The Optimal Curve

Adrian Garcia

June 15, 2022

Abstract

Throughout history, mathematicians continuously sought to challenge themselves with a variety of optimization problems with one such being the Brachistochrone problem or "shortest time" problem. Introduced in 1696 by mathematician Johann Bernoulli, this paper aims to answer such a problem using an analytical approach. To this end, the problem is reconstructed in the Cartesian coordinate system, and thus reintroduced as an optimization of a functional. Analytical results on examples including the optimal curve between $(0, 1)$ and $(1, 0)$, and the optimal curve's travel time compared to elementary functions such as a straight line, circle, and parabola suggests that the optimal curve we derived is, in fact, the solution to Bernoulli's problem. Additionally, analysis on how the results vary with the respective positions of the points A and B opens up possible use in practical applications.

1 Introduction

It is human nature to search for the optimal path in almost every aspect of life and even more so for mathematicians. As such, throughout history, mathematicians continuously sought to challenge themselves with a variety of optimization problems. One such problem being the Brachistochrone problem which can be seen in both natural and artificial applications. The Brachistochrone problem, introduced in 1696 by mathematician Johann Bernoulli, was first posed as follows, "Given two points A and B in a vertical plane, what is the curve traced out by a point acted on only by gravity, which starts at A and reaches B in the shortest time?" [3] and although simple on the surface, its solution is anything but.

In this paper, we will be answering Bernoulli's question and analyzing special cases; specifically, we will be analyzing when an object starts at a sufficiently small or large starting point A relative to an ending point B and determining how the optimal curve derived will behave.

To answer this question, in section 2 we will be providing a basis for the model in which we derive an integral expression for the travel time from points A to B as a function of the curve shape (2.1.1), initiate optimization (2.1.2), introduce the Beltrami Identity (2.2.1) and derive the optimal curve (2.2.2). In section 3 we will be answering Johann Bernoulli's original Brachistochrone problem in which we will display the optimal curve we derive (3.1.1), compare resulting travel times to other elementary functions (3.1.2) as well as study how the results vary with the respective positions of the points A and B (3.2). Finally, in section 4 we will summarize our findings, highlight caveats, and propose potential future work.

2 Model & Method

2.1 Model

2.1.1 Model Setup

For simplicity, we define point A as $(0, y_A)$ and point B as $(x_B, 0)$ such that both values y_A and x_B are positive and nonzero on the Cartesian coordinate system. Now, consider an arbitrary object of mass m sliding down path $y(x)$ without friction and only under the influence of gravity $g \approx 9.8\ \text{m/s}^2$.

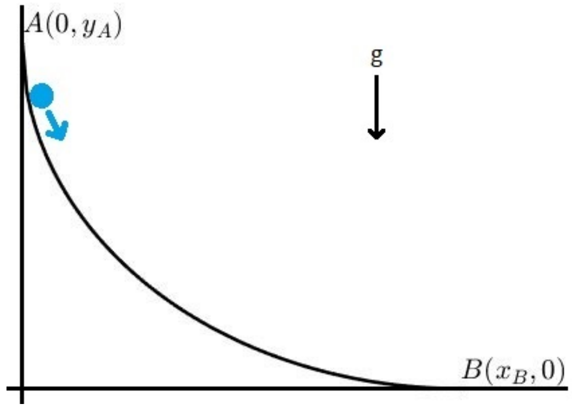

Figure 1: An object of mass m sliding down path $y(x)$ from point $A\,(0, y_A)$ to point $B\,(x_B, 0)$ only under the influence of gravity g.

In order to find the optimal curve $y(x)$ that produces the fastest travel time between points A and B, we must then look into minimizing the time integral

$$T = \int_0^T dt = \int_A^B \frac{dS}{v},$$

where T is the total travel time, dS is the infinitesimal arc length of the curve $y(x)$, and v is the velocity of the object at any given point. As the object is only under the influence of gravity, we may then assume the object's energy is conserved; that is, the sum of the potential and kinetic energy is the same at any point along the trajectory of the object. Hence

$$\begin{aligned} mgy_A &= mgy(x) + \frac{1}{2}mv^2, \\ v &= \sqrt{2g(y_A - y(x))}. \end{aligned}$$

Additionally, from the Pythagorean theorem, we know

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

As such, the optimal curve $y(x)$ that produces the fastest travel time between points A and B is then the function that minimizes the functional

$$T(y(x)) = \int_0^{x_B} \sqrt{\frac{1 + \left(\frac{dy}{dx}\right)^2}{2g(y_A - y(x))}}\, dx. \tag{1}$$

2.1.2 Optimization

Now, to minimize the functional $T(y(x))$, we must then consider the following. Let $y(x)$ be the optimal curve and define a function $\epsilon(x)$ such that $\epsilon(x) << y(x)$ and $\epsilon(x) = 0$ at both boundaries 0 and x_B. Then, by definition, as we are in the vicinity of the optimal curve $y(x)$, variations of $T(y(x))$ must be proportional to $\epsilon(x)^2$; hence

$$T(y(x) + \epsilon(x)) - T(y(x)) \propto \epsilon^2(x).$$

Now consider the following Taylor expansion of $T(y(x)+\epsilon(x))$ near $\epsilon(x)=0$

$$\begin{aligned}T(y+\epsilon)-T(y) &= \int_0^{x_B}\sqrt{\frac{1+[y'+\epsilon']^2}{2g(y_A-(y+\epsilon))}}\,dx-\int_0^{x_B}\sqrt{\frac{1+[y']^2}{2g(y_A-y)}}\,dx,\\ &= \sqrt{\frac{1}{2g}}\int_0^{x_B}\frac{\sqrt{1+[y']^2}\epsilon}{2(y_A-y)^{\frac{3}{2}}}\,dx+\int_0^{x_B}\frac{y'\epsilon'}{\sqrt{(y_A-y)(1+[y']^2)}}\,dx+O(\epsilon^2).\end{aligned}$$

In order to have $T(y(x)+\epsilon(x))-T(y(x))\propto\epsilon^2(x)$, the terms that are linear in $\epsilon(x)$ must be 0; thus

$$\sqrt{\frac{1}{2g}}\int_0^{x_B}\frac{\sqrt{1+[y']^2}\epsilon}{2(y_A-y)^{\frac{3}{2}}}\,dx+\int_0^{x_B}\frac{y'\epsilon'}{\sqrt{(y_A-y)(1+[y']^2)}}\,dx=0.$$

Now, using integration by parts on the second term of the integral we have

$$\int_0^{x_B}\frac{y'\epsilon'}{\sqrt{(y_A-y)(1+[y']^2)}}\,dx=\left[\frac{y'}{\sqrt{(y_A-y)(1+[y']^2)}}\epsilon\right]_0^{x_B}-\int_0^{x_B}\epsilon\frac{d}{dx}\left[\frac{y'}{\sqrt{(y_A-y)(1+[y']^2)}}\right]dx.$$

Since $\epsilon(x)$ is defined to be 0 at both boundaries 0 and x_B

$$\sqrt{\frac{1}{2g}}\int_0^{x_B}\epsilon\left[\frac{\sqrt{1+[y']^2}}{2(y_A-y)^{\frac{3}{2}}}-\frac{d}{dx}\left[\frac{y'}{\sqrt{(y_A-y)(1+[y']^2)}}\right]\right]dx=0,$$

and as this must be true for any function $\epsilon(x)$, the following must always be true

$$\frac{\sqrt{1+[y']^2}}{2(y_A-y)^{\frac{3}{2}}}=\frac{d}{dx}\left[\frac{y'}{\sqrt{(y_A-y)(1+[y']^2)}}\right]. \tag{2}$$

Therefore, to find the optimal curve $y(x)$ that produces the fastest travel time between points A and B we must solve the above ODE subject to the boundary conditions $y(0)=y_A$ and $y(x_B)=0$.

2.2 Method

2.2.1 The Beltrami Identity

Following from the previous section, we note equation (1) equates to a functional of the form

$$T(y(x))=\int_0^{x_B}F(x,y(x),y'(x)),$$

of which, does not explicitly have the independent variable x; hence, we may then use the Beltrami identity which states that if $\frac{\partial F}{\partial x}=0$, then equation (2) equates to the Beltrami identity

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

where C is a constant [4]. Thus

$$\sqrt{\frac{1+[y']^2}{2g(y_A-y)}}-\frac{[y']^2}{\sqrt{2g(y_A-y)(1+[y']^2)}}=C.$$

2.2.2 Optimal Curve Derivation

Subsequently, the equation above simplifies to

$$(y_A - y)(1 + [y']^2) = C_1, \qquad \left(C_1 = \frac{1}{2gC^2}\right)$$
$$\sqrt{\frac{y_A - y}{C_1 - (y_A - y)}}\, dy = dx.$$

This can then be integrated formally using Khan's choice of trigonometric substitutions [1]

$$y(\theta) = y_A - C_1 \sin^2\left(\frac{\theta}{2}\right),$$
$$dy = -C_1 \sin\left(\frac{\theta}{2}\right) \cos\left(\frac{\theta}{2}\right)\, d\theta,$$

which implies $x = \int \sqrt{\frac{y_A - y}{C_1 - (y_A - y)}}\, dy$ equates to

$$x(\theta) = \int \sqrt{\frac{C_1 \sin^2\left(\frac{\theta}{2}\right)}{C_1 - C_1 \sin^2\left(\frac{\theta}{2}\right)}} - C_1 \sin\left(\frac{\theta}{2}\right) \cos\left(\frac{\theta}{2}\right)\, d\theta,$$
$$= -\frac{C_1}{2}(\theta - \sin\theta) + C_2,$$

where C_2 is a constant. We then have the following parametric equations of the optimal curve

$$x(\theta) = -\frac{C_1}{2}(\theta - \sin\theta) + C_2,$$
$$y(\theta) = y_A - \frac{C_1}{2}(1 - \cos\theta).$$

Now, without loss of generality, in choosing θ_A to be 0 at point A and applying such boundary condition we have

$$\theta_A = 0 \implies x(0) = -\frac{C_1}{2}(0) + C_2 = 0 \implies C_2 = 0.$$

Therefore, the optimal curve that we require is of the form

$$x(\theta) = -\frac{C_1}{2}(\theta - \sin\theta), \tag{3}$$
$$y(\theta) = y_A - \frac{C_1}{2}(1 - \cos\theta). \tag{4}$$

Then, choosing θ_B to denote the θ value at point B, we have a total travel time of

$$T = \sqrt{\frac{1}{2g}} \int_{\theta_A = 0}^{\theta_B} \sqrt{\frac{1 + \left(\frac{dy/d\theta}{dx/d\theta}\right)^2}{(y_A - y)}}\, \frac{dx}{d\theta}\, d\theta,$$
$$= \sqrt{\frac{1}{2g}} \int_0^{\theta_B} \sqrt{\frac{1 + \frac{\sin^2\theta}{(1-\cos\theta)^2}}{\frac{C_1}{2}(1 - \cos\theta)}}\, \frac{-C_1}{2}(1 - \cos\theta)\, d\theta,$$
$$T(C_1, \theta_B) = -\sqrt{\frac{C_1}{2g}}\,\theta_B. \tag{5}$$

3 Results

3.1 Bernoulli's Brachistochrone Problem

3.1.1 Total Travel Time

So far, we have dealt with arbitrary y_A and x_B values, however, to answer Bernoulli's Brachistochrone problem we must look into an explicit example; thus, consider the case where point A is $(0, 1)$ and point B is $(1, 0)$. Then, to find C_1 and θ_B values we must evaluate the optimal curve as follows

$$x(\theta_B) = -\frac{C_1}{2}(\theta_B - \sin\theta_B) = 1,$$
$$y(\theta_B) = 1 - \frac{C_1}{2}(1 - \cos\theta_B) = 0,$$

and in using an equation solver for a nonlinear system of equations, such as Mathematica's `FindRoot` function, we see that $C_1 = 1.14583$ and $\theta_B = -2.41201$. Additionally, using equation (5) as well as the aforementioned C_1 and θ_B values, we have a total travel time of $T = 0.583$ seconds. Subsequently, Figure 2 shows the optimal curve for this particular choice of A and B. Note the optimal curve shown here has a relatively steep slope from 0 to 0.1 but then increasingly levels off slightly thereafter.

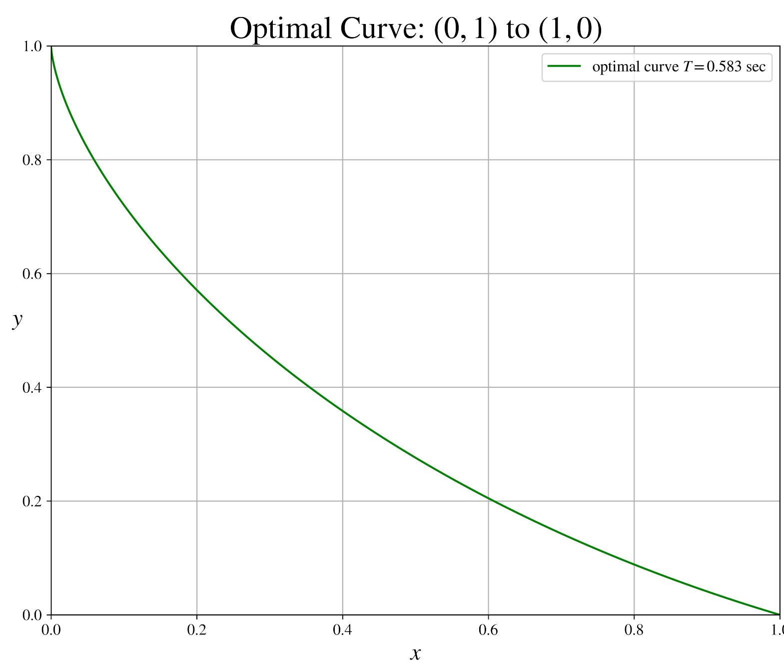

Figure 2: The optimal curve connecting point $A\,(0, 1)$ to point $B\,(1, 0)$ with a total travel time of $T = 0.583$ seconds.

3.1.2 Travel Time Comparison

To check if we indeed found the optimal curve that produces the fastest travel time, we now compare the results to other possible choices of curves between points A and B. Intuitively, one would assume either the path of shortest distance or the path of greater average velocity would yield the best total travel times. Thus, let us consider a straight line and circle equation to represent each path respectively, hence

$$y(x) = -x + 1,$$
$$1 = (x-1)^2 + (y-1)^2.$$

We will also consider a parabola in our comparison to diversify our possible choices of curves, hence

$$y(x) = (x-1)^2.$$

In using equation (1) to analytically solve for the travel time for the possible choices of curves, however, we see an increase in travel time compared to the optimal curve solution with times

Equation	T
$1 = (x-1)^2 + (y-1)^2$	0.592
$y(x) = (x-1)^2$	0.595
$y(x) = -x + 1$	0.639

suggesting the optimal curve was, in fact, derived. As such, Figure 3 shows the optimal curve, circle, parabola and straight line for points $(0, 1)$ and $(1, 0)$. It is significant to highlight in the figure that although the shape of the optimal curve is closer to the parabola's, the optimal curve's travel time is closer to the circle's. Additionally note the equation with the slowest travel time is the straight line despite being the path of shortest distance.

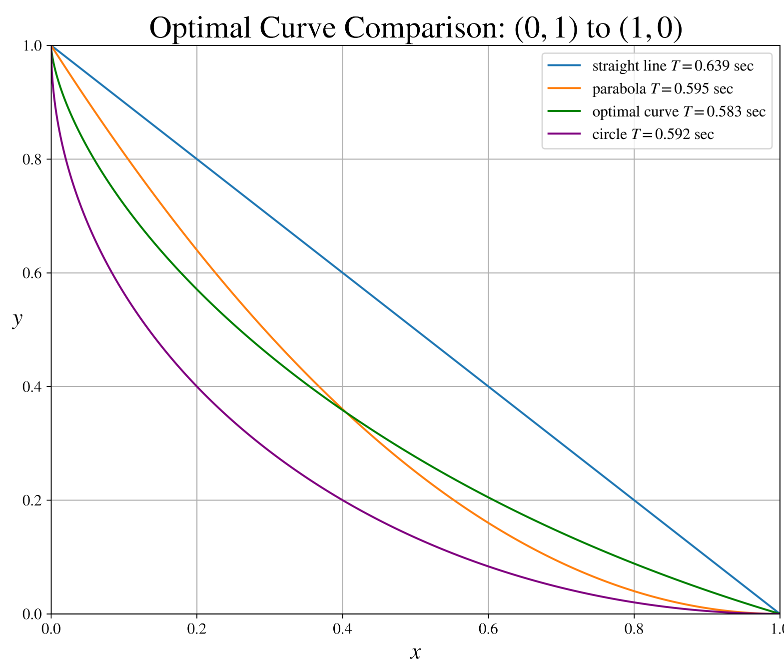

Figure 3: The total travel time of the optimal curve compared to a straight line, parabola, and circle from point $A\,(0, 1)$ to point $B\,(1, 0)$.

3.2 Varying Boundary Conditions

3.2.1 Sufficiently Small

We now consider the case where point A is $(0, 0.1)$ to $(0, 0.4)$ and point B is $(1, 0)$. Without loss of generality, we again use the steps taken in section 3.1.1 to find C_1, θ_B and T values relative to each point A. We then have

point A	C_1	θ_B	T
$(0, 0.1)$	0.331239	-5.11977	0.666
$(0, 0.2)$	0.357932	-4.59459	0.621
$(0, 0.3)$	0.397142	-4.17628	0.594
$(0, 0.4)$	0.449746	-3.81967	0.579

Following this, Figure 4 shows the optimal curve for the varying small starting points A to a fixed point B. Now comparing with Figure 2, note for small choices of starting points, the optimal curves in Figure 4 dip below the x-axis with starting point $(0, 0.1)$ having the largest dip. This suggests that once a starting point is sufficiently small relative to an ending point, the optimal curve must compensate by significantly increasing acceleration.

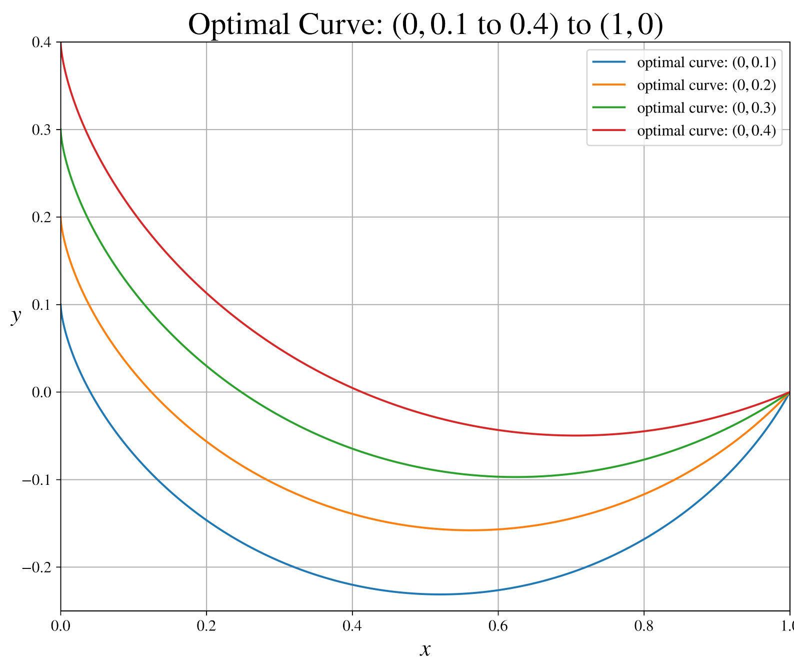

Figure 4: The optimal curve connecting various small point A's to a fixed point $B\,(1, 0)$

3.2.2 Sufficiently Large

For the case where point A is $(0, 7)$ to $(0, 10)$ and point B is $(1, 0)$, we have the following C_1, θ_B and T values relative to each point A

point A	C_1	θ_B	T
$(0, 7)$	156.661	-0.425978	1.204
$(0, 8)$	232.37	-0.373258	1.285
$(0, 9)$	329.413	-0.332108	1.362
$(0, 10)$	450.456	-0.299105	1.434

Subsequently, Figure 5 shows the optimal curve for the varying large starting points A to a fixed point B. Unlike Figure 4, however, Figure 5 stays above the x-axis similar to Figure 2 although with a straighter path. Suggesting a path of shorter distance is optimal in this case.

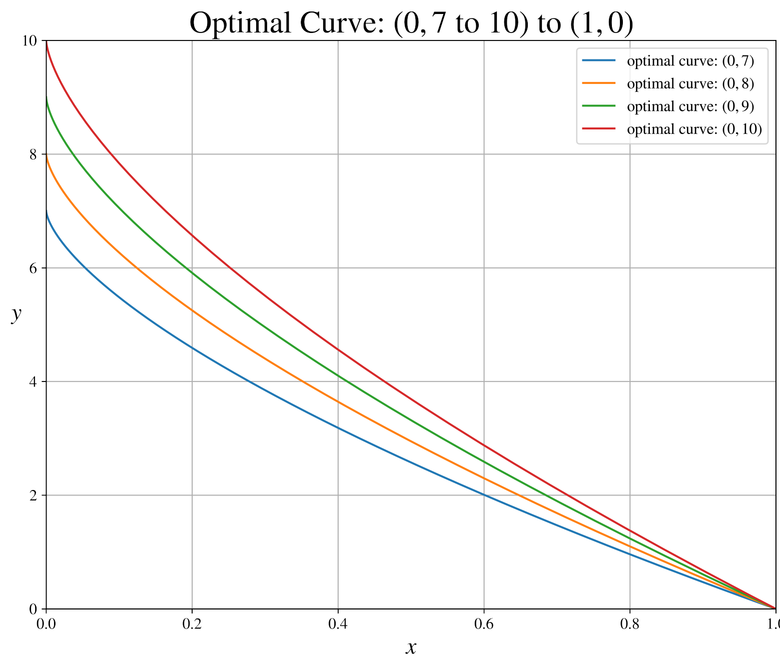

Figure 5: The optimal curve connecting various large point A's to a fixed point $B\,(1, 0)$

4 Discussion

Though lengthy in its derivation, we found that the optimal curve connecting point A to point B can be parametrized by the parametric equations

$$x(\theta) = -\frac{C_1}{2}(\theta - \sin\theta),$$
$$y(\theta) = y_A - \frac{C_1}{2}(1 - \cos\theta).$$

Directing our attention to Figure 5, we see as we increase the starting point A, the optimal curve resembles a straight line, suggesting a path of shorter distance is optimal in this scenario. By contrast, as we decrease the starting point A (Figure 4), the optimal curve resembles a circle that dips below the x-axis. This illustration also suggests that once a starting point is sufficiently small (relative to an ending point), the optimal curve must compensate by significantly increasing acceleration.

Moving away from the results and more towards the overall structure of the model, we see that there is indeed a few caveats. For instance, in regards to using this model for practical applications, we did not consider any sort of friction in the optimal curve derivation. Thus, future work answering how large of an impact friction has on the derivation of the optimal curve is in order. Nonetheless, the results we derived here proves there does exist an optimal path between two points, and analysis on such results opens up possible use in practical applications.

References

[1] Faculty of Khan. *The Brachistochrone Problem and Solution — Calculus of Variations*. URL: https://www.youtube.com/watch?v=zYOAUG8PxyM. (accessed: 02.20.2022).

[2] Erik Neumann. *Brachistochrone*. URL: https://www.myphysicslab.com/roller/brachistochrone-en.html. (accessed: 02.20.2022).

[3] J J O'Connor and E F Robertson. *The brachistochrone problem*. URL: https://mathshistory.st-andrews.ac.uk/HistTopics/Brachistochrone/. (accessed: 02.20.2022).

[4] Eric W. Weisstein. *Beltrami Identity*. URL: https://archive.lib.msu.edu/crcmath/math/math/b/b096.htm. (accessed: 02.20.2022).

[5] Eric W. Weisstein. *Brachistochrone Problem*. URL: https://archive.lib.msu.edu/crcmath/math/math/b/b355.htm. (accessed: 02.20.2022).