# Calculus of Variations: an Introduction to the Optimization of Functionals

Vedansh Gupta

July 8, 2026

**Abstract**

This paper describes the mathematical technique of calculus of variations starting with the derivation of key equations such as the Euler-Lagrange equation and the Beltrami Identity. These equations are then applied in the following scenarios: geodesic on a flat plane, solving the Brachistochrone problem, geodesic on a sphere, and finding the equation of the function formed by a string under the influence of gravity (catenary problem).

## 1 Introduction

We recall the methods of single-variable calculus, where finding the extremum value of a function $f(x)$ can be done by taking the derivative $f'(x)$ and equating it to 0. The value of $x$ obtained from $f'(x) = 0$ can be tested for making the function a maximum or minimum via the second derivative test where if $f''(x) > 0$, a minimum is obtained and if $f''(x) < 0$, a maximum is obtained.

This idea of extremals continues in the calculus of variations, however, instead of optimizing functions, we optimize functionals.

A functional is essentially a function that takes in a function input and returns a scalar output. We have already come across a functional in elementary calculus: the integral. An example may be in the form

$$I[y] = \int_{x_1}^{x_2} 3x^2[y(x)]^2 dx$$

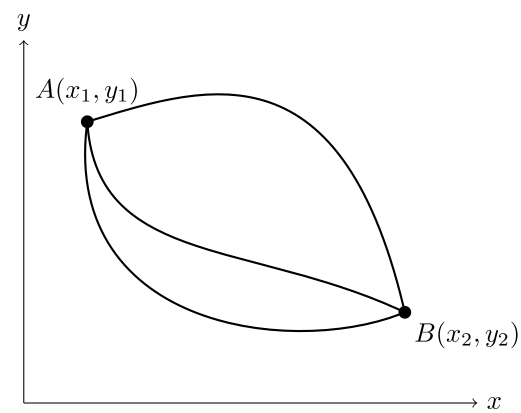
where the functional $I$ takes in a function $y(x)$ as an input and returns a scalar output given that the constants $x_1$ and $x_2$ are defined.

Generally, functionals are written in the following form

$$I[y] = \int_{x_1}^{x_2} F(x, y(x), y'(x))dx.$$

A functional takes on the inputs $x, y, y'$ as the values of $x$ and $y$ give the coordinates of a particular function in $\mathbb{R}^2$, and $y'$, which gives the function's gradient at a particular point, allows it to take on infinitely many shapes, allowing a family of functions to be obtained.

In the calculus of variations we find the optimal function $y(x)$ that gives a minimum value for the functional. Consider the following example of a use of calculus of variations




Given the points A and B, the diagram above shows 3 different paths that connect them. The calculus of variations can be used to find the function that gives the smallest distance between A and B. Intuitively, we know that this is just a straight line, but that is not sufficient to prove that no other function minimizes the distance between these two points. It is important to note that the boundary conditions $y(x_1) = y_1$ and $y(x_2) = y_2$ must be met as the minimizing function must pass through both A and B in order to minimise the distance between them.

The function between A and B can be written as

$$I[y] = \int_{x_1}^{x_2} ds$$

indicating many small changes in distance $ds$ from $x_1$ to $x_2$. Then we have using the Pythagorean theorem:

$$ds = \sqrt{(dx)^2 + (dy)^2}$$

$$ds = dx\sqrt{1 + \left(\frac{dy}{dx}\right)^2}$$

$$I[y] = \int_{x_1}^{x_2} \sqrt{(1 + \left(\frac{dy}{dx}\right)^2} dx = \int_{x_1}^{x_2} \sqrt{(1 + (y')^2} dx$$

But now a question arises on what to do from here. There needs to be some condition that we can use to find out the function that minimizes this functional. In elementary calculus, we could simply solve for the minimum $x$ value $x_{min}$ such that $f'(x_{min}) = 0$ is satisfied. We seek an analogous condition for functionals. This condition is the Euler-Lagrange equation, which is proved and further discussed in the next section.

## 2 Euler-Lagrange Equation

Suppose that we have a functional $I$ in the form

$$I[y] = \int_{x_1}^{x_2} F(x, y(x), y'(x)) dx$$

.

where $F$ is infinitely differentiable across $x$,$y(x)$, $y'(x)$ and the following boundary conditions are met for two boundary points X and Y with co-ordinates in $\mathbb{R}^2$ of $(x_1, y_1), (x_2, y_2)$, respectively. Let $y(x_1) = y_1$ and $y(x_2) = y_2$.

The idea of calculus of variations is used here to find the stationary value for the functional $F$ with respect to all perturbations of $y(x)$. We let this $y(x)$ value be an extremal for $F$.

Now a new $\eta(x)$ function is introduced which represents the deviation from the true extremum function $y(x)$. By adding this $\eta(x)$ function, we see what would happen if the extremum function is slightly perturbed. This perturbed function can then be checked on whether it satisfies optimality conditions.

It is important to note that

$$\eta(x_1) = \eta(x_2) = 0$$

as the boundary conditions must still be met and the deviation from the boundary points is not allowed to be greater than 0.

We can also define a new function $\bar{y}$ with a parameter $\epsilon$ which represents the family of curves that are possibilities for the path between the boundary points.

$$\bar{y}(x) = y(x) + \epsilon\eta(x)$$

Note that the boundary conditions still hold $\bar{y}(x_1) = y(x_1) + \epsilon\eta(x_1) = y_1$ and $\bar{y}(x_2) = y(x_2) + \epsilon\eta(x_2) = y_2$.

Our original functional $I[y]$ can be replaced with $I[\bar{y}]$ giving

$$I[y] = \int_{x_1}^{x_2} F(x, \bar{y}(x), \bar{y}'(x))dx.$$

To solve this integral, the independent variable $x$ is integrated out leaving $I$ as a function of $\epsilon$, therefore

$$I[\epsilon] = \int_{x_1}^{x_2} F(x, \bar{y}(x), \bar{y}'(x))dx.$$

If $I[\epsilon]$ was just a regular function, elementary calculus would require us to take $\frac{dI}{d\epsilon}$ and equate it to 0. Since $y(x)$ is a stationary function, the curve obtained from $\epsilon = 0$ should give the stationary value of $I[\epsilon]$. Therefore we have

$$\left.\frac{dI}{d\epsilon}\right|_{\epsilon=0} = \left.\frac{d}{d\epsilon}\int_{x_1}^{x_2} F(x, \bar{y}(x), \bar{y}'(x))\, dx\right|_{\epsilon=0} = \int_{x_1}^{x_2} \left.\frac{\partial}{\partial\epsilon} F(x, \bar{y}(x), \bar{y}'(x))\right|_{\epsilon=0} dx = 0.$$

The part inside the integral can be simplified using the multivariable chain rule in the following manner

$$\int_{x_1}^{x_2} \left.\left[\frac{\partial F}{\partial \bar{y}}\frac{\partial \bar{y}}{\partial \epsilon} + \frac{\partial F}{\partial \bar{y}'}\frac{\partial \bar{y}'}{\partial \epsilon}\right]\right|_{\epsilon=0} dx = 0$$

As defined earlier we still have $\bar{y} = y(x) + \epsilon\eta(x)$ and $\bar{y}' = y'(x) + \epsilon\eta'(x)$ so

$$\frac{\partial \bar{y}}{\partial \epsilon} = \eta(x), \frac{\partial \bar{y}'}{\partial \epsilon} = \eta'(x).$$

This allows us to rewrite the integral as

$$\int_{x_1}^{x_2} \left.\left[\frac{\partial F}{\partial \bar{y}}\eta(x) + \frac{\partial F}{\partial \bar{y}'}\eta'(x)\right]\right|_{\epsilon=0} dx = 0$$

We now proceed with integration by parts where $u = \frac{\partial F}{\partial \bar{y}'}$, $u' = \frac{d}{dx}[\frac{\partial F}{\partial \bar{y}'}]$,$v = \eta(x)$,$v' = \eta'(x)$ in the equation $\int_{x_1}^{x_2} uv' = [uv]_{x_1}^{x_2} - \int_{x_1}^{x_2} u'v$ giving

$$\int_{x_1}^{x_2} \left.\frac{\partial F}{\partial \bar{y}'}\right|_{\epsilon=0} \eta'(x)\, dx = \left[\left.\frac{\partial F}{\partial \bar{y}'}\right|_{\epsilon=0} \eta(x)\right]_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{d}{dx}\left[\left.\frac{\partial F}{\partial \bar{y}'}\right|_{\epsilon=0}\right]\eta(x)\, dx.$$

The boundary term is equal to 0 as $\eta(x_1) = \eta(x_2) = 0$ so

$$\int_{x_1}^{x_2} \left.\frac{\partial F}{\partial \bar{y}'}\right|_{\epsilon=0} \eta'(x)\, dx = -\int_{x_1}^{x_2} \frac{d}{dx}\left[\left.\frac{\partial F}{\partial \bar{y}'}\right|_{\epsilon=0}\right]\eta(x)\, dx.$$

Substituting the above equation into the original chain rule equation, we get

$$\int_{x_1}^{x_2} \left.\left[\frac{\partial F}{\partial \bar{y}}\eta(x) - \frac{d}{dx}\left(\frac{\partial F}{\partial \bar{y}'}\right)\eta(x)\right]\right|_{\epsilon=0} dx = 0.$$

We proceed by factoring out the $\eta(x)$ in the integral and rewriting $\bar{y}'(x)$ at $\epsilon = 0$ as $y'(x)$ and $\bar{y}(x)$ as $y(x)$

$$\int_{x_1}^{x_2} \left[ \frac{\partial F}{\partial y} - \frac{d}{dx} \left( \frac{\partial F}{\partial y'} \right) \right] \eta(x)\, dx = 0$$

Now we use an important idea called the fundamental lemma of calculus which states that for functionals in the form

$$\int_a^b f(x)\eta(x)dx$$

and for boundary conditions $\eta(a) = \eta(b) = 0$, the value of f(x) must be 0 for all x values from a to b

This suggests the following Euler-Lagrange equation

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

We note that this equation functions the same way as the first derivative equated to 0 to find the extremal in elementary calculus, where the Euler-Lagrange equation must be satisfied to find the extremal of a functional.

# 3 Beltrami Identity

The Beltrami Identity is a special variation of the Euler-Lagrange equation where the function $F$(integrand) in the functional $I$ doesn't explicitly depend on $x$.

$$I[y] = \int_{x_1}^{x_2} F(y(x), y'(x))dx$$

Hence, we can consider $\frac{\partial F}{\partial x} = 0$. We begin with the Euler-Lagrange equation discussed in the previous section

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

We can then proceed by multiplying the full equation by the term $y'(x)$ obtaining

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

Now we take the total derivative of $F$ in the form $\frac{dF}{dx}$. The function F depends on $x$, $y(x)$, and $y'(x)$ so there will be three terms in the following expression.

$$\frac{dF}{dx} = \frac{\partial F}{\partial x}\frac{dx}{dx} + \frac{\partial F}{\partial y}\frac{dy}{dx} + \frac{\partial F}{\partial y'}\frac{dy'}{dx}$$

$$\frac{dF}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}y' + \frac{\partial F}{\partial y'}y''$$

Rearranging to solve for the term $y'\frac{\partial F}{\partial y}$, we get

$$y'\frac{\partial F}{\partial y} = \frac{dF}{dx} - \frac{\partial F}{\partial x} - y''\frac{\partial F}{\partial y'}$$

Substituting this rearrangement into the Euler-Lagrange equation multiplied by $y'(x)$, we get

$$\frac{dF}{dx} - \frac{\partial F}{\partial x} - \left[y''\frac{\partial F}{\partial y'} + y'\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right] = 0.$$

The term in the square brackets can be simplified by using the reverse product rule giving the following equation

$$\frac{d}{dx}\left(y'\frac{\partial F}{\partial y'}\right) = y''\frac{\partial F}{\partial y'} + y'\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)$$

$$\frac{dF}{dx} - \frac{\partial F}{\partial x} - \frac{d}{dx}\left(y'\frac{\partial F}{\partial y'}\right) = 0$$

$$\frac{\partial F}{\partial x} = \frac{dF}{dx} - \frac{d}{dx}\left(y'\frac{\partial F}{\partial y'}\right)$$

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

As discussed earlier, the Beltrami Identity only applies when F does not explicitly depend on x so the term $\frac{\partial F}{\partial x} = 0$ giving

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

Now we can integrate both sides with respect to $x$ to get the following equation known as the Beltrami identity for some arbitrary integration constant $C$ determined using boundary conditions.

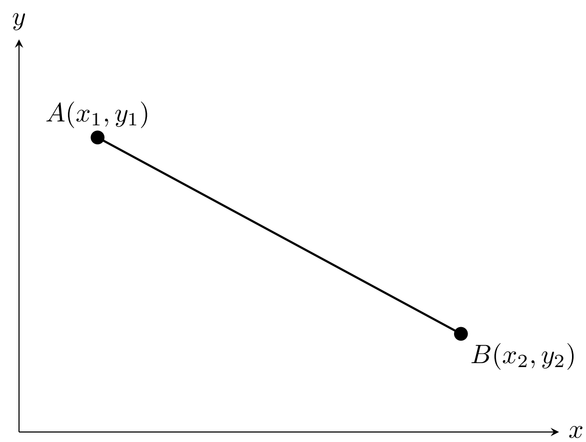
$$\boxed{F - y'\frac{\partial F}{\partial y'} = C}$$

As seen in later sections of this paper, the Beltrami identity serves as an extremely useful tool to optimize functions when they don't depend explicitly on $x$. In this case, the second order Euler-Lagrange equation can be sidelined completely through the use of Beltrami which proves to be simpler.

# 4 Geodesic on a Flat Plane

A geodesic is the shortest path that can be taken between two points on a curved space or a flat plane. In the introduction, it was stated that the functional mapping the path between two points on a flat plane is given by

$$I[y] = \int_{x_1}^{x_2} \sqrt{(1 + (y')^2} dx$$




Immediately we note that this functional only depends upon $y'(x)$ and not $x$ so the Beltrami identity can be used.

$$F - y' \frac{\partial F}{\partial y'} = C$$

We can solve for $\frac{\partial F}{\partial y'}$

$$\frac{\partial F}{\partial y'} = \frac{y'}{\sqrt{1 + (y')^2}}$$

Plugging the expression for $\frac{\partial F}{\partial y'}$ into the Beltrami identity above, we get

$$F - y' \frac{y'}{\sqrt{1 + (y')^2}} = F - \frac{(y')^2}{\sqrt{1 + (y')^2}} = C$$

Substituting the function $\sqrt{(1 + (y')^2}$ for $F$

$$\sqrt{(1 + (y')^2} - \frac{(y')^2}{\sqrt{1 + (y')^2}} = C$$

Multiplying both sides by $\sqrt{(1 + (y')^2}$, we get

$$1 + (y')^2 - (y')^2 = C\sqrt{(1 + (y')^2}.$$

Then expanding and solving for y', we get

$$1 = C\sqrt{(1+(y')^2}$$

$$1 = C^2(1+(y')^2)$$

$$1 = C^2 + C^2(y')^2$$

$$(y')^2 = \frac{1-C^2}{C^2}$$

$$y' = \frac{dy}{dx} = \frac{\sqrt{1-C^2}}{C} = M.$$

We note that the term $\frac{\sqrt{1-C^2}}{C}$ is just another constant as $C$ is a constant so we can call this constant $M$. Hence, the gradient of the geodesic between two points is a constant implying that the function of the geodesic is given by a straight line in the form $y = Mx + N$ for random constants M and N. To find the missing constants M and N, the boundary conditions are used which, stated in the introduction, were $y(x_1) = y_1$ and $y(x_2) = y_2$. So we have the following simultaneous equations to solve for the geodesic.

$$y_1 = Mx_1 + N$$

$$y_2 = Mx_2 + N$$

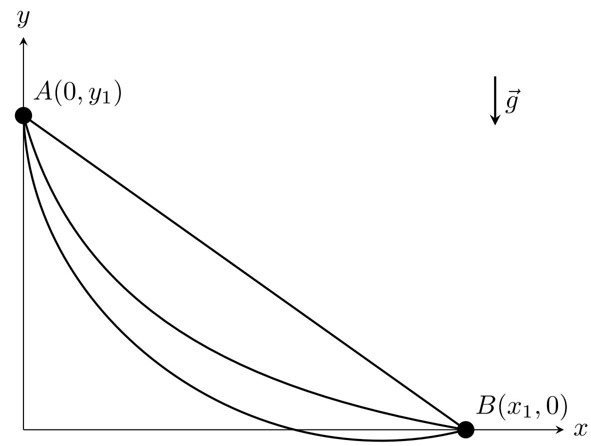
Therefore this proof, using the calculus of variations and the Beltrami identity, confirms that the geodesic between two points on a flat plane is a straight line.

# 5 The Brachistochrone Problem

We will now move on to the Brachistochrone problem proposed by Johann Bernoulli in 1696. The name comes from the Greek word 'Brachistos' meaning shortest and the word 'Chronos' meaning time. Therefore the word Brachistochrone suggests the path with the shortest time.

We now move on to the problem and its solution. We define two points A and B on an xy-plane with co-ordinates $(x_1, y_1)$ and $(x_2, y_2)$ respectively. We let $y_1 > y_2$ so that point A is above point B on this plane. We then introduce a ball with negligible mass and size that begins at point A and needs to travel to point B. The goal of this problem is to find the function between points A and B that allows this ball to travel from point A to point B in the shortest time possible under the influence of gravity.

Without loss of generality, we can place the point A on the Y-axis with co-ordinates $(0, y_1)$ and point B on the X-axis with co-ordinates $(x_1, 0)$. This is accurately shown on the diagram below with a few potential solutions to the Brachistochrone.




We call $y(x)$ the function that minimizes the time taken for the ball to fall from point A to B. We begin with the law of conservation of energy between gravitational potential energy and kinetic energy. For each height $y(x)$ we have

$$mgy(0) = mgy(x) + \frac{1}{2}mv^2.$$

y(0) can be rewritten as $y_1$ and then we can simplify and attempt to solve for $v$.

$$mgy_1 = mgy + \frac{1}{2}mv^2$$

$$\frac{1}{2}mv^2 = mg(y_1 - y)$$

Cancelling out the masses and isolating $v$

$$v = \sqrt{2g(y_1 - y)}.$$

As we know that $time = \frac{distance}{velocity}$, we can form the following small change equation.

$$dt = \frac{ds}{v}$$

We have already solved for $v$ and ds is simply just the infinitesimal arc length distance from A to B so this equation becomes

$$dt = \frac{\sqrt{1 + (y')^2}}{\sqrt{2g(y_1 - y)}}dx.$$

The overall time can be written using the functional below.

$$T[y] = \int_0^{x_1} \frac{\sqrt{1 + (y')^2}}{\sqrt{2g(y_1 - y)}}dx$$

Therefore, we have our $F(x, y, y')$

$$F(x, y, y') = \frac{\sqrt{1 + (y')^2}}{\sqrt{2g(y_1 - y)}}$$

We again note that $F$ doesn't directly depend upon $x$ so the Beltrami identity can be used. To reiterate this is

$$F - y' \frac{\partial F}{\partial y'} = C.$$

We now solve for $\frac{\partial F}{\partial y'}$

$$\frac{\partial F}{\partial y'} = \frac{y'}{\sqrt{2g(y_1 - y)} \cdot \sqrt{1 + (y')^2}}.$$

Substituting this into the Beltrami identity we have

$$\frac{\sqrt{1 + (y')^2}}{\sqrt{2g(y_1 - y)}} - \frac{(y')^2}{\sqrt{2g(y_1 - y)} \cdot \sqrt{1 + (y')^2}} = C.$$

Multiplying both sides by $\sqrt{1 + (y')^2}$ and $\sqrt{2g(y_1 - y)}$ we have

$$\begin{aligned} 1 + (y')^2 - (y')^2 &= C\sqrt{2g(y_1 - y)}\sqrt{1 + (y')^2} \\ 1 &= C\sqrt{2g(y_1 - y)}\sqrt{1 + (y')^2} \\ 1 &= C^2(2g(y_1 - y))(1 + (y')^2) \end{aligned}$$

We define a new constant $C_1 = \frac{1}{2gC^2}$, so the above equation becomes

$$C_1 = (1 + (y')^2)(y_1 - y).$$

Expanding this out

$$C_1 = y_1 - y + (y_1 - y)(y')^2.$$

Now we solve for $y'$

$$\begin{aligned} (y')^2 &= \frac{C_1 + y - y_1}{y_1 - y} \\ y' = \frac{dy}{dx} &= \sqrt{\frac{C_1 + y - y_1}{y_1 - y}} \end{aligned}$$

This can be rewritten as the following equation.

$$dx = \frac{\sqrt{y_1 - y}}{\sqrt{C_1 + y - y_1}} dy$$

We proceed by integrating both sides. The integral of $1dx$ is simply just $x$ but the right hand side gets a little tricky.

$$x = \int \frac{\sqrt{y_1 - y}}{\sqrt{C_1 - (y_1 - y)}} dy$$

In order to simplify this integral we can use the following trigonometric substitution.

$$y = y_1 - C_1 \sin^2\left(\frac{\theta}{2}\right)$$

Taking the derivative of $y$ with respect to $\theta$ we have

$$\frac{dy}{d\theta} = -C_1 \sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right).$$

$$dy = -C_1 \sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) d\theta.$$

Substituting $y$ and $dy$ back into the original integral we get

$$x = \int \frac{\sqrt{y_1 - y_1 + C_1 \sin^2\left(\frac{\theta}{2}\right)}}{\sqrt{C_1 - (y_1 - y_1 + C_1 \sin^2\left(\frac{\theta}{2}\right)}} - C_1 \sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) d\theta.$$

We then simplify the terms within the integral.

$$x = \int \frac{\sqrt{C_1 \sin^2\left(\frac{\theta}{2}\right)}}{\sqrt{C_1 - C_1 \sin^2\left(\frac{\theta}{2}\right)}} - C_1 \sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) d\theta.$$

Dividing by $C_1$ within the square roots and taking the $-C_1$ out of the integral, we get the following

$$x = -C_1 \int \frac{\sqrt{\sin^2\left(\frac{\theta}{2}\right)}}{\sqrt{1 - \sin^2\left(\frac{\theta}{2}\right)}} \sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) d\theta$$

$$x = -C_1 \int \frac{\sqrt{\sin^2\left(\frac{\theta}{2}\right)}}{\sqrt{\cos^2\left(\frac{\theta}{2}\right)}} \sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) d\theta$$

$$x = -C_1 \int \tan\left(\frac{\theta}{2}\right) \sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) d\theta$$

$$x = -C_1 \int \sin^2\left(\frac{\theta}{2}\right) d\theta$$

We now utilise the half angle identity for sin which is $\sin^2\left(\frac{\theta}{2}\right) = \frac{1-\cos(\theta)}{2}$

$$x = \frac{-C_1}{2} \int 1 - \cos(\theta) d\theta$$

Now we can finally eliminate the integral. $C_2$ is another arbitrary constant.

$$x = -\frac{C_1}{2}(\theta - \sin(\theta)) + C_2$$

Therefore the answer to the Brachistochrone problem is the two following parametric equations.

$$x = \frac{-C_1}{2}(\theta - \sin(\theta)) + C_2$$

$$y = y_1 - C_1 \sin^2\left(\frac{\theta}{2}\right) = y_1 - \frac{C_1}{2}(1 - \cos(\theta))$$

We can use the boundary conditions to further simplify these equations. We have $y = y_1$ when $x = 0$ and we have $y = 0$ when $x = x_1$. We take the first condition

$$y_1 = y_1 - \frac{C_1}{2}(1 - \cos(\theta))$$

$$0 = (1 - \cos(\theta))$$

$$\cos(\theta) = 1, \theta = 0.$$

When $\theta = 0$ and $x = 0$, we get $C_2$ is also 0 from the parametric equation for $x$.

This gives the parametric equations for a cycloid. The constant $\frac{-C_1}{2}$ is essentially just another constant that we can call $k$. Therefore, our final parametric equations are given below.

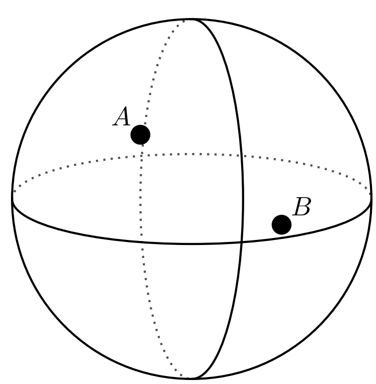
$$\boxed{x = k(\theta - \sin(\theta))}$$

$$\boxed{y = y_1 + k(1 - \cos(\theta))}$$

# 6 Geodesic on a Sphere

We now understand how the calculus of variations allows us to find geodesics on a flat plane. In this section we will find the geodesic on a sphere.

Suppose we have two points A and B on a sphere as shown in the diagram below. The goal of the problem is to find the equation of the function that minimizes their distance on the surface of the sphere.



We once again must use the arc length integral form for the shortest distance between the two points.

$$L = \int_A^B ds$$

As this sphere is in $\mathbb{R}^3$, $ds$ can be expressed in the following manner.

$$ds = \sqrt{(dx)^2 + (dy)^2 + (dz)^2}$$

$$L = \int_A^B \sqrt{(dx)^2 + (dy)^2 + (dz)^2}$$

To make this problem easier, we can convert everything into spherical coordinates. The conversions are given below where R is the radial distance from the origin, $\theta$ is the angle in the xy-plane,and $\phi$ is the angle from the positive z-axis.

$$x = R\cos\theta\sin\phi$$

$$y = R\sin\theta\sin\phi$$

$$z = R\cos\phi$$

In order to convert the terms $dx$,$dy$,$dz$ we can use the total derivative rule in multivariable calculus.

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

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

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

When we solve for all the partials, we can rewrite the $dx$,$dy$,$dz$ terms as the following.

$$dx = -R\sin\theta\sin\phi d\theta + R\cos\theta\cos\phi d\phi$$

$$dy = R\cos\theta\sin\phi d\theta + R\sin\theta\cos\phi d\phi$$

$$dz = -R\sin\phi d\phi$$

We now square each differential term individually to evaluate the arc length element $(ds)^2 = (dx)^2 + (dy)^2 + (dz)^2$. Expanding the squared expressions yields:

$$dx^2 = R^2\sin^2\theta\sin^2\phi\,(d\theta)^2 - 2R^2\sin\theta\cos\theta\sin\phi\cos\phi\,(d\theta)(d\phi) + R^2\cos^2\theta\cos^2\phi\,(d\phi)^2$$

$$dy^2 = R^2\cos^2\theta\sin^2\phi\,(d\theta)^2 + 2R^2\sin\theta\cos\theta\sin\phi\cos\phi\,(d\theta)(d\phi) + R^2\sin^2\theta\cos^2\phi\,(d\phi)^2$$

$$dz^2 = R^2\sin^2\phi\,(d\phi)^2$$

As we add these terms together, we observe that the cross-multiplication terms $(\pm 2R^2 \sin\theta\cos\theta\sin\phi\cos\phi\,(d\theta)(d\phi))$ in $(dx)^2$ and $(dy)^2$ cancel each other out:

$$(ds)^2 = R^2\sin^2\phi\left(\sin^2\theta+\cos^2\theta\right)(d\theta)^2 + R^2\cos^2\phi\left(\cos^2\theta+\sin^2\theta\right)(d\phi)^2 + R^2\sin^2\phi\,(d\phi)^2$$

Applying the pythagorean identity $\sin^2\theta+\cos^2\theta = 1$, the expression simplifies significantly:

$$(dx)^2+(dy)^2+(dz)^2 = R^2\sin^2\phi\,(d\theta)^2 + R^2\cos^2\phi\,(d\phi)^2 + R^2\sin^2\phi\,(d\phi)^2$$

Factoring out $R^2 d\phi^2$ from the last two terms and applying the pythagorean identity again, we obtain the standard line element on a sphere:

$$(dx)^2+(dy)^2+(dz)^2 = R^2\sin^2\phi\,(d\theta)^2 + R^2\,(d\phi)^2$$

Therefore, the infinitesimal arc length $ds = \sqrt{dx^2+dy^2+dz^2}$ that forms our functional integral simplifies conveniently to:

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

Therefore, we have the following functional $L$ and lagrangian in the form $F(\phi,\theta,\theta')$ below.

$$L[\theta] = \int_{\phi_A}^{\phi_B} R\sqrt{\sin^2\phi\left(\frac{d\theta}{d\phi}\right)^2+1}\,d\phi$$

Now that our functional is fully defined, we can use the Euler-Lagrange equation in the form

$$\frac{\partial F}{\partial\theta} - \frac{d}{d\phi}\left(\frac{\partial F}{\partial\theta'}\right) = 0.$$

Note that we use the Euler-Lagrange equation here rather than the Beltrami identity even though our Lagrangian $F$ doesn't explicitly depend on $\phi$. This is because the term $\frac{\partial F}{\partial\theta} = 0$ already so the equation can be simplified into a first integral quite simply. Therefore we have the following simplifications.

$$-\frac{d}{d\phi}\left(\frac{\partial F}{\partial\theta'}\right) = 0$$

We can integrate both sides with respect to $\phi$ where $k$ is some arbitrary constant.

$$\frac{\partial F}{\partial\theta'} = k$$

We then solve for $\frac{\partial F}{\partial\theta'}$ using the chain rule.

$$F = R\sqrt{\sin^2\phi\left(\frac{d\theta}{d\phi}\right)^2+1}$$

$$\frac{\partial F}{\partial \theta'} = R\left(\frac{1}{2}\right)(\sin^2\phi(\theta')^2+1)^{-\frac{1}{2}}(2\theta')(\sin^2\phi) = \frac{R\theta'\sin^2\phi}{\sqrt{\sin^2\phi(\theta')^2+1}}$$

Therefore we have

$$\frac{R\theta'\sin^2\phi}{\sqrt{\sin^2\phi(\theta')^2+1}} = k.$$

Squaring both sides gives

$$\frac{R^2(\theta')^2\sin^4\phi}{\sin^2\phi(\theta')^2+1} = k^2.$$

$$R^2(\theta')^2\sin^4\phi = k^2(\sin^2\phi(\theta')^2+1)$$

$$R^2(\theta')^2\sin^4\phi = k^2\sin^2\phi(\theta')^2+k^2$$

Now, as $R$ and $k$ are both constants we can call $\frac{k}{R}$ another constant $p$.

$$(\theta')^2\sin^4\phi = p^2\sin^2\phi(\theta')^2+p^2$$

$$(\theta')^2\left(\sin^4\phi - p^2\sin^2\phi\right) = p^2$$

Isolating for $\theta'$

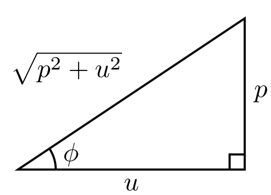
$$(\theta')^2 = \frac{p^2}{\sin^4\phi - p^2\sin^2\phi}$$

$$\frac{d\theta}{d\phi} = \theta' = \frac{p}{\sin\phi\sqrt{\sin^2\phi - p^2}}$$

We can then integrate both sides with respect to $\phi$ giving

$$\theta = \int \frac{p}{\sin\phi\sqrt{\sin^2\phi - p^2}}d\phi$$

To evaluate the above integral, we must use u-substitution where we can take $u = p\cot\phi$. Therefore, $\frac{du}{d\phi} = -p\csc^2\phi$, and $du = -p\csc^2\phi d\phi$. Now we need $\sin\phi$ and $\sin^2\phi$ and this can be seen simply on a right triangle below.




Therefore we have the following equations:

$$\sin\phi = \frac{p}{\sqrt{p^2+u^2}}$$

$$\sin^2\phi = \frac{p^2}{p^2+u^2}$$

Substituting all the terms back into the integral we get

$$\theta = -\int \frac{p}{\frac{p}{\sqrt{p^2+u^2}}\sqrt{\frac{p^2}{p^2+u^2}-p^2}} \frac{du}{p\csc^2\phi}.$$

Now we simplify

$$\theta = -\int \frac{1}{\frac{p}{\sqrt{p^2+u^2}}\sqrt{\frac{p^2}{p^2+u^2}-p^2}} du \sin^2\phi$$

$$\theta = -\int \frac{1}{\frac{p}{\sqrt{p^2+u^2}}\sqrt{\frac{p^2}{p^2+u^2}-p^2}} \frac{p^2}{p^2+u^2} du$$

$$\theta = -\int \frac{1}{\frac{p^2+u^2}{\sqrt{p^2+u^2}}\sqrt{\frac{1}{p^2+u^2}-1}} du$$

$$\theta = -\int \frac{1}{\sqrt{p^2+u^2}\sqrt{\frac{1-p^2-u^2}{p^2+u^2}}} du$$

$$\theta = -\int \frac{du}{\sqrt{1-p^2-u^2}}$$

As $p^2$ is a constant, $1-p^2$ will also be a constant that we can call $\alpha^2$

$$\theta = -\int \frac{du}{\sqrt{\alpha^2-u^2}}$$

Now we can get rid of the integral sign. $\theta_0$ is an arbitrary angle constant.

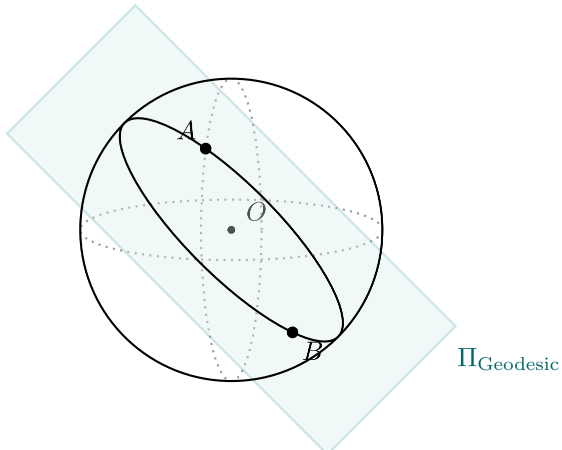
$$\theta = -\left(\arcsin\left(\frac{u}{\alpha}\right) + \theta_0\right)$$

Plugging back $u = p\cot\phi$

$$\theta = -\left(\arcsin\left(\frac{p\cot\phi}{\alpha}\right) + \theta_0\right)$$

If we take $\frac{p}{\alpha}$ as another constant $\beta$, our final geodesic takes the form below. We can solve for $\beta$, $\phi$, and $\theta_0$ from the boundary conditions.

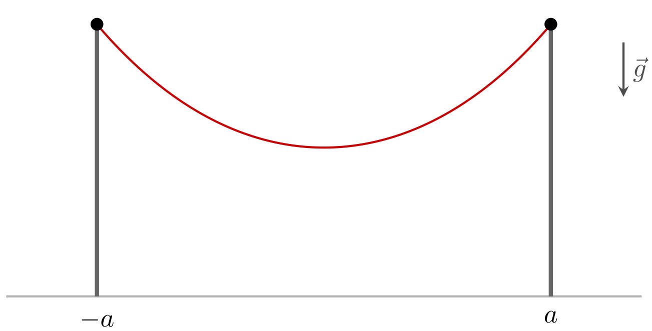
$$\boxed{\theta_0 - \theta = \arcsin(\beta\cot\phi)}$$




Visually, the geodesic on a sphere can be seen by intersecting a flat plane that passes through both points A and B. This makes up a shape called the great circle and the shortest path between two points on a sphere involves traveling along the circumference of the great circle. The diagram above highlights this geodesic.

# 7 Catenary Problem

The final problem we will be tackling in this paper is the catenary problem. Suppose that we have a string of fixed length L and we hang it between two poles from $x = -a$ to $x = a$ under the influence of gravity. Our length $L$ is greater than $2a$ so the string sags downwards rather than being horizontal. A catenary is essentially the shape that this string makes under the influence of gravity and the goal of this problem is to find that function which we can call $y(x)$. We assume that the string is uniform and has a constant density of $\rho$ throughout. We also assume that the points where the string meets with the two poles is $(-a, 0)$ and $(a, 0)$ as our boundary conditions for simplicity. This problem is depicting by the diagram below.



We begin with a constraint. As the string length is a fixed quantity $L$ we can write

$$L = \int_{-a}^{a} ds = \int_{-a}^{a} \sqrt{1 + (y')^2} dx$$

In the absence of external forces, physical systems tend to move towards formations that have lower potential energies. Therefore, the catenary would take the shape where the total gravitational potential energy ($u_g$) of the string is minimized.

If we take an infinitesimally small portion of the string we can say that portion's gravitational potential energy ($du_g$) is the product of the mass, acceleration due to gravity, and height of the string relative to the ground. The mass can be written as $\rho \cdot ds$ as the cross-sectional area of this string is considered negligible. The height is just $y$.

$$du_g = \rho ds g y$$

The total gravitational potential energy of the string is essentially just the sum of the gravitation potential energies of every small portion $ds$ so we have the integral below.

$$u_g = \int_{-a}^{a} \rho g y ds = \int_{-a}^{a} \rho g y \sqrt{1 + (y')^2} dx$$

Now we have the functional $u_g$ that we need to minimize and we have the Lagrangian $F(x, y, y') = \rho g y \sqrt{1 + (y')^2}$.

As we have a constraint for the length of the string, we can utilize a Lagrange multiplier $\lambda$ from multivariable calculus. We define a new functional $K$ such that

$$K = u_g + \lambda L$$

.

$$K = \int_{-a}^{a} \rho g y \sqrt{1 + (y')^2} dx + \lambda \int_{-a}^{a} \sqrt{1 + (y')^2} dx$$

$$K = \int_{-a}^{a} \left[ \rho g y \sqrt{1 + (y')^2} + \lambda \sqrt{1 + (y')^2} \right] dx$$

Our new Lagrangian is $G = \rho g y \sqrt{1 + (y')^2} + \lambda \sqrt{1 + (y')^2}$ that we use to optimise the functional $K$ in order to solve this problem.

As $G$ does not depend explicitly on $x$, the Beltrami identity can be used.

$$G - y' \frac{\partial G}{\partial y'} = C$$

We have $\frac{\partial}{\partial y'} \sqrt{1 + (y')^2} = \frac{y'}{\sqrt{1+(y')^2}}$, so plugging into Beltrami gives the following.

$$\rho g y \sqrt{1 + (y')^2} + \lambda \sqrt{1 + (y')^2} - y' \left[ \rho g y \frac{y'}{\sqrt{1 + (y')^2}} + \frac{\lambda y'}{\sqrt{1 + (y')^2}} \right] = C$$

Now we proceed by simplifying the equation.

$$\rho g y\sqrt{1+(y')^2}+\lambda\sqrt{1+(y')^2}-\rho g y\frac{(y')^2}{\sqrt{1+(y')^2}}-\frac{\lambda(y')^2}{\sqrt{1+(y')^2}}=C$$

$$\frac{\rho g y(1+(y')^2)+\lambda(1+(y')^2)-\rho g y(y')^2-\lambda(y')^2}{\sqrt{1+(y')^2}}=C$$

$$\frac{\rho g y+\lambda}{\sqrt{1+(y')^2}}=C$$

Squaring both sides, we obtain

$$\frac{(\rho g y+\lambda)^2}{1+(y')^2}=C^2.$$

Then, we attempt to isolate $y'$, giving

$$(\rho g y+\lambda)^2=C^2(1+(y')^2)$$

$$(y')^2=\frac{(\rho g y+\lambda)^2-C^2}{C^2}$$

$$\frac{dy}{dx}=y'=\sqrt{\frac{(\rho g y+\lambda)^2-C^2}{C^2}}=\sqrt{\frac{(\rho g y+\lambda)^2}{C^2}-1}.$$

We can then integrate by separating variables. $C_1$ is an arbitrary constant.

$$dx=\frac{1}{\sqrt{\frac{(\rho g y+\lambda)^2}{C^2}-1}}dy$$

$$x+C_1=\int\frac{1}{\sqrt{\frac{(\rho g y+\lambda)^2}{C^2}-1}}dy$$

To solve the right hand side integral, we need to use u-substitution where $\cosh u=\frac{\rho g y+\lambda}{C}$. If we take the derivative with respect to y on both sides, we obtain

$$\sinh u\frac{du}{dy}=\frac{\rho g}{C}.$$

Therefore

$$dy=\frac{C\sinh u}{\rho g}du.$$

Plugging this back into the integral, we obtain

$$x+C_1=\int\frac{\frac{C\sinh u}{\rho g}}{\sqrt{\cosh^2 u-1}}du$$

As we know that $\cosh^2 u - \sinh^2 u = 1$, $\cosh^2 u - 1 = \sinh^2 u$, so our integral simplifies even further.

$$x + C_1 = \int \frac{\frac{C \sinh u}{\rho g}}{\sinh u} du = \int \frac{C}{\rho g} du$$

Getting rid of the integral, we have

$$x + C_1 = \frac{C}{\rho g} u.$$

Since we set $\cosh u = \frac{\rho g y + \lambda}{C}$, $u = \cosh^{-1}\left[\frac{\rho g y + \lambda}{C}\right]$

$$x + C_1 = \frac{C}{\rho g} \cosh^{-1}\left[\frac{\rho g y + \lambda}{C}\right]$$

Now we need to solve for $y$ to get our function.

$$\frac{\rho g(x + C_1)}{C} = \cosh^{-1}\left[\frac{\rho g y + \lambda}{C}\right]$$

$$\cosh\left(\frac{\rho g(x + C_1)}{C}\right) = \frac{\rho g y + \lambda}{C}$$

$$y = \frac{C}{\rho g} \cosh\left[\frac{\rho g(x + C_1)}{C}\right] - \frac{\lambda}{\rho g}$$

There are still three unknown variables in the above equation($C, C_1, \lambda$). We can use the boundary conditions and the constraint introduced at the start of this section to solve for them. To reiterate, the boundary points are $(-a, 0)$ and $(a, 0)$.

From the condition $y = 0$ when $x = -a$, we have

$$\frac{\lambda}{\rho g} = \frac{C}{\rho g} \cosh\left[\frac{\rho g(-a + C_1)}{C}\right].$$

From the condition $y = 0$ when $x = a$,we have

$$\frac{\lambda}{\rho g} = \frac{C}{\rho g} \cosh\left[\frac{\rho g(a + C_1)}{C}\right].$$

There also exists an identity $\cosh x = \cosh(-x)$ and as the right hand side terms from each of the above two expressions are equal to $\frac{\lambda}{\rho g}$, we get $C_1 = 0$, as this is the only way the identity mentioned is followed.

Taking the second boundary condition equation and simplifying it by multiplying both sides by $\rho g$

$$\lambda = C \cosh\left[\frac{\rho g a}{C}\right]$$

If we plug this back into the original equation for $y$, we get

$$y = \frac{C}{\rho g} \cosh\left[\frac{\rho g x}{C}\right] - \frac{C}{\rho g} \cosh\left[\frac{\rho g a}{C}\right] = \frac{C}{\rho g}\left(\cosh\left[\frac{\rho g x}{C}\right] - \cosh\left[\frac{\rho g a}{C}\right]\right)$$

Now that we have eliminated $\lambda$ and $C_1$ from the final function, all that is left is to get ride of the constant $C$. We have our constraint that we derived at the start of the problem as $L = \int_{-a}^{a} \sqrt{1+(y')^2}dx$. If we take the derivative of the equation that we already have for $y$, we can plug that into the constraint equation to solve for $C$. We have the following derivate using the chain rule:

$$\frac{dy}{dx} = \frac{C}{\rho g}\frac{\rho g}{C} \sinh\left(\frac{\rho g x}{C}\right) = \sinh\left(\frac{\rho g x}{C}\right)$$

Plugging this into the constraint equation, we obtain

$$L = \int_{-a}^{a} \sqrt{1 + \sinh^2\left(\frac{\rho g x}{C}\right)}\, dx$$

Simplifying gives us

$$L = \int_{-a}^{a} \sqrt{\cosh^2\left(\frac{\rho g x}{C}\right)}\, dx = \int_{-a}^{a} \cosh\left(\frac{\rho g x}{C}\right)\, dx$$

$$L = \left[\frac{C}{\rho g} \sinh\left(\frac{\rho g x}{C}\right)\right]_{-a}^{a}$$

$$L = \frac{C}{\rho g}\left[\sinh\left(\frac{\rho g a}{C}\right) - \sinh\left(-\frac{\rho g a}{C}\right)\right]$$

Similar to the $\cosh x = \cosh(-x)$ identity mentioned earlier, there is also a $\sinh x = \sinh(-x)$ identity that we can use here that gives the following equation.

$$L = \frac{2C}{\rho g} \sinh\left(\frac{\rho g a}{C}\right)$$

Therefore the solution to the catenary problem is

$$\boxed{y = \frac{C}{\rho g}\left(\cosh\left[\frac{\rho g x}{C}\right] - \cosh\left[\frac{\rho g a}{C}\right]\right)}$$

where $C$ is found using the equation

$$\boxed{L = \frac{2C}{\rho g} \sinh\left(\frac{\rho g a}{C}\right)}$$

## References

[1] Charles Fox. *An Introduction to the Calculus of Variations*. Oxford University Press, Oxford, 1950.

[2] Israil M. Gelfand and Sergei V. Fomin. *Calculus of Variations*. Prentice-Hall, 1963.

[3] Herbert Goldstein, Charles Poole, and John Safko. *Classical Mechanics*. Addison-Wesley, 3rd edition, 2002.

[4] John McCuan. The hanging chain. Lecture notes, Georgia Institute of Technology, 2025.