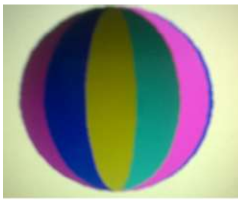
The equation of a great circle
Geodesics on a sphere



Ole Witt-Hansen 2019

Contents

1. Curve on a regular surface

Let there be given a continuous and differentiable surface $z = f(x, y)$, and let us assume that $r(t) = (x(t), y(t), z(t))$ is a parametric equation on the surface, which therefore fulfils $z(t) = f(x(t), y(t))$.
We challenge ourselves to find an (differential) equation for the shortest path on the surface between two given points.

This problem is formally solved in general in differential geometry, by using the metric fundamental form an covariant differentiation, but here we shall apply a more direct approach. (See e.g www.olewitthansen.dk Differential Geometry 1).

It is a classical problem from the calculus of variation, finding the minimum/maximum for the functional:

(1.1) $$I(y) = \int_a^b F(y', y, x)dx$$

The solution is the Euler-Lagrange differential equations.

(1.2) $$\frac{d}{dx}\frac{\partial F}{\partial y'} - \frac{\partial F}{\partial y} = 0$$

If F does not depend explicitly on x, this equation may be reduced to a first order differential equation:

(1.3) $$y'\frac{\partial F}{\partial y'} - F = Const$$

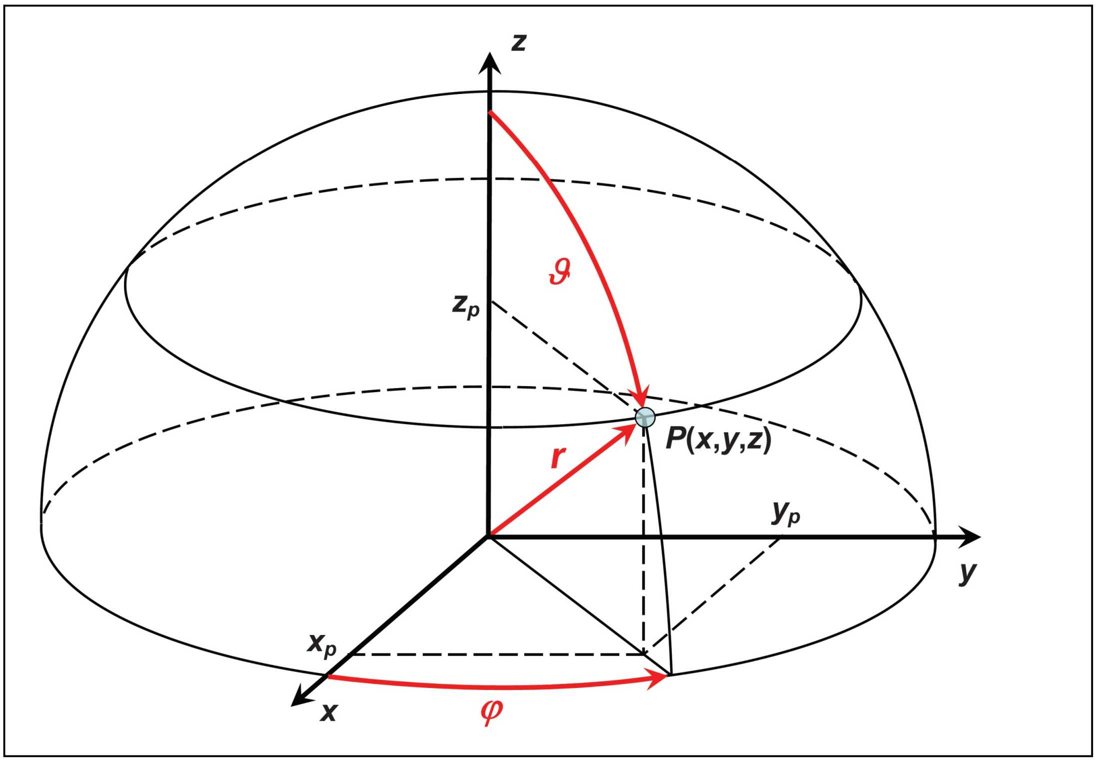
(See e.g www.olewitthansen.dk: Calculus of variation).
Some variation problems are stated with a side condition: $g(x,y) = k$, (e.g maximum volume within a given surface: The following theorem apply:

$f(x,y)$ has max/min in a point, given the side condition $g(x,y) = k$, if and only if the function $f(x,y) + \lambda g(x,y)$ has max/min in the point, and where λ is to be determined by the border conditions of the problem.

2. Shortest path between two points on a sphere

A great circle is a circle, the plane of which intersects the centre of the sphere.
It is common knowledge that a great circle arc connecting the two points is the shortest path between the two points.

We shall now prove analytically, when using the Euler-Lagrange equations, that this is indeed the case.
Using polar coordinates the condition $x^2 + y^2 + z^2 = r^2$ is automatically fulfilled, so we do not need a side condition.



The significance of the polar coordinates is to be seen in the figure above.
The coordinates for a point on a sphere, having radius r, has the polar coordinates:

(2.1) $$\vec{r} = (r\cos\varphi\sin\theta, r\sin\varphi\sin\theta, r\cos\theta)$$

The infinitesimal displacement on a sphere is then given by:

(2.2) $$ds^2 = r^2\sin^2\theta d\varphi^2 + r^2 d\theta^2 \quad \Rightarrow$$
$$(\frac{ds}{dt})^2 = r^2\sin^2\theta(\frac{d\varphi}{dt})^2 + r^2(\frac{d\theta}{dt})^2$$

If we consider a parametric equation for a curve on the sphere , and if we as usual designate differentiation with respect to t, with a period above the variable, the we get:

(2.3) $$\dot{s}^2 = r^2\sin^2\theta\dot{\varphi}^2 + r^2\dot{\theta}^2 .$$

To determine the shortest path between two points, we must determine the min for the functional.

(2.4) $$I = \int ds = \int_0^t \dot{s}dt = \int_0^t F dt = \int_0^t \sqrt{r^2\sin^2\theta\dot{\varphi}^2 + r^2\dot{\theta}^2}\,dt$$

To easy the calculations we put r =1, and determine the min for the functional: $\int_0^t \sqrt{\sin^2\theta\dot{\varphi}^2 + \dot{\theta}^2}\,dt$

The Euler-Lagrange equation $y'\frac{\partial F}{\partial y'} - F = Const$ then becomes two equations:

$$\dot{\theta}\frac{\partial F}{\partial \dot{\theta}} - F = c_1 \qquad \text{and} \qquad \dot{\varphi}\frac{\partial F}{\partial \dot{\varphi}} - F = c_2$$

(2.5) $$\frac{\partial F}{\partial \dot{\theta}} = \frac{\dot{\theta}}{\sqrt{\sin^2\theta\dot{\varphi}^2 + \dot{\theta}^2}} \qquad \text{and} \qquad \frac{\partial F}{\partial \dot{\varphi}} = \frac{\sin^2\theta\dot{\varphi}}{\sqrt{\sin^2\theta\dot{\varphi}^2 + \dot{\theta}^2}},$$

So the two equations yield:

(2.6) $$\frac{\dot{\theta}^2}{\sqrt{\sin^2\theta\dot{\varphi}^2 + \dot{\theta}^2}} - \sqrt{\sin^2\theta\dot{\varphi}^2 + \dot{\theta}^2} = c_1$$
$$\frac{\sin^2\theta\dot{\varphi}^2}{\sqrt{\sin^2\theta\dot{\varphi}^2 + \dot{\theta}^2}} - \sqrt{\sin^2\theta\dot{\varphi}^2 + \dot{\theta}^2} = c_2$$

Multiplying by the square root on both sides give:

(2.7) $$\dot{\theta}^2 - (\sin^2\theta\dot{\varphi}^2 + \dot{\theta}^2) = c_1\sqrt{\sin^2\theta\dot{\varphi}^2 + \dot{\theta}^2}$$

(2.8) $$\sin^2\theta\dot{\varphi}^2 - (\sin^2\theta\dot{\varphi}^2 + \dot{\theta}^2) = c_2\sqrt{\sin^2\theta\dot{\varphi}^2 + \dot{\theta}^2}$$

Since $\dot{s} = \sqrt{\sin^2\theta\dot{\varphi}^2 + \dot{\theta}^2}$, we may rewrite the two equations:

(2.9) $$\dot{\theta}^2 - \dot{s}^2 = c_1\dot{s}$$
$$\sin^2\theta\dot{\varphi}^2 - \dot{s}^2 = c_2\dot{s}$$

Notably, if we add the two equations, we may eliminate the two angular parameters.

$$\sin^2\theta\dot{\varphi}^2 + \dot{\theta}^2 - 2\dot{s}^2 = (c_1 + c_2)\dot{s} \quad \Leftrightarrow \quad \dot{s}^2 - 2\dot{s}^2 = (c_1 + c_2)\dot{s}$$

Which can be reduced to

$$-\dot{s}^2 = (c_1 + c_2)\dot{s}$$

And can immediately be solved.

(2.10) $$\dot{s}(\dot{s} + (c_1 + c_2) = 0 \quad \Leftrightarrow \quad \dot{s} = 0 \quad \vee \quad \dot{s} = -(c_1 + c_2) = c$$

An extremely simple equation, with has the solution, (since we assume that $s = 0$ when $t = 0$):

(2.11) $$s = ct$$

We shall then return to the two angular equations, where we insert the expression for s.

(2.12)
$$\dot{\theta}^2 - \dot{s}^2 = c_1\dot{s} \quad \Rightarrow \quad \dot{\theta}^2 - c^2 = c_1 c \quad \Rightarrow \quad \dot{\theta} = \pm c_\theta \quad \Rightarrow$$
$$\theta = \pm c_\theta t + \theta_0$$

And in a similar manner for the other equation

$$\sin^2\theta\dot{\varphi}^2 - \dot{s}^2 = c_2\dot{s} \quad \Rightarrow \quad \sin^2\theta\dot{\varphi}^2 = c^2 + c_2 c \quad \Rightarrow \quad \dot{\varphi} = \pm\frac{c_\varphi}{\sin\theta}$$

(2.13)
$$\dot{\varphi} = \pm\frac{c_\varphi}{\sin\theta} \quad \Rightarrow \quad \dot{\varphi} = \pm\frac{c_\varphi}{\sin(c_\theta t + \theta_0)}$$

To find φ, we must do an integral:

$$\int \frac{1}{\sin x}dx = \tfrac{1}{2}\int \frac{dx}{\sin\frac{x}{2}\cos\frac{x}{2}} = \tfrac{1}{2}\int \frac{\cos^2\frac{x}{2} + \sin^2\frac{x}{2}}{\sin\frac{x}{2}\cos\frac{x}{2}}dx =$$
$$\tfrac{1}{2}\int (\cot\tfrac{x}{2} + \tan\tfrac{x}{2})dx = \ln(|\sin\tfrac{x}{2}|) - \ln(|\cos\tfrac{x}{2}|) = \ln(|\tan\tfrac{x}{2}|) + x_0$$

Substituting $x = c_\theta t + \theta_0$ in the integral, we get:

$$\varphi = \pm\int \frac{c_\varphi}{\sin(c_\theta t + \theta_0)}dt = \pm\frac{c_\varphi}{c_\theta}\int \frac{1}{\sin x}dx = \pm\frac{c_\varphi}{c_\theta}\ln(|\tan\frac{c_\theta t + \theta_0}{2}|) + \varphi_0$$

The parametric equation for the great circle then becomes:

(2.14) $\theta = \pm c_\theta t + \theta_0$ and $\varphi = \pm\dfrac{c_\varphi}{c_\theta}\ln(|\tan\dfrac{c_\theta t + \theta_0}{2}|) + \varphi_0$

We "only" need to determine the constants in the equation.

Without reservations we may put $c_\theta = 1$. If the great circle intersects in (1, 0, 0), and if $\theta_0 = \dfrac{\pi}{2}$, we find: $\varphi = \pm c_\varphi \ln(|\tan\dfrac{\pi}{4}|) + \varphi_0 \quad \Rightarrow \quad \varphi = \varphi_0$, but φ_0 can moreover be chosen arbitrarily. If we chose $\varphi_0 = 0$, then the constant c_φ may be determined by the equation:

$$\dot{\varphi} = \pm\tfrac{1}{2}\frac{c_\varphi}{\sin(c_\theta t + \theta_0)},$$

For $t = 0$, and $\theta_0 = \frac{\pi}{2}$, we find $\dot{\varphi}_0 = c_\varphi$.
Since both angles run through great circle arcs, but θ runs through the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$, while φ runs through the interval $[0, 2\pi]$, we shall put $\dot{\varphi}_0 = 2\dot{\theta}_0 = 2c_\theta = c_\varphi \quad \Rightarrow \quad c_\varphi = 2$.
This results in the parametric equation:

(2.15) $$\theta = \pm t + \theta_0 \qquad \varphi = \pm 2\ln(|\tan\frac{t+\theta_0}{2}|)$$

A point on the great circle, then has the parametric equation:

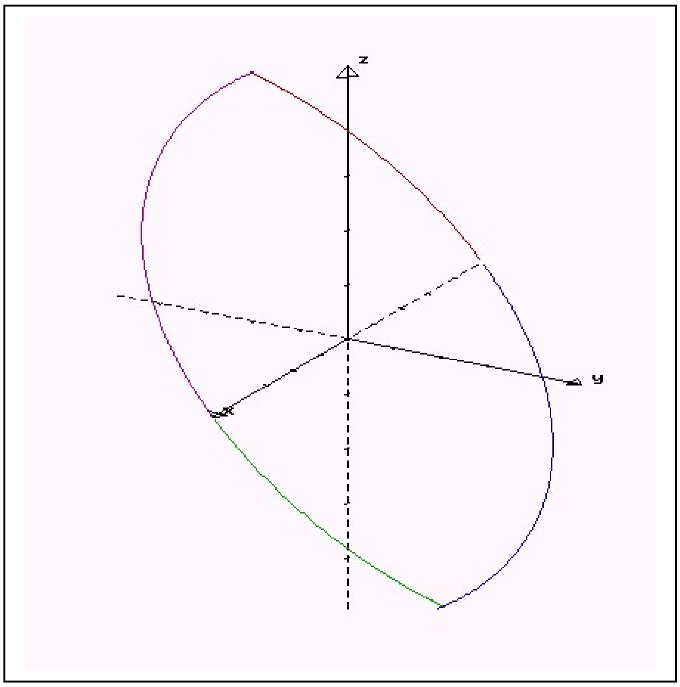
(2.16)
$$(x,y,z) = r(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$$
$$(x,y,z) = r(\sin(t+\theta_0)\cos(\ln(2\,|\tan\frac{t+\theta_0}{2}|)), \sin(t+\theta_0)\sin(\pm\ln(|\,2\tan\frac{t+\theta_0}{2}|)), \cos(t+\theta_0))$$

If we suppose that we have a great circle, which has an angle $\frac{\pi}{3}$ with the z-axis, then θ will run through the interval $[-\frac{\pi}{3},\frac{\pi}{3}]$, while φ will run through the interval $[0,2\pi]$.
It may be practical to invent a new parameter $t+\theta_0 \rightarrow t$, and the parametric equation becomes:

$$(x,y,z) = r(\sin(\pm t)\cos(\ln(|\,2\tan\frac{t}{2}|)), \sin(\pm t)\sin(\ln(|\,2\tan\frac{t}{2}|)), \cos(t))$$

When $\theta = t + \theta_0 = t + \frac{\pi}{2} = t_1$, $[-\frac{\pi}{3},\frac{\pi}{3}]$, then the new parameter will run through the same interval.

We shall the make a 3D graphical representation of the resulting great circle, where we divide the curve in 4 intervals.
As the graph below shows, we get a great circle. That two of the connection points do not connect properly, is a numerical technical thing.



3. Geometrical derivation of the equation of a great circle

All great circles intersect the equator in two points. If we choose the coordinate system, so that the one intersection point is $(r,0,0)$, (or $(1,0,0)$ if we choose $r = 1$), then the plane of the great circle that will pass through $\vec{r}_0 = (1,0,0)$ og $\vec{r} = (\cos\varphi_0 \sin\theta_0, \sin\varphi_0 \sin\theta_0, \cos\theta_0)$ will have a normal vector: $\vec{n} = \vec{r}_0 \times \vec{r}$.

(3.1) $$\vec{n} = \vec{r}_0 \times \vec{r} = \left(\begin{vmatrix} 0 & 0 \\ \sin\varphi_0 \sin\theta_0 & \cos\theta_0 \end{vmatrix}, \begin{vmatrix} 0 & 1 \\ \cos\theta_0 & \cos\varphi_0 \sin\theta_0 \end{vmatrix}, \begin{vmatrix} 1 & 0 \\ \cos\varphi_0 \sin\theta_0 & \sin\varphi_0 \sin\theta_0 \end{vmatrix} \right)$$

(3.2) $$\vec{n} = (0, -\cos\theta_0, \sin\varphi_0 \sin\theta_0)$$

We notice that $\vec{n} \perp \vec{r}_0 \quad \wedge \quad \vec{n} \perp \vec{r}$, as it should be.

If $\vec{n} = (a,b,c)$ and (x_0, y_0, z_0) is a fixed point in the plane, then an equation for the plane becomes:

(3.3) $$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

All great circle planes intersects (0,0,0), and the equation for the plane through (0,0,0) having the normal vector: $\vec{n} = (0, -\cos\theta_0, \sin\varphi_0 \sin\theta_0)$ and $\vec{r} = (x, y, z) = (\cos\varphi \sin\theta, \sin\varphi \sin\theta, \cos\theta)$, therefore have the equation:

(3.4) $$-\cos\theta_0 \sin\varphi \sin\theta + \sin\varphi_0 \sin\theta_0 \cos\theta = 0 \quad \Rightarrow$$

$$\sin\varphi = \frac{\sin\varphi_0 \sin\theta_0}{\cos\theta_0 \tan\theta} = \frac{\sin\theta_0}{\cos\theta_0 \tan\theta} \sin\varphi_0 = \frac{\tan\theta_0}{\tan\theta} \sin\varphi_0$$

(3.5) $$\sin\varphi = \frac{\tan\theta_0}{\tan\theta} \sin\varphi_0$$

Which may be considered as a parametric equation for a great circle.

If we put $c = \tan\theta_0 \sin\varphi_0$ and replace θ by *t,* the equation reads:

(3.5) $$\sin\varphi = \frac{c}{\tan t} \quad \text{and} \quad \cos\varphi = \pm\sqrt{1 - (\frac{c}{\tan t})^2}$$

If we chose $\theta_0 = \frac{\pi}{4}$ and $\varphi_0 = \frac{\pi}{6}$ (There are no limitations on the choice of φ_0), then the equations become simple:

(3.7) $$\sin\varphi = \frac{\tan\theta_0}{\tan\theta} \sin\varphi_0 = \frac{1}{2\tan\theta} \qquad \cos\varphi = \pm\sqrt{1 - \frac{1}{4\tan^2\theta}}$$

With this choice of initial values the we have the following constraints on the parameters:

$$|\frac{1}{2\tan\theta}|<1 \quad\Leftrightarrow\quad |\tan\theta|>\tfrac{1}{2} \quad\Leftrightarrow\quad \tan^2\theta>\tfrac{1}{4} \quad\Leftrightarrow\quad \tan\theta>\tfrac{1}{2} \;\vee\; \tan\theta<-\tfrac{1}{2}$$

Which have the solutions:

(3.6) $$0.4636+p\pi<\theta<\tfrac{\pi}{2}+p\pi \quad\vee\quad -\tfrac{\pi}{2}+p\pi<\theta<-0.4636+p\pi$$

Resulting in the 4 intervals:

$$0.4636<\theta<\tfrac{\pi}{2} \quad\vee\quad -\tfrac{\pi}{2}<\theta<-0.4636$$

$$0.4636+\pi<\theta<\tfrac{\pi}{2}+\pi \quad\vee\quad -\tfrac{\pi}{2}+\pi<\theta<-0.4636+\pi$$

Concerning the choice of sign for $\cos\varphi=\pm\sqrt{1-\frac{1}{4\tan^2\theta}}$, then it is a bit tricky, since it depends on the sign of φ, but we only know $\sin\varphi$, and not φ itself, and since $\sin\varphi=a$, as it is well known has two solutions in the interval $[0,2\pi]$, where one has a positive cosine, and the other has a negative cosine, one simply has to try, to get the right solution, Which is what I have done.
We shall then apply the parametric equation to a 3D graphic representation, where we have divided the interval in 4 parts
Below is the graphical representation of the solution. The normal vector to the circle is also marked.