Navigation on the spheroidal earth

Ed Williams

March 21, 2002

Adobe Acrobat version of this document

1 Spheroid Geometry

For many purposes, it is entirely adequate to model the earth as a sphere. However, in reality, the earth's mean sea-level surface is better approximated by a different geometric shape, an oblate spheroid– the surface created by rotating an ellipse about its polar axis. Compared to a sphere, an oblate spheroid is flattened at the poles. The earth's flattening is quite small, about 1 part in 300, and navigation errors induced by assuming the earth is spherical, for the most part, do not exceed this, and so for many purposes a spherical approximation may be entirely adequate. On a sphere, the commonly used coordinates are latitude and longitude, likewise on a spheroid, however on a spheroid one has to be more careful about what exactly one means by *latitude*.

In Figure 1 we depict a cross-section of the spheroid through the poles. The point O is the center of the earth. B is the North Pole. The major (equatorial) axis, OA, of the meridional ellipse has length a, the minor (polar) axis, OB has length b. A point P on the ellipse has coordinates $(a\cos\theta, b\sin\theta)$ where the angle $\angle AOP'$ is called the *reduced* or *parametric* latitude. The point P' is the point on the circumscribing circle (of radius a) the same distance from the polar axis as P. The angle $\angle AOP$ is called the *geocentric* latitude.

However, the latitude used in navigation and geodesy is the *geodetic* or *astronomical* latitude, which is defined to be the angle between the northerly horizon at P and the polar axis. It is equal to the angle ϕ in Figure 1.

Longitude, L, is defined in exactly the same way on the spheroid as on the sphere, namely as the angle between the meridian and the prime meridian. We use here the standard convention of North longitudes and East latitudes as positive.

In three dimensional Cartesian coordinates, points on the spheroid have coordinates $(a\cos\theta\cos L, a\cos\theta\sin L, b\sin\theta)$.

There are alternative ways of specifying the dimensions of the spheroid other than by its major and minor radii a and b. The *flattening*, f, is defined by $f = 1 - b/a$, and the *eccentricity* e by $e^2 = 1 - b^2/a^2$. For the WGS84 spheroid, $a = 6378.137$ km and $f = 1/298.257223563$. The eccentricity and flattening are

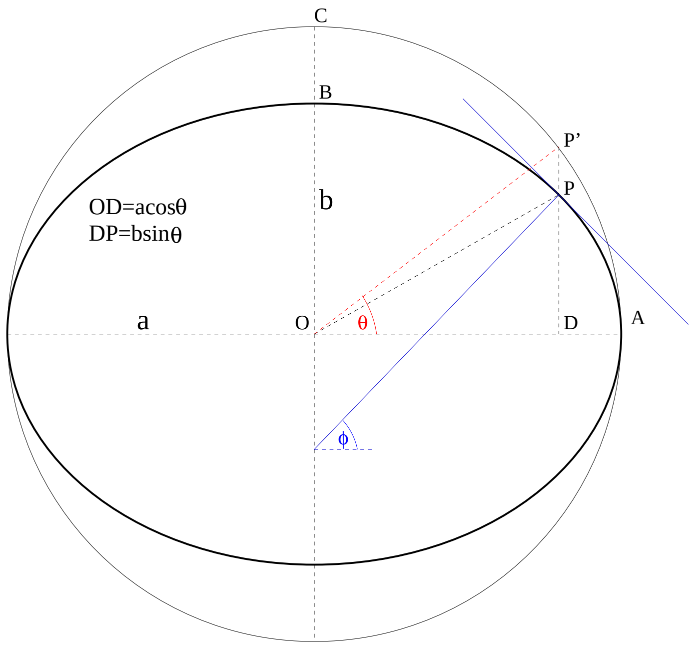

Figure 1: The meridional ellipse

thus related by:

$$e^2 = f(2 - f) \tag{1}$$

1.1 Differential geometry

A displacement of $\mathrm{d}\theta$ in parametric latitude along the meridional ellipse is illustrated in Figure 2. The poleward displacement is $\mathrm{d}(b \sin\theta) = b\cos\theta \mathrm{d}\theta$ and the equatorial component is $\mathrm{d}(a\cos\theta) = -a\sin\theta\mathrm{d}\theta$ From this we see that the geodetic and parametric latitudes are related by

$$\tan\phi = (a/b)\tan\theta \tag{2}$$

and that the displacement along the meridian is given by:

$$\begin{aligned}(a^2\sin^2\theta + b^2\cos^2\theta)^{1/2}\mathrm{d}\theta &= a(1 - e^2\cos^2\theta)^{1/2}\mathrm{d}\theta \\ &= a\frac{(1-e^2)}{(1 - e^2\sin^2\phi)^{3/2}}\mathrm{d}\phi = R_\phi\,\mathrm{d}\phi \end{aligned} \tag{3}$$

R_ϕ is the radius of curvature of the meridional arc at P. The radius of curvature in the perpendicular plane (i.e. in the plane of the parallel), R_L is given by

$$OD = a\cos\theta = a\cos\phi/(1 - e^2\sin^2\phi)^{1/2} \tag{4}$$

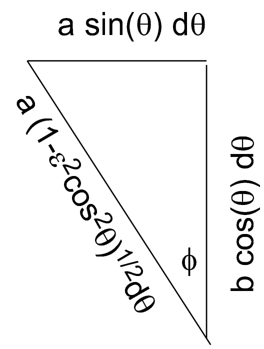

Figure 2: Triangle resulting from infinitesimal latitude change

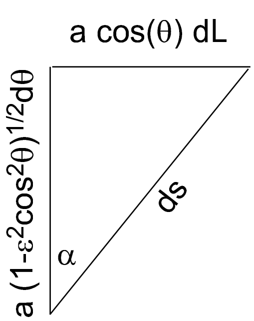

Figure 3: Triangle resulting from infinitesimal latitude and longitude changes

In Figure 3 we illustrate the result of a displacement of $\mathrm{d}\theta$ in parametric latitude and of $\mathrm{d}L$ in longitude, resulting in a Northerly displacement of $a(1-e^2\cos^2\theta)^{1/2}\mathrm{d}\theta$ and an Easterly displacement of $a\cos\theta\mathrm{d}L$, respectively. By Pythagoras' theorem the displacement distance $\mathrm{d}s$ is given by:

$$\mathrm{d}s^2 = a^2(\cos^2\theta\,\mathrm{d}L^2 + (1-e^2\cos^2\theta)\,\mathrm{d}\theta^2) \tag{5}$$

or, using eqn. (2),

$$\mathrm{d}s^2 = a^2\left(\frac{\cos^2\phi\,\mathrm{d}L^2}{1-e^2\sin^2\phi} + \frac{(1-e^2)^2\,\mathrm{d}\phi^2}{(1-e^2\sin^2\phi)^3}\right) \tag{6}$$

The true course α is given by:

$$\tan\alpha = \frac{\cos\theta}{(1-e^2\cos^2\theta)^{1/2}}\frac{\mathrm{d}L}{\mathrm{d}\theta} = \frac{\cos\phi(1-e^2\sin^2\phi)}{1-e^2}\frac{\mathrm{d}L}{\mathrm{d}\phi} \tag{7}$$

where we have again used Eqn. (2) to transform between reduced and geodetic latitude coordinates. Equations (5), (6) and (7) are the fundamental relations relating distances and directions on the spheroid at a point.

2 Rhumb Lines

Rhumb lines are paths of constant true course. They thus satisfy Eqn. (7) with α constant. This is most easily treated in geodetic coordinates. Integrating this equation, we obtain:

$$\begin{aligned} L|_{L1}^{L} &= \frac{\tan\alpha}{2}\log\left(\left(\frac{1+\sin\phi}{1-\sin\phi}\right)\left(\frac{1-e\sin\phi}{1+e\sin\phi}\right)^{e}\right)\Bigg|_{\phi1}^{\phi} \\ &= \tan\alpha\,\log\left(\tan(\phi/2+\pi/4)\left(\frac{1-e\sin\phi}{1+e\sin\phi}\right)^{e/2}\right)\Bigg|_{\phi1}^{\phi} \end{aligned} \tag{8}$$

giving the coordinates of points (ϕ, L) on a rhumb line with course α through $(\phi_1, L1)$.

Inverting this relation, to find the geodetic latitude ϕ given the longitude L, can most readily be done iteratively, using:

$$\phi = -\pi/2 + 2\tan^{-1}\left(\tan(\phi_1/2+\pi/4)\exp\left(\frac{L-L_1}{\tan\alpha}\right)\left(\frac{\left(\frac{1+e\sin\phi}{1-e\sin\phi}\right)}{\left(\frac{1+e\sin\phi_1}{1-e\sin\phi_1}\right)}\right)^{e/2}\right) \tag{9}$$

starting with $\phi = \phi_1$ on the RHS.

Combining equations (6) and (7) with α constant gives us a differential equation for the arc-length along the rhumb line:

$$\frac{\mathrm{d}s}{\mathrm{d}\phi} = \frac{a(1-e^2)}{\cos\alpha(1-e^2\sin^2\phi)^{3/2}} \tag{10}$$

Thus we find that s is given by

$$s = (M(\phi) - M(\phi_1))/\cos\alpha \tag{11}$$

where the function $M(\phi)$ is the distance from the equator to the ϕ parallel measured along a meridian, given by:

$$M(\phi) = a\int_0^{\phi} \frac{1-e^2}{(1-e^2\sin^2\phi')^{3/2}}\,\mathrm{d}\phi' \tag{12}$$

Formally, $M(\phi)$ can be expressed in terms of the elliptic integral of the second kind $E(\phi, e)$ by[1]

$$M(\phi) = a(E(\phi, e) - e^2\sin\phi\cos\phi/(1-e^2\sin^2\phi)^{1/2}) \tag{13}$$

Expanding to $O(e^6)$, $M(\phi)$ is approximately given by[4]:

$$\begin{aligned} M(\phi) = \quad & a \quad [(1 - e^2/4 - 3e^4/64 - 5e^6/256 - \ldots)\phi \\ & - \quad (3e^2/8 + 3e^4/32 + 45e^6/1024 + \ldots)\sin 2\phi \\ & + \quad (15e^4/256 + 45e^6/1024 + \ldots)\sin 4\phi \\ & - \quad (35e^6/3072 + \ldots)\sin 6\phi \\ & + \quad \ldots] \end{aligned} \tag{14}$$

Along a parallel, which is an E-W rhumb line, Eqns. (8) and (11) diverge, but since ϕ is constant, we have from Eqn. (6):

$$s = aR_L(L - L_1) = a\cos\phi\,(L - L_1)/(1-e^2\sin^2\phi)^{1/2} \tag{15}$$

A map with longitude as the x-axis and $M(\phi)$ as the y-axis has a Mercator[4] projection (with the equator as the standard parallel) on which rhumb lines plot as straight lines with the correct azimuth.

3 Geodesics on a spheroid

The shortest path between two points on a (smooth) surface is called a *geodesic curve* on the surface. On a flat surface the geodesics are straight lines, on a sphere they are *great circles*. Remarkably, the path taken by a particle sliding without friction on a surface will always be a geodesic. This is because a defining characteristic of a geodesic is that at each point on its path, the local center of curvature always lies in the direction of the surface normal, that is, in the direction of any constraint force required to keep the particle on the surface. There are thus no forces in the local tangent plane of the surface to *deflect* the particle from its geodesic path.

There is a general procedure, using the calculus of variations[2], to find the equation for geodesics given the metric of the surface (ie given Eqn. or (5)) (6). However, in this case, a simpler argument suffices.

Consider a particle of mass m sliding on the surface of a spheroid. The constraint forces, normal to the surface, do no work, so the particle's kinetic energy $mv^2/2$, and thus its speed v, remain constant. In addition, because of the spheroid's axisymmetry, all its surface normals pass through the polar axis. Thus the constraint forces have zero moment about the polar axis. The angular momentum of the particle around the polar axis is therefore conserved. Referring to Figs. 1 and 3 we can write this as $mv \sin\alpha \times OD$, where $vsin\alpha$ is the azimuthal component of the particle's velocity and OD is the distance of the particle from the polar axis OB. Thus, using Eqn. (4) we obtain:

$$\sin\alpha\cos\theta = \sin\alpha\cos\phi/(1 - e^2\sin^2\phi^2)^{1/2} = \text{constant} \tag{16}$$

If we refer to the azimuth of the geodesic as it crosses the equator ($\theta = \phi = 0$) as α_0, we can evaluate the above constant, obtaining:

$$\sin\alpha\cos\theta = \sin\alpha\cos\phi/(1 - e^2\sin^2\phi^2)^{1/2} = \sin\alpha_0 \tag{17}$$

This equation takes its simplest form when reduced latitudes, θ are used, so geodesic calculations are generally done in (θ, L) coordinates, with necessary conversions back and forth to geodetic coordinates being performed using Eqn. (2). In fact, the relationship between the azimuth α and the reduced latitude θ on a spheroidal geodesic is the same as on a spherical great circle. This sets up a correspondence between geodesics and great circles on an *auxiliary sphere*[3] with a common value of α_0. At each (reduced) latitude, θ the geodesics have the same azimuth, α. However, distances and longitude differences differ by $O(e^2)$ corrections. The great circle distances and longitude differences can be used as a first approximation to an iterative or perturbative evaluation of the corresponding quantities on the spheroidal geodesic.

3.1 Geodesic arc length

From Fig. 3 and Eqn. (17) we see that:

$$\cos\alpha = a(1 - e^2\cos^2\theta)^{1/2}\frac{\mathrm{d}\theta}{\mathrm{d}s} = \pm(\cos^2\theta - \sin^2\alpha_0)^{1/2}/\cos\theta \tag{18}$$

and thus

$$\frac{\mathrm{d}s}{\mathrm{d}\theta} = \pm a\frac{\cos\theta(1 - e^2\cos^2\theta)^{1/2}}{(\cos^2\alpha_0 - \sin^2\theta)^{1/2}} \tag{19}$$

with the sign being that of $\cos\alpha$. We now substitute $\sin\theta = \sin\sigma\cos\alpha_0$, where $a\sigma$ is the arc-length along the great circle on the auxiliary sphere, measured from where it crosses the equator in a Northerly direction, obtaining:

$$\frac{\mathrm{d}s}{\mathrm{d}\sigma} = a(1 - e^2\cos^2\theta)^{1/2} = b(1 + u^2\sin^2\sigma)^{1/2} \tag{20}$$

where $u^2 \equiv e^2\cos^2\alpha_0/(1 - e^2)$. By expanding in a power series in u^2 and integrating term by term[3, 5], we obtain s in the form:

$$s/b = \quad \sigma(1 + u^2/4 - 3u^4/64 + 5u^6/256 - 175u^8/16384 + \ldots)$$

$$
\begin{aligned}
&- \sin 2\sigma\,(u^2/8)(1 - u^2/4 + 15u^4/128 - 35u^6/512 + \ldots)\\
&- \sin 4\sigma\,(u^4/256)(1 - 3u^2/4 + 35u^4/64 - \ldots)\\
&- \sin 6\sigma\,(u^6/3072)(1 - 5u^2/4 + \ldots)\\
&- \sin 8\sigma (5u^8/131072)(1 - \ldots)\\
&- \ldots
\end{aligned} \tag{21}
$$

The distance between two points $s(\sigma_2, \alpha_0) - s(\sigma_1, \alpha_0)$, on a geodesic arc is best obtained, after differencing Eqn. (21), by using the identity $\sin(2n\sigma_2) - \sin(2n\sigma_1) = 2\cos(2n\sigma_m)\sin n(\sigma_2 - \sigma_1)$, where $\sigma_m = (\sigma_1 + \sigma_2)/2$. This avoids excessive loss of significant digits when the two points are close together.

Vincenty[5] has rearranged a subset of the resulting equations into nested forms more suitable for computation:

$$
\begin{aligned}
\tan\sigma_1 &= \tan\phi_1 / \cos\alpha_1\\
\sin\alpha_0 &= \cos\phi_1 \sin\alpha_1\\
u^2 &= e^2 \cos^2\alpha_0 / (1 - e^2)\\
A &= 1 + \frac{u^2}{16384}(4096 + u^2(-768 + u^2(320 - 175u^2)))\\
B &= \frac{u^2}{1024}(256 + u^2(-128 + u^2(74 - 47u^2)))\\
\sigma_m &= \sigma_1 + \sigma/2\\
\Delta\sigma &= B\sin\sigma\,(\cos 2\sigma_m + \frac{B}{4}(\cos\sigma\,(-1 + 2\cos^2 2\sigma_m)\\
&\quad -\frac{B}{6}\cos 2\sigma_m\,(-3 + 4\sin^2\sigma)(-3 + 4\cos^2 2\sigma_m))))\\
s &= bA(\sigma - \Delta\sigma)
\end{aligned} \tag{22}
$$

In equations (22) the origins of σ and s have been shifted from the equator to the initial point (1), where the reduced latitude is ϕ_1 and the azimuth of the geodesic is α_1.

3.2 Longitude difference

Again from Fig. 3 and Eqn. (17) we see that:

$$
\tan\alpha = \frac{\cos\theta}{(1 - e^2\cos^2\theta)^{1/2}} \frac{\mathrm{d}L}{\mathrm{d}\theta} \tag{23}
$$

which, in terms of the arc-length on the auxiliary sphere, σ, using the geodesic condition (17) becomes:

$$
\frac{\mathrm{d}L}{d\sigma} = \sin\alpha_0 \frac{(1 - e^2(1 - \cos^2\alpha_0\,\sin^2\sigma))^{1/2}}{(1 - \cos^2\alpha_0\,\sin^2\sigma)} \tag{24}
$$

In the spherical limit, $e = 0$, this is readily integrated to give $L = \lambda$, where $\tan\lambda = \sin\alpha_0 \tan\sigma$. λ is the longitude difference, corresponding the arc-length σ on the auxiliary sphere.

Expanding in powers of e^2 and integrating term by term, we thus obtain:

$$L = \lambda - e^2 \sin\alpha_0 (J_0\sigma + e^2 J_2 \sin 2\sigma + e^4 J_4 \sin 4\sigma + e^6 J_6 \sin 6\sigma + O(e^8)) \quad (25)$$

where

$$\begin{aligned}
J_0 &= \frac{1}{2}\left(1 + \frac{e^2}{8}(2-\mu) + \frac{e^4}{64}(8 - 8\mu + 3\mu^2) + \frac{5e^6}{1024}(16 - 24\mu + 18\mu^2 - 5\mu^3)\right) \\
J_2 &= \frac{\mu}{32}\left(1 + \frac{e^2}{2}(2-\mu) + \frac{15e^4}{256}(16 - 16\mu + 5\mu^2)\right) \\
J_4 &= \frac{\mu^2}{512}\left(1 + \frac{15e^2}{16}(2-\mu)\right) \\
J_6 &= \frac{5\mu^3}{24576} \\
\mu &= \cos^2\alpha_0
\end{aligned} \quad (26)$$

Vincenty[5] has again rearranged a subset of the resulting equations into nested forms more suitable for computation:

$$L = \lambda - (1-C) f \sin\alpha_0 (\sigma + C \sin\sigma(\cos 2\sigma_m + C\cos\sigma(-1 + 2\cos^2 2\sigma_m))) + O(f^4) \quad (27)$$

where

$$C = f\cos^2\alpha_0 (4 + f(4 - 3\cos^2\alpha_0))/16 \quad (28)$$

As in Eqn. (22), the origins of σ,λ and L have been shifted from the equator to the initial point (1). λ is then given by

$$\tan\lambda = \frac{\sin\sigma \sin\alpha_1}{\cos\phi_1 \cos\sigma - \sin\phi_1 \sin\sigma \cos\alpha_1} \quad (29)$$

on solving the spherical triangle on the auxiliary sphere.

4 Symbol Glossary

ϕ	Geodetic latitude
θ	Reduced latitude, defined by Eqn. (2)
L	Longitude (difference)
α	azimuths, clockwise from N (α_0 at equator)
s	arc-length along geodesic
λ	longitude (difference) on auxiliary sphere
σ	arc-length on auxiliary sphere
a	major (equatorial) radius of ellipse
b	minor (polar) radius of ellipse
f	flattening $b = a(1 - f)$
e	eccentricity $b^2 = a^2(1 - e^2)$

List of Figures

1 The meridional ellipse 2
2 Triangle resulting from infinitesimal latitude change 3
3 Triangle resulting from infinitesimal latitude and longitude changes 3

References

[1] Abramowitz, Milton, and Stegun, I. A. *Handbook of Mathematical Functions.* Washington, National Bureau of Standards. (reprinted by Dover Publications, Inc., New York, 1972).

[2] Arfken, G. "Calculus of Variations" Ch. 17 in *Mathematical Methods for Physicists, 3rd. Ed.* Orlando, Fl.: Academic Press, pp. 925-962, 1985.

[3] Rainsford, H. F. *Long Lines on the ellipsoid* Bulletin Géodésique, **30**, pp. 369-392, (1953) and Bulletin Géodésique, **37**, (Sept 1955)

[4] Snyder, John P. *Map Projections – A Working Manual* U. S. Geological Survey Professional Paper 1395. (reprinted by Navtech Seminars)

[5] *Direct and Inverse Solutions of Geodesics on the ellipsoid with Application of Nested Equations.* Survey Review, **XX11**, 176, (1975)