International Hydrographic Review, Monaco, LXX(1), March 1993

AREA COMPUTATION OF POLYGONS

by W.A. van GEIN and I. GILLISSEN [1]

Abstract

Area computation has always been an important task for land surveyors. Until recently the hydrographic surveyor could utilize approximative methods, as results up to 0.1 km² were sufficient for many purposes. Investigations resulted in computation algorithms, usable on the ellipsoid, with accuracies of the square metre magnitude.

INTRODUCTION

Offshore licence fees are often connected to the size of the area. Periodically there is an obligation to relinquish a certain percentage of the licence area. Due to several factors the area to be computed may have quite peculiar polygons.

The need for area computation may also arise in boundary delimitation especially when the subsoil contains valuable minerals.

DIFFERENT BOUNDARY LINES

Boundaries of licence area and agreements between countries are usually defined by certain lines between co-ordinates on a specified datum. Nowadays the different datums do not cause real difficulties as more and more generally accepted transformation data are available.

[1] Hydrographic Service, Royal Netherlands Navy, P.O. Box 90704, 2509 LS The Hague, Netherlands.

In contrast with this clarity with datums there is the fact that quite different definitions of boundary lines are used in licences and in agreements between countries.

Frequently, we read descriptions like:

a. geodesic
b. meridian
c. parallel circle
d. loxodrome
e. great circle
f. "straight line".

Some relevant properties of these lines within the context of this paper are:

Geodesic

The geodesic is defined on an ellipsoid to be the shortest distance between two given points. In the past, computations on the ellipsoid were hardly used due to the complexity of the algorithms and the limited capacity of the mathematical aids like logarithm tables and simple calculating machines. Today the availability of computers and adequate software help to solve these problems.

Meridian

On the ellipsoid all meridians are also geodesics and do not pose a problem.

Parallel circle

This line of constant latitude is to be regarded as a loxodrome as it cuts all meridians at right angles.

Loxodrome

Just like the meridian and parallel circle the loxodrome frequently is used for defining lines on charts for all kinds of purposes. This is because in Mercator's projection all three lines are represented as straight lines the meridians and parallel circles cutting at right angles and pointing respectively north-south and east-west while the loxodrome cuts all meridians at equal angles. However computations on the ellipsoid have disadvanges as differences in distances and azimuths compared to the geodesic may get quite large and are in fact unusable for geodetic computations.

Example:

Projected in Mercator's projection the parallel circle and the geodesic between two points at different latitudes and with a longitudinal difference of three degrees look like Figure 1.

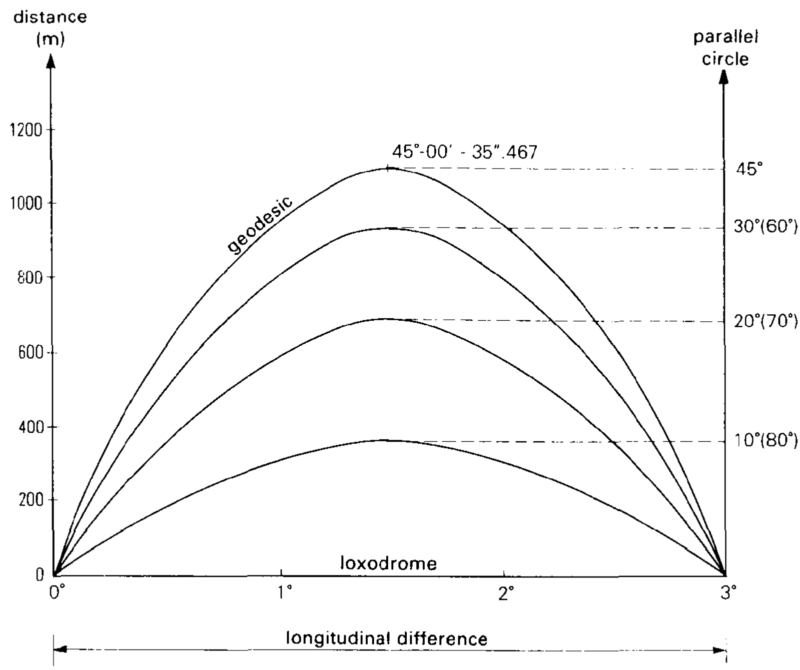

FIG. 1.- Distance between loxodrome and geodesic at different latitudes.

The maximum distance between both lines is situated at 45° latitude. It is also obvious that the azimuthal directions of both lines may differ quite a lot.

Great circle

The definition of a great circle is quite clear in relation to a sphere being the intersecting line of a plane through the centre of the sphere and the surface of the sphere. On a sphere you may perform straightforward computations using simple spherical trigonometric formulae. Using the results on an ellipsoid is something different.

Almost every time the description "Great Circle" (including capitals) is used the reference to the sphere used is missing. The Gaussian conformal sphere may be seen as an excellent geodetic replacement of the ellipsoid at relatively short distances from the chosen origin. Up to about 100 nautical miles there is no significant difference in distances and azimuthal directions between great circles on this sphere and the geodesics on the ellipsoid.

Example:

Positions: a) 50°-00'N,00° - 00'E and b) 51°-30'N,1° - 30'E.

	Gaussian conformal sphere (origin at 50°N)	Ellipsoid (ED 50)
distance:	197610.27 m	197610.29 m
azimuths:	31°-48'-47".446 32°-58'-29".623	31°-48'47".449 32°-58'29".625

The algorithms may be a little bit complicated but have been described excellently [JORDAN et al., 1959] and cause no problems with today's computers.

The great circle on a concentric sphere was and is being used by mariners.Probably this sphere is usually meant in the texts of many agreements. The straightforward algorithms are quite elegant in the case of positional computations as they are independent of the radius of the sphere. Using the most appropriate radius (semi major axis of the ellipsoid) and the quite simple distance correction of the ANDOYER-LAMBERT formula [ACIC.,1959] the differences from geodesics are quite small. However the azimuths are not usable on the ellipsoid. The differences in computed results between the great circle on the concentric sphere and the geodesic are not so large in comparison to the loxodrome (Fig.2).

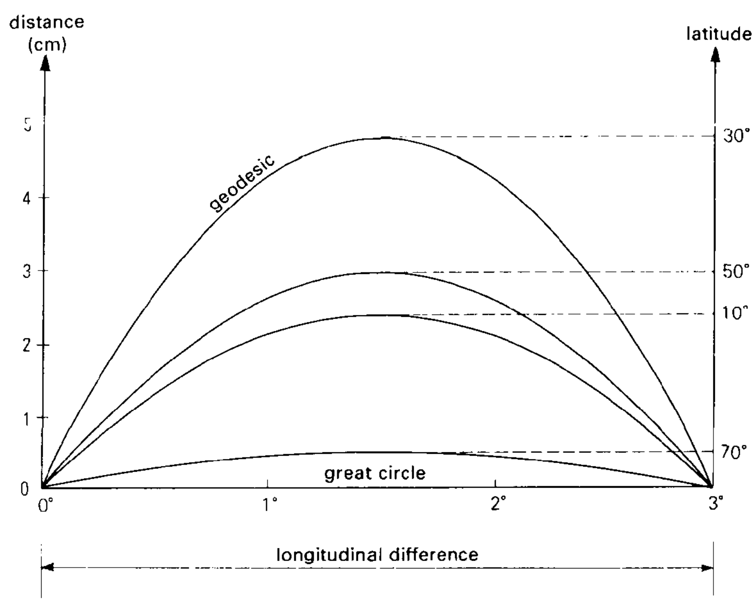

FIG. 2.- Distance between great circle and geodesic at different latitudes.

Straight line

The term *straight line* is even more vague because it refers to a chart most of the time and in agreement the charts hardly ever reveal the kind of projection and/or the ellipsoid used. The question arises whether a loxodrome, grid line or geodesic is meant.

AREA COMPUTATION

It is quite clear that the definitions of boundary lines of a certain area are crucial to the area computations. Until a few years ago several approximative algorithms could be used to provide reasonable results up to one square kilometre.

Quite recently the authors replaced their algorithms by the "Under geodesic"-method [DANIELSEN, 1989] but did not realise at the beginning that great differences could occur due to the different definitions of the boundary lines used. An obvious method seemed to be to approximate the non-geodesic lines by interpolating points along the lines and consider the lines between those points to be geodesics. This may be compared to the plotting of circles by drawing indefinite short lines between points on the circumference of the circle.

A question that arose was how many points are needed along a chosen line in order to approximate the defined line without losing accuracy in the area computation.

As an example the authors computed the area between the loxodrome and the geodesic at a latitude of 50 degrees along a longitudinal difference of 3 degrees. Results are given in Table 1 and Figure 3.

distance (m) between points (loxodrome)	number of interpolating points	computed area (km^2)	area diff. between steps (m^2)
800	271	154.650216	
		....	1602
400	540	.651818	
		....	401
200	1078	.652219	
		....	100
100	2153	.652319	
		....	25
50	4304	.652344	
		....	7
25	8606	.652351	
		....	1
12.5	17210	154.652352	

Table 1.Relations between numer of interpolating points and the accuracy of the computed area.

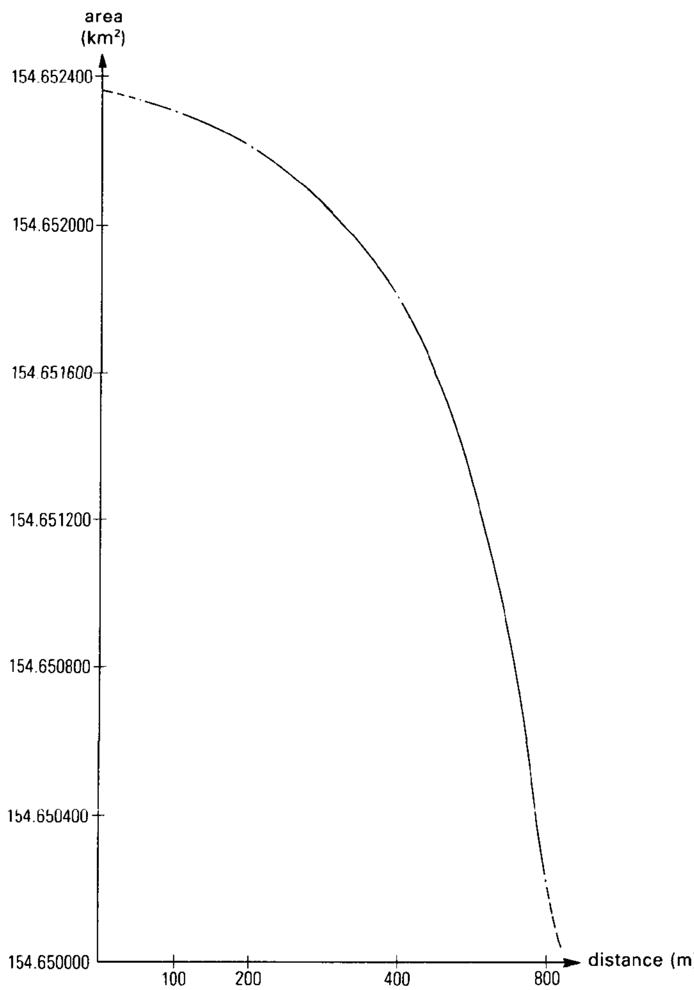

FIG. 3.- Relations between computed area and distances between interpolating points.

Similar results were found for the great circle approximation on the concentric sphere but with much smaller differences. This method produced quite satisfactory results in the end but appeared to be very time consuming for the authors as well as in computer time.

Looking for more direct and faster methods, a direct and much faster solution which gives the possibilities to compute any area bounded by mathematically defined lines was found. A more extensive description of this new approach may be found in GILLISSEN (1993).

Starting points for the GILLISSEN-method are:

1) the availability of a rather straightforward equal area projection [DIETZ et al.,1921]

2) well-known algorithms for the geodesic [VINCENTY, 1975], loxodrome and great circle [BOWRING, 1985] or any other mathematically defined line

3) combining these elements in a step-by-step process for computing polygon areas.

Polygon Areas

If the points and interpolating points of the polygon area on the ellipsoid are equivalently projected in a plane (Fig.4) and if the points $P_1, P_2, \ldots, P_n$ in the plane are defined by polar co-ordinates (ϱ_i, θ_i), ϱ is the radius vector, θ is the vectorial angle, then the area of polygon $P_1, P_2, \ldots, P_n$ =

$$\frac{1}{2}\left|\varrho_1\varrho_2\sin(\theta_2-\theta_1)+\varrho_2\varrho_3\sin(\theta_3-\theta_2)+\ldots\right.$$
$$\left.\ldots+\varrho_{n-1}\varrho_n\sin(\theta_n-\theta_{n-1})+\varrho_n\varrho_1\sin(\theta_1-\theta_n)\right|$$

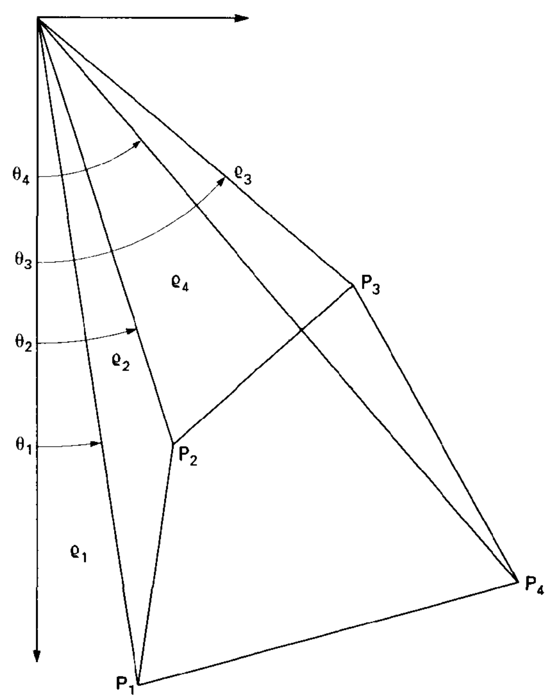

FIG. 4.- Polar co-ordinates of a polygon.

The formula is in absolute value so the points $P_1, P_2, \ldots, P_n$ may be entered counterclockwise or clockwise.

THE STEP-BY-STEP PROCESS

Between the endpoints of the polygon interpolation points are computed along the boundaries. Then all these points are mapped in the plane with the Albert's projection and the area is computed with formula (1). In general, the geodetic line, the loxodrome or any other line on the ellipsoid, will be curved lines after projection on a plane. To compensate for the curvature, interpolating points between the endpoints of the sides of the polygon will have to be computed. If a side of the polygon is defined by the geodesic between the polygon points P_i and P_{i+1}, the interpolating points between P_i and P_{i+1} are computed as follows:

1. Compute the azimuth from P_i to P_{i+1} of the geodesic [VINCENTY, 1975];
2. The distance from P_i to the interpolating point is the chosen step size;
3. Then the co-ordinates of the first interpolating point are computed directly from azimuth and distance;
4. The next interpolating point is computed in the same way from P_i with the same azimuth and twice the step size;
5. This process is repeated until P_{i+1} is reached.

If a side of the polygon is defined by the loxodrome between the endpoints the same procedure is applied except that the direct solution of the loxodrome is used for computing the interpolating points [BOWRING, 1985]. And for great circles the same process can be applied. As a matter of fact the same process may be used for any type of line as long as interpolating points can be computed. Even polygons drawn on a map in any projection can be computed by using the inverse mapping formulas for transformation of the endpoints and interpolating points to the ellipsoid and then with the Albert's projection back to the plane.

MAPPING FORMULAS

The equivalent Albert's projection employs a cone intersecting the ellipsoid at two parallels known as the standard parallels. In general the projection is used to represent a certain area and consequently the standard parallels are chosen accordingly. In this case the projection is used for computation purposes only and no relevant difference in the area computation caused by the choice of the standard parallels was found.

The meridians are straight lines in this projection, so if the endpoints of a side are on the same meridian, no interpolating points need being computed. The mapping equations for a point P given in latitude and longitude to P′ in vectorial angle and radius vector or $P(\varphi,\lambda)_{ellipsoid} \rightarrow P'(\varrho,\theta)_{plane}$ are:

$$\varrho = \sqrt{\varrho_1^2 + \frac{2a^2}{n}(1-e^2)(\beta_1-\beta)}$$

$$\theta = n\lambda$$

where ϱ is the radius vector,
θ is the vectorial angle,
n is the mapping factor for λ,
λ is the geodetic longitude,
φ is the geodetic latitude,
φ_1 is the geodetic latitude of standard parallel 1,
φ_2 is the geodetic latitude of standard parallel 2,
a is the equatorial radius of the ellipsoid,

$$\beta = \sin\varphi\left(1 + \frac{2e^2}{3}\sin^2\varphi + \frac{3e^4}{5}\sin^4\varphi + ...\right)$$

$$n = \frac{\left(\dfrac{\cos^2\varphi_1}{1-e^2\sin^2\varphi_1} - \dfrac{\cos^2\varphi_2}{1-e^2\sin^2\varphi_2}\right)}{2(1-e^2)(\beta_2-\beta_1)}$$

For standard parallel 1 :

$$\varrho_1 = \frac{a\cos\varphi_1}{n(1-e^2\sin^2\varphi_1)^{\frac{1}{2}}}.$$

ACCURACY AND CORRECTION

As the sides of the polygon are approximated by interpolating points, which after projection are as it were connected by straight lines for the area computation, an error is introduced the magnitude of which depends on the step size. Theoretically one would expect the error to decrease quadratically with a decreasing step size. Empirically this was confirmed; if we have the step sizes s_1 and s_2 with corresponding errors e_1 and e_2 and if

$$s_2 = \frac{1}{a}s_1 \quad then$$

$$e_2 = \frac{1}{a^2}e_1.$$

Knowing this the error can be computed and used as a correction:

Suppose the true area is A.
The computed area with step size s_1 is $A_c = A + e_1$.
The computed area with step size s_2 is $A_c = A + e_2$.

$$\Delta A_c = (A+e_2)-(A+e_1)=e_2-e_1.$$

Substituting $e_2 = 1/a^2\ e_1$ yields:

$$\frac{1}{a^2}e_1-e_1 = \Delta A_c \rightarrow e_1 = \frac{a^2}{1-a^2}\Delta A_c.$$

So if the area is computed twice, the second time with for instance half the step size, the error can be computed as

$$e_1 = -\frac{4}{3}\Delta A_c$$

and can be used as correction.

Using the step-by-step process with this correction method for areas - with geodesic boundaries- of any size there was no difference with the "Under geodesic" [DANIELSEN, 1989] method on the square meter level.

STEP-BY-STEP INTEGRATION

For non-geodesics the step-by step method may also be combined with "Under geodesic" [DANIELSEN, 1989] method. The Albert's projection and the area computation with polar co-ordinates must then be replaced by integration on the ellipsoid. In this combination the error formula (8) can also be applied as can be seen when using the results of Table 1, for instance with intervals of 800m and 400m respectively.

We find

$$\textbf{\textit{computed area}}\ (800m) = 154.650216\ km^2$$

$$e_1 = -\frac{4}{3}\times 0.001602 = -\ 0.002136\ km^2$$

$$\textit{area} = 154.652352\ km^2$$

which is the same result as the area computed with points at intervals of 12.5m.

Example 1

The area of the complete ellipsoid (HAYFORD) computed with the formula for the oblate ellipsoid is:

$$AREA = 2\pi a^2 + \pi\frac{b^2}{e}\ln\frac{1+e}{1-e} = 510100933.858376\ km^2.$$

This area computed with the corrected step-by-step method and step sizes of 500m and 250m gave exactly the same result.

Example 2

Polygon area ABCDEA enclosed by geodesics and sub-divided by loxodromes and great circles (concentric sphere) (Fig. 5).

Area		'under geodesic' (Danielsen)	'step-by-step' (Gillissen)
I - ECDE	:geodesics	3987.089841 km²	3987.089841 km²
II - ECE	:loxodrome (EC) and geodesic (CE) :geodesics (EC) and (CE)	— —	154.652353
III - AFCEA	:loxodrome (AF), great circle (FC) loxodrome (CE), geodesics (EA) :along geodesics	— 10112.486043	9967.231163 —
IV - AFCBA	:loxodrome (AF), great circle (FC), geodesics (CB) and (BA) : along geodesics	— 16088.372811 +	16078.975338 — +
Total area polygon ABCDEA =		30187.948695 km²	30187.948695 km²

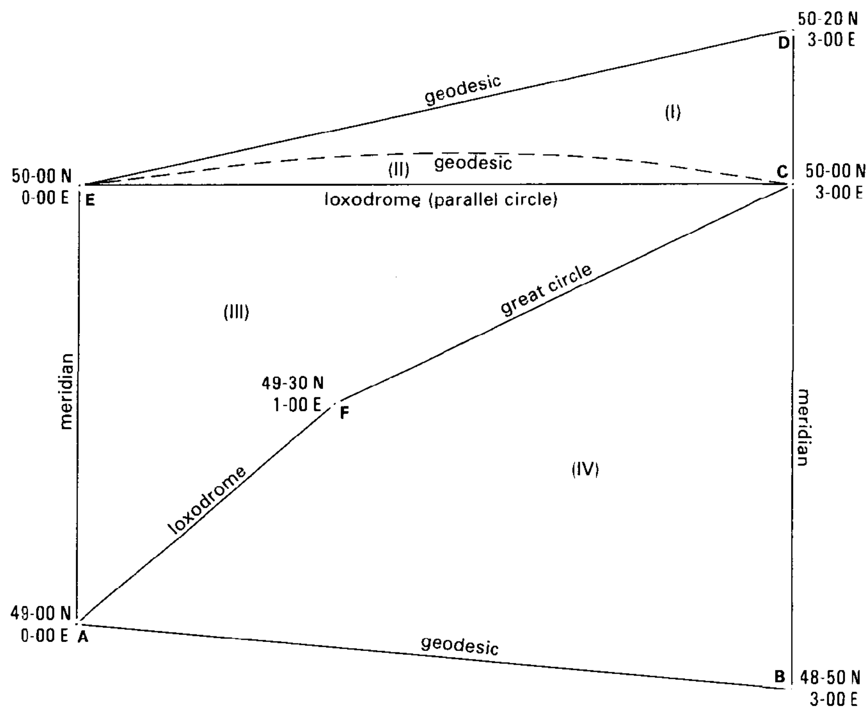

FIG. 5.- Polygon ABCDEA.

Example 3: Circular area on the ellipsoid

Centre : 50°N, 01°E
Radius : 10 000 m
Datum : European Datum 1950

Method:

A) 'Under geodesic':

number of interpolating points	distance (m) between points (circumference)	computed area area (km^2)	area diff. (m^2)
1) 150	419	314.067311	
			+ 68897
2) 300	210	314.136208	

Error computation:

area 1) = 314.067311

a = 2

e_1 = -4/3 x 68897 = - 91863 —> - 0.091863

Circle Area = 314.159174 -

B) Step-by-step [GILLISSEN]:
(computer printout)

Area computation, using Alberts projection, of polygon bounded by user defined sides:
1 = geodesic
2 = loxodrome
3 = great circle on conformal sphere
4 = great circle on concentric sphere
5 = arc of a circle

reference ellipsoid ED50							
Lat. Lon. centre	50	0	0.00	1	0	0.00	1
Polygon points:							
Lat. Lon.	50	5	23.64	1	0	0.00	5
Lat. Lon .	50	5	23.64	1	0	0.00	1

stepsize	=	200.0 m
nr. of interpolating points	=	315
On computing correction:		
stepsize	=	100.0 m
nr. of interpolating points	=	629

Area	=	314.153941 km²
Correction	=	.005233 km²
Corrected area	=	314.159174 km²

References

ACIC, 1959: Geodetic Distance and Azimuth Computations for Lines over 500 Miles, Technical Report No. 80, 1959. Aeronautical Chart and Information Centre, USA.

BOWRING, B.R., 1985: The Geometry of the Loxodrome on the Ellipsoid. The Canadian Surveyor, Vol. 39, No. 3, pp. 223-230.

DANIELSEN, J., 1989: The Area under the Geodesic. Survey Review, Vol. 32, No. 232, pp. 61-66.

DEETZ, H. and ADAMS, O.S., 1929: Elements of Map Projection. Spec. Publ. No. 68. Serial No. 146, Dept. of Commerce U.S. Coast & Geodetic Survey.

GILLISSEN, I., 1993: Area Computation of a Polygon on an Ellipsoid. To be published in April 1993 issue of Survey Review.

VINCENTY, T., 1975: Direct and Inverse Solutions of Geodesics on the Ellipsoid with Application of Nested Equations. Survey Review, Vol. XXII, No. 176, pp. 88-93.