either total station instruments and measure distances electronically or employ GNSS equipment, but the cited distances and the same basic principles described still apply.

■ 22.3 INITIAL POINT

Thomas Jefferson recognized the importance of property surveys and served as chairman of a committee to develop a plan for locating and selling the western lands. His report to the Continental Congress in 1784, adopted as an ordinance on

[2]At the time of this printing, the Manual of Instructions 2009 became available from the American Congress on Surveying and Mapping at http://www.acsm.net.

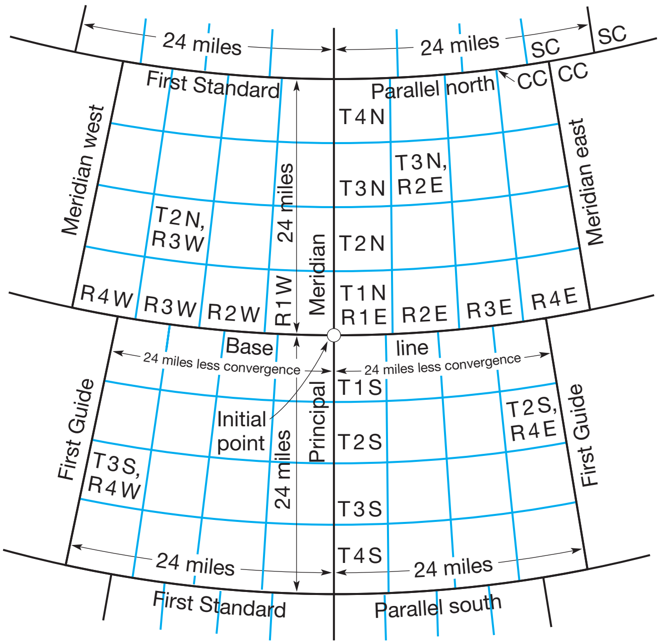


Figure 22.3 Survey of quadrangles. (Only a few of the standard corners and closing corners are identified.)

May 20, 1785, called for survey lines to be run and marked before land sales. Many of today's property disputes would have been eliminated if all property lines had been resurveyed and monuments checked and/or set before sales became final!

Subdivision of the public lands became necessary in many areas as settlers moved in and mining or other land claims were filed. The early hope that surveys would precede settlement was not fulfilled.

As settlers pressed westward, in each area where a substantial amount of surveying was needed, an *initial point* was established within the region to be surveyed. It was located by astronomic observations. The manual of 1902 was the first to specify an indestructible monument, preferably a copper bolt, firmly set in a rock ledge if possible and witnessed by rock bearings. In all, thirty-seven initial points have been set, five of them in Alaska. An initial point is illustrated near the center of Figure 22.3.

22.4 PRINCIPAL MERIDIAN

From each initial point, a true north-south line called a *principal meridian* (Prin. Mer. or PM) was run north and/or south to the limits of the area to be covered. Generally, a solar attachment—a device for solving mechanically the mathematics of the astronomical triangle—was used. Monuments were set for section and quarter-section corners every 40 ch, and at the intersections with all meanderable bodies of water (streams 3 ch or more in width, and lakes covering 25 acres or more).

The line was supposed to be within 3′ of the cardinal direction. Two independent sets of linear measurements were required to check within 20 lk (13.2 ft)/80 ch, which corresponds to a precision ratio of only 1/400. The allowable difference between sets of measurements is now limited to 7 lk/80 ch (precision ratio of 1/1140).

Areas within a principal meridian system vary greatly as depicted in Figure 22.1.

■ 22.5 BASELINE

From the initial point, a baseline was extended east and/or west as a true parallel of latitude to the limits of the area to be covered. As required on the principal meridian, monuments were set for section and quarter-section corners every 40 ch, and at the intersections with all meanderable bodies of water. Permissible closures were the same as those for the principal meridian.

Baselines are actually circular curves on the earth's surface, and were run with chords of 40 ch by the (1) solar method, (2) tangent method, or (3) secant method. These are briefly described as follows:

1. **Solar method.** An observation is made on the sun to determine the direction of astronomic north. A right angle is then turned off and a line extended 40 ch, where the process is repeated. The series of lines so established, with a slight change in direction every half mile, closely approaches a true parallel. Obviously, if the sun is obscured, the method cannot be used.
2. **Tangent method.** This method of laying out a true parallel is illustrated in Figure 22.4. A 90° angle is turned to the east or to the west, as may be required from an astronomic meridian, and corners are set every 40 ch. At the same time, proper offsets, which increase with increasing latitudes, are taken from *Standard Field Tables* issued by the BLM, and measured north from the tangent to the parallel. In the example shown, the offsets in links are 1, 2, 4, 6 1/2, 9, 12 1/2, 16 1/2, 20 1/2, 25 1/2, 31, and 37. The error resulting from taking right-angle offsets instead of offsets along the converging lines is negligible. The main objection to the tangent method is that the parallel

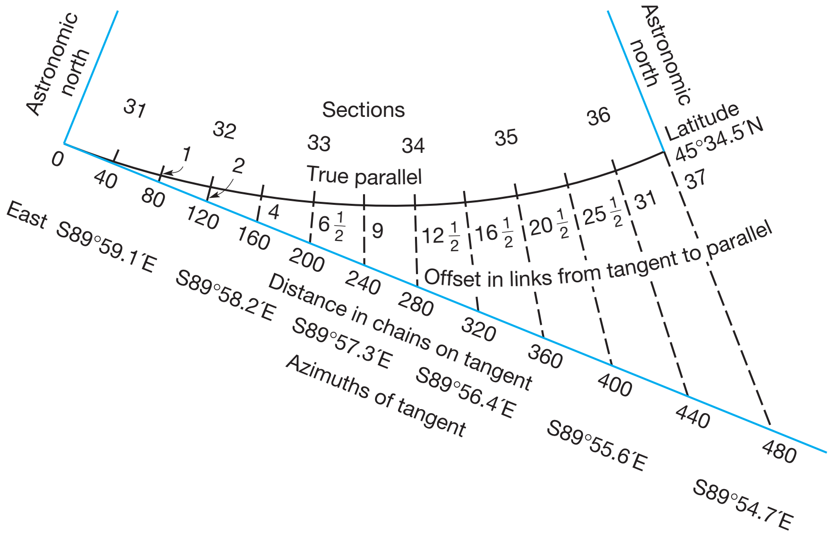


Figure 22.4 Layout of parallel by tangent method. (Adapted from 1973 *Manual of Surveying Instructions* © U.S. Department of the Interior, Bureau of Land Management.)

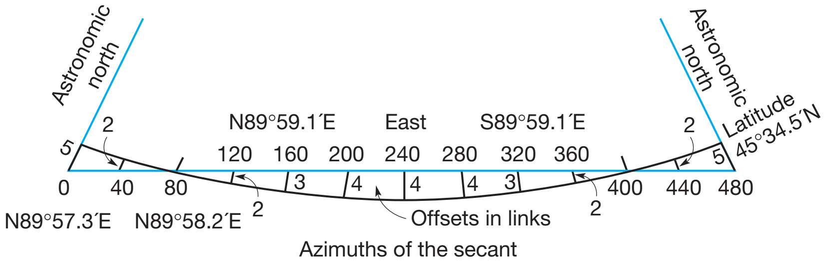


Figure 22.5
Layout of parallel by secant method. (Adapted from 1973 *Manual of Surveying Instructions* © U.S. Department of the Interior, Bureau of Land Management.)

departs considerably from the tangent, so both the tangent and the parallel must sometimes be brushed out to clear sight lines.

3. **Secant method.** This method of laying out an astronomic parallel is shown in Figure 22.5. It actually is a modification of the tangent method in which a line parallel to the tangent at the 3 mi (center) point is passed through the 1 and 5 mi points to produce minimum offsets.

Fieldwork includes establishing a point on the true meridian, south of the beginning corner, at a distance taken from the *Standard Field Tables* for the latitude of a desired parallel. The proper bearing angle from the same table is turned to the east or west from the astronomic meridian to define the secant, which is then projected 6 mi. Offsets, which also increase with increasing latitudes, are measured north or south from the secant to the parallel. Advantages of the secant method are that the offsets are small and can be measured perpendicular to the secant without appreciable error. Thus, the amount of clearing is reduced.

■ 22.6 STANDARD PARALLELS (CORRECTION LINES)

After the principal meridian and the baseline have been run, *standard parallels* (Stan. Par. or SP), also called *correction lines*, are run as true parallels of latitude 24 mi apart in the same manner as was the baseline. All 40 ch corners are marked. Standard parallels are shown in Figure 22.3. In some early surveys, standard parallels were placed at intervals of 30, 36, or 60 mi.

Standard parallels are numbered consecutively north and south of the baseline; examples are first standard parallel north and third standard parallel south.

■ 22.7 GUIDE MERIDIANS

Guide meridians (GM) are run due north (astronomic) from the baseline and the standard parallels at intervals of 24 mi east and west of the principal meridian, in the same manner as was the principal meridian, and with the same limits of error. Before work is started, the chain or tape must be checked by measuring 1 mi on the baseline or standard parallel. All 40-ch corners are marked.

Because meridians converge, a *closing corner* (CC) is set at the intersection of each guide meridian and standard parallel or baseline (see Figure 22.3). The distance from the closing corner to the *standard corner* (SC), which was set when the parallel was run, is measured and recorded in the notes as a check. Any error in the 24-mi-long guide meridian is put in the northernmost half mile.

Guide meridians are numbered consecutively east and west of the principal meridian; examples are first guide meridian west and fourth guide meridian east. Correction lines and guide meridians, established according to instructions, created *quadrangles* (or tracts) whose nominal dimensions are 24 mi on a side. These are shown in Figure 22.3.

■ 22.8 TOWNSHIP EXTERIORS, MERIDIONAL (RANGE) LINES, AND LATITUDINAL (TOWNSHIP) LINES

Division of a quadrangle, or tract, into townships is accomplished by running *range* (R) and *township* (T or Tp) lines.

Range lines are astronomic meridians through the standard township corners previously established at intervals of 6 mi on the baseline and standard parallels. They are extended north to intersect the next standard parallel or baseline and closing corners set (see Figures 22.3 and 22.6). Township lines are east-west lines that connect township corners previously established at intervals of 6 mi on the principal meridian, guide meridians, and range lines.

The angular amount by which two meridians converge is a function of latitude and the distance between the meridians. The linear amount of convergence is a function of the same two variables, plus the length that the meridians are extended. Formulas for angular and linear convergence of meridians (derived in various texts on geodesy), are as follows:

$$\theta = 52.13\ d\ \tan\phi \tag{22.1}$$

and

$$c = \frac{4}{3}\ Ld\ \tan\phi \text{ (slight approximation)} \tag{22.2}$$

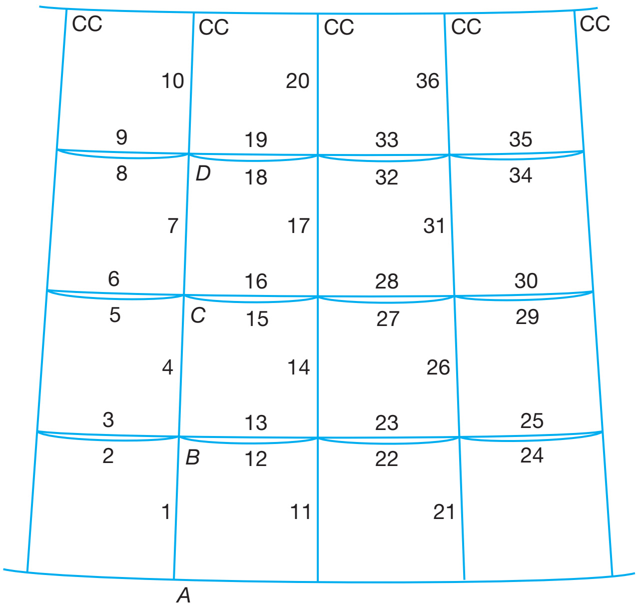


Figure 22.6 Order of running lines for the subdivision of a quadrangle into townships.

where θ is the angle of convergence (in seconds); d the distance between meridians (in miles), on a parallel; ϕ the mean latitude; c the linear convergence (in feet); and L the length of meridians (in miles).

Example 22.1

Compute the angular convergence at 40°25′ North latitude between two adjacent guide meridians.

Solution

By Equation (22.1) (guide meridians are 24 mi apart)

$$\theta = 52.13(24)\ \tan 40°25' = 1065'' = 17'45''$$

Example 22.2

Determine the distance that should exist between the standard corner and its closing corner (if there were no surveying errors) for a range line 12 mi east of the principal meridian, extended 24 mi north at a mean latitude of 43°10′.

Solution

By Equation (22.2)

$$c = \frac{4}{3}(24)\ (12)\ \tan 43°10' = 360.18 \text{ ft}$$

22.9 DESIGNATION OF TOWNSHIPS

A township is identified by a unique description based on the principal meridian governing it.

North and south rows of townships are called *ranges* and numbered in consecutive order east and west of the principal meridian as indicated in Figure 22.3.

East and west rows of townships are named *tiers* and numbered in order north and south of the baseline. By common practice, the term tier is usually replaced by the township in designating the rows.

An individual township is identified by its number north or south of the baseline, followed by the number east or west of the principal meridian. An example is Township 7 South, Range 19 East, of the Sixth Principal Meridian. Abbreviated, this becomes T 7 S, R 19 E, 6th PM.

22.10 SUBDIVISION OF A QUADRANGLE INTO TOWNSHIPS

The method to be used in subdividing a quadrangle into townships is fixed by regulations in the *Manual of Surveying Instructions*. Under the old regulations, township boundaries were required to be within 21 min of the cardinal direction.

Later this was reduced to 14 min to keep interior lines within 21 min of the cardinal direction.

The detailed procedure for subdividing a quadrangle into townships can best be described as a series of steps designed to ultimately produce the maximum number of regular sections with minimum unproductive travel by the field party. The order of running the lines is shown by consecutive numbers in Figure 22.6. Some details are described in the following steps:

1. Begin at the southeast corner of the southwest township, point *A*, after checking the chain or tape against a 1 mi measurement on the standard parallel.
2. Run north on the astronomic meridian for 6 mi (line 1 of Figure 22.6), setting alternate section and quarter-section corners every 40 ch. Set township corner *B*.
3. From *B*, run a random line (line 2 of Figure 22.6) due west to intersect the principal meridian. Set temporary corners every 40 ch.
4. If the random line has an excess or deficiency of 3 ch or less (allowing for convergence) and a falling north or south of 3 ch or less, the line is accepted. It is then corrected back (line 3), and all corners are set in their proper positions. Any excess or deficiency is thrown into the most westerly half mile. The method of correcting a random line having an excess of 1 ch and a north falling of 2 ch is shown in Figure 22.7.
5. If the random line misses the corner by more than the permissible 3 ch, all four sides of the township must be retraced.
6. The same procedure is followed until the southeast corner *D* of the most northerly township is reached. From *D*, range line 10 is continued as an astronomic meridian to intersect the standard parallel or baseline, where a closing corner is set. All excess or deficiency in the 24 mi is thrown into the most northerly half mile.
7. The second and third ranges of townships are run the same way, beginning at the south line of the quadrangle.
8. While the third range is being run, random lines are also projected to the east and corrected back and any excess or deficiency is thrown into the most westerly half mile. (On this line, all points may have to be moved diagonally to the corrected line, instead of just the last point, as in Figure 22.7.)

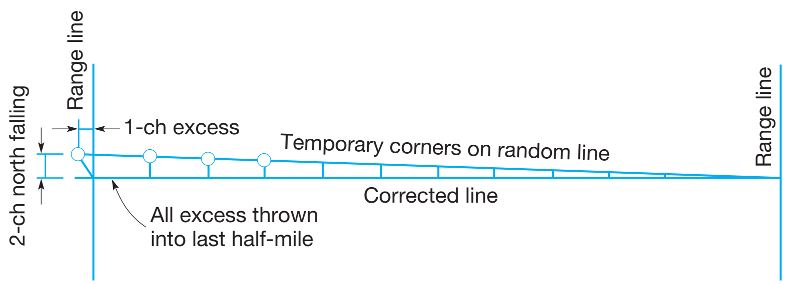


Figure 22.7 Correction of random line for excess and falling.

22.11 SUBDIVISION OF A TOWNSHIP INTO SECTIONS

Sections are now numbered from 1 to 36, beginning in the northeast corner of a township and ending in the southeast corner, as shown in Figure 22.8. The method used to subdivide a township can be described most readily as a series of steps to produce the maximum number of regular sections 1 mi on a side. Lines were run in the following order:

1. Set up at the southeast corner of the township, point *A*, and observe the astronomic meridian. Retrace the range line northward and the township line westward for 1 mi to compare the meridian, needle readings, and taped distances with those recorded.
2. From the southwest corner of section 36, run north *parallel with the east boundary of the township*. Set quarter-section and section corners on line 1 (Figure 22.8).
3. From the section corner just set, run a random line *parallel with the south boundary of the township* eastward to the range line. Set a temporary quarter corner at 40 ch.
4. If the 80 ch distance on the random line is within 50 lk, falling or distance, the line is accepted. The correct line is calculated and a quarter corner located at the *midpoint* of line *BC* connecting the previously established corner *C* and the new section corner *B*.
5. If the random line misses the corner by more than the permissible 50 lk, the township lines must be rechecked and the source of the error determined.

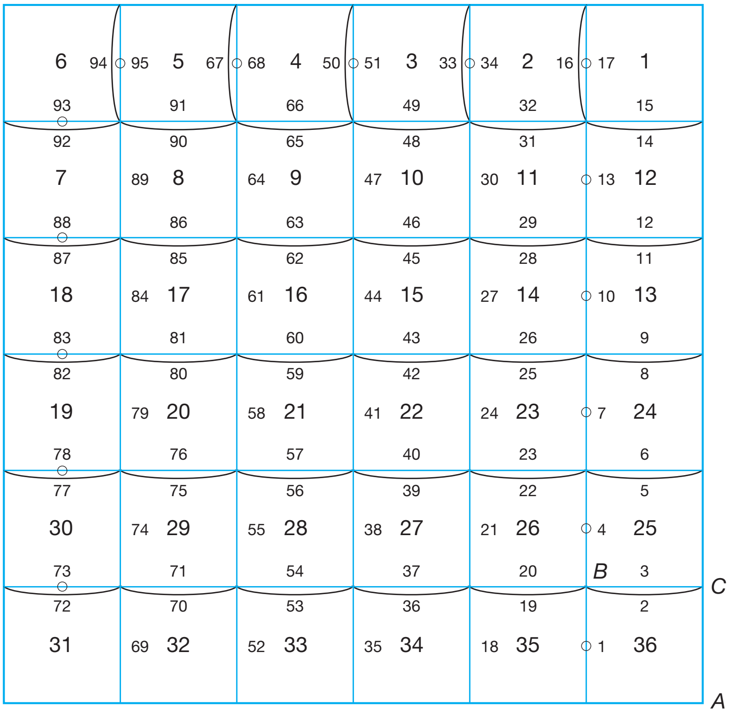


Figure 22.8
Order of running lines for the subdivision of a township into sections.

6. The east range of sections is run in a similar manner until the southwest corner of section 1 is reached. From this point a continuation of the northward line is run to connect with the north township line section corner. A quarter corner is set 40 ch from the south section corner (on line 17, corrected back by later manuals). All discrepancies in the 6 mi are thrown into the last half mile.
7. Successive ranges of sections across the township are run until the first four have been completed. All north-south lines are parallel with the township east side. All east-west lines are run randomly parallel with the south boundary line and then corrected back.
8. When the fifth range is being run, random lines are projected to the west as well as to the east. Quarter corners in the west range are set 40 ch from the east side of the section, with all excess or deficiency resulting from the errors and convergence being thrown into the most westerly half mile.
9. If the north side of the township is a standard parallel, the northward lines, which are run parallel with the east township boundary, are projected to the correction line and closing corners set. The distance to the nearest corner is measured and recorded.
10. Bearings of interior north-south section lines for any latitude can be obtained by applying corrections from tables for the convergence at a given distance from the east boundary.

By placing the effect of meridian convergence into the westernmost half mile of the township and all errors to the north and west, 25 regular sections nominally 1 mi^2 are obtained. Also, the south half of sections 1, 2, 3, 4, and 5; the east half of sections 7, 18, 19, 30, and 31; and the southeast quarter of section 6 are all regular size.

22.12 SUBDIVISION OF SECTIONS

A section was the basic unit of the General Land Office system but land was often patented in parcels smaller than a section. Local surveyors and others performed subdivision of sections as the owners took up the land. The BLM provides guidelines on the proper and intended way a section should be subdivided. To divide a section into quarter sections (nominally 160 acres), straight lines are run between opposite quarter-section corners previously established or reestablished. This rule holds whether or not the quarter-section corners are equidistant from the adjacent-section corners. Due principally to underground ore deposits, which caused large local attraction errors in compass directions, one-quarter section in a Wisconsin township contains 640 acres!

To divide a quarter section into quarter-quarter sections (nominally 40 acres), straight lines are run between opposite quarter-quarter-section corners established at the midpoints of the four sides of the quarter section. The same procedure is followed to obtain smaller subdivisions.

If the quarter sections are on the north or west side of a township, the quarter-quarter-section corners are placed 20 ch from the east or south quarter-section corners—or by single proportional measurement (see Section 22.19) on line if the total length on the ground is not equal to that on record.

22.13 FRACTIONAL SECTIONS

In sections made fractional by rivers, lakes, or other bodies of water, lots are formed bordering on the body of water and numbered consecutively through the section (see Section 8 in Figure 22.9). Boundaries of lots usually follow the quarter section and quarter-quarter-section lines, but extreme lengths or narrow widths are avoided, as are areas of fewer than 5 acres or more than 45 acres.

Quarter sections along the north and west boundaries of a township, made irregular by discrepancies of measurements and convergence of the range lines, are usually numbered as lots (see Figure 22.9). Lot lines are not actually run in the field. Like quarter-section lines, they are merely indicated on the plats by *protraction* (subdivisions of parcels on paper only). Areas needed for selling the lots are computed from the plats.

22.14 NOTES

Specimen field notes for each of the several kinds of lines to be run are shown in various instruction manuals. Actual recording had to closely follow the model sets. The original notes, or copies of them, are maintained in a land office in each state for the benefit of all interested persons.

22.15 OUTLINE OF SUBDIVISION STEPS

Pertinent points in the subdivision of quadrangles into townships, and townships into sections, are summarized in Table 22.1.

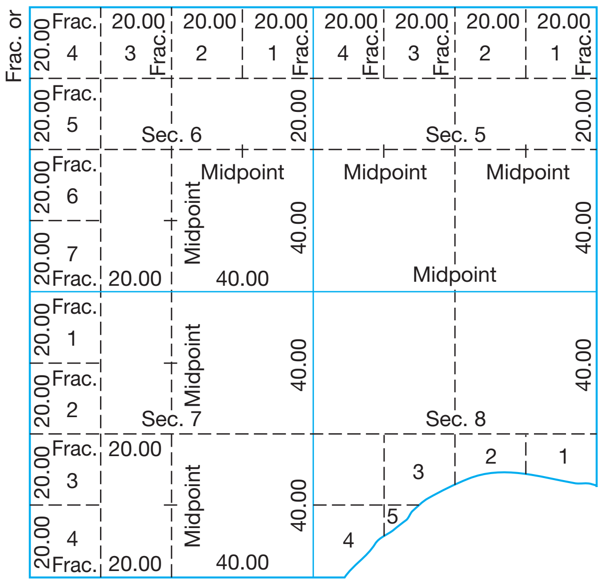


Figure 22.9
Subdivision of regular and fractional sections.

TABLE 22.1 SUBDIVISION STEPS

Item	Subdivision of a Quadrangle	Subdivision of a Township
Starting point	SE corner of SW township	SW corner of SE section (36)
Meridional lines		
Name	Range line	Section line
Direction	Astronomic North	North, parallel with east range line
Length	6 mi = 480 ch	1 mi = 80 ch
Corners set	Quarter-section and section corners at 40 and 80 ch alternately	Quarter-section corner at 40 ch; section corner at 80 ch
Latitudinal lines		
Name	Township line	Section line
Direction of random	True east-west parallel	East, parallel with south side of section
Length	6 mi less convergence	1 mi
Permissible error	3 ch, length or falling	50 lk, length of falling
Distribution of error		
Falling	Corners moved proportionately from random to true line	Corner moved proportional from random to true line
Distance	All error thrown into west quarter section	Error divided equally between quarter sections

(*Work repeated until north side of area is reached. Subdivision of last area on the north of the range of townships and sections follows.*)

Case I. When Line on the North Is a Standard Parallel

Item	Subdivision of Quadrangle	Subdivision of Township
Direction of line	Astronomic north	North, parallel with east range line
Distribution of error in length	Placed in north quarter section	Placed in north quarter section
Corner placed at end	Closing corner	Closing error
Permissible errors	Specified in *Manual of Surveying Instructions*	Specified in *Manual of Surveying Instructions*

Case II. When Line on the North Is Not a Standard Parallel

Item	Subdivision of a Quadrangle	Subdivision of a Township
Direction of line	No case	Random north and correct back to section corner already established
Distribution of error in length		Same as Case I

TABLE 22.1 (CONTINUED)

Item	Subdivision of a Quadrangle	Subdivision of a Township
[Other Ranges of Townships and Sections Continued Until All but Two are Laid Out.]		
Location of last two ranges	On east side of tract	On west side of township
Next-to-last range subdivided	As before	As before
Last range		
Direction of random	True east	Westerly, parallel with south side of section
Nominal length	6 mi less convergence	1 mi less convergence
Correction of temporary corners	Corners moved proportionately from random to true line	Corners moved proportionately from random to true line
Distribution of closure error	Corners moved westerly (or easterly) to place error in west quarter section	Corners are placed on the true line so total error falls in west quarter section

22.16 MARKING CORNERS

Various materials were approved and used for monuments in the original surveys. These included pits and mounds, stones, wooden posts, charcoal, and bottles. A zinc-coated, alloyed iron pipe with 2 1/2 in. outside diameter, 30 in. long, with brass cap is now standard. The bottom end of the pipe is split for several inches to form flanges that help hold it in place in the ground. Substitutions for this standard monument are permitted when authorized. Specially manufactured markers are now becoming commonplace. One type has a breakaway top so that if it is accidentally hit, for example by a plow or bulldozer, the upper part of the stake will break off but the lower part will remain in place. Another type uses a small, attached magnetic device to aid in recovery with a metal detector. In rock outcrops, a 3 1/4 in. diameter brass tablet with 3 1/2 in. stem is specified.

Stones and posts were marked with one to six notches on one or two faces. The arrangements identify a monument as a particular section or township corner. Each notch represents 1 mi of distance to a township line or corner. Quarter-section monuments were marked with the fraction “1/4” on a single face. In prairie country, where large stones and trees were scarce, a system of pits and mounds was used to mark corners. Different groupings of pits and mounds, 12 in. deep and 18 in. square, designated corners of several classes. However, unless some other type of mark perpetuated these markers, these corners were unfortunately lost in the first plowing.

22.17 WITNESS CORNERS

Whenever possible, monuments were witnessed by two or three adjacent objects such as trees and rock outcrops. *Bearing trees* were blazed on the side facing the corner and marked with scribing tools.

When a regular corner fell in a creek, pond, swamp, or other place where it was impractical to place a mark, *witness corners* (WC) were set on a line(s) leading to the corner. Letters WC were added to all witness corner markers, and the witness corners were also witnessed.

22.18 MEANDER CORNERS

A *meander corner* (MC) was established on survey lines intersecting the bank of a stream having a width greater than 3 ch, or a lake, bayou, or other body of water of 25 acres or more. The distance to the nearest section corner or quarter-section corner was measured and recorded in the notes. A monument was set and marked MC on the side facing the water, and the usual witness noted. If practical, the line was carried across the stream or other body of water by triangulation to another corner set in line on the farther bank.

A traverse joined successive meander corners along the banks of streams or lakes and followed as closely as practical the sinuosities of the bank. Calculating the position of the new meander corner and comparing it with its known position on a surveyed line checked the traverse.

Meander lines follow the mean high-water mark and are used only for plotting and protraction of the area. They are not boundaries defining the limits of property adjacent to the water.

22.19 LOST AND OBLITERATED CORNERS

A common problem in resurveys of the public lands is the replacement of lost or obliterated corners. This difficult task requires a combination of experience, hard work, and ample time to reestablish the location of a wooden stake or post, set perhaps 150 years ago, and with all witness trees long since cut or burned by apathetic owners.

An *obliterated corner* is one for which there are no remaining traces of the monument or its accessories, but whose location has been perpetuated or can be recovered beyond reasonable doubt. The corner may be restored from the acts or testimony of interested landowners, surveyors, qualified local authorities, witnesses, or from written evidence. Satisfactory evidence has value in the following order:

1. Evidence of the corner itself.
2. Bearing trees or other witness marks.
3. Fences, walls, or other evidence showing occupation of the property to the lines or corners.
4. Testimony of living persons.

A *lost corner* is one whose position cannot be determined beyond reasonable doubt, either from traces of the original marks or from acceptable evidence or testimony that bears on the original position. It can be restored only by rerunning lines from one or more independent corners (existing corners that were established at the same time and with the same care as the lost corner). *Proportionate measurements* distribute the excess or deficiency between a recently measured distance d

separating the nearest found monuments that straddle the lost point and the record distance D given in the original survey notes between these monuments. Then the distance x from one of the found monuments required to set the lost point is calculated by proportion as $x = X(d/D)$, where X is the record distance from that monument.

Single-proportionate measurement follows the procedure just described and is used to relocate lost corners that have a specific alignment in one direction only. These include standard corners on baselines and standard parallels, intermediate section corners on township boundaries, all quarter-section corners, and meander corners established originally on lines carried across a meanderable body of water. Corners on true lines of latitude such as baselines and standard parallels must be offset (south) from the proportion line to maintain the curvature of the latitudinal line.

Example 22.3

Figure 22.10 illustrates a lost quarter-section corner a on the line between sections 2 and 3. Section corners b and c have been found. The record distances for lines ba and ac are 40.00 and 39.57 ch, respectively. The observed distance between found corners b and c was 5246.25 ft. Describe the process for restoring lost corner a.

Solution

1. Since point a is a quarter-section corner, it is replaced on line bc by single-proportionate measurement.
2. Distance ba that must be laid off from section corner b to restore lost corner a is

$$(ba) = \frac{40}{79.57} \times 5246.25 = 2637.30 \text{ ft}$$

Double-proportionate measurements are used to establish lost corners located originally by specific alignment in two directions, such as interior section corners and corners common to four townships. The general procedure for single-proportionate measurement is used, but in two directions. It will establish two temporary points: one on the north-south line and another on the east-west line. The lost corner is then located where lines run from the two points, in the cardinal directions of north-south and east-west, intersect.

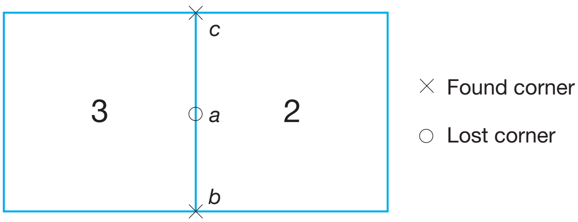


Figure 22.10 Example of single-proportionate measurement.

Example 22.4

Figure 22.11 illustrates lost corner *a* that is common to sections 22, 23, 26, and 27. Corners *b, c, d*, and *e* have been found. Record and measured distances are as follows:

Record		Measured	
Line	**Distance (ch)**	**Line**	**Distance (ft)**
ba	40.00	*bd*	7925.49
Ca	40.00	*ec*	5293.24
Da	79.20		
Ea	39.72		

Describe the process of restoring lost corner *a*.

Solution

1. Corner *a* is an interior section corner, which is constrained in alignment in two directions. Thus, it must be restored by double-proportionate measurement.
2. First establish a temporary point *f* by laying off distance *bf* along line *bd*, where *bf* is computed as

$$bf = \frac{40.00}{119.20}7925.49 = 2659.56 \text{ ft}$$

3. Then locate temporary point *g* by laying of distance *eg* along line *ec*, where *eg* is computed as

$$eg = \frac{39.72}{79.72}5293.24 = 2637.32 \text{ ft}$$

4. Establish point *h*, the restored lost corner, where an east-west line through *f* intersects a north-south line through *g*.

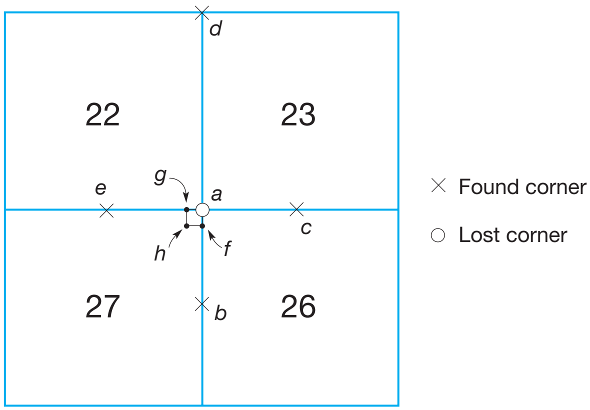


Figure 22.11 Example of double-proportionate measurement.

When the original surveys were run, *topographic calls* (distances along each line from the starting corner to natural features such as streams, swamps, and ridges) were recorded. Using the recorded distances to any of these features found today and applying single- or double-proportionate measurements to them, may help locate an obliterated corner or produce a more reliable reestablished lost corner.

■ 22.20 ACCURACY OF PUBLIC LAND SURVEYS

The accuracy required in the early surveys was a very low order. Frequently it fell below what the notes indicated. A small percentage of the surveys were made by unscrupulous surveyors drawing on their imaginations in the comparative comfort of a tent; no monuments were set, and the notes serve only to confuse the situation for present-day surveyors and land owners. A few surveyors threw in an extra chain length at intervals to assure a full measure.

Many surveys in one California county were fraudulent. In another state, a meridian 108 mi long was run without including the chain handles in its 66 ft length. When discovered, it was rerun without additional payment.

The poor results obtained in various areas were primarily caused by the following:

1. Lack of trained personnel; some contracts were given to persons with no technical training.
2. Very poor equipment by today's standards.
3. Work done by contract at low prices.
4. Surveys made in piecemeal fashion as the Indian titles and other claims were extinguished.
5. Conflicts with native people.
6. Swarms of insects, dangerous animals, and reptiles.
7. Lack of appreciation for the need to do accurate work.
8. Erratic or missing field inspection.
9. Magnitude of the problem.

In general, considering the handicaps listed, the work was reasonably well done in most cases.

■ 22.21 DESCRIPTIONS BY TOWNSHIP SECTION AND SMALLER SUBDIVISION

Description by the U.S. Public Land Surveys System offers a means of defining boundaries uniquely, clearly, and concisely. Several examples of acceptable descriptions are listed.

Sec. 6, T 7 S, R 19 E, 6th PM.
Frac. Sec. 34, T 2 N, R 5 W, Ute Prin. Mer.

The SE$\frac{1}{4}$, NE$\frac{1}{4}$, Sec. 14, Tp. 3 S, Range 22 W, SBM [San Bernardino Meridian].

$$E\frac{1}{2} \text{ of } N\frac{1}{4} \text{ of Sec. 20, T 15 N, R 10 E, Indian Prin. Mer.}$$

$$\text{E 80 acres of } NE\frac{1}{4} \text{ of Sec. 20, T 15 N, R 10 E, Indian Prin. Mer.}$$

Note that the last two descriptions do not necessarily describe the same land. A California case in point occurred when the owner of a southwest quarter section, nominally 160 acres but actually 162.3 acres, deeded the westerly portion as "the West 80 acres" and the easterly portion as "the East 1/2."

Sectional land that is privately owned may be partitioned in any legal manner at the option of the owner. The metes-and-bounds form is preferable for irregular parcels. In fact, metes and bounds are required to establish the boundaries of mineral claims and various grants and reservations.

Differences between the physical and legal (or record) ground locations and areas may result because of departures from accepted procedures in description writing, loose and ambiguous statements, or dependence on the accuracy of early surveys.

■ 22.22 BLM LAND INFORMATION SYSTEM

As noted in Section 22.1, approximately one third of the United States land area remains in the public domain. The Bureau of Land Management (BLM) is responsible for managing a significant portion of this vast acreage. Rapid access to accurate information related to this public land is in demand now more than ever before. As an example of its importance, consider that millions of dollars worth of oil and mineral royalties can he gained or lost by relatively small changes in property boundaries.

To assist in the monumental task of managing this enormous quantity of diverse land, the BLM and Forest Service provide business solutions for the management of cadastral records and land parcel information in a Land Information System (LIS—see Chapter 28) called the *National Integrated Land Survey* (NILS).[3] The LIS can be used to aid in making resource management decisions such as processing applications for mineral leases, designating utility corridors, issuing land use permits, locating wildlife habitat improvements, preparing timber sales, evaluating alternatives in environmental assessments and land use plans, and generating reports and maps.

The NILS concept provides users with tools to manage land records and cadastral data. The NILS project has four major components: (1) *survey management*, (2) *measurement management*, (3) *parcel management*, and (4) the *GeoCommunicator*. The survey management component consists of an integrated set of automation objects that will be embedded into compatible data collection software packages. This software will support the capture and input of feature measurements, and metadata (see Section 28.8) directly into a GIS database format. The measurement management system will allow users to further enhance the data set by performing a weighted least-squares adjustment (see Chapter 16) of

[3]Information on this project can be obtained at http://www.blm.gov/wo/st/en/prog/more/nils.html.

new features. This will enable creation of a higher-quality network database in both PLSS and metes-and-bounds environments. The parcel management system will provide a process for managing land records and cadastral data stored in the database model. The GeoCommunicator is a website for sharing information about data and activities of interest to land managers. This system allows users to discover information that meets their needs with the goal of GeoCommunicator to facilitate data sharing and collaborative efforts among land managers.

As part of the NILS project, the BLM created a *Geographic Coordinate Data Base* (GCDB).[4] This database contains a digital layer of information on the U.S. Public Land Surveys System (PLSS) and provides the positional components necessary for correlating all other information in the LIS. Included in the GCDB are geographic coordinates of all PLSS corners and estimates of their reliability, identifications of the surveyors who set the corners, the types of corners set and dates placed, any records of resurveys, a full record of ownership of each parcel, ownership of abutting parcels, and much more information. This effort was started in 1989 and continues today. The GCDB files for many townships are available for download via BLM websites. The national GCDB website is http://www.blm.gov/wo/st/en/prog/more/gcdb.html.

22.23 SOURCES OF ERROR

Some of the many sources of error in retracing the public land surveys are:

1. Discrepancy between length measured with an early surveyor's chain and one obtained with present-day equipment.
2. Change in magnetic declination, local attraction, or both.
3. Lack of agreement between field notes and actual measurements.
4. Change in watercourses.
5. Nonpermanent objects used for corner marks.
6. Loss of witness corners.

22.24 MISTAKES

Some typical mistakes in the retracement of boundaries in public land surveys are:

1. Failing to follow the general rules and special instructions of procedure governing the original survey.
2. Neglecting to calibrate measured lengths against record distances for marks in place.
3. Treating corners as lost when they are actually obliterated.
4. Resetting corners without exhausting every means of relocating the original corners.
5. Failing to recognize that restored corners must be placed in their original locations regardless of deviations by the original surveyor from the general rules and special instructions.

[4]More information on the GCDB can be found on the BLM web page http://www.blm.gov/wo/st/en/prog/more/gcdb.html.

PROBLEMS

Asterisks (*) indicate problems that have partial answers given in Appendix G.

22.1* Convert 65.44 chains to feet.

22.2 What two factor originally governed the survey and disposition of the public lands in the U.S.?

22.3 Who was Rufus Putnam?

22.4 Describe the tangent method of establishing a standard parallel.

22.5 Why are the boundaries of the public lands established by duly appointed surveyors unchangeable, even though incorrectly set in the original surveys?

22.6 What is the convergence in feet of meridians for the following conditions:
(a)* 24 mi apart, extended 24 mi, at mean latitude 45°20′ N.
(b) 24 mi apart, extended 24 mi, at mean latitude 34°25′ N.

22.7 What is the angular convergence, in seconds, for the two meridians defining a township exterior at a mean latitude of:
(a) 42°00′ N?
(b) 34°00′ N?

22.8 What is the nominal distance in miles between the following?
(a)* First Guide Meridian East, and the west Range Line of R8E.
(b) SE corner of Sec. 14, T 6 S, R 5 E, Indian PM, and the NW corner of Sec. 23, T 6 S, R 3 E, Indian PM.

22.9 Discuss when meander corners are to be set in a public land survey.

Sketch and label pertinent lines and legal distances, and compute nominal areas of the parcels described in Problems 22.10 through 22.12.

22.10 E 1/2, SE 1/4, Sec. 6, T 2 S, R 3 E, Salt River PM.

22.11 SW 1/4, NW 1/4, Sec. 15, T 1 N, R 2 E, Fairbanks PM.

22.12 NE 1/4, SE 1/4, SE 1/4, Sec. 30, T 1 S, R 4 E, 6th PM.

22.13 What are the nominal dimensions and acreages of the following parcels:
(a) NW 1/4, NE 1/4, Sec. 28.
(b) S 1/2, Sec. 3.
(c) SE 1/4, NE 1/4, SW 1/4, Sec. 16.

22.14 How many rods of fence are required to enclose the following:
(a)* A parcel including the NE 1/4, NE 1/4, Sec. 32, and the NW 1/4, NW 1/4, Sec. 33, T 2 N, R 3 E?
(b) A parcel consisting of Secs. 8, 9, and 16 of T 2 N, R 1 W?

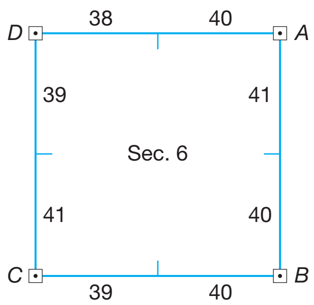
22.15 What lines of the U.S. public land system were run from astronomic observations?

22.16 In subdividing a township, which section line is run first? Which last?

22.17 Corners of the SE 1/4 of the NW 1/4 of Sec. 22 are to be monumented. If all section and quarter-section corners originally set are in place, explain the procedure to follow, and sketch all lines to be run and corners set.

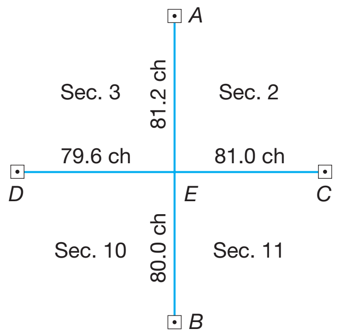
22.18 The quarter-section corner between Secs. 15 and 16 is found to be 40.28 ch from the corner common to Secs. 9, 10, 15, and 16. Where should the quarter-quarter-section corner be set along this line in subdividing Sec. 15?

22.19 As shown in the figure, in a normal township the exterior dimensions of Sec. 6 on the west, north, east, and south sides are 80, 78, 81, and 79 ch, respectively. Explain with a sketch how to divide the section into quarter sections. (See the following figure.)



Problem 22.19

22.20 The problem figure shows original record distances. Corners A, B, C, and D are found, but corner E is lost. Measured distances are $AB = 10{,}602.97$ and $CD = 10{,}718.03$ ft. Explain how to establish corner E. (See the following figure.)



Problem 22.20

To restore the corners in Problems 22.21 through 22.24, which method is used, single proportion or double proportion?

22.21* Township corners on guide meridians; section corners on range lines.

22.22 Section corners on section lines; township corners on township lines.

22.23 Quarter-section corners on range lines.

22.24 Quarter-quarter-section corners on section lines.

22.25 Why are meander lines not accepted as the boundaries defining ownership of lands adjacent to a stream or lake?

22.26 What is a meander corner?

22.27 Explain the difference between “obliterated corner” and “lost corner.”

22.28 Explain the value of using proportionate measurements from topographic calls in relocating obliterated or lost corners?

22.29 Why are the areas of many public lands sections smaller than the nominal size?

22.30 Visit the NILS website and briefly describe the four components of NILS.

22.31 Visit the BLM website at http://www.blm.gov/wo/st/en/prog/more/nils.html, and prepare a paper on the NILS project.

BIBLIOGRAPHY

Foster, R. W. 2008. "A National Cadastre for the U.S.?" *Point of Beginning* 34 (No. 3): 46.

Hansen, S. 2009. "Why a Federal Surveying Manual is Relevant to the States." *The American Surveyor* 6 (No. 8): 64.

Hedquist, B. 2006. "The National Integrated Land System." *Surveying and Land Information Science* 66 (No. 4): 279.

Kent, G. 2009. "Retracement Surveys and Undocumented Corners (Part 1 of 2)." *The American Surveyor* 6 (No. 9): 54.

Stahl, J. B. 2009. "Marking a Point." *Point of Beginning* 35 (No. 3): 28.

U.S. Department of the Interior, Bureau of Land Management. 1973. *Manual of Surveying Instructions*. Washington, DC: U.S. Government Printing Office.

U.S. Department of Interior, Bureau of Land Management. 1974. *Restoration of Lost or Obliterate Corners and Subdivision of Sections*. Washington, DC: U.S. Government Printing Office.

Zimmer, R. and S. Kirkpatrick. 2009. "GIS Data Integration with the GCDB." *The American Surveyor* 6 (No. 5): 48.

23 Construction Surveys

■ 23.1 INTRODUCTION

Construction is one of the largest industries in the United States, and thus surveying, as the basis for it, is extremely important. It is estimated that 60% of all hours spent in surveying are on location-type work, giving line and grade. Nevertheless, insufficient attention is frequently given to this type of survey.

An accurate control, topographic survey, and site map are the first requirements in designing streets, sewer and water lines, and structures. Surveyors then lay out and position these facilities according to the design plan. A final "as-built" map, incorporating any modifications made to the design plans, is prepared during and after construction and filed. Such maps are extremely important, especially where underground utilities are involved, to assure that they can be located quickly if trouble develops, and that they will not be disturbed by later improvements.

Construction surveying involves establishing both *line* and *grade* by means of stakes and reference lines that are placed on the construction site. These guide the contractor so that proposed facilities are constructed according to a plan. Placement of the stakes is most often done by making the fundamental observations of horizontal distances, horizontal and vertical angles, and differences in elevation using the basic equipment and methods described in earlier chapters of this text. However, the global navigation satellite system (GNSS) is also being used with increasing frequency for construction surveys (see Section 23.10). Other specialized equipment, including laser alignment devices and reflectorless electronic distance measuring equipment (see Section 23.2) are also used which greatly facilitates construction surveying.

All surveyors, engineers, and architects who may be involved with planning, designing, and building constructed facilities should be familiar with the fundamental procedures involved in construction surveying. In smaller projects, traditional methods of staking out the construction details are still performed. However, in larger projects, machine control using robotic total stations and/or GNSS surveying equipment has replaced many of the staking requirements. This chapter describes procedures applicable for some of the more common types of construction projects using both traditional methods and machine control methods. Chapters 24, 25, and 26 cover the subjects of horizontal curves, vertical curves, and volume computations, respectively. These topics are all pertinent to construction surveys, particularly those for transportation routes.

Common steps for surveying engineers in any construction project consist of (1) placement of horizontal and vertical control, (2) a topographic survey used in the creation of an existing surface map, (3) staking of the design, which may include physically staking the design on the ground to guide the equipment operators or calibration of the equipment and uploading and maintenance of the multiple design surfaces into a machine control system, (4) periodic checks on the layout of the design determination of quantities moved or placed during the construction process, and (5) a final as-built survey of the project. Construction surveying is perhaps best learned on the job and consists in adapting fundamental principles to the undertaking at hand. Since each project may involve unique conditions and present individual problems, coverage in this text is limited to a discussion of the fundamentals.

■ 23.2 SPECIALIZED EQUIPMENT FOR CONSTRUCTION SURVEYS

As noted above, the placement of stakes for line and grade to guide construction operations is often accomplished using the surveyor's standard equipment—levels, tapes, total station instruments, and GNSS receivers. However, there are some additional instruments that improve, simplify, and greatly increase the speed with which certain types of construction surveying are accomplished. *Visible laser-beam* alignment instruments and *reflectorless total stations* are among these. These are described briefly in the subsections that follow.

23.2.1 Visible Laser-Beam Instruments

The fundamental purpose of laser instruments is to create a visible line of known orientation, or a plane of known elevation, from which measurements for line and grade can be made. Two general types of lasers are described here.

Single-beam lasers, as shown in Figures 23.1 and 23.2, project visible reference lines ("string lines" or "plumb lines") that are utilized in linear and vertical alignment applications such as tunneling, sewer pipe placement, and building construction. The instrument shown in Figure 23.1 is a single-beam-type laser that has been combined with a total station instrument. This combination provides flexibilities that are convenient for a variety of construction layout applications. The laser beam is projected collinear with the instrument's line of sight, a feature that facilitates aligning it in prescribed horizontal alignments and/or

Figure 23.1 Laser-beam incorporated with a total station instrument. (Courtesy Leica Geosystems AG.)

Figure 23.2 Rotating laser providing finished grades to machine control system in grader. (Courtesy Topcon Positioning Systems.)

along planned grade lines. The instrument can be used to project string lines for distances up to about 1000 m. With the zenith angle set to either 90° or 270°, if the total station instrument is rotated about its vertical axis, the laser will generate a horizontal plane. Also if it is turned about its horizontal axis, the laser will define a vertical plane.

The instrument shown in Figure 23.2 projects a visible laser-beam a distance of 5 m below and 100 m above the instrument along the plumb line. These instruments are useful for alignment of objects in vertical structures. A similar type of single-beam laser projects a visible laser-beam at a selected grade—a device that is especially useful in aligning pipelines.

Rotating-beam lasers are merely single-beam lasers with spinning optics that rotate the beam in azimuth, thereby creating planes of reference. They expedite the placement of grade stakes over large areas such as airports, parking lots, and subdivisions, and are also useful for topographic mapping.

Figure 23.3 shows a rotating-beam-type laser. It projects a beam up to 350 m while rotating at 600 rpm. One or more receivers attached to grade rods or staffs can pick up the laser signal. The instrument is self-leveling and is quickly set up. If somehow bumped out of level, the laser beam shuts off and does not come back on until it is releveled. It can be operated with the laser plane oriented horizontally for setting footings, floors, etc., or the beam can be turned 90° and used vertically for plumbing walls or columns.

Because laser beams are not readily visible to the naked eye in bright sunlight, special detectors attached to a hand-held rod are often used. To lay out horizontal planes with either of these devices, the height of the instrument above datum, *HI*, must be established. Then the height on a graduated rod that a reference mark or detector must be set is the difference between the *HI* and the plane's required elevation.

Figure 23.3
Sokkia LP30 laser plane level. (Courtesy Sokkia Corporation.)

(a) (b)

Figure 23.4 Pulsed laser instruments: (a) total station and (b) hand-held EDM instrument. (Courtesy Leica Geosystems AG.)

23.2.2 Reflectorless EDMs

As described in Chapter 8, total station instruments simultaneously observe horizontal and vertical angles as well as slope distances. Their built-in microcomputers reduce the observed slope distances to their horizontal and vertical components and display the results in real time. These features make total stations very convenient for construction stakeout.

Some total stations include electronic distance measurement (EDM) that does not require a reflector. Rather, these instruments employ a laser light, which is capable of measuring distances of up to 100 m or more to objects without the use of a reflector. Alternatively, they can be used with reflectors or reflective sheets that can be applied to surfaces, a procedure that enables them to measure longer distances. Figure 23.4(a) shows a total station instrument equipped with this technology, and Figure 23.4(b) depicts a similar hand-held pulsed laser EDM instrument. Both devices are useful in observing distances to inaccessible locations, a feature that is particularly useful in assembling and checking the placement of structural members in bridges, buildings, and other large fabricated objects as well as establishing control to support laser scanning (see Sections 17.9.5 and 23.12).

■ 23.3 HORIZONTAL AND VERTICAL CONTROL

The importance of a good framework of horizontal and vertical control in a project area cannot be overemphasized. It provides the basis for positioning structures, utilities, roads, etc., in each of the stages of planning, design, and construction.

Too often surveyors and engineers have skimped on establishing a suitable network of control points, and they have also failed to preserve them through proper monumentation, references, and ties.

The surveyor in charge should receive copies of the plans well in advance of construction to become familiar with the job and have time to "tie out" or "transfer" any established control points that might be destroyed during building operations. The methods shown in Figures 9.4 and 9.5 are especially applicable and should be used with intersection angles as close to 90° as possible.

On most projects, additional horizontal and vertical control is required to supplement any control already available in the job area. The control points must be:

1. Convenient for use, that is, located sufficiently close to the item being built so that work is minimized and accuracy enhanced in transferring alignment and grade.
2. Far enough from the actual construction to ensure working room for the contractor and to avoid possible destruction of stakes.
3. Clearly marked and understood by the contractor in the absence of a surveyor.
4. Supplemented by guard stakes to deter removal and referenced to facilitate restoring them. Contracts usually require the owner to pay the cost of setting initial control points and the contractor to replace damaged or removed ones.
5. Suitable for securing the accuracy agreed on for construction layout (which may be to only the nearest foot for a manhole, 0.01 ft for an anchor bolt, or 0.001 ft for a critical feature).

Construction stakes can be set in their required horizontal positions by making observations of horizontal angles or horizontal distances from established control points. Radial stakeout by angle and distance from one control point is often most expedient, but the choice will depend on the project's nature and extent. Frequent checks should be made on points set. This can be done with observations from other control stations or by checking distances from nearby points to verify their correctness in position.

Grade stakes and reference elevations are most often set using a leveling instrument whose *HI* has been established by differential leveling. For convenience, enough benchmarks are generally placed on construction sites so that at least one is readily accessible at any location in the area. Then the *HI* of the level can be established with just a single backsight to the benchmark. After grade stakes are set, a closing foresight is taken back to the benchmark for a check. However, this practice can be dangerous since the instrument operator will have a tendency to expect the closing foresight to agree with the initial backsight, and therefore could read it carelessly. As a result, a serious mistake in the initial backsight could go undetected, resulting in a faulty setting of grade stakes. Therefore, even though it requires more time, it is recommended that level circuits for setting grade stakes always begin on one benchmark and close on another.

■ 23.4 STAKING OUT A PIPELINE

Pipelines are used to carry water for human consumption, storm water, sewage, oil, natural gas, and other fluids. Pipes, which carry storm runoff, are called *storm sewers*; those which transport sewage, *sanitary sewers*. Flow in these two types of sewers is usually by gravity, and therefore their alignments and grades must be set carefully. Flow in pipes carrying city water, oil, and natural gas is generally under pressure, so usually they need not be aligned to as high an order of accuracy.

In pipeline construction, trenches are usually opened along the required alignment to the prescribed depth (slightly undercut if pipe bedding is required), the pipe is installed according to plan, and the trench backfilled. Pipeline grades are fixed by a variety of existing conditions, topography being a critical one. A profile like that of Figure 5.12 is usually used to analyze the topography and assist in designing the grade line for each pipe segment. To minimize construction difficulties and costs, excavation depths are minimized, but at the same time a certain minimum cover over the pipeline must be maintained to protect it from damage by heavy loading from above and to prevent freezing in cold climates. Minimum grades also become an important design factor for pipes under gravity flow. Accordingly, a grade of at least 0.5% is recommended for storm sewers, but slightly higher grades are needed for sanitary sewers. In designing pipe grade lines, other existing underground elements often must be avoided, and due regard must also be given to the grades of connecting lines and the vertical clearances needed to construct manholes, catch basins, and outfalls.

Prior to staking a pipeline, the surveyor and contractor should discuss details of the project. An understanding must be reached concerning the planned trench width, where the installation equipment will be placed, and how and where the excavated material will be stockpiled. Then a reference offset line can be appropriately established that will (1) meet the contractor's needs, (2) be safe from destruction, and (3) not interfere with operations.

The alignment and grade for the pipeline are taken from the plans. An offset reference line parallel to the required centerline is established, usually at 25- or 50-ft stations when the ground is reasonably uniform. Marks should be closer together on horizontal and vertical curves than on straight segments. For pipes of large diameter, stakes may be placed for each pipe length—say, 6 or 8 ft. On hard surfaces where stakes cannot be driven, points are marked by paint, spikes, or scratch marks.

Either *batter boards* or laser beams guide precise alignment and grade for pipe placement. Figure 23.5 shows one arrangement of a batter board for a sewer line. It is constructed using 1×4 in., 1×6 in., 2×4 in., or boards nailed to 2×4 in. posts, which have been pointed and driven into the ground on either side of the trench. Depending upon conditions, these may be placed at 50 ft, 25 ft, or any other convenient distance along the sewer line. The top of the batter board is generally placed a full number of feet above the *invert* (flow line or lower inside surface) of the pipe. Nails are driven into the board tops so a string stretched tightly between them will define the pipe centerline. A graduated pole or special rod, often called a *story pole,* is used to measure the required distance from the

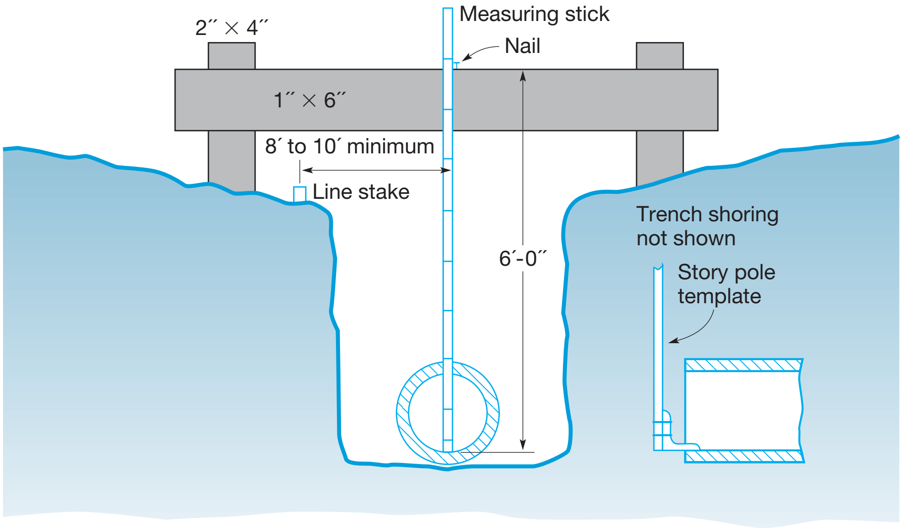


Figure 23.5 Batter board for sewer line.

string to the pipe invert. Thus, the string gives both line and grade. It can be kept taut by hanging a weight on each end after wrapping it around the nails.

In Figure 23.5, instead of a fixed batter board, a two-by-four carrying a level vial can be placed on top of the offset-line stake, which has been set at some even number of feet above the pipe's invert elevation. Measurement is then made from the underside of the leveled two-by-four with a tape or graduated pole to establish the flow line.

If laser devices are used for laying pipes, the beam is oriented along the pipe's planned horizontal alignment and grade, and the trench opened. Then with the beam set at some even number of feet above the pipe's planned invert, measurements can be made using a story pole to set the pipe segments. Thus, the laser beam is equivalent to a batter-board string line. On some jobs that have a deep wide cut, the laser instrument is set up in the trench to give line and grade for laying pipes. And, if the pipe is large enough, the laser beam can be oriented inside it.

■ 23.5 STAKING PIPELINE GRADES

Staking pipeline grades is essentially the reverse of running profiles, although in both operations the centerline must first be marked and stationed in horizontal location. The actual profiling and staking are on an offset line.

Information conveyed to the contractor on stakes for laying pipelines usually consists of two parts: (1) giving the depth of cut (or fill), normally only to the nearest 0.1 ft, to enable a rough trench to be excavated; and (2) providing precise grade information, generally to the nearest 0.01 ft, to guide in the actual placement of the pipe invert at its planned elevation. Cut (or fill) values for the first part are vertical distances from ground elevation at the offset stakes to the pipe invert. After the pipe's grade line has been computed and the offset line run, cuts (or fills) can be determined by a leveling process, illustrated in Figure 23.6 and

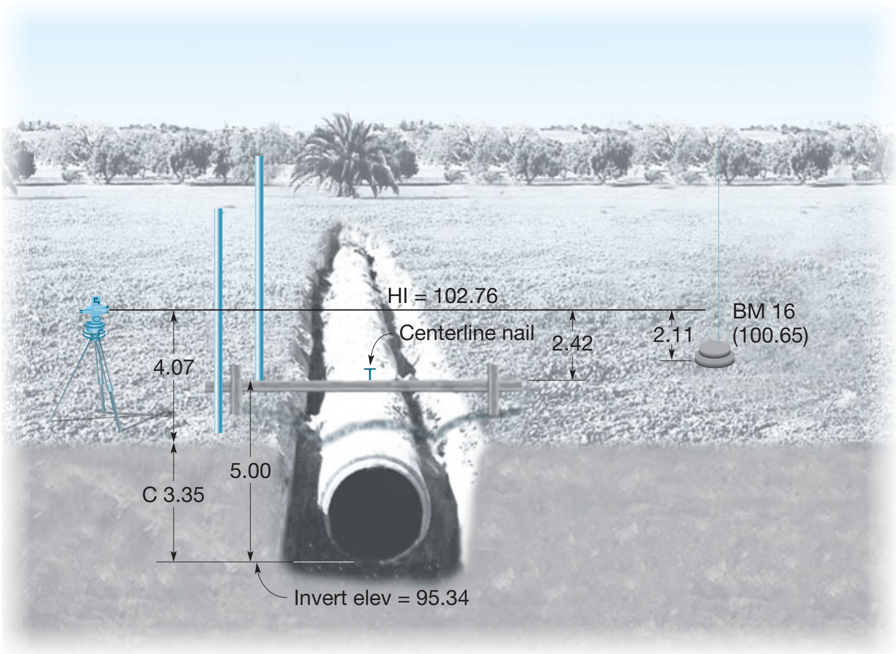


Figure 23.6 Leveling process to determine cut or fill and set batter boards for laying pipelines.

the corresponding field notes given in Plate B.6 in Appendix B. The process is summarized as follows:

1. List the stations staked on the pipeline (column 1 of the field notes).
2. Compute the flow line or invert elevation at each station (column 6).
3. Set up the level and get an *HI* by reading a plus sight on a BM; for example, $HI = 2.11 + 100.65 = 102.76$ (see Plate B.6 in Appendix B and Figure 23.6).
4. Obtain the elevation at each station from a rod reading on the ground at every stake (column 4)—for example, 4.07 at station 1 + 00 (see Plate B.6 and Figure 23.6)—and subtract it from the *HI* (column 5); for example, $102.76 - 4.07 = 98.69$ at station 1 + 00.
5. Subtract the pipe elevation from the ground elevation to get cut (+) or fill (−) (column 7); for example, $98.69 - 95.34 =$ C 3.35 (see Plate B.6 in Appendix B and Figure 23.6).
6. Mark the cut or fill (using a permanent marking felt pen or keel) on an offset stake facing the centerline; the station number is written on the other side.

In another variation, which produces the same results, *grade rod* (difference between *HI* and pipe invert) is computed, and *ground rod* (reading with rod held at stake) is subtracted from it to get cut or fill. For station 1 + 00, grade rod = $102.76 - 95.34 = 7.42$ and $7.42 - 4.07 =$ C 3.35.

After the trench has been excavated based on cuts and fills marked on the stakes, batter boards are set. Marks needed to place them can be made with a pencil or felt pen on the offset stakes during the same leveling operation used to obtain cut and fill information. Figure 23.6 also illustrates the process. Suppose that at station 1 + 00, the batter board will be set so its top is exactly 5.00 ft above the pipe invert. The rod reading necessary to set the batter board is obtained by subtracting the pipe invert elevation plus 5.00 ft from the *HI*; thus $102.76 - (95.34 + 5.00) = 2.42$ ft (see Figure 23.6). The rod is held at the stake and adjusted in vertical position by commands from the level operator until a rod reading of 2.42 ft is obtained; then a mark is made at the rod's base on the stake. (To facilitate this process, a rod target or a colored rubber band can be placed on the rod at the required reading.) The board is then fastened to the stake with its top at the mark using nails or C clamps, and a carpenter's level is used to align it horizontally across the trench. A nail marking the pipe centerline is set by measuring the stake's offset distance along the board.

If a laser is to be employed, this same leveling procedure can be used to establish the elevation of the laser beam at some desired vertical offset distance above the pipe's invert. The procedure is used to establish the height of the laser instrument and also to set another identical offset elevation at a station forward on line. Then the laser beam is aimed at that target to establish the required grade line.

■ 23.6 STAKING OUT A BUILDING

The first task in staking out a building is to locate it properly on the correct lot by making measurements from the property lines. Most cities have an ordinance establishing setback lines from the street and between houses to improve appearance and provide fire protection.

Stakes may be set initially at the exact building corners as a visual check on the positioning of the structure, but obviously such points are lost immediately when excavation is begun on the footings. A set of batter boards and reference stakes, placed as shown in Figure 23.7, is therefore erected near each corner, but

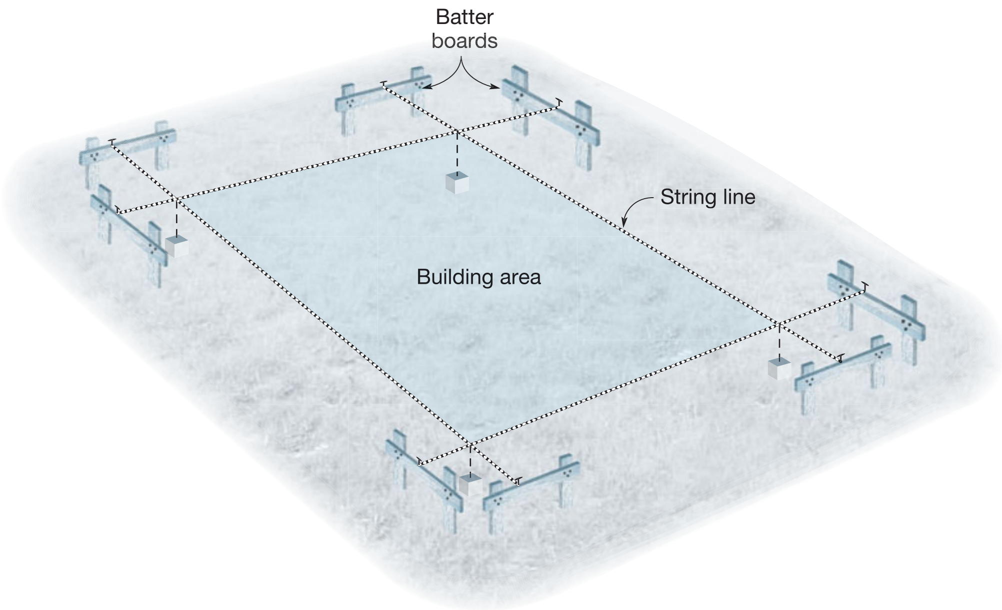


Figure 23.7 Batter boards for building layout.

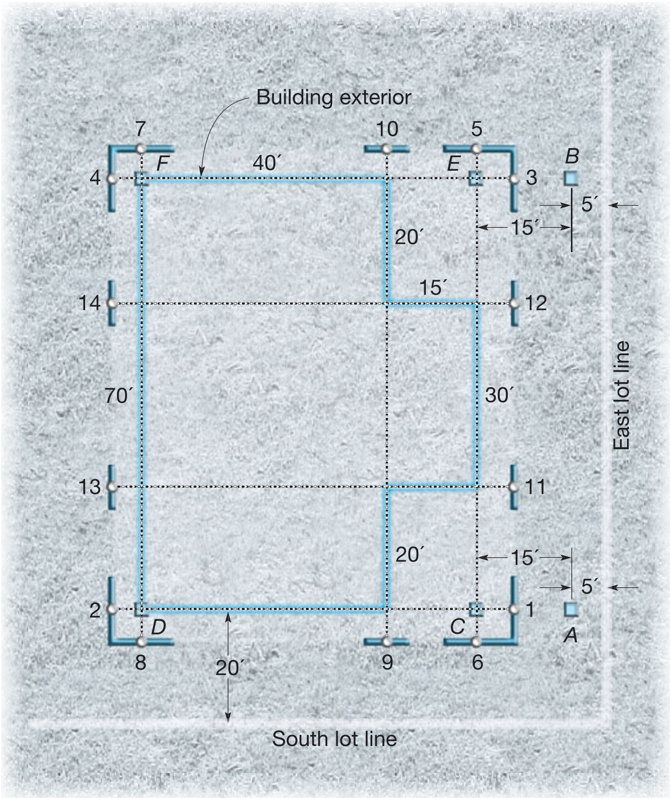


Figure 23.8 Location of building on lot and batter-board placement.

out of the way of construction. The boards are nailed a full number of feet above the footing base or at first-floor elevation. (The procedure of setting boards at a desired elevation was described in the preceding section.) Nails are driven into the batter board tops so that strings stretched tightly between them define the outside wall or form line of the building. The layout is checked by measuring diagonals and comparing them with each other (for symmetric layouts) or to their computed values. Figure 23.8 illustrates the placement on a lot and staking of a slightly more complicated building. The following are recommended steps in the procedure:

1. Set hubs *A* and *B* 5.00 ft inside the east lot line, with hub *A* 20.00 ft from the south lot line and hub *B* 70.00 ft from *A*. Mark the points precisely with nails.
2. Set a total station instrument over hub *A*, backsight on hub *B*, and turn a clockwise angle of 270° to set batter-board nails 1 and 2 and stakes *C* and *D*.
3. Set the instrument over hub *B*, backsight on hub *A*, and turn a 90° angle. Set batter-board nails 3 and 4 and stakes *E* and *F*.
4. Measure diagonals *CF* and *DE* and adjust if the error is small or restake if large.
5. Set the instrument over *C* backsight on *E*, and set batter-board nail 5. Plunge the instrument and set nail 6.