The Brachistochrone Curve: The Problem of Quickest Descent

Yutaka Nishiyama

Abstract

This article presents the problem of quickest descent, or the Brachistochrone curve, that may be solved by the calculus of variations and the Euler-Lagrange equation. The cycloid is the quickest curve and also has the property of isochronism by which Huygens improved on Galileo's pendulum.

Keywords : Brachistochrone curve, Law of energy conservation, Calculus of variations, Euler-Lagrange equation, Cycloid, Isochronism, Huygens's pendulum

1. Which is the quickest path?

Suppose there is an incline such as that shown in Figure 1. When a ball rolls from A to B, which curve yields the shortest duration? Let's assume that we have three hypotheses: a straight line, a quadratic, and a cycloid. The shortest path from A to B is the straight line, so one might think that the straight path is the fastest, but in fact it is surprisingly slow. It's better to select a path which has a downward drop in order to accelerate the ball in the first phase, so that it rolls quickly. The ball arrives earlier on the quadratic path than on the straight line path. However, increasing the degree of the function causes the ball to travel more slowly on the flat section.

It is said that Galileo (1564–1642) first presented this problem. It is also known that the cycloid is the curve which yields the quickest descent. This time I will discuss this problem, which may be handled under the field known as the calculus of variations, or variational calculus in physics, and introduce the charming nature of cycloid curves.

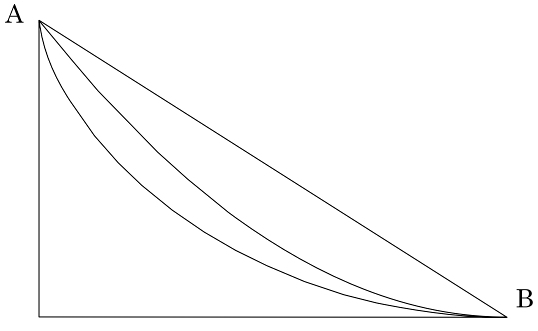

Figure 1. Which path yields the shortest duration?

2. Model construction and numerical computation

Before obtaining the form of the curve analytically, let's try some numerical calculations in order to gain a rough understanding of the problem. I calculated the arrival time for several different curves on a computer using some spreadsheet software. Using the coordinate frame shown in Figure 2, the ball was assumed to roll from a point at a height of y_0. The only force acting on the ball is the force due to gravity, mg, and from the law of energy conservation, the sum of the potential energy and the kinetic energy is constant, so when the ball is at a height of y, the speed v may be obtained as follows.

$$mgy_0 = \frac{1}{2}mv^2 + mgy \tag{2.1}$$

$$v = \sqrt{2g(y_0 - y)} \tag{2.2}$$

When the shape of the curve is fixed, the infinitesimal distance ds may be found, and dividing this by the velocity v yields the infinitesimal duration dt. If an infinitesimal duration such as this dt is integrated, the result is the time until arrival.

$$ds = \sqrt{dx^2 + dy^2} \tag{2.3}$$

$$dt = \frac{ds}{v} \tag{2.4}$$

$$T = \int dt \tag{2.5}$$

A piecewise curve with 100 divisions, a height of 2 meters and a width of π (=3.14) meters was used to produce numerical data. The results revealed an arrival time of 1.189s for the straight line, 1.046s for the quadratic, 1.019s for a cubic curve, 1.007s for an ellipse and 1.003s for the cycloid. The straight line was the slowest, and the curved line was the quickest. The difference between the ellipse and the cycloid was slight, being only 0.004s.

The arrival times were confirmed with a computer, but this lacks a sense of reality, which made me want to build an actual model. I wanted to make a large model, but considering the cost of construction and storage I considered a cut-down model. I found some plywood with horizontal and vertical dimensions of 30×45 cm in a DIY store. The block was 1.2 cm thick,

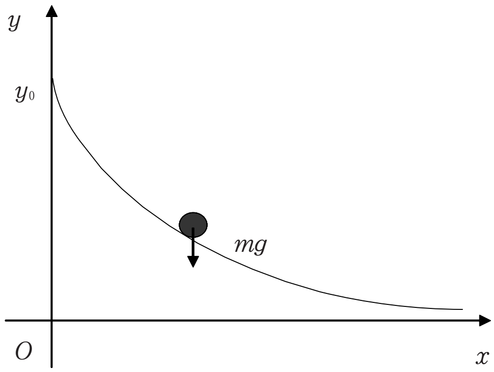

Figure 2. Gravity is the only external force

and since *pachinko* balls (used in the popular Japanese version of pinball) are 1.1 cm in diameter, this was sufficient. On the computer, the time for the straight line was 0.445s, for the quadratic it was 0.391s, and for the cycloid it was 0.375s. Performing these experiments in reality, the difference between the cycloid and the straight line was clear, but the difference between the cycloid and the quadratic required appropriate caution. The difference in arrival times corresponded to those for a single *pachinko* ball.

3. The calculus of variations and functional integrals

The numerical calculations above were made after the form of the curve was known, but let's think about the problem of minimizing the arrival time in the case that the form of the curve is not known. Let us denote the starting and final locations by A and B respectively. The integral from the time-step when the ball is at the starting location, t_A, and the time-step of arrival, t_B, is the duration of motion, T.

$$T=\int_{t_A}^{t_B} dt \tag{3.1}$$

Let us investigate the expression of this dt using x, y and y'. If the infinitesimal element is taken as ds, then the following relationship may be established by Pythagoras' theorem.

$$ds=\sqrt{dx^2+dy^2}=\sqrt{1+\left(\frac{dy}{dx}\right)^2}\,dx \tag{3.2}$$

The speed of the ball, v, may be found by taking the time derivative of the distance along the curve. This may be written as follows.

$$v=\frac{ds}{dt}=\frac{ds}{dx}\frac{dx}{dt}=\sqrt{1+\left(\frac{dy}{dx}\right)^2}\,\frac{dx}{dt} \tag{3.3}$$

Equation (3.1) may be rewritten as follows, using Equations (3.2) and (3.3).

$$T=\int_{t_A}^{t_B} dt=\int_{x_A}^{x_B}\frac{dt}{ds}\frac{ds}{dx}dx=\int_0^x \frac{\sqrt{1+\left(\frac{dy}{dx}\right)^2}}{v}dx \tag{3.4}$$

If the y coordinate is taken as being in the downwards direction, then the distance fallen, y, and the speed, v, must obey the principle of energy conservation so the equation

$$\frac{1}{2}mv^2=mgy \tag{3.5}$$

is satisfied. Rearranging yields,

$$v=\sqrt{2gy}\,. \tag{3.6}$$

Substituting this into Equation (3.4), and writing y' for $\frac{dy}{dx}$ yields the following equation.

$$T=\int_0^x \sqrt{\frac{1+\left(\frac{dy}{dx}\right)^2}{2gy}}\,dx=\int_0^x\sqrt{\frac{1+y'^2}{2gy}}\,dx \tag{3.7}$$

How should one go about minimizing T according to this equation? The selection of the integrand

$$L(x, y, y')=\sqrt{\frac{1+y'^2}{2gy}}, \tag{3.8}$$

in order to minimize T is a problem in the calculus of variations.

4. The Euler-Lagrange equation

Now, given a function $L(x, y, y')$, let's think about the problem of finding the extremal value of the integral,

$$I=\int_{x_1}^{x_2} L(x, y, y')\,dx, \tag{4.1}$$

by setting the function $y(x)$. I is called a functional. This expresses the meaning that, in comparison to a normal function, it is a "function of a function". Suppose we have a function which is slightly offset from $y(x)$, the function we are seeking,

$$Y(x)=y(x)+\varepsilon\delta(x). \tag{4.2}$$

Its integral is

$$I(\varepsilon)=\int_{x_1}^{x_2} L(x, Y, Y')\,dx. \tag{4.3}$$

Consider the condition according to which it takes its extreme value,

$$\left.\frac{dI(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=0}=0. \tag{4.4}$$

Y and Y' both depend on ε. Bearing in mind that $\frac{dY}{d\varepsilon}=\delta(x)$ and $\frac{dY'}{d\varepsilon}=\delta'(x)$, and taking the derivative yields

$$\left.\frac{dI(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=0}=\int_{x_1}^{x_2}\left.\left(\frac{\partial L}{\partial Y}\delta(x)+\frac{\partial L}{\partial Y'}\delta'(x)\right)\right|_{\varepsilon=0}dx. \tag{4.5}$$

Integrating the second term on the right hand side by parts yields

$$\int_{x_1}^{x_2}\frac{\partial L}{\partial Y'}\delta'(x)\,dx=\left.\frac{\partial L}{\partial Y'}\delta(x)\right|_{x_1}^{x_2}-\int_{x_1}^{x_2}\frac{d}{dx}\left(\frac{\partial L}{\partial Y'}\right)\delta(x)\,dx. \tag{4.6}$$

For x_1 and x_2, $\delta(x)=0$ so the first term on the right hand side of this equation is zero. The condition for the extremal value thus becomes

$$\left.\frac{dI(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=0}=\int_{x_1}^{x_2}\left(\frac{\partial L}{\partial y}-\frac{d}{dx}\frac{\partial L}{\partial y'}\right)\delta(x)\,dx=0. \tag{4.7}$$

Since $\delta(x)$ is arbitrary, the following condition determining $y(x)$ may be obtained.

$$\frac{\partial L}{\partial y}-\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)=0 \tag{4.8}$$

This formula is known as Euler's equation, or alternatively the Euler-Lagrange equation. The calculus of variations originates in Fermat's principle which expresses how the path of a beam of light varies as it passes through media with different refractive indices. This operates according to the principle that the path is selected in order to minimize the passage time.

5. Solving Euler's equation

Now then, $L=\sqrt{\frac{1+y'^2}{2gy}}$ may be substituted into the Euler-Lagrange equation $\frac{\partial L}{\partial y}-\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)=0$, but x is not explicitly contained in L, *i.e.*, it is a function of y and y' alone. The following transformation of Euler's equation may therefore be used.

$$L-y'\left(\frac{\partial L}{\partial y'}\right)=C \tag{5.1}$$

Substituting for L in this equation,

$$\sqrt{\frac{1+y'^2}{2gy}}-y'\left(\frac{y'}{\sqrt{2gy(1+y'^2)}}\right)=\frac{1}{\sqrt{2gy(1+y'^2)}}=C \tag{5.2}$$

Squaring both sides, the equation may be rearranged as follows. Since the right hand side is constant, so we may write it as $2A$.

$$y(1+y'^2)=\frac{1}{2gC^2}=2A \tag{5.3}$$

Equation (5.3) is rearranged as follows.

$$y'=\sqrt{\frac{2A-y}{y}} \tag{5.4}$$

The domain of the curve is taken as

$$2A\geq y\geq 0. \tag{5.5}$$

The initial condition is taken as $y=0$ when $\theta=0$. At this point y may be rewritten as the following parametrical expression using a change of variable.

$$y=A-A\cos\theta \tag{5.6}$$

This change of variable (5.6) may seem somewhat sudden. Rather than determining the nature of the function according to the calculus of variations, in this case it was already known that the cycloid is the curve of quickest descent because research on cycloids has been developing for a considerable length of time. It is sufficient to understand that this curve was taken as a hypothesis and the solution was obtained using the calculus of variations.

If both sides of Equation (5.6) are differentiated then the result is as follows.

$$dy=A\sin\theta d\theta=2A\cos\frac{\theta}{2}\sin\frac{\theta}{2}d\theta \tag{5.7}$$

It is possible rewrite Equation (5.7) using the parametric expression for y as follows.

$$y'=\sqrt{\frac{2A-y}{y}}=\sqrt{\frac{A+A\cos\theta}{A-A\cos\theta}}=\sqrt{\frac{\cos^2\frac{\theta}{2}}{\sin^2\frac{\theta}{2}}}=\frac{\cos\frac{\theta}{2}}{\sin\frac{\theta}{2}} \tag{5.8}$$

This can be written as follows, by multiplying both sides by dx.

$$dy=\frac{\cos\frac{\theta}{2}}{\sin\frac{\theta}{2}}dx \tag{5.9}$$

The relationship between x and θ may be found by collating Equations (5.7) and (5.9), and eliminating dy.

$$dx=2A\sin^2\frac{\theta}{2}d\theta=A(1-\cos\theta)d\theta \tag{5.10}$$

Integrating both sides,

$$x=A(\theta-\sin\theta)+D. \tag{5.11}$$

For the initial condition $x=0$ when $\theta=0$, the constant of integration is $D=0$. Finally, the parametric expression for the curve of quickest descent is as follows.

$$x=A(\theta-\sin\theta) \tag{5.12}$$

$$y=A(1-\cos\theta) \tag{5.13}$$

6. The cycloid

A cycloid may be expressed as the trajectory of a point fixed on the circumference of a circle with radius a, when the circle rolls along a straight line. When the angle of the circle's rotation is θ, the coordinates on the curve are as follows.

$$x=a(\theta-\sin\theta) \tag{6.1}$$

$$y=a(1-\cos\theta) \tag{6.2}$$

The derivatives are $\frac{dx}{d\theta}=a(1-\cos\theta)$ and $\frac{dy}{d\theta}=a\sin\theta$. Bearing in mind that $\cos\theta=1-\frac{y}{a}$, we may write

$$\left(\frac{dy}{dx}\right)^2=\frac{(dy/d\theta)^2}{(dx/d\theta)^2}=\frac{2a-y}{y}. \tag{6.3}$$

This is the differential equation of the cycloid, and it should be noted that it is equivalent to the previously stated Equation (5.4). It is common for the explanation of the cycloid given in high-school mathematics textbooks to state no more than that it is the trajectory of a point on a bicycle wheel. It's a shame that the exceptional property that it is the curve of quickest descent is rarely explained.

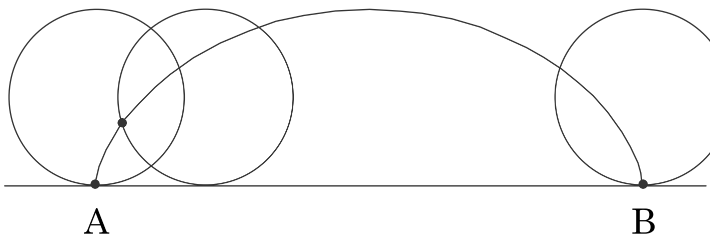

Figure 3. The cycloid

7. Isochronism

It was Galileo who discovered the isochrosism of pendulums. Supposing the length of a given pendulum is l, the gravitational constant is g, and the equilibrium point of the pendulum is at the origin, then when the swing angle of the pendulum is θ, the equation of the pendulum's motion is as follows.

$$ml\frac{d^2\theta}{dt^2}=-mg\sin\theta \tag{7.1}$$

Hypothesizing that when the angle of swing is small, $\sin\theta\approx\theta$, the equation may be simplified as follows.

$$l\frac{d^2\theta}{dt^2}=-g\theta \tag{7.2}$$

The period is

$$T=2\pi\sqrt{\frac{l}{g}}. \tag{7.3}$$

I explained that the cycloid was the curve of quickest descent, but it has one more exceptional property, it is isochronic. Whether a ball is rolled from the point A shown in Figure 4, or from the intermediate point C, the time taken to arrive at point B is the same.

Question 1

Prove that for balls placed on a cycloid curve, even if their positions differ, the time taken to reach the lowest point is the same.

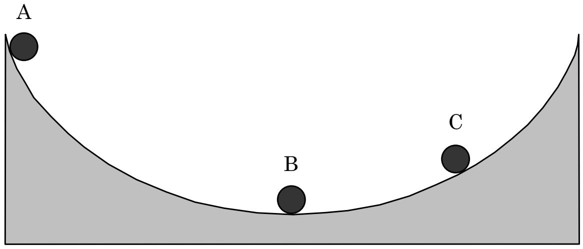

Figure 4. The isochonism of cycloids

When the swing of Galileo's pendulum grows large, the isochronism breaks down and the cycle time grows longer. If the pendulum moves back and forth on a cycloid, isochronism should be satisfied. Huygens (1629-1695) implemented an isochronic pendulum as follows. Two cycloids are described ($0\leq\theta\leq4\pi$), as shown in Figure 5. When a rope of length $4a$ (the length of the cycloid curve is $8a$) with point B at its center is pulled from point D, the trajectory P forms a cycloid. This kind of curve is known as an involution.

Question 2

Prove that the cycloid's involution is indeed a cycloid.

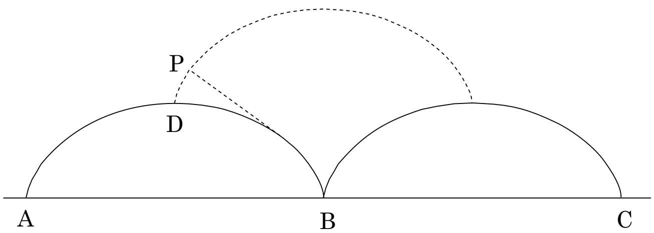

Figure 5. The cycloid's involution is a cycloid

The cycloid pendulum devised by Huygens is the same as Figure 5 flipped vertically with the central half removed (Figure 6). With this pendulum, even when the swing is large, isochronism is maintained. Clocks employing this principle are more accurate than Galileo clocks. The length of the pendulum l, is half the length of the pendulum's swing cycle, $4a$. The period, T is

$$T = 2\pi\sqrt{\frac{4a}{g}}. \tag{7.4}$$

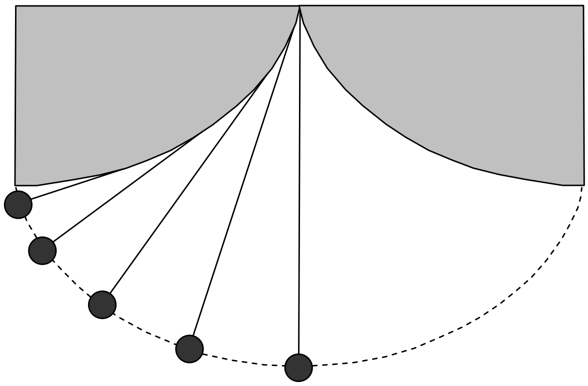

Figure 6. Huygens's pendulum

Nowadays actual models of the Brachistochrone curve can be seen only in science museums. But we should not forget that the problem of quickest descent mathematically developed the study of the cycloid and the calculus of variations, and contributed to the improvement of pendulums.

Reference

Takakuwa, S. (2003). *Bibun Hoteishiki to Henbunho* [Differential equations and the calculus of variations], Tokyo: Kyoritsu.