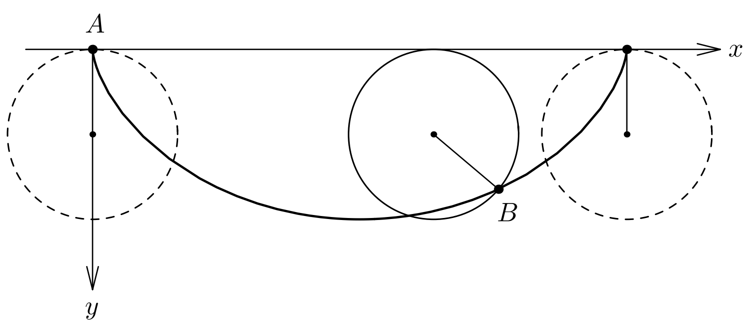
YET ANOTHER ELEMENTARY SOLUTION OF THE BRACHISTOCHRONE PROBLEM

GARY BROOKFIELD

In 1696 Johann Bernoulli issued a famous challenge to his fellow mathematicians:

> Given two points A and B in a vertical plane, find the curve connecting the two points such that an object, starting with zero velocity at A, slides without friction along the curve to B in the least possible time.

Such a curve is called a **brachistochrone**. Newton, Leibniz, l'Hôpital, Jakob Bernoulli (Johann's brother) and the challenger were able to show that a brachistochrone is a segment of a cycloid arc. By a cycloid arc we mean the curve traced out by a point on the rim of a disk as it rolls once along a line. The graph shows the cycloid arc formed by a disk rolling underneath a horizontal line, which is the orientation appropriate for our problem.



Since the object starts with zero velocity at A, this point is at one end of the cycloid arc. With the coordinate system shown, the cycloid arc is given by the equations

$$x = R(\theta - \sin\theta) \quad \text{and} \quad y = R(1 - \cos\theta), \tag{1}$$

with $0 \leq \theta \leq 2\pi$, R being the radius of the disk.

The brachistochrone problem is considered to be the beginning of the calculus of variations [3, 4], and a modern solution [8] would make use of general methods from that branch of mathematics: the Euler, Lagrange and Jacobi tests, the Weierstrass excess function and more. Even so, many solutions which avoid the calculus of variations have been published [1, 6, 2]. The solution we present here amounts to

Date: September 14, 2006.

little more than a change of coordinate systems, and is general enough that we prove that the cycloid arc yields the minimum travel time, not just among curves that are smooth, but also among curves that have loops and corners.

To begin, we set up a cartesian coordinate system for the vertical plane containing A and B as above, with the x-axis horizontal and the positive y-axis pointing down. The coordinates of the sliding object (x, y) are functions of time t on an interval $[0, T]$ such that $A = (0, 0) = (x(0), y(0))$ and $B = (x(T), y(T))$. The number T is, of course, the travel time of the object and the quantity we want to minimize.

Since we assume that there is no friction, the sum of the kinetic energy and the gravitational potential energy, $E = \frac{1}{2}mv^2 - mgy$, is a constant of the motion. Here v is the velocity of the object, m is the mass of the object and $g = 9.8 \text{ m/s}^2$ is the acceleration due to gravity at the Earth's surface. By construction, we have $y = 0$ and $v = 0$ at A, so $E = 0$, and

$$2gy = v^2 \tag{2}$$

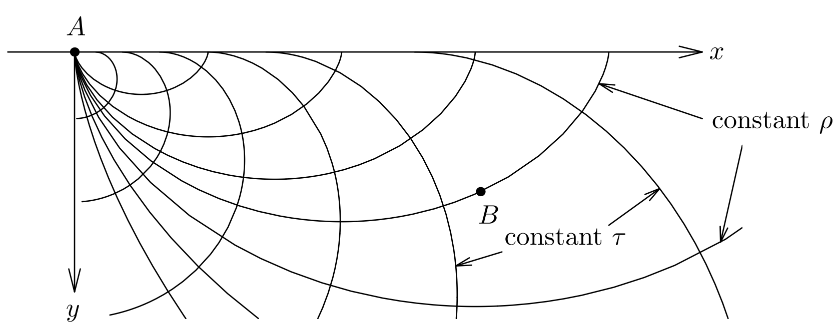
throughout the object's motion.

In what follows, we require that $y > 0$ and $x > 0$ except at A and (possibly) B. It is plausible that any trajectory which minimizes travel time will satisfy these conditions. Indeed, $y \geq 0$ follows from (2).

We now introduce new coordinates ρ and τ which are related to x and y by

$$x = \rho\tau - \rho^2 \sin\frac{\tau}{\rho} \quad \text{and} \quad y = \rho^2\left(1 - \cos\frac{\tau}{\rho}\right) \tag{3}$$

where $0 < \rho$ and $0 \leq \tau \leq 2\pi\rho$. These equations are just (1) with $R = \rho^2$ and $\theta = \tau/\rho$. In particular, for a fixed $\rho > 0$, the curve parametrized by τ is a cycloid arc made by rolling a disk of radius $R = \rho^2$ along the x-axis. The graph shows several of these cycloid arcs, as well as some constant τ curves, and makes plausible the fact, which we prove later, that (3) represents a change of coordinate systems.



We now suppose that all relevant trajectories of the object are given by functions ρ and τ of time on the interval $[0, T]$ that determine the cartesian coordinates (x, y) of the object by (3). Notice that, since $\rho > 0$, the point A has zero τ-coordinate, and so $\tau(0) = 0$. We write $\dot{x}$, $\dot{y}$, $\dot{\tau}$ and $\dot{\rho}$ for the derivatives of these functions with respect to time. Using the chain rule we can express $\dot{x}$ and $\dot{y}$ in terms of $\dot{\tau}$ and $\dot{\rho}$.

$$\dot{x} = \frac{\partial x}{\partial \tau}\,\dot{\tau} + \frac{\partial x}{\partial \rho}\,\dot{\rho} = \left(\rho - \rho\cos\frac{\tau}{\rho}\right)\dot{\tau} + \left(\tau + \tau\cos\frac{\tau}{\rho} - 2\rho\sin\frac{\tau}{\rho}\right)\dot{\rho}$$

$$\dot{y} = \frac{\partial y}{\partial \tau}\,\dot{\tau} + \frac{\partial y}{\partial \rho}\,\dot{\rho} = \left(\rho\sin\frac{\tau}{\rho}\right)\dot{\tau} + \left(2\rho - 2\rho\cos\frac{\tau}{\rho} - \tau\sin\frac{\tau}{\rho}\right)\dot{\rho}$$

With a bit of calculation, (2) can be also be written in terms of $\dot{\tau}$ and $\dot{\rho}$.

$$\begin{aligned} 2gy &= v^2 = \dot{x}^2 + \dot{y}^2 \\ &= 2\rho^2\left(1 - \cos\frac{\tau}{\rho}\right)\dot{\tau}^2 \\ &\qquad + 2\left(4\,\rho^2\left(1 - \cos\frac{\tau}{\rho}\right) - 4\,\rho\,\tau\,\sin\frac{\tau}{\rho} + \tau^2\left(1 + \cos\frac{\tau}{\rho}\right)\right)\dot{\rho}^2 \\ &= 2y\,\dot{\tau}^2 + 4\left(2\rho\,\sin\frac{\tau}{2\rho} - \tau\,\cos\frac{\tau}{2\rho}\right)^2\dot{\rho}^2 \end{aligned} \tag{4}$$

Using this equation it is now easy to solve the brachistochrone problem. The term in $\dot{\rho}^2$ is nonnegative, so $2y\dot{\tau}^2 \le 2gy$ and, since $y > 0$ except at A and (possibly) B, we have $\dot{\tau} \le \sqrt{g}$ except perhaps at $t = 0$ and $t = T$. Integrating this inequality on the interval $[0, T]$ gives

$$\tau(T) = \int_0^T \dot{\tau}\;dt \le \int_0^T \sqrt{g}\;dt = \sqrt{g}\;T, \tag{5}$$

or $\tau(T) \le \sqrt{g}\;T$. Thus the time taken for the object to travel from A to B is bounded below (except for the factor $\sqrt{g}$) by the τ-coordinate of B.

The obvious way of obtaining this minimum travel time is to set $\dot{\tau} = \sqrt{g}$ and $\dot{\rho} = 0$, since then (4) holds and we get equality in (5). This, of course, means that ρ is a constant and the path of the object is a cycloid arc.

Since ρ is a constant, we have $\dot{\theta} = \dot{\tau}/\rho = \sqrt{g}/\rho = \sqrt{g/R}$, and so the object's motion is the same as a point on the rim of a disk of radius R rolling along the x-axis with constant angular velocity $\omega = \dot{\theta} = \sqrt{g/R}$.

For example, suppose that the points A and B are L units apart at the same elevation. Then A and B are the end points of a cycloid arc made by one complete rotation of a disk of radius $R = L/2\pi$. Thus $\omega = \dot{\theta} = \sqrt{2\pi g/L}$ and the time taken to travel from A to B is $T = 2\pi/\omega = \sqrt{2\pi L/g}$. If L is 100 meters and $g = 9.8\ \mathrm{m/s}^2$, then $T \approx 8$ seconds—faster than the world record time for sprinters over the same distance.

There are some important questions left unanswered by our discussion so far. Is there a cycloid arc joining A and B? If so, is this cycloid arc unique? Which points (x, y) in the plane can be expressed in the form (3) for some τ and ρ? Are τ and ρ uniquely determined by x and y? Is the cycloid arc the only way to get equality in (5), and hence minimum travel time?

To answer these questions, we first prove some simple trigonometric inequalities.

Lemma 1. *The following inequalities hold for* $0 < \theta < 2\pi$*:*

(1) $0 < \theta - \sin\theta$

(2) $0 < \sin\dfrac{\theta}{2} - \dfrac{\theta}{2}\cos\dfrac{\theta}{2}$

(3) $0 < 2(1-\cos\theta) - \theta\sin\theta$

(1) *and* (2) *hold also when* $\theta = 2\pi$.

Proof. The function $f(x) = \theta - \sin\theta$ derivative $f'(\theta) = 1 - \cos\theta$ which is positive on $(0, 2\pi)$, and so f is strictly increasing. Since $f(0) = 0$, this implies $f(\theta) > 0$ on $(0, 2\pi]$. Similarly, the function $g(x) = \sin(\theta/2) - (\theta/2)\cos(\theta/2)$ has derivative $g'(\theta) = (\theta/4)\sin(\theta/2)$ which is positive on $(0, 2\pi)$. Since $g(0) = 0$, this implies $g(\theta) > 0$ on $(0, 2\pi]$. The remaining inequality can be obtained by multiplying (2) by $2\sin(\theta/2)$, which is positive on $(0, 2\pi)$, and then using the double angle identities for the sine and cosine functions. □

Lemma 2. *If* $x > 0$ *and* $y \geq 0$*, then there are unique* $R > 0$ *and* $0 \leq \theta \leq 2\pi$ *satisfying* (1)*, as well as unique* $\rho > 0$ *and* $0 \leq \tau \leq 2\pi\rho$ *satisfying* (3).

Proof. The function

$$h(\theta) = \frac{1-\cos\theta}{\theta - \sin\theta}$$

is defined on $(0, 2\pi]$ by Lemma 1(1) and has the derivative

$$h'(\theta) = \frac{2(\cos\theta - 1) + \theta\sin\theta}{(\theta - \sin\theta)^2}.$$

By Lemma 1(3), $h'(\theta) < 0$ on $(0, 2\pi)$ and so h is strictly decreasing on $(0, 2\pi]$.

We also have $h(2\pi) = 0$, and, by l'Hôpital's rule, $\lim_{\theta\to 0^+} h(\theta) = \infty$. Since $y/x \geq 0$ and h is continuous on $(0, 2\pi]$, the Intermediate Value Theorem guarantees the existence of some θ in $(0, 2\pi]$ such that $h(\theta) = y/x$. Since h is strictly decreasing, θ is unique. Now it is easy to check that θ and $R = x/(\theta - \sin\theta)$ is the unique solution of (1), and that $\rho = \sqrt{R}$ and $\tau = \theta\rho$ is the unique solution of (3). □

Since each point (x, y) with $x > 0$ and $y \geq 0$ corresponds to uniquely determined ρ and τ, the equations in (3) represent a change of coordinate systems for the first quadrant of the xy-plane. So long as B is in this quadrant, its ρ coordinate determines a unique cycloid arc of the form (1) that passes through it. Moreover, if the object remains in this quadrant, its motion can be described by functions ρ and τ of time.

We claimed at the beginning of this note that our proof shows that the cycloid arc yields the minimum travel time among curves that may have loops or corners. Since we are using two functions of time, x and y, to describe possible paths of the object, loops are not a problem. At a corner in the path, however, the derivatives of x, y, ρ and τ may not exist. Since we have so far implicitly assumed that these functions are differentiable, we now need to see whether our discussion can be generalized to functions which are not differentiable everywhere. For our arguments to work, we need that the integral in (5) exists and that τ is the indefinite integral of $\dot{\tau}$. This happens if and only if τ is **absolutely continuous**. See [7, Chapter 5] for the definition and properties of absolutely continuous functions—including the fact that such functions are differentiable almost everywhere and are the indefinite integrals of their derivatives.

It is then natural to suppose that τ and ρ are both absolutely continuous, and that (4) holds almost everywhere. With these assumptions, we can prove that the cycloid arc is the only brachistochrone. If the concept of absolute continuity is unfamiliar, the reader can show that the argument below works with the stronger assumption that τ and ρ have continuous, or piecewise continuous, derivatives.

Suppose that, for some absolutely continuous functions τ and ρ, equality is attained in (5):

$$\int_0^T \dot{\tau}\ dt = \int_0^T \sqrt{g}\ dt.$$

Since (4) holds almost everywhere, we also have $\dot{\tau} \leq \sqrt{g}$ almost everywhere. These conditions imply that $\dot{\tau} = \sqrt{g}$ almost everywhere. Plugging this result into (4), and using the fact, from Lemma 1(2), that the coefficient of $\dot{\rho}^2$ is nonzero, we get that $\dot{\rho} = 0$ almost everywhere. This implies that ρ is a constant function, and hence, that the minimum travel time is attained only by the cycloid arc.

References

[1] D. C. Benson, *An Elementary Solution of the Brachistochrone Problem,* Amer. Math. Monthly, **76**, (1969), 890-894.

[2] G. A. Bliss, *Calculus of Variations,* Math. Assoc. Amer., (1925).

[3] H. Ehlichson, *Johann Bernoulli's brachistochrone solution using Fermat's principle of least time,* Eur. J. Phys. , **20**, (1999), 299-304.

[4] H. H. Goldstine, *A History of the Calculus of Variations from the 17th Through the 19th Century,* Springer Verlag , Berlin, (1980).

[5] Nils Johnson, *The Brachistochrone Problem,* College Math J., **35**, No. 3, (2004), 192-197.

[6] W. Hrusa and J. L. Troutman, *Elementary Characterization of Classical Minima,* Amer. Math. Monthly, **88**, (1981), 321-327.

[7] H. Royden, *Real Analysis, 3rd ed.*, Prentice Hall, (1988).

[8] F. Y. M. Wan, *An Introduction to the Calculus of Variations and its Applications,* Chapman & Hall, New York, (1995).

[9] Jim Zeng, *A Note on the Brachistochrone Problem,* College Math J., **27**, No. 3, (1996), 206-208.

Department of Mathematics, California State University, Los Angeles CA 90032-8204

E-mail address: gbrookf@calstatela.edu