

Geometric Reduction of Measured Lines

By T. Vincenty

Abstract: This paper is a review of known procedures for use with geometric reductions of lines to the ellipsoid and with related transformations. Emphasis is placed on those aspects of the problem which are often ignored or misunderstood, with sufficient information given about the magnitudes of errors that can result from using mathematical approximations. Reduction relating to the influence of the atmosphere are not included.

Introduction

It could be said that all there is to know about geometric reductions of measured distances to the ellipsoid has already been written and that nothing new can be added to it. This may be true, but it can also be said that many people are still misinformed about this subject. This paper is therefore a restatement of correct procedures to be used for this purpose, without unnecessarily complicating the solution.

It is known, or should be, that before a measured distance can be reduced to any projection plane it must first be reduced to the ellipsoid, also known as "spheroid". This process is described usually as "reduction to sea level", which in general is incorrect because the two surfaces, sea level and ellipsoid, are not the same but may be separated by tens of meters. This separation is known as geoid height, geoidal height, or undulation. It varies slowly from point to point. An error of 30 m in geoid height introduces an error of 4.7 parts per million (ppm) in geodetic distance, that is in the length of the line on the ellipsoid, which error can be much larger when the measured line is very steep.

In the past, reductions to sea level were not only justified but necessary because the NAD 1927 framework, like all other major geodetic networks, was computed and adjusted by the development method which ignores geoid heights for the simple reason that they are unknown. On the other hand, geodetic data on NAD 1983 will be published with inclusion of geoid heights at all points for which horizontal coordinates are given. For the purpose of this paper the term "height" will be used to denote geodetic height of the point to which the measurement refers, that is the sum of elevation above mean sea level (or simply "elevation") and geoid height, not elevation alone. Also, the term "surveyor" embraces members of all professions that use surveying.

Reduction to the Ellipsoid

Two well-known methods of reducing measured distance to the ellipsoid will be given. Both use the great circle distance on a sphere of a given radius to describe the length of the geodetic line approximately but, with a proper choice of the radius, with adequate accuracy for all practical geodetic purposes.

T. Vincenty, a geodesist, has recently retired from the Geodetic Research and Development Laboratory, National Geodetic Survey, NOS, NOAA, Rockville, Maryland. His mailing address is: P.O. Box 208, Washington Grove, Maryland 20880.

First Method

Let L_o denote the measured slope distance (with atmospheric and instrumental corrections applied), h_1 and h_2 the heights of its endpoints (with heights of the instruments above the ground included), Δh the difference in heights, h_m the mean height, R_a the radius of curvature in the direction of the line (see next section), and s the length of the geodetic line, or geodetic distance. Then the "horizontal" distance at mean height is given by

$$L_h = (L_o^2 - \Delta h^2)^{1/2} \quad (1)$$

The horizontal distance is converted to the length of the chord on the ellipsoid by

$$L_c = L_h R_a / (R_a + h_m) \quad (2)$$

and finally

$$s = L_c + L_c^3 / (24R_a^2). \quad (3)$$

This method is an approximation but its errors are imperceptible in most practical situations. The small term of (3) is often omitted over short lines such as under 10 km. If this term has been included with atmospheric reductions, it should be removed from one place or the other, so that it would not be applied twice.

Second Method

This method avoids the approximations of the first method except one: the assumption that the geodetic line is an arc on a sphere. This assumption, however, is totally harmless even over very long lines.

The chord distance is obtained by

$$L_c = R_a \left[\frac{L_o^2 - \Delta h^2}{(R_a + h_1)(R_a + h_2)} \right]^{1/2} \quad (4)$$

and then the geodetic distance by

$$s = 2 R_a \sin^{-1} [L_c / (2 R_a)] \quad (5)$$

Reverse Case

It is sometimes necessary to compute slope distance from geodetic distance and heights, as for example between a radar antenna and a distant calibration target, after completion of a survey to connect them. Then equations (1)-(3) and (4)-(5) are easily rearranged for inverse computations as follows.

First Method

$$L = s - s^3 / (24R_a^2) \quad (3.1)$$

$$L_h = L_c (R_a + h_m) / R_a \quad (2.1)$$

$$L_o = (L_h^2 + \Delta h^2)^{1/2} \quad (1.1)$$

Second Method

$$L_c = 2 R_a \sin[s / (2R_a)] \quad (5.1)$$

$$L_o = [L_c^2 (1 + h_1/R_a)(1 + h_2/R_a) + \Delta h^2]^{1/2}. \quad (4.1)$$

Determination of Radius of Curvature

At any point on the ellipsoid the radius of curvature varies with the starting azimuth of the geodetic line. Two special cases of it are the principal radii M (in the meridian, or north-south direction) and N (in the prime vertical, or east-west direction). In general, this radius is given by

$$\begin{aligned} R_a &= MN / (M \sin^2 a + N \cos^2 a) \\ &= N / (1 + e'^2 \cos^2 \varphi \cos^2 a) \\ &= \frac{a}{(1 - e^2 \sin^2 \varphi)^{1/2} (1 + e'^2 \cos^2 \varphi \cos^2 a)} \end{aligned} \quad (6)$$

where a is equatorial radius (major semiaxis) of the ellipsoid, φ is latitude of the point, e is first eccentricity, e' is second eccentricity, and a is azimuth of the normal section. For the purpose of line reductions it doesn't matter much which azimuth is meant by a (geodetic or grid), as long as it is reckoned from north or south in either direction, nor is it important at which point of the line R_a is to be computed, as will be shown presently.

A glance at equation (2) shows that, if $h_m = 0$ (which could occur at a low elevation if geoid height is negative), then it doesn't matter what value of R_a is used in it because it drops out, leaving $L_c = L_h$. However, even if the elevations are very low, the chord-to-arc correction by (3), which is implicit in (5), requires a value of R_a that is correct to two significant figures, for example the equatorial radius or the mean radius of the Earth. To investigate a general case, differentiate (2) with respect to R_a to obtain the approximation

$$dL_c = (L_h h_m / a^2) dR_a.$$

If $h_m = 2000$ m and $L_h = 15000$ m, an error of 1000 m in R_a results in an error of less than 1 mm in reduced distance. The maximum allowable error in R_a corresponding to an error in chord distance not exceeding 0.1 ppm can be obtained from (7) as

$$dR_a = (a^2 / h_m) \cdot 10^{-7}.$$

This means that even at very high elevations an error of 1500 m in the radius of curvature for the line is acceptable in high precision surveys. It also means that if $h_m = 500$ m, a standard

radius of the Earth used in the above computations will produce an error in distance of less than 0.5 ppm.

It will now be shown to what extent the determination of R_a is affected by inaccurate values of latitude and azimuth. Expanding equation (6) to terms in e^2 and differentiating it with respect to ϕ and a , we get

$$dR_a = a e^2 \sin \phi \cos \phi (1 + 2 \cos^2 a) d\phi + 2a e^2 \cos^2 \phi \sin a \cos a da. \quad (9)$$

The maximum error caused by an error in latitude occurs when $\phi = 45^\circ$ and $a = 0$, amounting to 1130 m per degree. The maximum error due to inaccuracy of azimuth occurs at the equator if $a = 45^\circ$ and amounts to 750 m per degree.

This shows that for the purpose of reductions of lines such as can be measured in practice the errors in latitude and azimuth are not at all of major concern to surveyors. It is for this reason that various geodetic agencies customarily tabulate the values of R_a or their logarithms at intervals of 1° of latitude and 5° of azimuth to be used with visual interpolation, if any. In fact, tables computed for the Clarke 1866 ellipsoid may be safely used with reductions to the GRS80 ellipsoid (used for NAD83), except in high precision surveys at very high elevations, since the values of R_a for the two ellipsoids are the same within 500 m.

How does all this affect the surveyor who does his computations in plane coordinates? It tells him that the determination of the radius of curvature for the line is a trivial matter indeed.

Equation (6) can be written as

$$R_a = \frac{U}{1 + V \cos^2 a} \quad (10)$$

and the values of U and V can be precomputed on a calculator once and for all and tabulated at 1° intervals of latitude to the nearest 100 m and five decimals respectively, for use anywhere in the world without interpolation. The latitude of the survey area can be obtained from an atlas. One set of U , V values will suffice for practically any survey.

Alternatively, these values can be tabulated against northing values on the central meridian of the zone. This is acceptable because the latitude of a point can differ from the corresponding footpoint latitude by only some minutes.

As for the value of a , this can be taken as meaning the grid azimuth, perhaps only with addition of convergence in places far away from the central meridian in Lambert projection zones and at high elevations.

The Role of Geoid Heights

From equation (4) we have the approximation

$$L = L_0 [1 - h_m / R_a - (\Delta h / L_0)^2 / 2]. \quad (11)$$

Differentiation of (11) with respect to the heights gives

$$DL_c = -(L_o/R_a)(dh_1 + dh_2)/2 - (\Delta h/L_o)(dh_2 - dh_1). \quad (12)$$

Now let S denote the sea level distance and H the elevation above sea level. If h is replaced by H in equations (1)-(3) and (4)-(5) then sea level distance is obtained as the result. Further, let geoid height be denoted by its customary symbol N (no relation to the same symbol in (6)), $N = h - H$, $\Delta N = N_2 - N_1$, $N_m = (N_1 + N_2)/2$, and $\Delta H = H_2 - H_1$. Then from (12) we have

$$s = S - S N_m/R_a - \Delta H \Delta N/S. \quad (13)$$

Equations (12) and (13) show that the error in reduced distance due to omission of geoid heights is composed of two parts. The first part is caused by convergence of ellipsoid normals and becomes significant when the values of N are large. The second part is due to a local slope of the geoid with respect to the ellipsoid which can vary from nothing to tens of seconds. It can amount to several parts per million over steep lines. Thus it is clear that not only the average value of geoid height but also the difference in geoid heights between the ends of the line can be important in precise surveys. Any shortcuts in their application are left to the judgment of the users.

Mark to-Mark Distance

Permanent records of measured distances are kept most conveniently in terms of slope distances between the marks on the ground. Geodetic and sea level distances may be subject to change if elevations that were used for their reductions have over the years been re-established with greater accuracies, and it may be difficult to trace their previous values.

Let H_1, H_2 be elevations of station marks and h'_1, h'_2 the heights of EDM instrument and its target. Put $\Delta H = H_2 - H_1$, $\Delta h' = h'_2 - h'_1$, and $h'_m = (h'_1 + h'_2)/2$. Then the mark-to-mark distance is given by

$$\begin{aligned} L_g &= [L_o^2 + \Delta H^2 - (\Delta H + \Delta h')^2]^{1/2} - h'_m L_o/R \\ &= (L_o^2 - 2\Delta h' \Delta H - \Delta h'^2)^{1/2} - h'_m L_o/R, \end{aligned} \quad (14)$$

where R is a standard radius of the Earth. This equation is derived by ordinary geometry, first under the assumption that the normals at the two points are parallel, which gives the first term, then adding the second term to allow for their convergence.

Equation (14) can be written approximately as

$$L_g = L_o(1 - h'_m/R - \Delta h' \Delta H/L_o^2), \quad (15)$$

which shows that the accuracy of the reduction is not sensitive to changes in elevations. For nearly horizontal lines the last term of (15) is nearly zero, and if the measurements are performed from near the ground (eye level) then L_o and L_g differ by less than 0.3 ppm. Note the analogy between this equation and (13).

Example The following example is based on the GRS80 ellipsoid ($a = 6378137$ m, $e^2 = 0.006694380$, $e^2 Z = 0.006739497$).

	<i>Point 1</i>	<i>Point 2</i>
H	1000.00 m	1700.00 m
N	20.00	20.50
h_g	1020.00	1720.50
h'	5.30	1.50
h	1025.30	1722.00
	35°	
U	6385200	(by (6))
V	0.00452	(by (61))
a	234°	
R a	6375200	(by (10))
L $_o$	15000.0000	
L $_g$	15000.1689	(by (14))
L $_h$	14983.8116	(by (1))
L $=$	14980.5837	(by (21) or (4))
s	14980.5872	(by (3) or (5))
S	14980.6581	(like s, using H)
s	14980.5871	(by (13))

If R $_o$ is changed by + 1000 m, s is computed as 14980.5877 m, which agrees with what can be predicted by (7). Note also that in this example the mark-to-mark distance is longer than the distance between the instrument and the target.

Concluding Remarks

Errors in geometric reductions of distances can be classified as resulting from two categories of the approximations, (1) determination of the radius of the sphere corresponding to the line and (2) the heights of the endpoints of the line above the ground, the geoid, and the ellipsoid.

The first source of error is of little concern because the required radius can be obtained with sufficient precision by crude methods. Its determination from state plane coordinates presents no particular problem. In low elevations even a standard value of the radius of the Earth will be satisfactory.

Errors in the three kinds of heights affect the accuracy of the result identically. In all cases they can be split into two parts, one depending on the mean value of the heights and the other on the difference in heights. Both can be important.

The equations of this paper have been constructed so as to give precisions beyond what is required in most engineering surveys. Any shortcuts and omissions to suit particular needs are left to the decision of the users.

References

- Bomford, G. (1980), *Geodesy*, 4th ed., Clarendon Press. Oxford, pp. 82-85.
 Clark, D. (1963), *Plane and Geodetic Surveying for Engineers*, Vol. 2. 5th ed., Constable & Company Ltd., London, pp. 298-300.

U.S. Dept. of the Army (1969). *Surveying Computer's Manual*, TM S-237, Washington, D.C., pp. 112, 124 - 126.

Vincenty, T. (1975), "A Note on the Reduction of Measured Distances to the Ellipsoid," *Survey Review*, Vol. 23, No.175, pp. 40-42.