SOME ASPECTS OF ELLIPSOIDAL DISTANCE ENUMERATION IN GIS APPLICATIONS

NĚKTERÉ ASPEKTY VÝPOČTU VZDÁLENOSTÍ NA ELIPSOIDU V APLIKACÍCH GIS

Vladimír HOMOLA

doc. Dr. Ph.D., Institute of Geological Engineering, Faculty of Mining and Geology,
VŠB - Technical University of Ostrava
17. listopadu 15, 708 33 Ostrava - Poruba, tel. (+420) 59 699 4367
e-mail: vladimir.homola@vsb.cz

Abstract

The article is the introductory study for one of the fundamental geo-informatics operations - the reciprocal position calculation of Earth localities. It is known that, in general, this calculation leads up to an elliptic integral calculation. In a computer environment, the algorithms are determined by data storage precision. Furthermore, repeated calculation with numerous data results in deceleration of all the applications. In special cases, the accurate application of ellipsoidal trigonometry can be replaced by a procedure that is more elementary. This study discusses some of these procedures and demonstrates their application in a given task.

Abstrakt

Článek je přípravnou studií pro jednu ze základních operací v geoinformatice - zjišťování vzájemné polohy míst na zemském povrchu. Jak známo, obecně vedou výpočty k výpočtu eliptických integrálů. Algoritmy realizované v počítačovém prostředí jsou však determinovány přesností uložení dat, opakované výpočty na značném množství dat navíc neúměrně zpomalují aplikaci. V některých specielních případech však lze nahradit doslovnou aplikaci elipsoidální trigonometrie jednoduššími postupy. Článek některé z nich diskutuje a odůvodňuje jejich použití v konkrétní úloze.

Key words: Geoinformatics, ellipsoid, elliptic integral, geographical coordinates, distance.

1 INTRODUCTION

In geo-informatics, the spatial component of data holds the information about the position on the Earth's surface. Due to the irregular shape of the Earth's body, a reference ellipsoid is used to approximate the Earth's surface. Actual position on the Earth's surface is transformed to a position on the ellipsoid surface, given by well-known geographic latitude and geographic longitude. Values of latitude and longitude are joined to the given ellipsoid dimensions. WGS-84 is the mostly used reference ellipsoid in global positioning systems, in navigation systems, etc.

In geo-sciences today, during field study, the predominant amount of spatial data is delivered by GPS. The distance between two points given by geographic coordinates is consequently the basis of geo-solid numeric descriptions: perimeter, volume, and the others. In next example, having a navigation track log, it is possible to reconstruct all the way driven through: actual and average speeds, total distance, and direction for instance. The two-point distance is here the total base of related quantities enumeration.

Two examples of distance usage given above are slightly different. Track log two-point distance is running to hundreds of meters, maximally. On the other hand, the dimension of geological rock unit can be up to hundreds of kilometres. However, in this case, the curvature of ellipsoidal surface cannot be omitted.

Mathematically, the exact way to calculate the distance between two points leads to solve the set of elliptic integrals (see below). This advancement seems to be unnecessarily complicated in such cases like the track log mentioned above. Let us discuss now some aspects of elliptic calculation, and its substitution by sphere and plane, respectively.

The conclusions of this article will be consecutively utilized in the software application handling the track log generated by a real navigation system.

2 TRACK LOG GENERATED

To clarify the mathematical reflection, the brief description of navigation system's track log is now given.

Primarily, the PDA's ***iGo 2006 Plus*** navigation software (developed by ***Nav-n-go Kft.***) is used in co-operation with an external GPS unit. This software generates the track log into a very simple structured GPX file (see http://www.topografix.com/GPX/1/1, for example). The file's XML scheme looks like the next figure:

```
xml                      version="1.0" encoding="UTF-8"
gpx
   xmlns                 http://www.topografix.com/GPX/1/0
   version               1.0
   creator               iGO 2006
   xmlns:xsi             http://www.w3.org/2001/XMLSchema-instance
   xsi:schemaLocation    http://www.topografix.com/GPX/1/0 http://www.topografix.com/GPX/1/0/gpx.xsd
   trk
      trkseg
         trkpt
            lat           50.64061499
            lon           13.82354665
            time
               #text      2007-11-02T10:30:15Z
      trkseg
         trkpt
            lat           50.64033663
            lon           13.82351494
            time
               #text      2007-11-02T10:30:25Z
```

Fig. 1 XML scheme of GPS file

As seen from Fig. 1, the data generated by the mentioned software are composed into a quaternion [Latitude; Longitude; Date; Time]. The memory amount is about 90 bytes for each position record. Suppose the traced distance through all the Czech Republic, i.e. about 500 [km]. Suppose the average speed (keeping all the road rules) about 50 [km/hrs]. So 10 [hrs] of contiguous recording are needed. If the interval of scanning is 1 [sec], then the track log contains about 36 000 records. The total amount of memory needed is about 3.5 [MB] max.

Finally, total range of today's memory card takes gigabytes. Four megabytes required to store the most voluminous track log, it is only a little bit of card capacity. It means that the optional recording time step may be set to 1 second without the risk of memory overflow. On the other side, the distance corresponding to 1-second stretch represents 50 meters (Skoda, 180 km/h), or 70 meters (Ferrari, 252 km/h), or 250 meters (Boeing, 900 km/h).

3 REFERENCE ELLIPSOID

The reference ellipsoid is a solid surface obtained by rotating a plane ellipse around its semi-minor axis. It serves primarily as a coordinate system holder. Although geocentric rectangular coordinates are mathematically exact Cartesian system, the geodetic latitude, geodetic longitude, and elevation form the most common Earth coordinate system used. For reminder, the longitude is expressed as degrees from <-180°; +180°> with the Prime Meridian at 0°, and latitude as degrees from <-90°; +90°> with the Equator at 0°.

To simplify the problem, let points A and B fall to the same meridian. We will determine the length of the meridian arc between A and B, s = s (A, B), with intention to replace this elliptic arc by simpler curve: circle arc, or line segment.
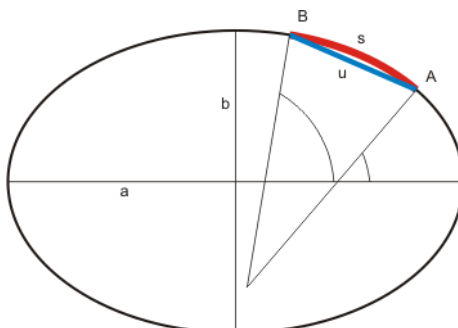

Fig. 2 Elliptic, circle, and line segments

Parametrical formulas for ellipse point are

$$x = a \, \cos \varphi$$
$$y = b \, \sin \varphi$$

Since, in general, the length of curve arc is defined as

$$s = \int_{\varphi_1}^{\varphi_2} \sqrt{\dot{x}^2 + \dot{y}^2} \; d\varphi$$

specifically for an ellipse with semi-axes **a** and **b** are

$$\dot{x} = -a \, \sin \varphi$$

$$\dot{y} = +b \, \cos \varphi$$

then the length of elliptic arc is

$$s = \int_{\varphi_1}^{\varphi_2} \sqrt{a^2 \sin^2 \varphi + b^2 \cos^2 \varphi} \; d\varphi$$

Letting

$$k^2 = \frac{b^2 - a^2}{b^2} \; < 1$$

is **s** (without decrease of universality it can be assumed **a**<**b**) given by

$$s = b \int_{\varphi_1}^{\varphi_2} \sqrt{1 - k^2 \sin^2 \varphi} \; d\varphi$$

Of course, there is

$$E(k,\varphi) = \int_0^{\varphi} \sqrt{1 - k^2 \sin^2 \phi} \; d\phi$$

the incomplete elliptic integral of the second kind. The required elliptic arc **s** is then given as

$$s = b(E(k,\varphi_2) - E(k,\varphi_1))$$

The incomplete elliptic integral can be expressed as a power series (see reference [1])

$$E(k,\varphi) = \sum_{n=0}^{\infty} \frac{\binom{-1/2}{n}}{n!} \; k^{2 \cdot n} \; S_{2 \cdot n}(\varphi) \tag{1}$$

where

$$S_{2 \cdot n}(\varphi) = \int_0^{\varphi} \sin^{2 \cdot n} \phi \, d\phi$$

and

$$\int_0^{\varphi} \sin^{2 \cdot n} \phi \, d\phi = 2^{-2 \cdot n}\left[\binom{2 \cdot n}{n} \varphi + Q(n,\varphi)\right]$$

and, finally, letting

$$Q(n,\varphi) = \sum_{i=1}^{n} (-1)^i \binom{2i}{n-i} \frac{\sin(2 i \varphi)}{i}$$

Let us discuss the evaluation of a given distance exactly in this way. It will be performed by some computer application. Standardly, common usable data model is related to the definition of *double-precision floating-point* numeric data. Excel, Basic (***double***), SQL (***float***), use such a data model. The precision is about 17 (decimal) valid digits. Unfortunately - using the ***double*** data model - the series (1) convergent rapidly. The member

$$\frac{\binom{-1/2}{n}}{n!} k^{2 \cdot n}$$

is in the sum of members in (1) un-significant at present for n=4. The following table lists the leading ten members (the value of k corresponds to the WGS-84 ellipsoid):

Tab. 1: Some members of series (1)

n	$P=k^{2n}$	$R=\binom{-1/2}{n}$	P R	n	$P=k^{2n}$	$R=\binom{-1/2}{n}$	P R
1	4,5E-05	-0,5000	-2,2E-05	6	8,1E-27	-0,0210	-1,7E-28
2	2,0E-09	-0,1300	-2,5E-10	7	3,6E-31	-0,0160	-5,8E-33
3	9,0E-14	-0,0630	-5,6E-15	8	1,6E-35	-0,0130	-2,1E-37
4	4,0E-18	-0,0390	-1,6E-19	9	7,3E-40	-0,0110	-8,0E-42
5	1,8E-22	-0,0270	-4,9E-24	10	3,3E-44	-0,0093	-3,0E-46

For example, applying this process to evaluate the length of the Earth's quadrant by applications using double precision data (Excel is one of them), the result differs about 16.7 [km] (= 0,16%) from reality.

In fact, there exist some algorithms to evaluate the incomplete integral of second kind (arithmetic – geometric mean algorithm, for example) eliminating the mentioned rapid convergence. However, the price is more complicated program code, and enumeration that is more difficult; keep in mind, the aim is an enumeration of thousands of such values!

4 SPHERICAL SURFACE

To approximate the elliptic surface by the surface of sphere in a near neighbourhood of point P, the casual sphere is that with the same curvature. Exactly, it has the same radius as the ellipsoidal radius of curvature in tangent point P.

Of course, there are three radii of curvature on the ellipsoidal surface in point P. The first one is the transverse (or normal) radius of curvature, usually marked as N. The second one is the meridian radius of curvature, usually marked as M. The third one is the geometric mean of both M and N, usually marked as R.

Letting a as semi-major axis, b as semi-minor axis of the reference ellipsoid, there are

$$\varepsilon^2 = 1 - \frac{b^2}{a^2}$$

$$W^2 = 1 - \varepsilon^2 \sin^2 \varphi$$

$$N = \frac{a}{W}$$

$$M = (1 - \varepsilon^2) \frac{a}{W}$$

$$R^2 = M\, N$$

where: ε - the first eccentricity, W – the first geodetic function. It is important to keep in mind, that all of above given geodetic quantities are the functions of latitude φ: W = W(φ), and so on.

The following figure shows the situation on ellipsoid to introduce its replacement by sphere:

$$\varepsilon = \sqrt{\frac{a^2 - b^2}{a^2}}$$

$$N = \frac{a}{\sqrt{1 - \varepsilon^2 \sin^2 \varphi}}$$

$$x = N \cos\varphi \cos\lambda$$

$$y = N \cos\varphi \sin\lambda$$

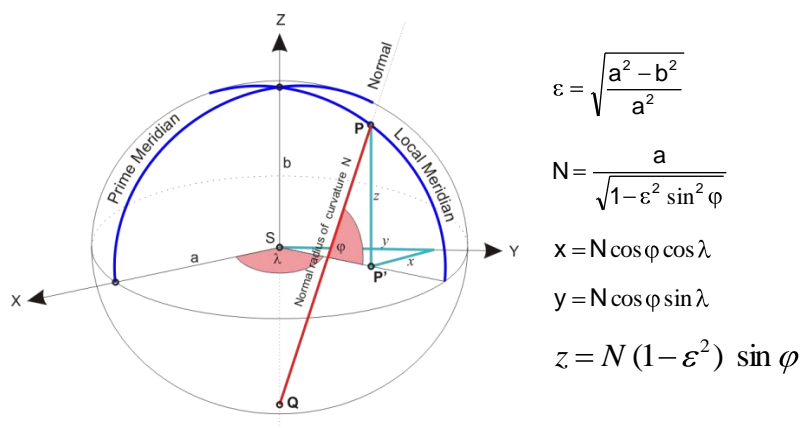$$z = N(1 - \varepsilon^2)\sin\varphi$$

Fig. 3 The ellipsoidal surface point P and appropriate formulas

Let us verify the individual values of some possibilities by direct calculation.

It is known, one *angular* meridian second is approximately 30 [m]. The speed 180 [km/h] equals to 50 [m/s]. It means, having the sample interval like one *time* second, it can be sure that no more than 60 [m] will be driven away. This distance is app. two *angular* seconds in the position of the Czech Republic. Let us compare the alternate calculation methods of two-point (A and B) distance having the same latitude coordinate, i.e. along the **meridian**. Let the *angular* distance between A and B is 2 seconds.

*Rem.: The distance enumeration along the **parallel** is no problem. The cross section through the reference ellipsoid, and the plane normal to the ellipsoid rotation axis, is a circle having the radius R (φ). The length of circle arc is the ratio of ΔAB and 2 × π.*

The meanings of the symbols used in the next text are as follows:

- u length of line segment AB,
- s length of elliptic arc AB,
- P centre of meridian segment AB,
- Rs length of radius vector of point P,
- N normal radius of curvature in point P,
- C centre of ellipse,
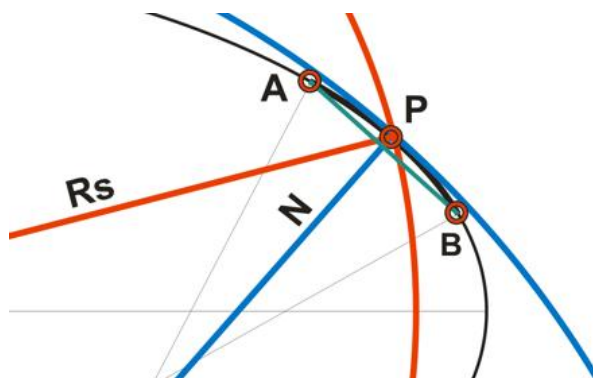- O intersection point of normal in point P and the line of semi-minor axis.

Fig. 4 Different segments connecting A and B

Now the comment on the next Fig. 5: Let s be enumerated by mathematical formulas given in chapter 3. Let s(AG) be enumerated by the same formulas but using the method of arithmetic - geometric mean. Let u be the result of the Pythagoras' theorem. Let s(N) be the length of circle arc having the radius N and the central angle AOB (Fig. 4: blue arc). Let s(R) be a length of circle arc having the radius Rs and the central angle ACB (Fig. 4: red arc). Let all enumerations be performed for B's latitudes 10°, 20°, ... , 89°.
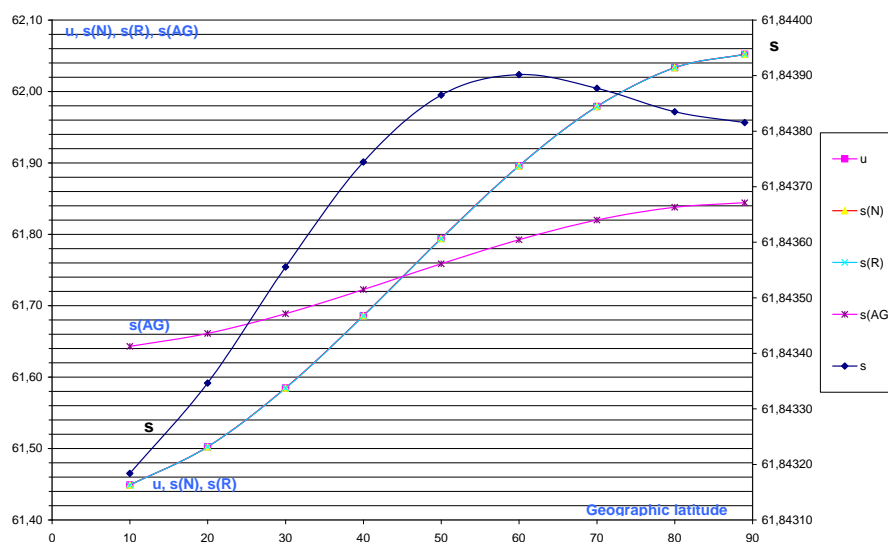
Fig. 5 Length of 2" arcs

All the enumerations were realized by Excel formulas. Certainly, the own functions were used for the elliptic integral solution. The line chart (Fig. 5) is the best way to demonstrate the result.

Primarily, it is evident the line s documents the missing members in the power series (1) - see chapter 3. The conclusion is that this method is not usable in the double-precision data environment.

The next interesting fact: The enumerations of straight-line distance, and both spherical distances, lead to practically equal results. It looks like the acknowledgement of the hypothesis that the plane can replace the nearest neighbourhood of the spherical point.

Finally, the most interesting is the s(AG) curve, being enumerated by the arithmetic - geometric mean method. It is not clearly seen in the chart, but considers: The radius of polar tangent circle equals to the semi-minor axis b of ellipsoid. The radius of equatorial tangent circle equals to the semi-major axis a of ellipsoid. The length of circle arc having the radius R and the central angle α is given by R α. In view of that, the ratio of arcs (on polar and equatorial spheres) with the same central angle should be the same as the ratio of both semi-axes. For WGS-84, the ratio a/b equals to 1.003364. For the s(AG) curve, the ratio of arcs having latitudes 0° and 90° equals to 1.003326 - excellent correspondence with the presumption. Thence is close to thesis, that s(AG) is the best correspondence to reality.

The mentioned ratio, for other three above given curves, is about 1.0101. Comparing them with s(AG), they are distributed in ±19 [cm] band along 60 [m] distance - it represents ±0.3%. For the Czech Republic latitude value, the difference is about ±0.06%, i.e. 3.5 [cm] along 60 [m] distance.

Related to one-second track log, it must be kept in mind that lowering the angular velocity, the line segment AB approximates the ellipsoidal surface more and more.

5   THE PLANE

Due to the results of both previous chapters, it is legitimate to consider the plane as an elliptic surface replacement in a near environment of the tangent point.

To approximate the surface by plane, it is appropriate to select the local tangent plane and set the secondary coordinate system in a natural way. Let Z-axis be equal to the normal both plane and elliptic surfaces. X-axis and Y-axis belong to the tangent plane with the expected orientation: from south to north (Y-axis), from west to east (X-axis).
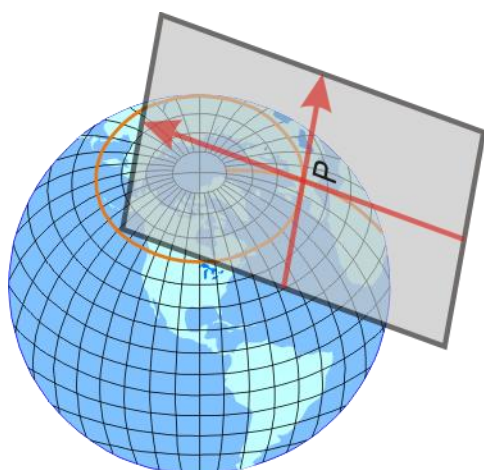
The next figures show this situation:
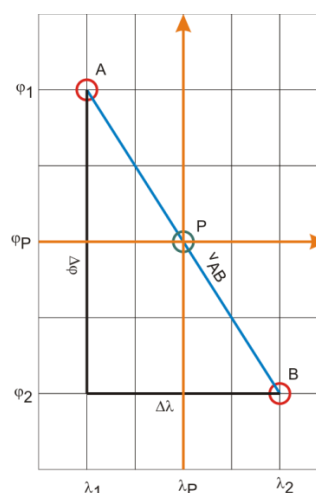
Fig. 6 Local tangent plane

Fig. 7 Secondary coordinate system and the formulas

$$\Delta\lambda = \lambda_2 - \lambda_1$$

$$\Delta\varphi = \varphi_2 - \varphi_1$$

$$\lambda_P = \frac{\lambda_1 + \lambda_2}{2}$$

$$\varphi_P = \frac{\varphi_1 + \varphi_2}{2}$$

The distance of two near points A and B can be evaluated as follows.

In common, lines of grid (see Fig. 7) are very complicated curves. But for small area, the grid can be considered as the regular rectangular net. Both the latitude and longitude differences are about ones of angular seconds. Having one angular second of latitude expressed in length units, and the same for one angular second of longitude, the Pythagoras' theorem can be applied directly.

For this reason, let the ellipsoidal surface be replaced by a tangent circle with the radius equal to the normal radius of curvature N. Then, the length of one-second arc along the meridian is given as

$$L_m^{1s} = \frac{\pi \cdot N}{180 \cdot 3600}$$

The radius of parallel circle having the latitude as φ is $N \cdot \cos(\varphi)$. Then, the length of one-second arc along the parallel is given as

$$L_p^{1s} = \frac{\pi \cdot N \cdot \cos(\varphi)}{180 \cdot 3600}$$

More universal task is now given. It solves both the radius of parallel circle and the radius of tangent sphere belonging to the point with a given latitude. This function can be called directly from Excel formulas.

```
Public Function RadiusAtLatitudeL _
    (ByVal qLatitude As Double, _
     Optional ByVal qA As Double = cSemiAxisA, _
     Optional ByVal qFr As Double = cFlatRec) _
  As Double()

  Dim V(0 To 1) As Double    ' Resulting array: (0)-for parallel (1)-for sphere
  Dim a As Double            ' Semi-major axis
  Dim b As Double            ' Semi-minor axis
  Dim psi As Double          ' Latitude in radians
  Dim eps2 As Double         ' eps2 = 1 - b^2/a^2
  Dim CosPsi As Double       ' cos (psi)
  Dim Sin2psi As Double      ' sin^2 (psi)
  Dim W As Double            ' First geodetic function
  Dim N As Double            ' Normal radius of curvature
  Dim rR As Double           ' Radius of parallel circle

  a = qA
  If qA = cSemiAxisA And qFr = cFlatRec Then
     b = cSemiAxisB
     eps2 = cEps2
  Else
     b = qA * (qFr - 1#) / qFr
     eps2 = 1# - (b * b) / (a * a)
  End If

  psi = qLatitude * cDegToRad
  CosPsi = Cos(psi)
  Sin2psi = 1 - CosPsi * CosPsi

  W = Sqr(1 - eps2 * Sin2psi)
```

```
    N = a / W
    rR = N * CosPsi

    V(0) = rR                        ' Radius of parallel circle
    V(1) = N                         ' Radius of sphere

  RadiusAtLatitudeL = V

End Function
```

This function may be used for any other reference ellipsoid. If second and third parameters are missing, WGS-84 is used. The constants of identifier like cXX are declared at the public level as follows:

```
' Constants - WGS-84:
Public Const cSemiAxisA As Double = 6378137            ' Semi-major axis [m]
Public Const cSemiAxisB As Double = 6356752.31424518   ' Semi-minor axis [m]

Public Const cFlatRec As Double = 298.257223563        ' Reciprocal flattening

Public Const cEccentricity As Double = 521854.008423385  ' Eccentricity [m]

Public Const cEps As Double = 0.081819190842622        ' First eccentricity
Public Const cEps2 As Double = cEps * cEps             ' Square of 1. ecc.

Public Const cPi As Double = 3.14159265358979          ' Ludolph number
Public Const cDegToRad As Double = cPi / 180#          ' Degrees to radians
```

6 CONCLUSIONS

The article primarily presents a mathematically exact way to determine the distance between two points on the Earth's surface. In this case, it alerts to the insufficient precision of computer's data models. Giving all the formulas needed, very complicated application programming can be expected. The next step is checking some simpler surface than the ellipsoid is. Experimenting with a sphere, it was demonstrated that there is only an elementary difference between the circle arc and the line segment in a small surface area. Therefore, the last step is the distance evaluation in a plane environment, but keeping the geographical, not Cartesian coordinates. Finally, the sufficient but simple and rapid function returning values needed for a distance evaluation is given.

REFERENCES

[1] REKTORYS et al.: *Přehled užité matematiky*. 1st ed. Praha : SNTL, 1963. 1136 pp.

[2] ŠKRÁŠEK, J. & TICHÝ, Z.: *Základy aplikované matematiky II*. 1st ed. Praha : SNTL, 1986. 896 pp.

RESUMÉ

Článek se zabývá některými aspekty určování vzdáleností na zemském povrchu. V době GPS, navigačních systémů a mnohých úloh geoinformatiky je to jedna ze základních úloh, která přitom podstatně ovlivňuje výsledek. Nejde totiž o samostatný, jednorázový výpočet vzdálenosti dvou bodů, ale o opakované výpočty nad tisíci a desetitisíci daty.

Určujícím aspektem je to, že Země je idealizovaně pojímaná jako rotační elipsoid. Ten je nositelem geografické soustavy souřadné, v níž jsou dodávány prostorové složky geo-dat. Výpočet vzdáleností na elipsoidu vede k eliptickým integrálům, které nejsou triviálně řešitelné. Vyjádření nekonečnými řadami, byť matematicky přesné, vede v počítačovém prostředí k obtížně eliminovatelným nepřesnostem díky konečné, relativně malé přesnosti zobrazení racionálních čísel. Článek to ukazuje na nepříjemně rychle konvergující posloupností členů konkrétní mocninné řady.

Proto je v dalším kroku nutné opustit požadavek obecnosti. V řadě případů - a jako zástupce takové třídy úloh byl zvolen např. záznam trasy navigačního programu - však jde o vzdálenosti bodů nepříliš vzdálených. Povrch elipsoidu tak velkého by mohl být aproximován jednodušší plochou. Nabízí se koule. Článek upřesňuje, která (z několika možností) je ta koule, která může být pro danou úlohu použita. Při porovnávání délky eliptického oblouku a oblouku na kouli však bylo zjištěno, že chyba vzniklá náhradou koulí se téměř neliší od chyby vzniklé náhradou za úsečku.

Následuje tedy rozbor náhrady tečnou rovinou. Uvádí se příklad v ní zavedené soustavy souřadné včetně potřebných transformačních vztahů. V závěru článku je uvedena funkce odevzdávající poloměry tečné koule na dané zeměpisné šířce a rovnoběžkové kružnice na této šířce. Obě dvě hodnoty se pak použijí pro výpočet vzdálenosti. Funkce je v rámci uvedených intervalů přesná a velmi rychlá při použití: neobsahuje cykly ani podmínky. Uvedený článek je tak teoretickým základem aplikace rozebírající a zpracovávající skutečné záznamy trasy navigačních programů.