Extended Essay

May 2022

Shortest Distances on the Sphere

Research Question: What is the shortest distance between two points on the surface of a sphere?

By Kulkarni Venugopal Vasant

Mathematics

Word Count: 3916

Contents

1 Introduction **1**

2 Geodesics on a Two-Dimensional Plane **2**

2.1 Arc Length 3

2.2 Functionals 6

2.3 Finding Extremals of Functionals 6

2.4 Applying the Euler-Lagrange Equation 11

3 Geodesic on a Sphere **14**

3.1 Spherical Coordinates 15

3.2 Arc Length on a Sphere 17

3.3 Applying the Euler-Lagrange Equation 20

3.4 Visualising Solutions 24

3.5 Great Circles 27

4 Spherical Geometry **29**

4.1 Deriving the Shortest Path Equation 29

4.2 Evaluating the Shortest Path Distance 34

5 Conclusion **35**

Works Cited **37**

1 Introduction

As an Indian living in Singapore, I frequently travelled between Singapore's Changi Airport and Mumbai's Chhatrapati Shivaji Maharaj International Airport. Figure 1 shows in red the route of an Air India flight travelling between these two airports. Staring at the flight map, one might notice that the path this flight takes is not quite a straight line connecting the two locations. It appears as a curve compared to the straight line shown in black. If airlines are trying to minimise the distance being travelled to save on fuel and time, why do these paths appear curved?

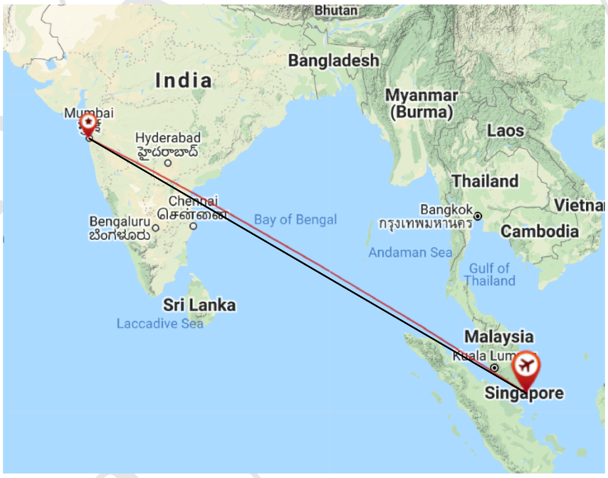


Figure 1. Flight Path from Air India: Interactive Route Map.

In actuality, these curved lines are the shortest paths between two locations, as seen in figure 2. It is just that the Earth is an oblate spheroid, a slightly flattened sphere, and the map is a flat projection of its curved surface. As a result, the true shortest path between two points on the spheroidal Earth does not translate into the straight line between the same points on a flat map.

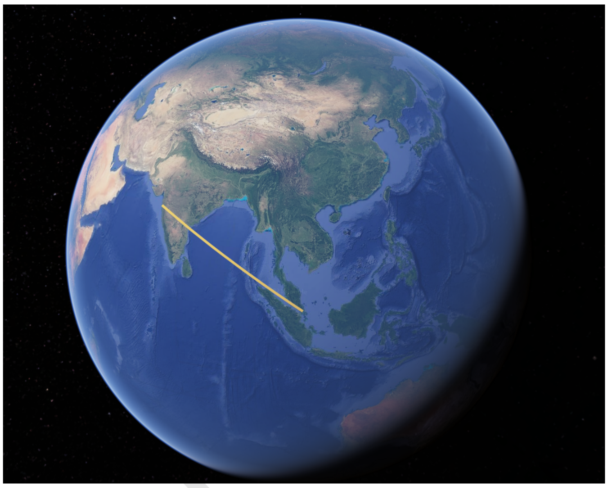


Figure 2. Shortest Path between Singapore and Mumbai .

Nonetheless, while it is well-known that the shortest path between two points on a flat plane is a straight line, it is not so obvious what it is between two points on a curved surface such as a spherical approximation of Earth. This leads us to the research question and main focus of this extended essay: *What is the shortest distance between two points on the surface of a sphere?*

To answer this question, we must first find the geodesic on the surface of a sphere.

Definition 1.1. Geodesic lines or geodesics are the locally shortest paths on a surface (Geodesic Line).

To do this, we must use the Euler-Lagrange Equation, an important tool in the Calculus of Variations. A secondary objective of this essay is to discuss the intuition behind this equation.

Finally, based on our solution to the geodesic problem on the surface of a sphere, we will develop the tools to calculate the shortest distance between two points on a sphere given their coordinates.

2 Geodesics on a Two-Dimensional Plane

Let us first take a look at the geodesic problem on a two-dimensional plane: What is the shortest path between two points, A and B, on a plane?

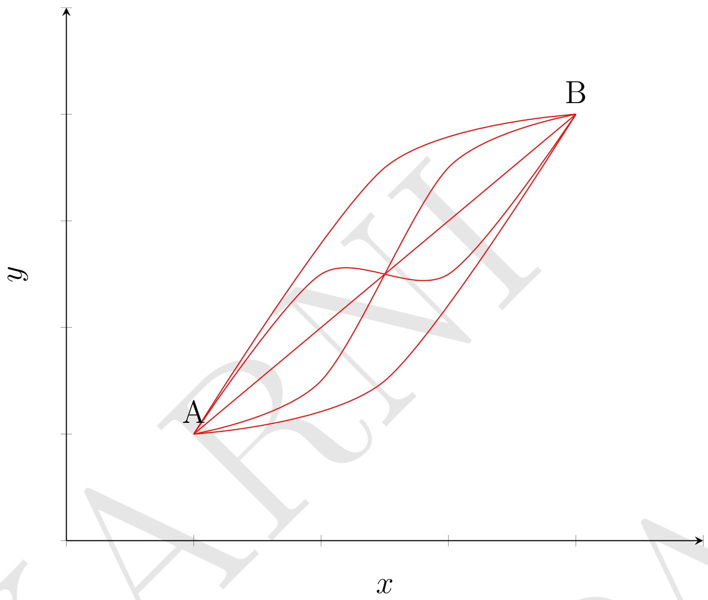


Figure 3. A few of the possible paths between two points A and B on a Cartesian plane.

This is a deceptively simple problem. Even a child may be able to conjecture that the answer is a straight line connecting the two.

Conjecture 2.1. *The shortest path between two points, A and B, on a plane is a straight line of the form $y = mx + c$ passing through them.*

Interestingly, however, the proof is rather mathematically intensive. It incorporates several important concepts that will form the framework that we will use to tackle the harder geodesic problem for the sphere.

2.1 Arc Length

In order to determine which path is the shortest, we must first have an expression for the arc lengths for any path. Let us consider the points A and B with coordinates (A_x, A_y) and (B_x, B_y) on a Cartesian plane respectively. Let path $y(x)$ be a continuous function where $A_x \leq x \leq B_x$ connecting points A and B. This means the $y(A_x) = A_y$ and $y(B_x) = B_y$. It must be also differentiatiable on the interval $[A_x, B_x]$.

We can begin by dividing the interval $[A_x, B_x]$ into n equal smaller intervals of width $\Delta x = (B_x - A_x) \div n$. As seen in figure 5, the resulting line segments approximate the function $y(x)$. For $n = 6$:

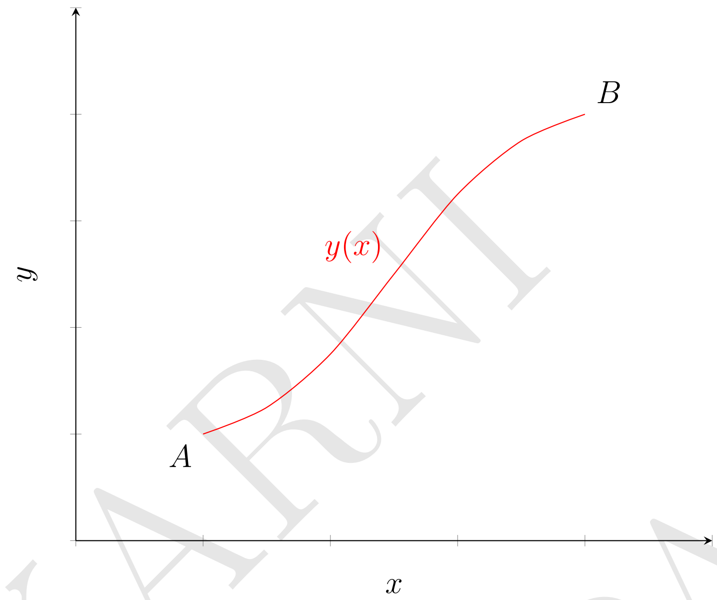


Figure 4. Function $y(x)$ connecting points A and B.

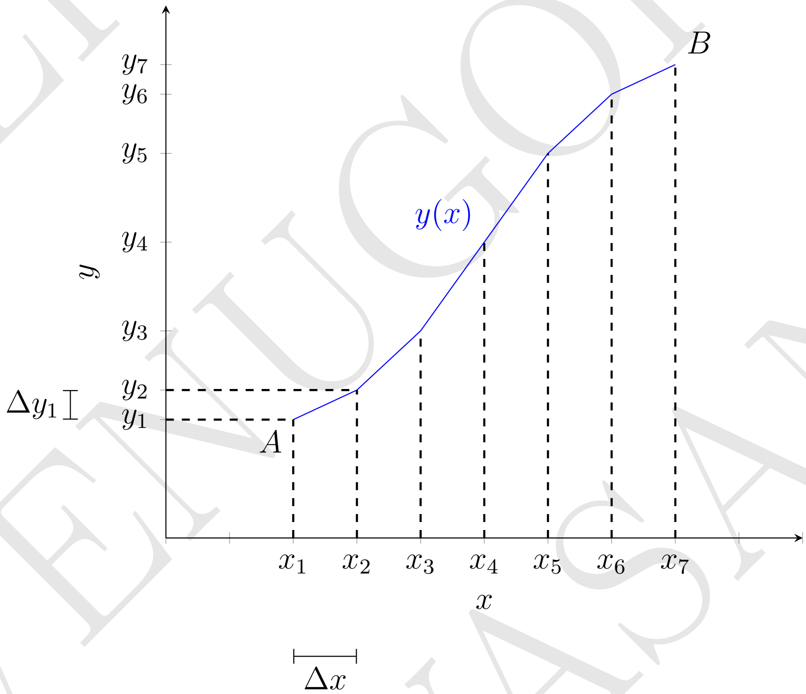


Figure 5. Dividing interval into $n = 6$ subintervals.

By Pythagoras' theorem, the length, s_i, of the line segment in the interval $[x_i, x_{i+1}]$

is given by:

$$s_i = \sqrt{\Delta x^2 + {\Delta y_i}^2} \tag{2.1}$$

$$= \Delta x \sqrt{1 + \left(\frac{\Delta y_i}{\Delta x}\right)^2}. \tag{2.2}$$

and the approximate total arc length, L, is given by the sum of all the line segments:

$$L \approx \sum_{i=1}^{n} s_i. \tag{2.3}$$

Notice that as a larger n is chosen, the width of the subintervals Δx becomes narrower and the approximation by the line segments become closer to the original curve. Hence, in the limit as n approaches infinity, we obtain the exact arc length of the function:

$$L = \lim_{n\to\infty} \sum_{i=1}^{n} s_i \tag{2.4}$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} \Delta x \sqrt{1 + \left(\frac{\Delta y_i}{\Delta x}\right)^2}. \tag{2.5}$$

In the limit, $\frac{\Delta y_i}{\Delta x}$ becomes the derivative of $y(x)$ at x_i, $y'(x_i) = \left.\frac{dy}{dx}\right|_{x_i}$. Furthermore, by the definition of a Riemann Integral, the sum becomes the following integral:

$$L = \int_{A_x}^{B_x} \sqrt{1 + (y'(x))^2}\, dx. \tag{2.6}$$

Thus, we obtain a formula for the arc length of the function $y(x)$ in the interval $[A_x, B_x]$.

Note that $ds = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx$ is known as the *length differential*. Generally, finding the arc length in other situations involves finding the expression for this length differential ds and then integrating it between the appropriate bounds. In this case, it was:

$$L = \int_{A_x}^{B_x} ds, \tag{2.7}$$

2.2 Functionals

L is an example of a functional.

Definition 2.2. A *functional* is a correspondence that maps functions to real numbers (Gelfand and Fomin 1).

They can be thought of as a functions of functions. This relationship can be denoted as

$$L[y] = \int_{A_x}^{B_x} \sqrt{1 + (y'(x))^2} dx. \tag{2.8}$$

The square brackets show that in this case L is a functional of function y.

2.3 Finding Extremals of Functionals

Since we are interested in finding the shortest path connecting points A and B, we are trying to find the function $y(x)$ which minimises the value of the arc length functional $L[y]$. Analogous to how functions have stationary points, functionals have stationary functions known as extremals. These extremals can be minima or maxima.

Definition 2.3 (Extremals)**.** A function $f(x)$ is an *extremal* of functional $J[f]$ if arbitrarily small perturbations to $f(x)$ do not result in a change in the value of $J[f]$.

While stationary points of a function are obtained by equating its derivative to zero, extremals of functionals can be determined with the aforementioned Euler-Lagrange Equation.

Theorem 2.4 (Euler-Lagrange Equation)**.** *Let $J[y]$ be a functional of the form*

$$\int_{A_x}^{B_x} F(x, y, y')dx, \tag{2.9}$$

defined on the set of functions $y(x)$ which have continuous first derivatives on the interval $[x_1, x_2]$ and satisfies the boundary conditions $y(A_x) = A_y$ and $y(B_x) = B_y$. Then every $y(x)$ that extremizes $J[y]$ satisfies the second-order ordinary differential equation (Gelfand

and Fomin 15)

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) - \frac{\partial F}{\partial y} = 0. \tag{2.10}$$

Proof. Let us first consider a discretised problem where the closed interval $[A_x, B_x]$ is divided in n equal subintervals of width $\Delta x = \frac{B_x - A_x}{n}$. We will denote the endpoints of each interval by x_i. Any continuous function $y(x)$ can now be approximated by line segments connecting coordinates $(x_1, y_1), (x_2, y_2), (x_3, y_3), ..., (x_{n+1}, y_{n+1})$, where $y_i = y(x_i)$.

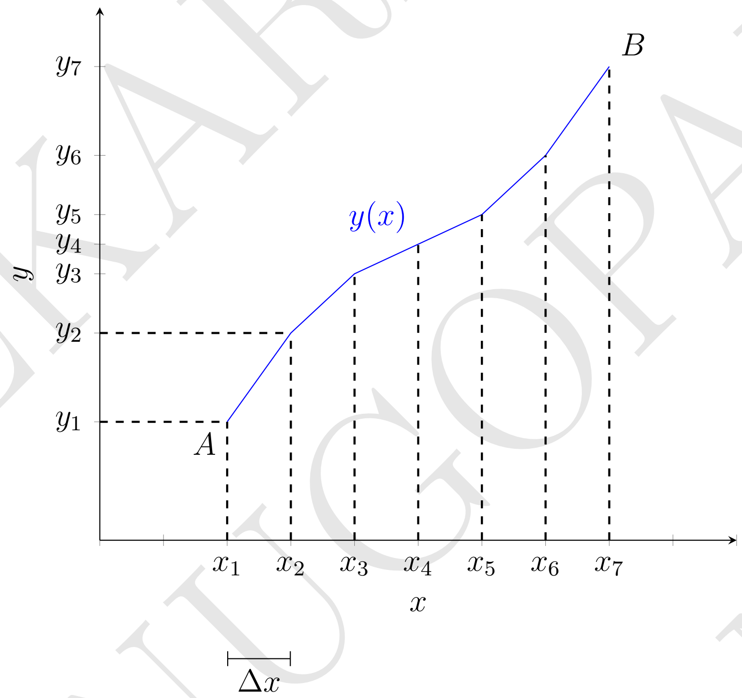


Figure 6. Discretising an arbitrary function $y(x)$ with $n = 6$ equal subintervals.

Now the functional $J[y]$ can be approximated by the sum

$$J[y] \approx \sum_{i=1}^{n+1} F\left(x_i, y_i, \frac{y_{i+1} - y_i}{\Delta x}\right). \tag{2.11}$$

Following definition 2.2, we will analyse the effects of an arbitrarily small perturbation to $y(x)$ on $J[y]$. In our discretised function, let us first consider the effect of a perturbation δy at a particular chosen y_k on the value of $J[y]$. This perturbation is depicted in the figure below as the change from the blue graph to the green graph.

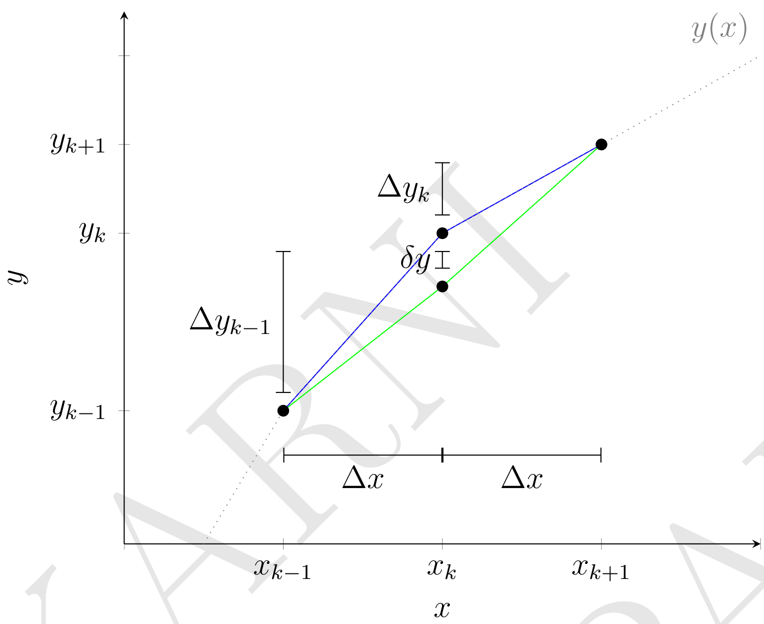


Figure 7. A perturbation δy at y_k.

Since function F is dependent on x, y and y', we must analyse the effects of this perturbation on them and how they contribute to a change in the value of $J[y]$.

Firstly, we note that the perturbation results in a change in the value of y_k. The resulting change in the value of $J[y]$ can be expressed as

$$\frac{\partial J}{\partial y_k}\delta y. \tag{2.12}$$

Intuitively, the partial derivative of J with respect to y_k gives us the change in the value of J for an infinitesimal change in the value of y_k. Hence, multiplying it with the small perturbation δy gives us the resultant change in J.

Secondly, we note that the perturbation results in a change in the derivatives y' around x_k. This can be seen in figure 7 as the change in the gradients between the blue graph and the green graph. In the limit $n \to \infty$, Δx vanishes and the gradient between x_i and x_{i+1} becomes the derivative $y'(x_i)$. Hence, the initial derivative to the left of x_k is

$$y'(x_{k-1}) = \lim_{\Delta x \to 0} \frac{\Delta y_{k-1}}{\Delta x}, \tag{2.13}$$

and to the right of x_k is

$$y'(x_k) = \lim_{\Delta x \to 0} \frac{\Delta y_k}{\Delta x}. \tag{2.14}$$

According to the figure, after the perturbation these become

$$y'(x_{k-1}) = \lim_{\Delta x \to 0} \frac{\Delta y_{k-1} - \delta y}{\Delta x}, \tag{2.15}$$

and

$$y'(x_k) = \lim_{\Delta x \to 0} \frac{\Delta y_k + \delta y}{\Delta x}, \tag{2.16}$$

respectively.

The change in the derivatives is then given by the difference between the derivatives at a certain point before and after perturbation:

$$\Delta y'(x_{k-1}) = \lim_{\Delta x \to 0} \frac{\delta y}{\Delta x}, \tag{2.17}$$

and

$$\Delta y'(x_k) = \lim_{\Delta x \to 0} \frac{-\delta y}{\Delta x}, \tag{2.18}$$

The effect of these change in the derivatives y' on $J[y]$ can now be expressed as:

$$\frac{\partial J}{\partial y'}\Big|_{x_{k-1}} \Delta y'(x_{k-1}) + \frac{\partial J}{\partial y'}\Big|_{x_k} \Delta y'(x_k) \tag{2.19}$$

The intuition here is similar to earlier. The partial derivatives of J with respect to y' gives the change in the value of J for an infinitesimal change in the value of y' evaluated at a certain x_i. Hence, multiplying this with the change in y' at a certain x_i caused by the perturbation δy gives us the change in the value of J due to the perturbation. Since y' changes on both side of the perturbed point, we add the terms accounting for the changes on each side.

Substituting equations 2.17 and 2.18 into 2.19:

$$\frac{\partial J}{\partial y'}\Big|_{x_{k-1}} \lim_{\Delta x \to 0} \frac{\delta y}{\Delta x} + \frac{\partial J}{\partial y'}\Big|_{x_k} \lim_{\Delta x \to 0} \frac{-\delta y}{\Delta x}. \tag{2.20}$$

Factorising:

$$-\lim_{\Delta x \to 0} \frac{\delta y}{\Delta x} \left(\frac{\partial J}{\partial y'}\Big|_{x_k} - \frac{\partial J}{\partial y'}\Big|_{x_{k-1}} \right). \tag{2.21}$$

Since $x_k = x_{k-1} + \Delta x$, we can see that this limit beautifully reduces to a derivative:

$$-\delta y \frac{d}{dx} \left(\frac{\partial J}{\partial y'} \right) \Big|_{x_{k-1}}. \tag{2.22}$$

Combining expressions 2.22 and 2.12, we find the total effect of perturbation δy on $J[y]$, δJ, is given by:

$$\delta J = \left(\frac{\partial J}{\partial y_k} - \frac{d}{dx} \left(\frac{\partial J}{\partial y'} \right) \Big|_{x_{k-1}} \right) \delta y. \tag{2.23}$$

From definition 2.2, we recall that a function $f(x)$ is an extremal if arbitrarily small perturbations do not result in a change in the value of $J[f]$. Hence, for a function $y(x)$ to be an extremal, we can expect that $\delta J = 0$ for a small perturbation δy at any point (x_k, y_k) along the function, or $\frac{\delta J}{\delta y} = 0$. Returning to continuous functions, this leaves us with the Euler-Lagrange equation

$$\frac{\partial J}{\partial y} - \frac{d}{dx} \left(\frac{\partial J}{\partial y'} \right) = 0, \tag{2.24}$$

proving theorem 2.3. □

Do note that this proof is today considered an unconventional one. It combines ideas by Nakkiran and Kot (29), though it appears to have been first published in some form by Euler (Hanc). The conventional proof employed in literature was found by Joseph-Louis Lagrange and uses more rigorous techniques in the calculus of variations. However, I feel that this method allows us to obtain a better intuition for each term in the equation. We can now understand that the term $\frac{\partial J}{\partial y}$ accounts for changes in the value of $y(x)$ and the term $\frac{d}{dx}\left(\frac{\partial J}{\partial y'}\right)$ accounts for changes in the derivatives around any x.

2.4 Applying the Euler-Lagrange Equation

Going back to the problem we were considering at the beginning of this section,

$$L[y] = \int_{A_x}^{B_x} \sqrt{1 + (y'(x))^2} dx. \tag{2.25}$$

With $F = \sqrt{1 + (y'(x))^2}$, we can now apply the Euler-Lagrange equation to find the path $y(x)$ which extremises the arc length:

$$\frac{d}{dx}\left[\frac{\partial}{\partial y'}\left(\sqrt{1 + (y'(x))^2}\right)\right] - \frac{\partial}{\partial y}\left(\sqrt{1 + (y'(x))^2}\right) = 0. \tag{2.26}$$

Since F is independent of y:

$$\frac{d}{dx}\left[\frac{\partial}{\partial y'}\left(\sqrt{1 + (y'(x))^2}\right)\right] = 0. \tag{2.27}$$

Differentiating F with respect to y':

$$\frac{d}{dx}\left(\frac{1}{2\sqrt{1 + (y'(x))^2}} \times 2y'(x)\right) = 0 \tag{2.28}$$

$$\frac{d}{dx}\left(\frac{y'(x)}{\sqrt{1 + (y'(x))^2}}\right) = 0. \tag{2.29}$$

Integrating both sides:

$$\int \frac{d}{dx}\left(\frac{y'(x)}{\sqrt{1 + (y'(x))^2}}\right) dx = \int 0 dx \tag{2.30}$$

$$\frac{y'(x)}{\sqrt{1 + (y'(x))^2}} = C_1, \tag{2.31}$$

where C_1 is an arbitrary constant.

Squaring both sides:

$$\frac{(y'(x))^2}{1+(y'(x))^2} = C_1^2. \tag{2.32}$$

Making $y'(x)$ the subject of the equation:

$$(y'(x))^2 = C_1^2(1+(y'(x))^2) \tag{2.33}$$

$$(y'(x))^2 = C_1^2 + C_1^2(y'(x))^2 \tag{2.34}$$

$$(y'(x))^2 - C_1^2(y'(x))^2 = C_1^2 \tag{2.35}$$

$$(y'(x))^2(1-C_1^2) = C_1^2 \tag{2.36}$$

$$(y'(x))^2 = \frac{C_1^2}{1-C_1^2} \tag{2.37}$$

$$y'(x) = \frac{C_1}{\sqrt{1-C_1^2}}. \tag{2.38}$$

Recalling that $y'(x)$ is the derivative of $y(x)$, we can see that equation 2.38 is the differential equation

$$\frac{dy}{dx} = \frac{C_1}{\sqrt{1-C_1^2}}. \tag{2.39}$$

This can be solved by the separation of variables:

$$\int dy = \int \frac{C_1}{\sqrt{1-C_1^2}}dx. \tag{2.40}$$

Since the term $\frac{C_1}{\sqrt{1-C_1^2}}$ is a constant,

$$y = \frac{C_1}{\sqrt{1-C_1^2}}x + C_2, \tag{2.41}$$

where C_2 is an arbitrary constant.

We can already notice that the solution is in the form $y = mx + b$, a straight line, thus proving our conjecture. For completeness, let us find expressions for the constants. These can be obtained by substituting the boundary conditions $y(A_x) = A_y$ and $y(B_x) = B_y$.

$$A_y = \frac{C_1}{\sqrt{1 - C_1^2}} A_x + C_2, \tag{2.42}$$

and

$$B_y = \frac{C_1}{\sqrt{1 - C_1^2}} B_x + C_2. \tag{2.43}$$

Subtracting equation 2.42 from 2.43:

$$B_y - A_y = \frac{C_1}{\sqrt{1 - C_1^2}} (B_x - A_x). \tag{2.44}$$

Thus:

$$\frac{C_1}{\sqrt{1 - C_1^2}} = \frac{B_y - A_y}{B_x - A_x}. \tag{2.45}$$

Notice that the right hand side is the formula for the gradient.

Substituting equation 2.45 back into 2.42 and solving for C_2 yields:

$$C_2 = A_y - \frac{B_y - A_y}{B_x - A_x} A_x. \tag{2.46}$$

Substituting equations 2.46 and 2.45 into 2.42, we obtain:

$$y = \frac{B_y - A_y}{B_x - A_x} x + A_y - \frac{B_y - A_y}{B_x - A_x} A_x. \tag{2.47}$$

This simplifies to the final solution

$$y = A_y + \frac{B_y - A_y}{B_x - A_x}(x - A_x). \tag{2.48}$$

As such, we have found the geodesic equation for a two-dimensional plane. In doing so, we have encountered arc length, functionals, and the Euler-Lagrange Equation. Together, these three form the framework by which we will approach the geodesic problem on the sphere. Note that we have not rigorously proved the solution is a minima. However, intuitively, the maxima would be a line of infinite length. Therefore, it is reasonable to assume that this solution is the minima.

It is interesting to think about the shortest path between two points on a two-dimensional plane being the *definition* of a straight line. Can the concept of a line or its straightness even be defined? And if it can, and this is the definition, then what does it mean to *prove* that the shortest path between two points on a plane is a straight line?

Euclid's first axiom is that "A straight line segment can be drawn joining any two points" (Weisstein, Euclid's Postulates). Here, the definitions of a "straight", "line segment", and even "point" are assumed. These concepts are undefined, yet they are the building blocks of Euclidean geometry. For this essay, let us define straight lines to mean geodesics of a certain surface - in this case, straight lines are the geodesics of the two-dimensional plane. As we will discover later in this essay, this generalisation will lead to useful analogies in a different type of geometry - spherical geometry.

3 Geodesic on a Sphere

Now, let us consider the geodesic problem on a sphere of radius R. Mathematically, such a sphere is described by the following equation:

$$x^2 + y^2 + z^2 = R^2. \tag{3.1}$$

Note that a sphere is a three-dimensional surface, and hence each point is described by three coordinates instead of the two in the previous section. Let us define two points A and B with Cartesian coordinates (A_x, A_y, A_z) and (B_x, B_y, B_z) respectively that lie on this sphere.

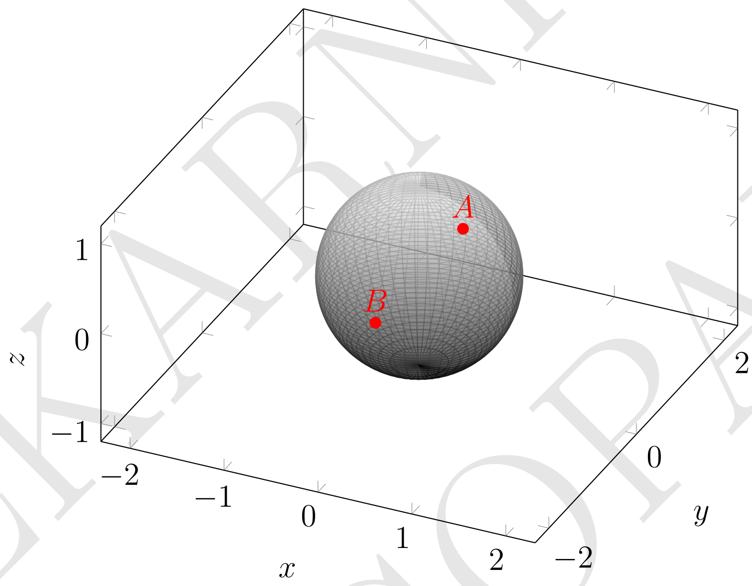


Figure 8. Geodesic Problem on a Sphere.

3.1 Spherical Coordinates

Using spherical coordinates are necessary to obtain the geodesic on a sphere. In this system, points are described by the azimuthal angle θ, the polar angle ϕ and the distance from the origin r (Spherical Coordinates). The azimuthal angle is defined as the angle in the x-y plane anti-clockwise from the positive x-axis with $0 \leq \theta < 2\pi$. The polar angle is defined to be the angle from the positive z-axis with $0 \leq \phi < \pi$. This is depicted in the figure below.

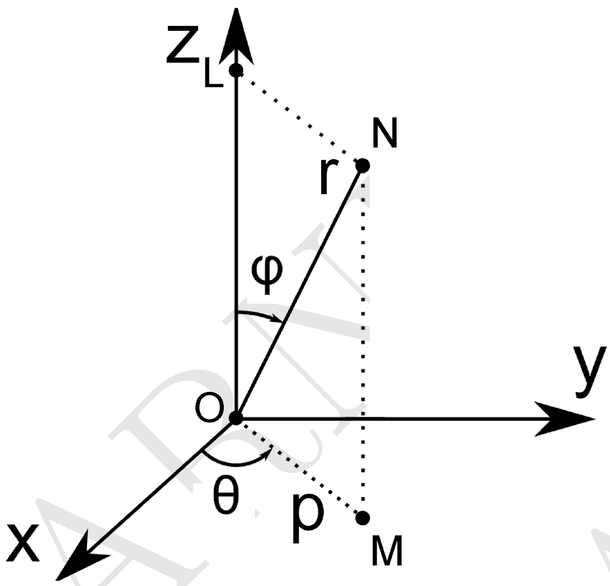


Figure 9. Spherical Coordinate System. Image modified from Miettinen.

From the diagram we can obtain equations to transform between cartesian coordinates and spherical coordinates.

By the definition of cosine,

$$\frac{z}{r} = \cos\phi \tag{3.2}$$

$$z = r\cos\phi. \tag{3.3}$$

Letting p be the length of r projected on the x-y plane, by trigonometry,

$$\frac{x}{p} = \cos\theta \tag{3.4}$$

$$x = p\cos\theta, \tag{3.5}$$

and,

$$\frac{y}{p} = \sin\theta \tag{3.6}$$

$$y = p\sin\theta. \tag{3.7}$$

Since $\phi = \angle LON = \angle ONM$, by the alternate angle theorem,

$$\frac{p}{r} = \sin\phi \tag{3.8}$$

$$p = r\sin\phi, \tag{3.9}$$

Substituting p into equations 3.5 and 3.7, we obtain our three equations,

$$x = r\cos\theta\sin\phi. \tag{3.10}$$

$$y = r\sin\theta\sin\phi. \tag{3.11}$$

$$z = r\cos\phi. \tag{3.12}$$

3.2 Arc Length on a Sphere

Let us construct an expression for the arc length on a sphere. We can again start by considering an arbitrary continuous path function s that connects points A and B along the surface of the sphere. By Pythagoras' theorem, the length differential ds between two neighbouring points on the path would be

$$ds = \sqrt{dx^2 + dy^2 + dz^2}. \tag{3.13}$$

To obtain the arc length S, we integrate this element between points A and B:

$$S = \int_A^B ds \tag{3.14}$$

$$= \int_A^B \sqrt{dx^2 + dy^2 + dz^2} \tag{3.15}$$

One might notice that this integral can give the arc length of any path in three-dimensional space, not only ones on the surface of a sphere. A key observation is that every point on the sphere's surface is exactly distance R from the origin. Consequently, every point in any path s on the sphere's surface must also be exactly R from the origin.

Recall that the distance from the origin is a component of the spherical coordinate system. We can now see that expressing these paths in spherical coordinates reduces three variables to two as the radius component is constant, $r = R$.

One way to express ds in terms of spherical coordinates is to differentiate equations 3.10, 3.11 and 3.12 via the multivariable chain rule.

Differentiating x with respect to θ and ϕ:

$$dx = \frac{\partial x}{\partial \theta} d\theta + \frac{\partial x}{\partial \phi} d\phi \tag{3.16}$$

$$dx = -R \sin\theta \sin\phi d\theta + R \cos\theta \cos\phi d\phi \tag{3.17}$$

Differentiating y with respect to θ and ϕ:

$$dy = \frac{\partial y}{\partial \theta} d\theta + \frac{\partial y}{\partial \phi} d\phi \tag{3.18}$$

$$dy = R \cos\theta \sin\phi d\theta + R \sin\theta \cos\phi d\phi \tag{3.19}$$

Differentiating z with respect to θ and ϕ:

$$dz = \frac{\partial z}{\partial \theta} d\theta + \frac{\partial z}{\partial \phi} d\phi \tag{3.20}$$

$$dz = -R \sin\phi d\phi \tag{3.21}$$

Squaring equations 3.17, 3.19 and 3.21:

$$\begin{gathered} dx^2 = R^2 \sin^2\theta \sin^2\phi d\theta^2 + R^2 \cos^2\theta \cos^2\phi d\phi^2 - \\ 2R^2 \sin\theta \sin\phi \cos\theta \cos\phi d\theta d\phi \end{gathered} \tag{3.22}$$

$$\begin{gathered} dy^2 = R^2 \cos^2\theta \sin^2\phi d\theta^2 + R^2 \sin^2\theta \cos^2\phi d\phi^2 + \\ 2R^2 \sin\theta \sin\phi \cos\theta \cos\phi d\theta d\phi \end{gathered} \tag{3.23}$$

$$dz^2 = R^2 \sin^2 \phi d\phi \tag{3.24}$$

Thus, by substituting equations 3.22, 3.23 and 3.24 into the Pythagoras' Theorem:

$$\begin{aligned} ds^2 = R^2 \sin^2 \theta \sin^2 \phi d\theta^2 + R^2 \cos^2 \theta \cos^2 \phi d\phi^2 - \\ 2R^2 \sin \theta \sin \phi \cos \theta \cos \phi d\theta d\phi + R^2 \cos^2 \theta \sin^2 \phi d\theta^2 + \\ R^2 \sin^2 \theta \cos^2 \phi d\phi^2 + 2R^2 \sin \theta \sin \phi \cos \theta \cos \phi d\theta d\phi + R^2 \sin^2 \phi d\phi \end{aligned} \tag{3.25}$$

Simplifying and factorising out R^2:

$$\begin{aligned} ds^2 = R^2(\sin^2 \theta \sin^2 \phi d\theta^2 + \cos^2 \theta \cos^2 \phi d\phi^2 + \\ \cos^2 \theta \sin^2 \phi d\theta^2 + \sin^2 \theta \cos^2 \phi d\phi^2 + \sin^2 \phi d\phi) \end{aligned} \tag{3.26}$$

Grouping terms containing $d\theta^2 \sin^2 \phi$ and $d\phi^2 \cos^2 \phi$:

$$\begin{aligned} ds^2 = R^2[d\theta^2 \sin^2 \phi(\sin^2 \theta + \cos^2 \theta) + \\ d\phi^2 \cos^2 \phi(\cos^2 \theta + \sin^2 \theta) + \sin^2 \phi d\phi^2] \end{aligned} \tag{3.27}$$

By the Pythagorean trigonometric identity $\sin^2 \alpha + \cos^2 \alpha = 1$, this simplifies to:

$$ds^2 = R^2[d\theta^2 \sin^2 \phi + d\phi^2 \cos^2 \phi + \sin^2 \phi d\phi^2] \tag{3.28}$$

Grouping terms containing $d\phi^2$ again:

$$ds^2 = R^2[d\theta^2 \sin^2 \phi + d\phi^2(\cos^2 \phi + \sin^2 \phi)] \tag{3.29}$$

This further simplifies to:

$$ds^2 = R^2[d\theta^2 \sin^2 \phi + d\phi^2] \tag{3.30}$$

Factorising $d\phi^2$ out of the expression:

$$ds^2 = R^2 d\phi^2[\sin^2\phi\left(\frac{d\theta}{d\phi}\right)^2 + 1] \tag{3.31}$$

Square rooting both sides:

$$ds = Rd\phi\sqrt{\sin^2\phi\left(\frac{d\theta}{d\phi}\right)^2 + 1} \tag{3.32}$$

Substituting ds into the integral in equation 3.15:

$$S = R\int_{A_\phi}^{B_\phi}\left[\sqrt{\sin^2\phi\left(\frac{d\theta}{d\phi}\right)^2 + 1}\right]d\phi \tag{3.33}$$

Note that the step in equation 3.31 is significant. It expresses the integrand in the form $F(\phi, \theta, \theta')$, where $\theta' = \frac{d\theta}{d\phi}$. We can now see that arc length S is a functional of θ. This also enables the use of the Euler-Lagrange Equation to find the extremals of $S[\theta]$. Also note that the bounds are now the polar angles of A and B.

3.3 Applying the Euler-Lagrange Equation

We now have to find the extremals of the functional

$$S[\theta] = R\int_{A_\phi}^{B_\phi}\left[\sqrt{(\theta'(\phi))^2\sin^2\phi + 1}\right]d\phi \tag{3.34}$$

Applying the Euler-Lagrange equation for $F = \sqrt{(\theta'(\phi))^2\sin^2\phi + 1}$:

$$\frac{d}{d\phi}\left[\frac{\partial}{\partial\theta'}\left(\sqrt{(\theta'(\phi))^2\sin^2\phi + 1}\right)\right] - \frac{\partial}{\partial\theta}\left(\sqrt{(\theta'(\phi))^2\sin^2\phi + 1}\right) = 0. \tag{3.35}$$

Since F is independent of θ, we have

$$\frac{d}{d\phi}\left[\frac{\partial}{\partial\theta'}\left(\sqrt{(\theta'(\phi))^2\sin^2\phi + 1}\right)\right] = 0. \tag{3.36}$$

Differentiating F with respect to θ':

$$\frac{d}{d\phi}\left(\frac{1}{2\sqrt{(\theta'(\phi))^2\sin^2\phi+1}}\times 2\theta'(\phi)\sin^2\phi\right)=0 \tag{3.37}$$

$$\frac{d}{d\phi}\left(\frac{\theta'(\phi)\sin^2\phi}{\sqrt{(\theta'(\phi))^2\sin^2\phi+1}}\right)=0. \tag{3.38}$$

Integrating both sides:

$$\int\frac{d}{d\phi}\left(\frac{\theta'(\phi)\sin^2\phi}{\sqrt{(\theta'(\phi))^2\sin^2\phi+1}}\right)d\phi=\int 0d\phi \tag{3.39}$$

$$\frac{\theta'(\phi)\sin^2\phi}{\sqrt{(\theta'(\phi))^2\sin^2\phi+1}}=C_1, \tag{3.40}$$

where C_1 is an arbitrary constant.

Squaring both sides:

$$\frac{(\theta'(\phi))^2\sin^4\phi}{(\theta'(\phi))^2\sin^2\phi+1}=C_1^2. \tag{3.41}$$

Making $\theta'(\phi)$ the subject:

$$(\theta'(\phi))^2\sin^4\phi=C_1^2((\theta'(\phi))^2\sin^2\phi+1) \tag{3.42}$$

$$(\theta'(\phi))^2\sin^4\phi=C_1^2(\theta'(\phi))^2\sin^2\phi+C_1^2) \tag{3.43}$$

$$(\theta'(\phi))^2\sin^4\phi-C_1^2(\theta'(\phi))^2\sin^2\phi=C_1^2 \tag{3.44}$$

$$(\theta'(\phi))^2[\sin^4\phi-C_1^2\sin^2\phi]=C_1^2 \tag{3.45}$$

$$(\theta'(\phi))^2=\frac{C_1^2}{\sin^4\phi-C_1^2\sin^2\phi}. \tag{3.46}$$

Square rooting both sides:

$$\sqrt{(\theta'(\phi))^2}=\sqrt{\frac{C_1^2}{\sin^4\phi-C_1^2\sin^2\phi}} \tag{3.47}$$

$$\theta'(\phi)=\frac{C_1}{\sqrt{\sin^4\phi-C_1^2\sin^2\phi}}. \tag{3.48}$$

We can now solve the differential equation by separation of variables.

$$\frac{d\theta}{d\phi} = \frac{C_1}{\sqrt{\sin^4\phi - C_1^2\sin^2\phi}} \tag{3.49}$$

$$\int d\theta = \int \frac{C_1}{\sqrt{\sin^4\phi - C_1^2\sin^2\phi}} d\phi. \tag{3.50}$$

In order to solve this integral, we must first express in a form more suitable for substitution. Factoring out $\sin^4\phi$:

$$\theta = \int \frac{C_1}{\sqrt{\sin^4\phi\left(1 - C_1^2\frac{1}{\sin^2\phi}\right)}} d\phi \tag{3.51}$$

$$\theta = \int \frac{C_1}{\sin^2\phi\sqrt{1 - C_1^2\frac{1}{\sin^2\phi}}} d\phi \tag{3.52}$$

$$\theta = \int \frac{C_1\csc^2\phi}{\sqrt{1 - C_1^2\csc^2\phi}} d\phi. \tag{3.53}$$

Recalling that $\frac{d}{dx}\cot x = -\csc^2 x$, let us try to express the denominator in terms of cot so can we make this substitution.

Dividing the Pythagorean trigonometric identity $\sin^2\alpha + \cos^2\alpha = 1$ by $\sin^2\alpha$, we have

$$1 + \cot^2\alpha = \csc^2\alpha. \tag{3.54}$$

Substituting into equation 3.53, we have

$$\theta = \int \frac{C_1\csc^2\phi}{\sqrt{1 - C_1^2(1 + \cot^2\phi)}} d\phi \tag{3.55}$$

$$\theta = \int \frac{C_1\csc^2\phi}{\sqrt{1 - C_1^2 - C_1^2\cot^2\phi}} d\phi. \tag{3.56}$$

Now, letting

$$u = \cot\phi, \tag{3.57}$$

we have

$$du = -\csc^2 \phi d\phi \tag{3.58}$$

$$-\frac{du}{\csc^2 \phi} = d\phi \tag{3.59}$$

Substituting equations 3.57 and 3.59 into equation 3.56:

$$\theta = -\int \frac{C_1}{\sqrt{1 - C_1^2 - C_1^2 u^2}} du. \tag{3.60}$$

Here we must notice that the form of the expression inside the integral is similar to the derivative of arcsin, $\frac{d}{dx} \arcsin = \frac{1}{\sqrt{1-x^2}}$. We can make use of this indentity by first factoring out $(1 - C_1^2)$ in the square root.

$$\theta = -\int \frac{C_1}{\sqrt{(1 - C_1^2)\left(1 - \frac{C_1^2}{1-C_1^2} u^2\right)}} du \tag{3.61}$$

$$\theta = -\int \frac{C_1}{\sqrt{(1 - C_1^2)}\sqrt{\left(1 - \frac{C_1^2}{1-C_1^2} u^2\right)}} du. \tag{3.62}$$

Since $\frac{C_1}{\sqrt{1-C_1^2}}$ is constant,

$$\theta = -\frac{C_1}{\sqrt{1 - C_1^2}} \int \frac{1}{\sqrt{1 - \frac{C_1^2}{1-C_1^2} u^2}} du \tag{3.63}$$

$$\theta = -\frac{C_1}{\sqrt{1 - C_1^2}} \int \frac{1}{\sqrt{1 - \left(\sqrt{\frac{C_1^2}{1-C_1^2}} u\right)^2}} du \tag{3.64}$$

Thus, by linear substitution and the derivative of arcsin, the integral in the above equation yields

$$\theta = -\frac{C_1}{\sqrt{1 - C_1^2}} \left(\frac{\arcsin \sqrt{\frac{C_1^2}{1-C_1^2}} u}{\sqrt{\frac{C_1^2}{1-C_1^2}}} + C_2 \right), \tag{3.65}$$

where C_2 is a constant.

Simplifying:

$$\theta = -\frac{C_1}{\sqrt{1-C_1^2}} \frac{\arcsin \sqrt{\frac{C_1^2}{1-C_1^2}} u}{\sqrt{\frac{C_1^2}{1-C_1^2}}} - \frac{C_1}{\sqrt{1-C_1^2}} C_2 \tag{3.66}$$

$$\theta = -\frac{C_1}{\sqrt{1-C_1^2}} \frac{\arcsin \sqrt{\frac{C_1^2}{1-C_1^2}} u}{\frac{C_1}{\sqrt{1-C_1^2}}} - \frac{C_1}{\sqrt{1-C_1^2}} C_2 \tag{3.67}$$

$$\theta = -\arcsin \sqrt{\frac{C_1^2}{1-C_1^2}} u - \frac{C_1}{\sqrt{1-C_1^2}} C_2. \tag{3.68}$$

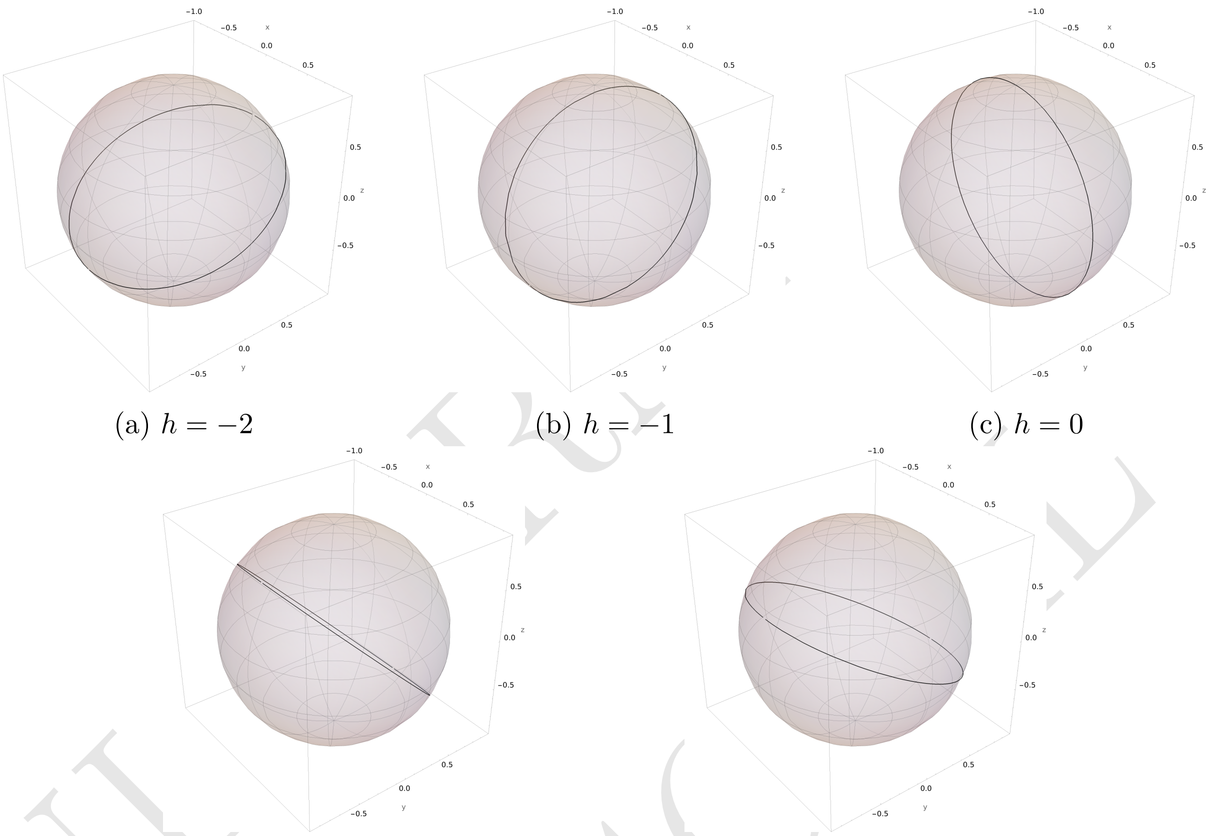
Let $h = \sqrt{\frac{C_1^2}{1-C_1^2}}$ and $k = \frac{C_1}{\sqrt{1-C_1^2}} C_2$. Substituting back $u = \cot \phi$:

$$\theta = -\arcsin(h \cot \phi) - k, \tag{3.69}$$

where h and k are constants depending on the boundary conditions, in this case the points A and B.

3.4 Visualising Solutions

Due to the combination of nested trigonometry and the use of spherical coordinates, the above solution can be difficult to interpret. Plotting solutions for various values of h and k would allow us to better understand the geodesic on the sphere.

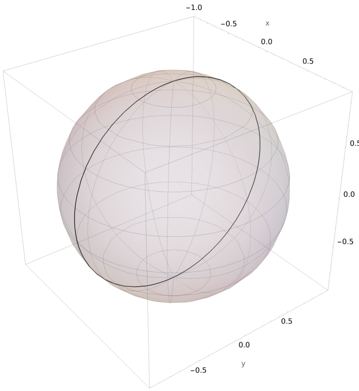


(a) $h = -2$ (b) $h = -1$ (c) $h = 0$

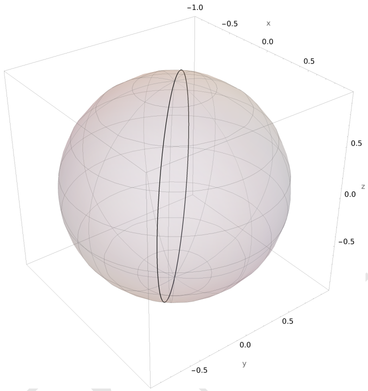
(d) $h = 1$ (e) $h = 2$

Figure 10. Sphere Geodesic Solution Plotted for Different Values of h on a Unit Sphere using Wolfram Cloud.

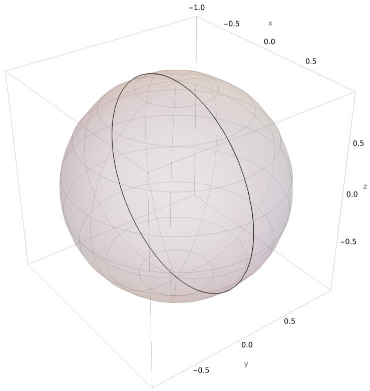
In figure 10, we can see that the constant h tilts the solution around a lateral axis, where negative values results in the solution tilting in the opposite direction.



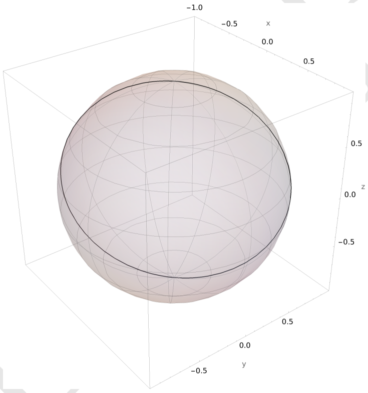
(a) $k = -\frac{\pi}{2}$



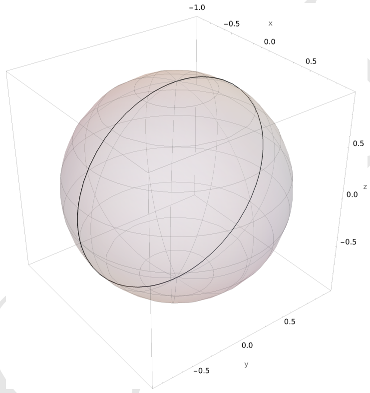
(b) $k = -\frac{\pi}{4}$



(c) $k = 0$



(d) $k = \frac{\pi}{4}$



(e) $k = \frac{\pi}{2}$

Figure 11. Sphere Geodesic Solution Plotted for Different Values of k using Wolfram Cloud.

In figure 11, we can see that the constant k rotates the solution around a vertical axis. Intuitively, this is because k is a constant term in the solution, and the coordinate of θ represents a rotation about the vertical axis.

Here, we can appreciate the difference between the shortest path and the geodesic. The two points split the geodesic into two paths, one shorter and one longer. Both of these paths are, by definition, locally shortest paths - at each point on the geodesic, infinitesimal movements in either direction along the geodesic are the shortest possible. Hence, we must note that geodesics are not necessarily the shortest paths but shortest paths necessarily are geodesics.

3.5 Great Circles

From the plots in the previous section, it appears that the geodesic is a circular path of the same radius as the sphere on which the chosen points A and B lie. The constants h and k, dependent on these chosen points ensure the geodesic is oriented such that it passes through them. Such a circle that contains the diameter of a certain sphere is termed a *great circle*.

Definition 3.1 (Great Circle)**.** A *great circle* is the intersection of a sphere and a plane passing through its centre (Karttunen).

Thus, from the observations above, we conjecture the following.

Conjecture 3.2. *The geodesic of a sphere is the arc of a great circle.*

Proof. Rearranging equation 3.69,

$$-\theta - k = \arcsin(h \cot \phi) \tag{3.70}$$

$$\sin(-\theta - k) = h \cot \phi \tag{3.71}$$

Expanding the left hand side with compound angle identity $\sin(\alpha + \beta) = \sin\alpha\cos\beta + \sin\beta\cos\alpha$,

$$\sin(-\theta)\cos(-k) + \sin(-k)\cos(-\theta) = h \cot \phi \tag{3.72}$$

By negative angle identities $\sin(-\alpha) = -\sin\alpha$ and $\cos(-\alpha) = \cos\alpha$,

$$-\sin\theta\cos k - \sin k \cos\theta = h \cot \phi \tag{3.73}$$

By the definition of the cot function,

$$-\sin\theta\cos k - \sin k \cos\theta = h\frac{\cos\phi}{\sin\phi} \tag{3.74}$$

Rearranging,

$$-\sin\phi(\sin\theta\cos k + \sin k\cos\theta) = h\cos\phi \tag{3.75}$$

$$-\sin\phi(\sin\theta\cos k + \sin k\cos\theta) - h\cos\phi = 0 \tag{3.76}$$

$$\sin\phi(\sin\theta\cos k + \sin k\cos\theta) + h\cos\phi = 0 \tag{3.77}$$

$$\sin\phi\sin\theta\cos k + \sin\phi\cos\theta\sin k + h\cos\phi = 0. \tag{3.78}$$

One might notice the similarities between the above terms and the cartesian to spherical coordinate transformation equations in section 3.1. Rearranging equations 3.10, 3.11 and 3.12,

$$\frac{x}{r} = \cos\theta\sin\phi. \tag{3.79}$$

$$\frac{y}{r} = \sin\theta\sin\phi. \tag{3.80}$$

$$\frac{z}{r} = \cos\phi. \tag{3.81}$$

Substituting equations 3.80, 3.81 and 3.82 into 3.78,

$$\frac{y}{r}\cos k + \frac{x}{r}\sin k + h\frac{z}{r} = 0 \tag{3.82}$$

$$x\sin k + y\cos k + hz = 0. \tag{3.83}$$

Since coefficients $\sin k$, $\cos k$, and h are constants, we are left with an equation in the form of $ax + by + cz = 0$. This is the equation of a plane passing through the origin. Thus, we have proved that the geodesic on a sphere must lie on the intersection of the surface of the sphere and a plane passing through its centre. This is, by definition 3.1, a great circle. □

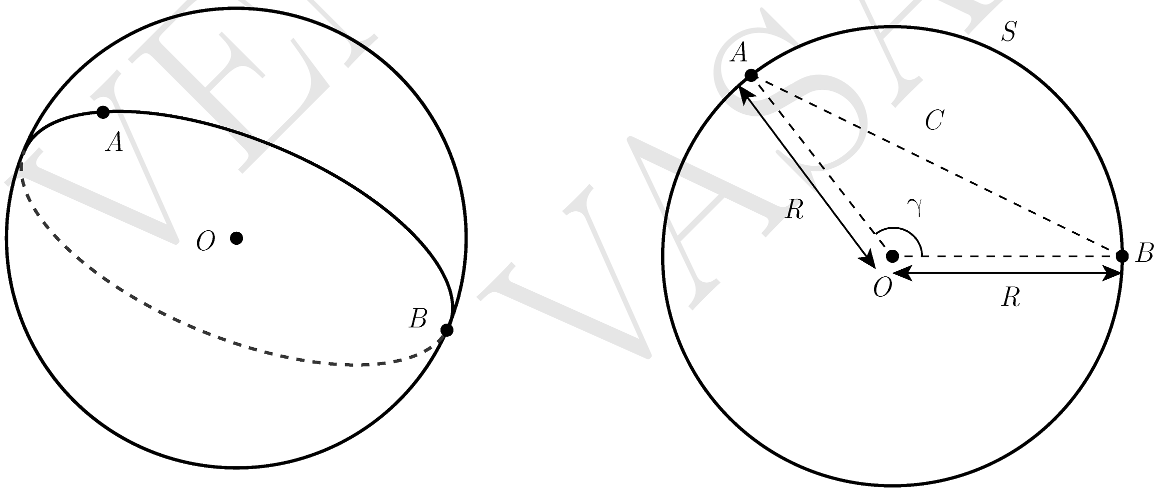
4 Spherical Geometry

With the equation of the geodesic 3.69, we could evaluate the shortest path by simply substituting it back into the arc length functional 3.34. However, this would result in a rather unwieldy integral. Instead, now that we know the geodesic on a sphere is a great circle, we can employ spherical geometry.

Staying consistent with our definition of straight lines to mean geodesics of a certain surface, as discussed in Section 2, we have great circles on the surface of a sphere analogous to straight lines on a two-dimensional plane. This notion will allow us to define relationships of geometries on the surface of a sphere, much like Euclid developed his axioms based on the notion of a line.

4.1 Deriving the Shortest Path Equation

Let us consider the great circle slice that intersects with both points A and B as shown in figure 12a below. As we know that this is the great circle, we know that it has the same radius R. Our aim is to derive an expression for the length of the shorter arc S, which is the shortest path distance, in terms of R and the coordinates of points A and B.



(a) Great Circle Passing Through Points A and B on a Sphere

(b) Great Circle Slice

Figure 12. Diagrams made with Adobe Illustrator.

We can express arc length S in terms of the angle γ the arc subtends at the centre of the great circle, as shown above in figure 12b.

$$S = R\gamma. \tag{4.1}$$

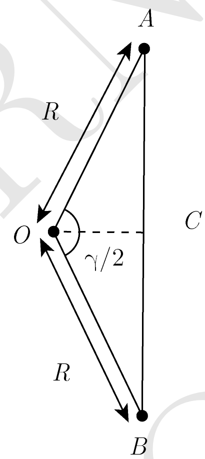


Figure 13. Isosceles triangle AOB made with Adobe Illustrator.

Constructing an isosceles triangle AOB as shown above in figure 13, we can observe that the angle γ can be expressed in terms of the chord length C by the definition of sine.

$$\sin\frac{\gamma}{2} = \frac{C}{2R}. \tag{4.2}$$

Noticing that length C is the Euclidean distance between points A and B, by the Pythagorean theorem we have

$$C^2 = (B_x - A_x)^2 + (B_y - A_y)^2 + (B_z - A_z)^2. \tag{4.3}$$

While one could proceed by substituting C into equation 4.2, let us convert the Cartesian coordinates to geographic coordinates (i.e. longitudes and latitudes) to make it more practical for finding the shortest distance between locations on a spherical approximation of the Earth. The difference between the geographic coordinate system and the spherical coordinate system is that the angle ϕ is the latitude measured from the equator instead of the azimuth angle measured from the positive z-axis. To account for this, we can replace ϕ in the original transformation equations with $\frac{\pi}{2} - \phi$. We get

$$x = r\cos\theta\sin(\frac{\pi}{2} - \phi). \tag{4.4}$$

$$y = r\sin\theta\sin(\frac{\pi}{2} - \phi). \tag{4.5}$$

$$z = r\cos(\frac{\pi}{2} - \phi), \tag{4.6}$$

which simplify respectively to

$$x = r\cos\theta\cos\phi. \tag{4.7}$$

$$y = r\sin\theta\cos\phi. \tag{4.8}$$

$$z = r\sin\phi. \tag{4.9}$$

Hence, converting the Euclidean distance equation to geographic coordinates,

$$\begin{aligned} C^2 = {} & (R\cos B_\theta\cos B_\phi - R\cos A_\theta\cos A_\phi)^2 + \\ & (R\sin B_\theta\cos B_\phi - R\sin A_\theta\cos A_\phi)^2 + (R\sin B_\phi - R\sin A_\phi)^2. \end{aligned} \tag{4.10}$$

Expanding,

$$\begin{aligned} C^2 = {} & R^2\cos^2 B_\theta\cos^2 B_\phi + R^2\cos^2 A_\theta\cos^2 A_\phi - \\ & 2R^2\cos B_\theta\cos B_\phi\cos A_\theta\cos A_\phi + R^2\sin^2 B_\theta\cos^2 B_\phi + \\ & R^2\sin^2 A_\theta\cos^2 A_\phi - 2R^2\sin B_\theta\cos B_\phi\sin A_\theta\cos A_\phi + \\ & R^2\sin^2 B_\phi + R^2\sin^2 A_\phi - 2R^2\sin B_\phi\sin A_\phi. \end{aligned} \tag{4.11}$$

Dividing both sides by R^2,

$$\begin{aligned}\frac{C^2}{R^2} = {} & \cos^2 B_\theta \cos^2 B_\phi + \cos^2 A_\theta \cos^2 A_\phi - \\ & 2\cos B_\theta \cos B_\phi \cos A_\theta \cos A_\phi + \sin^2 B_\theta \cos^2 B_\phi + \\ & \sin^2 A_\theta \cos^2 A_\phi - 2\sin B_\theta \cos B_\phi \sin A_\theta \cos A_\phi + \sin^2 B_\phi + \\ & \sin^2 A_\phi - 2\sin B_\phi \sin A_\phi. \end{aligned} \tag{4.12}$$

Factorising,

$$\begin{aligned}\frac{C^2}{R^2} = {} & \cos^2 B_\phi(\cos^2 B_\theta + \sin^2 B_\theta) + \cos^2 A_\phi(\cos^2 A_\theta + \sin^2 A_\theta) - \\ & 2\cos A_\phi \cos B_\phi(\cos B_\theta \cos A_\theta + \sin B_\theta \sin A_\theta) + \\ & + \sin^2 B_\phi + \sin^2 A_\phi - 2\sin B_\phi \sin A_\phi. \end{aligned} \tag{4.13}$$

By the Pythagorean identity and the compound angle identity, $\cos(\alpha - \beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta)$,

$$\begin{aligned}\frac{C^2}{R^2} = {} & \cos^2 B_\phi + \cos^2 A_\phi - \\ & 2\cos A_\phi \cos B_\phi \cos(B_\theta - A_\theta) + \\ & + \sin^2 B_\phi + \sin^2 A_\phi - 2\sin B_\phi \sin A_\phi \end{aligned} \tag{4.14}$$

$$\frac{C^2}{R^2} = 2 - 2\cos A_\phi \cos B_\phi \cos(B_\theta - A_\theta) - 2\sin B_\phi \sin A_\phi. \tag{4.15}$$

Here we might notice some similarities between the left hand side above and the right hand side of equation 4.2. Diving both sides by four,

$$\frac{C^2}{4R^2} = \frac{1}{2} - \frac{1}{2}\cos A_\phi \cos B_\phi \cos(B_\theta - A_\theta) - \frac{1}{2}\sin B_\phi \sin A_\phi \tag{4.16}$$

$$\left(\frac{C}{2R}\right)^2 = \frac{1}{2} - \frac{1}{2}\cos A_\phi \cos B_\phi \cos(B_\theta - A_\theta) - \frac{1}{2}\sin B_\phi \sin A_\phi. \tag{4.17}$$

Squaring equation 4.2 we have,

$$\sin^2\frac{\gamma}{2} = \left(\frac{C}{2R}\right)^2. \tag{4.18}$$

Thus we have,

$$\sin^2\frac{\gamma}{2} = \frac{1}{2} - \frac{1}{2}\cos A_\phi \cos B_\phi \cos\left(B_\theta - A_\theta\right) - \frac{1}{2}\sin B_\phi \sin A_\phi. \tag{4.19}$$

This is, in fact, an alternative form of the Haversine formula in spherical trigonometry, historically used by sailors navigating Earth (Goodwin 736):

$$\operatorname{haversin}\gamma = \operatorname{haversin}\left(B_\phi - A_\phi\right) + \cos A_\phi \cos B_\phi \operatorname{haversin}\left(B_\theta - A_\theta\right), \tag{4.20}$$

where $\operatorname{haversin}\alpha = \sin^2\frac{\alpha}{2}$.

Hence, angle γ is given by,

$$\gamma = 2\arcsin\left(\sqrt{\frac{1}{2} - \frac{1}{2}\cos A_\phi \cos B_\phi \cos\left(B_\theta - A_\theta\right) - \frac{1}{2}\sin B_\phi \sin A_\phi}\right). \tag{4.21}$$

Finally, substituting γ back into equation 4.1,

$$S = 2R\arcsin\left(\sqrt{\frac{1}{2} - \frac{1}{2}\cos A_\phi \cos B_\phi \cos\left(B_\theta - A_\theta\right) - \frac{1}{2}\sin B_\phi \sin A_\phi}\right), \tag{4.22}$$

or alternatively,

$$S = 2R\arcsin\left(\sqrt{\operatorname{haversin}\gamma}\right). \tag{4.23}$$

Thus, we have obtained a general equation for length S in terms of the geographical coordinates of points A and B.

4.2 Evaluating the Shortest Path Distance

Let us now use the derived formula to evaluate the shortest path distance between the two airports in the Introduction - Changi airport, Singapore and Chhatrapati Shivaji Maharaj International Airport, Mumbai.

Location	**Latitude**, ϕ (°)	**Longitude**, θ (°)
Changi Airport, Singapore	1.359167	103.989441
Chhatrapati Shivaji Maharaj International Airport, Mumbai	19.097403	72.874245

Table 1. Coordinates of Singapore Changi Airport, Singapore and Chhatrapati Shivaji International Airport, Mumbai from LatLong.

Substituting the above values into equation 4.22, we have

$$S = 2(6371008.7714)\arcsin\left(\sqrt{\frac{1}{2} - \frac{1}{2}\sin 19.097403^\circ \sin 1.359167^\circ - \frac{1}{2}\cos 1.359167^\circ \cos 19.097403^\circ \cos\left(72.874245^\circ - 103.989441^\circ\right)}\right), \tag{4.24}$$

where the value of R used is the mean radius of the Earth (Moritz).

This evaluates to

$$S = 3920695.080\text{m}. \tag{4.25}$$

Note that this value assumes a spherical Earth, while the Earth is an oblate spheroid. A literature value which takes into account this flattening is 3919211m (Karney). This corresponds to an accuracy of roughly 99.96214%.

5 Conclusion

Thus, we have reached an answer for the research question *"What is the shortest distance between two points on the surface of a sphere?"* The shortest distance between two points on the surface of a sphere is the shorter arc on a great circle of the sphere that intersects both said points. The general expression for this shortest distance in terms of geographic coordinates is

$$S = 2R \arcsin\left(\sqrt{\frac{1}{2} - \frac{1}{2}\cos A_\phi \cos B_\phi \cos\left(B_\theta - A_\theta\right) - \frac{1}{2}\sin B_\phi \sin A_\phi}\right). \tag{5.1}$$

It is incredible that a question as naive as this requires tools from several subfields of mathematics to answer rigorously. We began by formulating an expression for arc lengths, an exercise in single variable calculus. However, this quickly led to the calculus of variations - namely the idea of *functionals*, which are in a sense functions of functions, and the effects of infinitesimal perturbations on their outputs. Subsequently, we derived the Euler-Lagrange Equation to aid us in finding extremals of arc lengths while also gaining valuable insight into the second-order differential equation. We then applied it to the arc length functional for the surface of a sphere to find the geodesic equation before proving that it is a great circle. Finally, with this information, we used simple circular and three-dimensional geometry to find our desired expression.

The calculus of variations, differential equations and geometry are foundational aspects of differential geometry - the study of the geometry of curves, surfaces, and manifolds (the higher-dimensional analogs of surfaces) (Henderson). While in this essay we have only looked at the shortest path and distance on a sphere, a similar approach can be used to find the shortest path on other surfaces such as an oblate spheroid, a more accurate surface to model the Earth than the sphere. In this essay, we primarily use coordinate systems to define surfaces, namely the spherical coordinate system for the surface of a sphere. However, in future work, a general method of defining surfaces and finding geodesics on those surfaces could be explored. Furthermore, these ideas in differential geometry can also be extended beyond the standard two or three spatial dimensions as explored in this essay. For example, a key problem in General Relativity is finding geodesics on a single four-dimensional manifold which combines the three dimensions of space and the dimension of time (Schutz 172).

This leaves us wondering, what other applications are there of differential geometry?

Works Cited

Google Earth. earth.google.com/.

.

Adobe Illustrator. 2020.

Air India: Interactive Route Map. *Air India | Interactive Route Map*, 2018. www.airindia.in/RouteMap/RouteMap.aspx.

Gelfand, I. M. and S. V. Fomin. *Calculus of Variations by I.M. Gelfand and S.V. Fomin.* Prentice-Hall, 1964.

Geodesic Line. *Geodesic line - Encyclopedia of Mathematics*, 2012. encyclopediaofmath.org/wiki/Geodesic_line.

Goodwin, H. B. "The Haversine in Nautical Astronomy". *Naval Institute Proceedings*, vol. 36, 1910. babel.hathitrust.org/cgi/pt?id=njp.32101043277217&view=1up&seq=783.

Hanc, Jozef. "The original Euler's calculus-of-variations method: Key to Lagrangian mechanics for beginners". www.eftaylor.com/pub/HancEulerEJP.pdf.

Henderson, David W. Differential Geometry. *Encyclopædia Britannica*, www.britannica.com/science/differential-geometry.

Karney, Charles F. F. *Geodesic lines, circles, envelopes in Google Maps*, geographiclib.sourceforge.io/scripts/geod-google.html.

Karttunen, Hannu. *Fundamental Astronomy*. Springer, 2003.

Kot, Mark. *A First Course in the Calculus of Variations*. American Mathematical Society, 2014.

LatLong. Chhatrapati Shivaji International Airport, Mumbai. *Lat Long Finder*, www.latlong.net/place/chhatrapati-shivaji-international-airport-bom-vile-parle-east-mumbai-india-2610.html.

———. Singapore Changi Airport, Singapore. *Lat Long Finder*, www.latlong.net/place/singapore-changi-airport-singapore-27493.html.

Miettinen, Arttu. Spherical coordinates¶. *Spherical coordinates - pi2 documentation*, 2011. pi2-docs.readthedocs.io/en/latest/spherical_coordinates.html.

Moritz, H. "Geodetic Reference System 1980". *Journal of Geodesy*, vol. 74, no. 1, 2000, pp. 128–133. doi:10.1007/s001900050278.

Nakkiran, Preetum. Geometric Derivation of Euler-Lagrange Equation. *Geometric derivation of Euler-Lagrange equation*, preetum.nakkiran.org/lagrange.html.

Schutz, Bernard F. *A First Course in General Relativity.* Cambridge UP, 2009.

Weisstein, Eric W. Euclid's Postulates. *from Wolfram MathWorld*, mathworld.wolfram.com/EuclidsPostulates.html.

———. Spherical Coordinates. *from Wolfram MathWorld*, mathworld.wolfram.com/SphericalCoordinates.html.

Wolfram Cloud. *Wolfram Cloud*, www.wolframcloud.com/.