GEODETIC INVERSE SOLUTION BETWEEN ANTIPODAL POINTS

T. Vincenty
DMAAC Geodetic Survey Squadron
F. E. Warren AFB, Wyoming 82001

ABSTRACT

The solution given in this paper supplements standard geodetic inverse methods which fail in antipodal situations. A modification of the proposed formula can be used for the solution around the back side of the ellipsoid in analogous cases.

INTRODUCTION

The classical methods of solving the inverse geodetic problem which are based on the Besselian approach (hereafter referred to as standard methods) involve an iterative determination of λ, the longitude on a conformal sphere. It is well known that these methods fail in cases which have been described as antipodal or nearly antipodal, that is when the attempted solution gives the absolute value of λ as greater than π.

Thien (1967) gives an antipodal solution limited to points on the equator, also quoted by Rapp (1975). The method of Saito (1970), apparently designed for general use, is valid only for cases when the end points of the line are in antipodal latitudes ($\phi_2 = -\phi_1$). The proposed method, derived by rearranging certain equations of a previous paper (Vincenty, 1975), is completely general. It will be shown that the approach on which the proposed antipodal solution is based can also be used for the analogous case of solution around the back side of the ellipsoid between points which are close together.

NOTATION

a, b,	major and minor semiaxes of the ellipsoid.
f,	flattening = $(a-b)/a$.
ε,	the square of second eccentricity = $(a^2-b^2)/b^2$.
ϕ,	geodetic latitude, positive north of the equator.
L,	difference in geodetic longitude, positive east.
s,	geodesic distance; also length of any geodetic line.
α_1, α_2,	azimuths of the geodesic, clockwise from north; α_2 in the direction P_1P_2 produced.
α,	azimuth of the geodesic at the equator.
U,	reduced latitude, defined by $\tan U = (1-f) \tan \phi$.
λ,	difference in longitude on an auxiliary sphere.
σ,	angular distance P_1P_2 on the sphere.
σ_m,	angular distance on the sphere from the midpoint of the line to its intersection with the equator.

ANTIPODAL REGION

We consider a geodetic line starting from point P_1 in northern hemisphere in latitude ϕ_1 and in azimuth $\alpha_1 = \pi/2$. As we follow this line, its azimuth changes in accordance with Clairaut's equation

$$\cos U_i \sin \alpha_i = \text{constant} \tag{1}$$

or in this case

$$\sin \alpha_i = \cos U_1/\cos U_i. \tag{2}$$

The line will cross the equator in azimuth α given by

$$\sin \alpha = \cos U_1 \tag{3}$$

until it reaches the antipodal region at point P_a at which $\phi_a = -\phi_1$, $\lambda = \sigma = \pi$, and $\alpha_a = \pi/2$. Up to this point this geodetic line defines the shortest distance on the surface of the ellipsoid and it is therefore a geodesic. It can be shown that, up to terms in f^3, the geodetic longitude of this point is

$$L_a = \pi - f \cos U \, \pi \left\{1 - \frac{1}{16} f \sin^2 U[4 + f(4 - 3 \sin^2 U)]\right\}, \tag{4}$$

where $U = U_1 = -U_2$. This is the maximum geodetic longitude of the end of a geodesic which starts in azimuth $\alpha_1 = \pi/2$. Now if we select point P_2 such that $\phi_2 = \phi_a$ or $\phi_2 \approx \phi_a$ and $\pi > L > L_a$, the shortest geodetic line between P_1 and P_2 must start in an azimuth different from $\pi/2$ because of the flattening of the ellipsoid. There are three possibilities:

If $\phi_1 + \phi_2 > 0$, $\pi/2 > \alpha_1 > 0$.

If $\phi_1 + \phi_2 < 0$, $\pi > \alpha_1 > \pi/2$.

If $\phi_1 + \phi_2 = 0$, $\lambda = \sigma = \pi$ and there are two solutions for azimuths.

We now follow a geodetic line starting in latitude $\phi_1 > 0$ and in azimuth $\pi/2 > \alpha_1 > 0$. The line will reach its maximum latitude, then turn in the south-easterly direction, cross the equator, and continue towards the antipodal latitude $\phi_a = -\phi_1$. We can compute point coordinates along this line and verify the results by inverse solutions using standard methods, until the line reaches the antipodal region at P_o.

If we extend this line to latitude ϕ_a and select any point P_2 on it such that $\phi_2 > \phi_a$, we will find from the direct solution that $\lambda < \pi$ and $\sigma < \pi$, in the inverse solution λ will oscillate between values greater and smaller than π without converging to any value. We can conclude that there is a region within which any point may be thought of as being antipodal with respect to P_1, even though this is not true in the geometric sense. In the eastern hemisphere (reckoned from P_1) it must extend in longitude from L_a as given by eq. (4) to π, and it must be bounded by two curves symmetric with respect to ϕ_a, corresponding to $\pi/2 \geq \alpha_1 \geq 0$ and $\pi \geq \alpha_1 \geq \pi/2$. Its mirror image in the western hemisphere follows logically for values of α_1 in the third and fourth quadrants. The precise limits of the antipodal region have not been established.

ANTIPODAL SOLUTION

If during the initial computation of λ the result is greater than π in absolute value, a different method must be used. A method that works well in general is described below.

Use $L' = \pi - L$ if $L > 0$ or $L' = -\pi - L$ if $L < 0$. Adopt as initial approximations $\lambda' = 0$, $\cos^2\alpha = 1/2$, $\cos 2\sigma_m = 0$, and $\sigma = \pi - |U_1 + U_2|$. (The use of L' and λ' instead of L and λ is designed to avoid loss of significant figures in computations.) Iterate the following equations:

$$C = \frac{1}{16} f\cos^2\alpha[4 + f(4 - 3\cos^2\alpha)] \tag{5}$$

$$\cos 2\sigma_m = \cos\sigma - 2\sin U_1 \sin U_2/\cos^2\alpha \tag{6}$$

$$D = (1-C)\, f\{\sigma + C\sin\sigma[\cos 2\sigma_m + C\cos\sigma(-1 + 2\cos^2 2\sigma_m)]\} \tag{7}$$

$$\sin\alpha = (L' - \lambda')/D \tag{8}$$

$$\sin\lambda' = \sin\alpha\sin\sigma/(\cos U_1 \cos U_2) \tag{9}$$

$$\sin^2\sigma = (\cos U_2 \sin\lambda')^2 + (\cos U_1 \sin U_2 + \sin U_1 \cos U_2 \cos\lambda')^2 \tag{10}$$

Iteration is stopped when the change in $\sin\alpha$ from the previous value is smaller than a specified value.

We now obtain

$$\sin\alpha_1 = \sin\alpha/\cos U_1 \tag{11}$$

$$\cos\alpha_1 = \pm(1 - \sin^2\alpha_1)^{1/2}, \tag{12}$$

the sign of $\cos\alpha_1$ being negative if

$$\cos U_1 \sin U_2 + \sin U_1 \cos U_2 \cos\lambda' < 0. \tag{13}$$

The azimuth at P_2 is computed by

$$\tan\alpha_2 = \frac{\sin\alpha}{-\sin U_1 \sin\sigma + \cos U_1 \cos\sigma\cos\alpha_1}. \tag{14}$$

The azimuth at P_1 is best determined from its tangent as given by eq. (11) and (12). Here, as in eq. (14), the signs of the numerator and of the denominator determine the quadrant. Most computers can evaluate arc tangent in the proper quadrant on this basis, making it unnecessary to do so by program instructions.

Geodesic distance is obtained by a formula given by Helmert (1880), here rewritten in a form which is more convenient for use with computers.

$$E = (1 + \varepsilon \cos^2\alpha)^{1/2} \tag{15}$$

$$F = (E-1)/(E+1) \tag{16}$$

$$A = (1 + \frac{1}{4} F^2)/(1 - F) \tag{17}$$

$$B = F(1 - \frac{3}{8} F^2) \tag{18}$$

$$\Delta\sigma = B \sin \sigma\Big\{\cos 2\sigma_m + \frac{1}{4} B[\cos \sigma(-1 + 2 \cos^2 2\sigma_m) - \frac{1}{6} B \cos 2\sigma_m(-3 + 4 \sin^2\sigma)(-3 + 4 \cos^2 2\sigma_m)]\Big\} \tag{19}$$

$$s = bA(\sigma - \Delta\sigma) = (1 - f)aA(\sigma - \Delta\sigma) \tag{20}$$

The special case when $\phi_2 = -\phi_1$ can be solved without iteration. In this case $\sigma = \pi$ and $\lambda' = 0$, therefore we have from (8)

$$L'/f\pi = \sin\alpha(1 - C) = Q \tag{21}$$

Substituting $\cos^2\alpha = 1 - \sin^2\alpha$, we can write (21) as

$$Q = a_1 \sin\alpha + a_3 \sin^3\alpha + a_5 \sin^5\alpha \tag{22}$$

in which

$$a_1 = 1 - f/4 - f^2/16$$

$$a_3 = f/4 - f^2/8$$

$$a_5 = 3f^2/16$$

Reversing the series, we obtain from (22)

$$\sin\alpha = b_1 Q + b_3 Q^3 + b_5 Q^5 \tag{23}$$

in which

$$b_1 = 1 + f/4 + f^2/8$$

$$b_3 = 1 - b_1$$

$$b_5 = 0$$

Geodesic distance is given by

$$s = bA\pi = (1 - f)aA\pi \tag{24}$$

This shows that in special antipodal cases when $\phi_2 = -\phi_1$ the values of α and s do not depend on latitude but only on difference in longitude.

ANALYSIS OF STABILITIES

Questions may arise as to stabilities of eq. (6), (8), and (12). These will now be investigated.

The second term of eq. (6) cannot become indeterminate, since there are no equatorial lines in antipodal solutions. If $\cos\alpha$ is very nearly zero, any error in $\cos 2\sigma_m$ is unimportant because it will vanish when multiplied by the very small value of $f^2\cos^2\alpha/4$ and eq. (7) will give a correct result. (In non-antipodal cases it is only necessary to exclude division by zero, for example by setting this term a priori to zero and replacing it with a computed value if $\cos^2\alpha$ is not zero.)

Eq. (8) can be written in the mathematically equivalent form

$$\sin\alpha = (\lambda - L)/D \qquad (8a)$$

but this would cause a loss of significant figures in the result. This is because at least the first two digits of λ and of L are the same, therefore the same number of digits would be lost in subtraction. Now if we convert L to L', the new value will be correct to the same number of decimal places as used with L. Then λ', as computed by eq. (9), will be given to more significant figures and to more decimals than L', and two or more significant figures will be gained in $\sin\alpha$.

Eq. (12) becomes unstable as α_1 approaches $\pm\pi/2$. This can occur only if P_2 almost coincides with P_a ($\phi_a = -\phi_1$, L_a as given by eq. (4)), in which case the slightest change in the position of P_2 causes a large change in α_1. Conversely, this means that any error in α_1 due to loss of significance in computing its cosine by the Pythagorean formula will not cause a larger positional error than what can be expected in any other case, that is an error due only to the finite number of digits that the computer can handle. Therefore the instability of eq. (12) is harmless in practice.

SOLUTION AROUND THE BACK SIDE OF THE ELLIPSOID

A geodetic line starting at P_1 in latitude ϕ_1 and in azimuth α_1 reaches $\phi_a = -\phi_1$ at P_a in azimuth such that $\sin \alpha_a = \sin \alpha_1$ and $\cos \alpha_a = -\cos \alpha_1$. At this point $\sigma_a = \lambda_a = \pi$ and the geodetic longitude is

$$\begin{aligned} L_a &= \pi[1 - f(1-C)\sin \alpha] \\ &= \pi[1 - f(1-C)\cos U_1 \sin \alpha_1]. \end{aligned} \tag{25}$$

If we extend the line in azimuth α_a by geodesic distance P_1P_a, it will reach point P_t where $\phi_t = \phi_1$, $L_t = 2L_a$, $\alpha_t = \alpha_1$, and $\sigma_t = \lambda_t = 2\pi$.

If we select point P_2 on this line such that the spherical distance $P_1P_aP_2$ is greater than π, the geodetic line $P_1P_aP_2$ is not a geodesic but a line around the back side of the ellipsoid. The shortest line P_1P_2 is a geodesic starting in azimuth which in general is different from $\alpha_1 \pm \pi$.

The length of a geodetic line around the back side of the ellipsoid and the azimuths at its ends can be computed by a standard inverse formula (e.g. Vincenty, 1975) after modifying it so that $\sin \sigma$ is given a negative sign ($2\pi \geq \sigma > \pi$) and the azimuths are changed by $\pm \pi$. This method fails when the two points are close together, that is when they are within a certain neighborhood region, in which case λ will oscillate between a positive and a negative value. When this happens we can obtain a solution by using the equations given in this paper with some modifications. As in the antipodal case, the object is to obtain $\sin \alpha$ by iterating eq. (5) to (10); however, eq. (8a) should be used instead of (8), since L and λ will be closer to 0 than to π, and

σ should be obtained from

$$\sin^2\sigma = (\cos U_2 \sin \lambda)^2 + (\cos U_1 \sin U_2 - \sin U_1 \cos U_2 \cos \lambda)^2. \qquad (10a)$$

The sign of $\cos \alpha_1$ is negative if

$$\cos U_1 \sin U_2 - \cos U_1 \cos U_2 \cos \lambda > 0. \qquad (13a)$$

The initial approximations for this solution can be $\lambda = 0$, $\cos^2\alpha = 1/2$, $\cos 2\sigma_m = 0$, and $\sigma = 2\pi - |U_1 - U_2|$.

It should be noted that in general we have an infinite number of geodetic lines joining two points (Helmert, 1880), since we may specify

$$2n\pi \geq \sigma > (2n-1)\pi \qquad (26)$$

where n denotes the number wraps around the ellipsoid. To each value of n there corresponds a different solution. For the present purpose n = 1.

The inverse solution around the back side of the ellipsoid may be verified by any standard direct formula. It is of interest to note that in the direct solution it is not necessary to split the line into two or more sections such that each would satisfy the condition $\sigma < \pi$. In fact, there is no theoretical limit as to the length of the line that may be used in the direct solution to compute point coordinates on it as it continues around the ellipsoid indefinitely. In practice the limitations are of numerical nature.

COMPUTATIONAL EXAMPLES

Examples of results of antipodal and back side solutions are given in Tables 1 and 2. The azimuths are given to only three decimals of a second, since more precision is not needed to duplicate the positions by the direct solution within 0.5 mm when σ is equal to π or 2π or nearly so.

The results given in Table 1 were computed using a FORTRAN program and an IBM 7040/94 computer, and those of Table 2 were obtained by a Wang 720 desk calculator program.

The following ellipsoid parameters were used:

$$a = 6378388.000 \text{ m}, \quad 1/f = 297.$$

TABLE 1. Antipodal Solutions

ϕ_1	41°41'45".88000	0°00'00".00000	30°00'00"000.00	60°00'00".00000
ϕ_2	-41 41 46.20000	0 00 00.00000	-30 00 00.00000	-59 59 00.00000
L	179 59 59.44000	179 41 49.78063	179 40 00.00000	179 50 00.00000
α_1	179 58 49.163	30 00 00.000	39 24 51.806	29 11 51.070
α_2	0 01 10.838	150 00 00.000	140 35 08.194	150 49 06.868
s	20 004 566.7228	19 996 147.4169	19 994 364.6069	20 000 433.9629

TABLE 2. Back Side Solutions

ϕ_1	41°41'45".88000	00°00'00".00000	30°00'00".00000	60°00'00".00000
ϕ_2	41 41 46.20000	0 00 00.00000	30 00 00.00000	59 59 00.00000
L	0 00 00.56000	0 18 10.21937	0 20 00.00000	0 10 00.00000
α_1	180 00 35.423	194 28 47.448	198 30 47.488	344 56 31.727
α_2	180 00 35.423	194 28 47.448	198 30 47.488	344 56 59.622
s	40 009 143.3208	40 004 938.2722	40 004 046.7114	40 006 087.0024

REFERENCES

Helmert, F. R., Die mathematischen und physikalischen Theorien der höheren Geodäsie, B. G. Teubner, Leipzig, 1880 (translated by Aeronautical Chart and Information Center, St Louis, 1964).

Kivioja, L. A., Computation of geodetic direct and indirect problems by computers accumulating increments from geodetic line elements, Bull. Geod., No. 99, 55-63, 1971.

Lambert, W. D., The distance between two widely separated points on the surface of the earth, J. Wash. Acad. Sci., Vol, 32, No. 5, 125-130, 1942.

Lewis, E. A., Parametric formulas for geodesic curves and distances on slightly oblate earth, AF Cambridge Research Laboratories, 1963.

Rapp, R. H., Geometric geodesy, Vol, II, The Ohio State Univ., 1975.

Saito, T., The computation of long geodesics on the ellipsoid by non-series expanding procedure, Bull. Geod., No. 98, 341-374, 1970.

Sodano, E. M., General non-iterative solution of the inverse and direct geodetic problems, Bull. Geod., No. 75, 69-89, 1965.

Thien, G., A solution to the inverse problem for nearly-antipodal points on the equator of the ellipsoid of revolution, M.S. thesis, The Ohio State Univ., 1967.

Vincenty, T., Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations, Survey Rev., No. 176, 88-93, 1975.