The Brachistochrone Problem Revisited

Pablo V. Negrón–Marrero*

Department of Mathematics
University of Puerto Rico
Humacao, PR 00791-4300

Abstract

The brachistochrone problem is posed as a problem of the calculus of variations with differential side constraints, among piecewise smooth parametrized curves satisfying appropriate initial and boundary conditions. In the classical brachistochrone problem, the initial speed is taken to be zero with initial angle of inclination of $-\pi/2$. In this paper we consider more general initial conditions. In particular, we show that the initial value problems in which the initial angle of inclination is given, or that in which the initial speed (kinetic energy) is specified, both have solutions. If a certain compatibility condition between the initial angle of inclination and speed is not satisfied, we show that there is a minimum time of descent but no minimizer of the time functional.

Key words: brachistochrone, calculus of variations, differential constraints.

1 Introduction

The brachistochrone problem consists of finding the curve, joining two (non–vertical) given points, along which a bead of given mass falls under the influence of gravity in the minimum time[1]. This problem was first posed by Johann Bernoulli in 1696 and solved that same year by Newton, Leibniz, the Bernoulli brothers Johann and Jacob, and de L'Hôpital. The solution of the brachistochrone problem was pivotal to the development of the now very important branch of analysis called the calculus of variations. Since then variations of the brachistochrone problem have been presented in many books and papers (see e.g. [2], [3], [5], [8]) but the great majority of these expositions assume from the start

*pnm@math.uprh.edu

[1]The curve as well as the initial velocity of the bead are assumed to lie in the vertical plane containing the two given points.

that the solution curve is a function of x in the xy plane. Although this assumption turns out to be correct, these derivations do not expose a whole variety of problems associated to the brachistochrone problem, in particular those concerning the specification of initial conditions which are not discussed in standard differential equations textbooks.

In this paper we discuss the solution of the brachistochrone problem along the lines of the treatments in [1] and [4]. The problem is posed as one of the calculus of variations with differential side constraints, among smooth parametrized curves satisfying appropriate initial and boundary conditions. The equation of motion of the mass along the tangential direction to the curve, becomes now a differential side condition in the variational formulation (cf. (4a)). A major advantage of this approach as compared to the standard derivations based on conservation of energy, is that one does not commit itself to any specific parametrization of the curve of descent, like time or arc–length. A suitable parameter can then be chosen for the solution of the corresponding Euler–Lagrange equations. Even though this formulation is more appropriate for the problems with friction discussed in [1] and [4], it allows us to preserve all the structure of the classical problem that we need for the discussion regarding the specification of initial conditions.

In Section (4) we show that the Euler–Lagrange equations for the resulting variational formulation can be solved explicitly. The variational process leads naturally to the use of the angle of inclination of the tangent to the curve from the horizontal as the parametrization for the solution curve and time integrand. The resulting extremals are sections of a cycloid[2] with the size of the cycloid and the initial angle of inclination as parameters to be determined from the initial and boundary conditions. It is not immediately obvious that these extremals can be chosen so as to satisfy the given initial and boundary conditions. In this paper we show that the following two problems have unique solutions: the initial value problem in which (in addition to the boundary conditions) the initial angle of inclination is given (Section (5)), or that in which the initial speed (kinetic energy) is specified (Section (6)). Even though we do not discuss in this paper the numerical aspects of constructing the curves of minimum descent, we describe briefly the procedures used in the numerical examples.

The specification of the initial velocity of the particle, which amounts to specifying both the initial speed and initial angle of inclination, in general leads to an inconsistent problem[3], unless the *compatibility condition* (cf. (30)) is satisfied. When the compatibility condition is not satisfied, we show in Section (7) that there is a minimum time of descent but there is no minimizer, i.e., there is no curve that yields the minimum time.

Notation: We denote derivatives with respect to time t by the usual "dot" notation,

[2]The cycloid is also the solution of the problem of the tautochrone, that is finding the curve with property that the time of descent from any point on the curve to a lower (fixed) point is independent of the initial point. This property was proved by Huygens (1659) and used for the constructions of the cycloidal pendulum clocks.

[3]In the classical brachistochrone problem the initial angle of inclination is taken to be $-\pi/2$. We show that in this case an initial speed of zero is the only possible selection that leads to a consistent problem.

$\dot{x}(t)$, etc.. Those derivatives with respect to any other parameter not equal to time will be denoted by primes.

2 The Variational Formulation

We let a particle slide from $(0,0)$ to (a,b) (where $a > 0$ and $b \leq 0$) along the curve $(x(t), y(t))$ under the influence of gravity. We let $\theta(t)$ be the angle of inclination of the tangent to the curve from the horizontal. If $v(t)$ denotes the speed of the particle at $(x(t), y(t))$, then we have that

$$\begin{aligned} \dot{x}(t) &= v(t)\cos\theta(t), & \text{(1a)} \\ \dot{y}(t) &= v(t)\sin\theta(t), & \text{(1b)} \\ \dot{v}(t) &= -g\sin\theta(t). & \text{(1c)} \end{aligned}$$

We can eliminate $\theta(t)$ from these equations upon recalling that

$$v(t) = \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}. \tag{2}$$

Thus we now have that

$$\begin{aligned} v(t) &= \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}, & \text{(3a)} \\ v(t)\dot{v}(t) &= -g\dot{y}(t). & \text{(3b)} \end{aligned}$$

Note that (3b) is equivalent to conservation of energy. We let $t = \hat{t}(\tau)$ be a reparametrization of the curve in terms of a parameter τ where we assume that:

$$\hat{t}'(\tau) = \frac{\mathrm{d}\hat{t}(\tau)}{\mathrm{d}\tau} > 0.$$

If we let $\bar{x}(\tau) = x(\hat{t}(\tau))$, then it follows now that

$$\hat{t}'(\tau)\dot{x}(\hat{t}(\tau)) = \bar{x}'(\tau), \quad \text{etc..}$$

Using these expressions we can write (3a), (3b) as:

$$\begin{aligned} \hat{t}'(\tau)\bar{v}(\tau) &= \sqrt{\bar{x}'(\tau)^2 + \bar{y}'(\tau)^2}, & \text{(4a)} \\ \bar{v}(\tau)\bar{v}'(\tau) &= -g\bar{y}'(\tau). & \text{(4b)} \end{aligned}$$

The problem now is to minimize the time integral:

$$\int_{\tau_1}^{\tau_2} \hat{t}'(\tau)\mathrm{d}\tau, \tag{5}$$

among *piecewise smooth*[4] functions subject to the differential constraints (4a), (4b), and to the boundary conditions:

$$\bar{x}(\tau_1) = \bar{y}(\tau_1) = 0, \quad \bar{x}(\tau_2) = a\ ,\ \bar{y}(\tau_2) = b, \tag{6a}$$
$$\hat{t}(\tau_1) = 0. \tag{6b}$$

3 The Lagrange Multiplier Rule with Differential Constraints

Let $\mathbf{y} : [a,b] \to \mathbb{R}^n$, $f : [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, and $\phi : [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ where $m < n$. We consider the problem of minimizing

$$I(\mathbf{y}) = \int_a^b f(\tau, \mathbf{y}(\tau), \mathbf{y}'(\tau))\, \mathrm{d}\tau, \tag{7}$$

among smooth functions $\mathbf{y}$ subject to the *differential side conditions*:

$$\phi(\tau, \mathbf{y}(\tau), \mathbf{y}'(\tau)) = 0, \quad a < \tau < b, \tag{8}$$

and to the boundary conditions:

$$\mathbf{y}(a) = \mathbf{y}_a, \quad \mathbf{y}(b) = \mathbf{y}_b, \tag{9}$$

We now have (see e.g. [6], [7]):

Theorem 3.1. *Let* $\mathbf{y}$ *be a piecewise smooth function minimizing* (7) *subject to* (8) *and* (9). *Let* f, ϕ *be* C^1 *functions and* $\mathrm{D}_{\mathbf{y}'}\phi$ *be of full rank* m. *Then there exists a piecewise continuous nonzero function* $\boldsymbol{\lambda} : [a,b] \to \mathbb{R}^m$ *such that*

$$L_{\mathbf{y}'}(\tau, \mathbf{y}(\tau), \mathbf{y}'(\tau), \boldsymbol{\lambda}(\tau)) = \int_a^\tau L_{\mathbf{y}}(\xi, \mathbf{y}(\xi), \mathbf{y}'(\xi), \boldsymbol{\lambda}(\xi))\, \mathrm{d}\xi + \mathbf{c}, \tag{10}$$

for some constant vector $\mathbf{c} \in \mathbb{R}^n$ *and where*

$$L(\tau, \mathbf{y}, \mathbf{y}', \boldsymbol{\lambda}) = -f(\tau, \mathbf{y}, \mathbf{y}') + \boldsymbol{\lambda} \cdot \phi(\tau, \mathbf{y}, \mathbf{y}'). \tag{11}$$

The function L in the theorem is called the *Lagrangian* while $\boldsymbol{\lambda}$ is the *Lagrange multiplier* or *costate variable*. If $\mathbf{y}(\cdot) \in C^1$ and $\boldsymbol{\lambda}(\cdot)$ is continuous, then the right hand side of (10) is a differentiable function of τ. Thus upon differentiating on both sides,

[4] A continuous function is called *piecewise smooth* on $[a,b]$ if it is differentiable except possibly at a finite number of points. At the points of discontinuity of the first derivative, both the left and right derivatives exist.

we get that the so called *Euler–Lagrange equations*[5] for the problem of minimizing (7) subject to (8) and (9) are given by:

$$\frac{\mathrm{d}}{\mathrm{d}\tau} L_{\mathbf{y}'}(\tau, \mathbf{y}(\tau), \mathbf{y}'(\tau), \boldsymbol{\lambda}(\tau)) = L_{\mathbf{y}}(\tau, \mathbf{y}(\tau), \mathbf{y}'(\tau), \boldsymbol{\lambda}(\tau)), \tag{12}$$

in addition to (8) and (9). If some components of $\mathbf{y}(a)$ or $\mathbf{y}(b)$ are not specified in (9), then we would have the so called *natural boundary conditions* corresponding to the unspecified components. (See e.g. [6], [7].)

4 The Euler–Lagrange Equations

We return now to the problem posed at the end of Section (2). Let

$$\begin{aligned} \mathbf{y}(\tau) &= (\bar{x}(\tau), \bar{y}(\tau), \bar{v}(\tau), \hat{t}(\tau)), && (13a)\\ \boldsymbol{\lambda}(\tau) &= (\bar{\sigma}(\tau), \bar{\lambda}(\tau)), && (13b)\\ L(\tau, \mathbf{y}, \mathbf{y}', \boldsymbol{\lambda}) &= \hat{t}' + \bar{\sigma}\left(\hat{t}'\bar{v} - \sqrt{(\bar{x}')^2 + (\bar{y}')^2}\right) \\ &\quad + \bar{\lambda}\left(\bar{v}\bar{v}' + g\bar{y}'\right). && (13c) \end{aligned}$$

It follows now from Theorem (3.1) and assuming that $\mathbf{y}(\cdot) \in C^1$ and $\boldsymbol{\lambda}(\cdot)$ is continuous, that the Euler–Lagrange equations for the problem of minimizing (5) subject to (6), and the differential constraints (4), are given by:

$$1 + \bar{\sigma}(\tau)\bar{v}(\tau) = c_1, \quad \bar{v}(\tau)\bar{\lambda}'(\tau) = \hat{t}'(\tau)\bar{\sigma}(\tau), \tag{14a}$$

$$\frac{\bar{\sigma}(\tau)\bar{x}'(\tau)}{\hat{t}'(\tau)\bar{v}(\tau)} = c_2, \quad \frac{\bar{\sigma}(\tau)\bar{y}'(\tau)}{\hat{t}'(\tau)\bar{v}(\tau)} = g\bar{\lambda}(\tau) - c_3, \tag{14b}$$

together with (4), where c_1, c_2, c_3 are some constants of integration. Note that the time parametrization $\hat{t}(\tau)$ is part of the variational process (cf. $(14a)_2$) and as such is completely determined up to a constant.

The boundary conditions are given by (6) and the *natural boundary conditions*[6] by:

$$\left.\frac{\partial L}{\partial \bar{v}'}\right|_{\tau=\tau_2} = 0, \quad \left.\frac{\partial L}{\partial \hat{t}'}\right|_{\tau=\tau_2} = 0,$$

which are equivalent (assuming $\bar{v}(\tau_2) \neq 0$) to:

$$\bar{\lambda}(\tau_2) = 0, \quad 1 + \bar{\sigma}(\tau_2)\bar{v}(\tau_2) = 0, \tag{15}$$

[5]Equation (10) is the integral form of the first order necessary condition for a minimizer and is satisfied at every point of the domain of the extremal curve. Equation (12) however is satisfied only at the points where $\mathbf{y}'$ and $\boldsymbol{\lambda}$ are continuous. Equations (10) and (12) are equivalent if one specifies that (12) holds except at the points of discontinuity of $\mathbf{y}'$ or $\boldsymbol{\lambda}$.

[6]These boundary conditions correspond to the unspecification of $\bar{v}(\tau_2)$ and $\hat{t}(\tau_2)$ respectively.

the second of which implies that $c_1 = 0$ in (14). In terms of the parameter τ we have that (1a) and (1b) can be written as:

$$\bar{x}'(\tau) = \hat{t}'(\tau)\bar{v}(\tau)\cos\bar{\theta}(\tau), \quad \bar{y}'(\tau) = \hat{t}'(\tau)\bar{v}(\tau)\sin\bar{\theta}(\tau). \tag{16}$$

It follows now that (4a) is satisfied and that (4b) and (14) reduce to:

$$\begin{aligned}
\bar{v}'(\tau) &= -g\hat{t}'(\tau)\sin\hat{\theta}(\tau), && \text{(17a)}\\
1+\bar{\sigma}(\tau)\bar{v}(\tau) &= 0, && \text{(17b)}\\
\bar{v}(\tau)\bar{\lambda}'(\tau) &= \hat{t}'(\tau)\bar{\sigma}(\tau), && \text{(17c)}\\
\bar{\sigma}(\tau)\cos\bar{\theta}(\tau) &= c_2, && \text{(17d)}\\
\bar{\sigma}(\tau)\sin\bar{\theta}(\tau) &= g\bar{\lambda}(\tau) - c_3. && \text{(17e)}
\end{aligned}$$

It follows immediately that

$$\bar{\sigma}(\tau) = \frac{c_2}{\cos\bar{\theta}(\tau)}, \qquad g\bar{\lambda}(\tau) = c_3 + c_2\tan\bar{\theta}(\tau), \tag{18a}$$

$$\bar{v}(\tau) = -\frac{1}{c_2}\cos\bar{\theta}(\tau), \qquad \hat{t}'(\tau) = -\frac{1}{c_2^2}\cos^2\bar{\theta}(\tau)\bar{\lambda}'(\tau). \tag{18b}$$

Note that we can use now θ as a parameter and using (16) and the second equation in (18b) we get that

$$\bar{x}'(\theta) = 4\gamma\cos^2\theta, \quad \bar{y}'(\theta) = 4\gamma\sin\theta\cos\theta, \quad \gamma = \frac{1}{4gc_2^2}. \tag{19}$$

Assuming that the initial angle of inclination is $\theta_1 \in [-\pi/2, \pi/2)$, we get from the first equation in (6a) and (19) that

$$\bar{x}(\theta) = \gamma(2(\theta-\theta_1) + \sin 2\theta - \sin 2\theta_1), \quad \bar{y}(\theta) = \gamma(\cos 2\theta_1 - \cos 2\theta), \tag{20}$$

which are the parametric equations for a cycloid. If $\theta_2 > \theta_1$ is the final angle of inclination, we get from the second and third equations in (6a) that $\bar{y}(\theta_2)/\bar{x}(\theta_2) = b/a$, that is

$$\frac{\cos 2\theta_1 - \cos 2\theta_2}{2(\theta_2-\theta_1) + \sin 2\theta_2 - \sin 2\theta_1} = \frac{b}{a}. \tag{21}$$

If such θ_2 is found, then γ in (20) can be obtained for instance from $\bar{x}(\theta_2) = a$ which gives that

$$\gamma = \frac{a}{2(\theta_2-\theta_1) + \sin 2\theta_2 - \sin 2\theta_1}. \tag{22}$$

Note that if the initial velocity of the particle is given, then both v_0 (the initial speed) and the initial angle θ_1 are specified. It follows now from the first equation in (18b) that c_2 is completely determined. This fixes γ, the size of the cycloid, according to the third equation in (19). The final angle θ_2 depends only on θ_1, a, b and can be obtained from (21). (See Proposition (5.1).) Finally equation (22), which is the condition that the cycloid contains the target point (a, b), must be satisfied and acts in this case as a compatibility condition for the values of γ and θ_2 found so far. Thus the problem with the initial velocity of the particle specified leads to an over–determined system which in general is inconsistent.

We consider now two cases in which these equations have solutions, namely when the initial angle of inclination is specified, or when the initial speed, or equivalently, the initial kinetic energy is specified.

5 Initial Angle of Inclination Specified

We assume that the initial angle of inclination $\theta_1 \in [-\pi/2, \pi/2)$ is given. We now give conditions under which (21) has solutions.

Proposition 5.1. *Let $\theta_1 \in [-\pi/2, \pi/2)$, $a > 0$, and $b \in \mathbb{R}$. Then equation (21) has a solution $\theta_2 \in [\theta_1, \pi/2]$ which is unique provided θ_1 is such that*

$$\tan\theta_1 \leq \frac{b}{a} \leq \frac{1+\cos 2\theta_1}{\pi - 2\theta_1 - \sin 2\theta_1}. \tag{23}$$

Proof: Let

$$w(\theta) = \frac{\cos 2\theta_1 - \cos 2\theta}{2(\theta-\theta_1) + \sin 2\theta - \sin 2\theta_1}.$$

We show below that $w(\cdot)$ is strictly increasing on $[\theta_1, \pi/2]$ from which the result of the proposition would follow since

$$\begin{aligned} w(\theta_1) &= \lim_{\theta\to\theta_1^+} \frac{\sin 2\theta}{1+\cos 2\theta} = \tan\theta_1, \\ w(\pi/2) &= \frac{1+\cos 2\theta_1}{\pi - 2\theta_1 - \sin 2\theta_1} \geq 0, \end{aligned}$$

where the last inequality holds for $\theta_1 \in [-\pi/2, \pi/2)$.

By direct computation and after some simplifications one gets that

$$\frac{\mathrm{d}w}{\mathrm{d}\theta} = 8h(\theta)\left[2(\theta-\theta_1) + \sin 2\theta - \sin 2\theta_1\right]^{-2},$$

where

$$\begin{aligned} h(\theta) &= (\theta - \theta_1 + \sin\theta\cos\theta - \sin\theta_1\cos\theta_1)\sin\theta\cos\theta \\ &\quad -(\cos^2\theta_1 - \cos^2\theta)\cos^2\theta, \\ &= \left[(\theta-\theta_1 - \sin\theta_1\cos\theta_1)\sin\theta + \sin^2\theta_1\cos\theta\right]\cos\theta, \\ &\equiv q(\theta)\cos\theta. \end{aligned}$$

Since $\cos\theta \geq 0$ for $\theta \in [\theta_1, \pi/2]$, $\theta_1 \in [-\pi/2, \pi/2)$ (being zero only at $\pi/2$, and at θ_1 when $\theta_1 = -\pi/2$), we need only to check the sign of $q(\theta)$. We consider two subcases:

1. **Case** $\theta_1 \in [-\pi/2, 0)$**:** Since $\theta - \theta_1 - \sin\theta_1 \cos\theta_1 \geq 0$ for $\theta \in [\theta_1, \pi/2]$, $\theta_1 \in [-\pi/2, 0)$ (equal to zero only at $\theta = \theta_1 = -\pi/2$), it follows that $q(\theta) > 0$ for $\theta \in [0, \pi/2]$. For $\theta \in [\theta_1, 0]$, first note that

$$
\begin{aligned}
q'(\theta) &= (\theta - \theta_1 - \sin\theta_1 \cos\theta_1)\cos\theta + \cos^2\theta_1 \sin\theta, \\
q''(\theta) &= -(\theta - \theta_1 - \sin\theta_1 \cos\theta_1)\sin\theta + (1 + \cos^2\theta_1)\cos\theta.
\end{aligned}
$$

Since $q(\theta_1) = 0$, $q'(\theta_1) = 0$, and $q''(\theta) \geq 0$ for $\theta \in [\theta_1, 0]$, it follows that $q(\theta) > 0$ for $\theta \in (\theta_1, 0]$. Combining this with the result on $[0, \pi/2]$, we get that $w(\cdot)$ is strictly increasing.

2. **Case** $\theta_1 \in [0, \pi/2)$**:** Note that

$$
\begin{aligned}
q(\theta) &= (\theta - \theta_1)\sin\theta + \sin\theta_1(\sin\theta_1 \cos\theta - \cos\theta_1 \sin\theta), \\
&= (\theta - \theta_1)\left[\sin\theta - \sin\theta_1 \frac{\sin(\theta - \theta_1)}{\theta - \theta_1}\right].
\end{aligned}
$$

Now for $\theta \in [\theta_1, \pi/2]$, $\theta_1 \in [0, \pi/2)$, we have that $\sin\theta \geq \sin\theta_1 \geq 0$, and that

$$
\frac{\sin(\theta - \theta_1)}{\theta - \theta_1} \leq 1,
$$

from which it follows that $q(\theta) \geq 0$ with equality only at $\theta = \theta_1$. Hence $w(\cdot)$ is strictly increasing in this case as well.

□

From the expression for $\bar{v}(\cdot)$ in (18b), we see that the initial speed is given by the formula

$$
\bar{v}(\theta_1) = -\frac{1}{c_2}\cos\theta_1 = 2\sqrt{g\gamma}\cos\theta_1. \tag{24}
$$

Thus the initial speed can be zero only if the initial angle of inclination is $\theta_1 = -\pi/2$. Furthermore, note that the inequality (23) is automatically satisfied when $b \leq 0$ for $\theta_1 = -\pi/2$. Thus only final points with negative y–coordinate are accesible by the extremals (20) in this case. This should be intuitively clear since in this case the initial speed must be zero, and by conservation of energy, one can not reach with an extremal a point above the positive axis.

To completely determine (20) one would proceed in two steps:

1. Given a, b, θ_1, solve equation (21) for θ_2.

2. Compute the constant γ from equation (22).

This gives essentially an algorithm to compute the cycloid which we implemented in MATLAB. We show in Figure (1) the corresponding brachistochrone curves for $a = 6$,

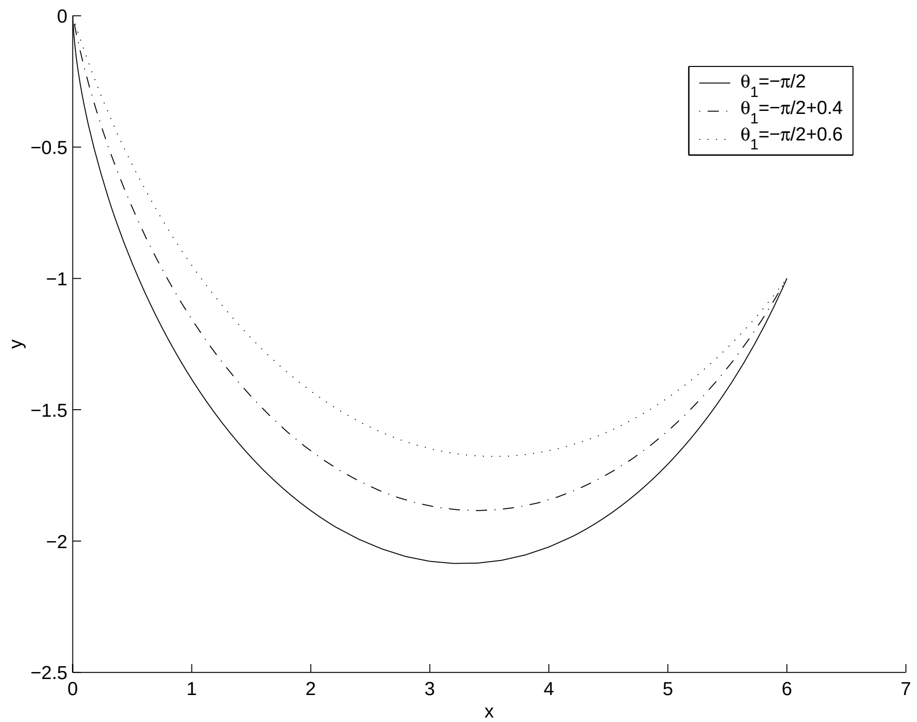

Figure 1: Graphs of the brachistochrones (cycloids) for initial angles of inclination $\theta_1 = -\pi/2, -\pi/2 + 0.4, -\pi/2 + 0.6$ and final point $a = 6$, $b = -1$.

$b = -1$ and the cases $\theta_1 = -\pi/2, -\pi/2 + 0.4, -\pi/2 + 0.6$. The corresponding initial speeds are 0, 2.5692, and 3.9234 with units consistent with those of a, b. The times of descent (c.f. (25)) were respectively 1.5508, 1.2478, and 1.0799.

We close this section with some additional results and properties.

Proposition 5.2. *For any given $a > 0$ and $b \in \mathbb{R}$, the solution θ_2 of* (21) *has the form $\theta_2 = \hat{\theta}_2(\theta_1)$ where $\hat{\theta}_2 : [-\pi/2, \theta_1^*) \to \mathbb{R}$ is a C^1 decreasing function with $\hat{\theta}_2(\theta_1) \geq \theta_1$ and $\hat{\theta}_2(\theta_1^*) = \theta_1^*$ where $\theta_1^* = \tan^{-1}(b/a)$. The time of descent along the curve* (20) *is given by the formula:*

$$T(\theta_1) = 2\sqrt{\frac{a(\hat{\theta}_2(\theta_1) - \theta_1)^2}{g(2(\hat{\theta}_2(\theta_1) - \theta_1) + \sin 2\hat{\theta}_2(\theta_1) - \sin 2\theta_1)}}. \tag{25}$$

Furthermore

$$\lim_{\theta_1 \to \theta_1^*} \bar{v}(\theta_1) = \infty, \quad \lim_{\theta_1 \to \theta_1^*} T(\theta_1) = 0. \tag{26}$$

Proof: Let $w(\theta_2, \theta_1)$ be given by the left hand side of (21). Since $w(\theta_2, \theta_1) = w(\theta_1, \theta_2)$, it follows from the proof of the previous proposition that:

$$\frac{\partial w}{\partial \theta_2} > 0, \quad \frac{\partial w}{\partial \theta_1} > 0. \tag{27}$$

Let $\tilde{\theta}_1$ satisfy (23) and let $\tilde{\theta}_2$ be the corresponding solution given by Proposition (5.1), i.e., $w(\tilde{\theta}_2, \tilde{\theta}_1) = b/a$. It follows now from first inequality in (27) and the Implicit Function Theorem that there exists a C^1 function $\hat{\theta}_2(\cdot)$ such that $\hat{\theta}_2(\tilde{\theta}_1) = \tilde{\theta}_2$ and

$$w(\hat{\theta}_2(\theta_1), \theta_1) = \frac{b}{a},$$

for all θ_1 in a maximal interval containing $\tilde{\theta}_1$. Since

$$\hat{\theta}_2'(\theta_1) = -\frac{\dfrac{\partial w}{\partial \theta_1}(\hat{\theta}_2(\theta_1), \theta_1)}{\dfrac{\partial w}{\partial \theta_2}(\hat{\theta}_2(\theta_1), \theta_1)},$$

it follows from (27) that $\hat{\theta}_2(\cdot)$ is decreasing. The statement concerning $\theta_1^* = \tan^{-1}(b/a)$ follows from the observation that when $\theta_1 = \theta_1^*$, then $\theta_2 = \theta_1^*$ is the unique solution of (23) predicted by Proposition (5.1).

Formula (25) follows from (5), (18b), (22), and the result above concerning $\hat{\theta}_2(\theta_1)$. Since $\hat{\theta}_2(\theta_1) \searrow \theta_1^*$ as $\theta_1 \nearrow \theta_1^*$, it follows from equation (22) that $\gamma \to \infty$ as $\theta_1 \nearrow \theta_1^*$. This together with (24) implies that $(26)_a$ holds.

Finally to show $(26)_b$, note that by Taylor's Theorem applied to $\sin 2\theta$ about $\theta = \theta_1$, we get that

$$2(\hat{\theta}_2(\theta_1) - \theta_1) + \sin 2\hat{\theta}_2(\theta_1) - \sin 2\theta_1 = 2(\hat{\theta}_2(\theta_1) - \theta_1)(1 + \cos 2\xi(\theta_1)),$$

where $\xi(\theta_1)$ is between θ_1 and $\hat{\theta}_2(\theta_1)$. Thus (25) can now be written as

$$T(\theta_1) = 2\sqrt{\frac{a(\hat{\theta}_2(\theta_1) - \theta_1)}{g(1 + \cos 2\xi(\theta_1))}}.$$

Since $\hat{\theta}_2(\theta_1) \searrow \theta_1^*$ as $\theta_1 \nearrow \theta_1^*$, we get that as $\theta_1 \to \theta_1^*$, then $\xi(\theta_1) \to \theta_1^*$, and since $\theta_1^* \in (-\pi/2, \pi/2)$ we have that $(26)_b$ follows. Note that according to $(26)_a$ this is achieved only as the initial speed becomes infinite. $\square$

6 Initial Kinetic Energy Specified

We assume now that the initial speed $v_0 > 0$ is given. After eliminating γ from them, equations (21) and (22) now have to be solved for θ_1, θ_2. Once these values are determined, we can get γ from (22) and the solution curve from (20).

Using the first equation in (18b) evaluated at θ_1 and the third equation in (19), we can write (22) as

$$2\theta_2 + \sin 2\theta_2 = 2\theta_1 + \sin 2\theta_1 + \left(\frac{2ag}{v_0^2}\right)(1 + \cos 2\theta_1). \tag{28}$$

Even though θ_1 is unknown we can still apply Propositions (5.1) and (5.2) to write the above equation as

$$2\hat{\theta}_2(\theta_1) + \sin 2\hat{\theta}_2(\theta_1) = 2\theta_1 + \sin 2\theta_1 + \left(\frac{2ag}{v_0^2}\right)(1 + \cos 2\theta_1). \tag{29}$$

Since $\hat{\theta}_2(-\pi/2) \geq -\pi/2$ and $\theta \mapsto 2\theta + \sin 2\theta$ is increasing, we have that

$$2\hat{\theta}_2(-\pi/2) + \sin 2\hat{\theta}_2(-\pi/2) \geq 2(-\pi/2) + \sin 2(-\pi/2).$$

Thus the left hand side of (29) is greater than the right hand side for $\theta_1 = -\pi/2$. On the other hand since $\hat{\theta}_2(\theta_1^*) = \theta_1^*$ where $\theta_1^* = \tan^{-1}(b/a)$, and $1 + \cos 2\theta_1^* > 0$, we get that the right hand side of (29) is greater than the left hand side for $\theta_1 = \theta_1^*$. Hence equation (29) has a solution $\theta_1 \in [-\pi/2, \theta_1^*]$. With this θ_1 we then proceed to solve equation (21) as before for θ_2. This proves the following:

Proposition 6.1. *Let $v_0 > 0$, $a > 0$, and $b \in \mathbb{R}$. Then the system* (21)–(22) *has a solution pair (θ_1, θ_2) with $\theta_2 \geq \theta_1$.*

The numerical procedure described above to solve (21) and (28) for (θ_1, θ_2) given a, b, v_0, is better implemented in practice by a *blocked fixed point iteration.* Namely:

1. Given a, b, v_0 and an initial approximation to θ_1,
   (a) Solve equation (21) for θ_2 given θ_1.
   (b) With the θ_2 from part (a), solve equation (28) for θ_1.
   (c) Repeat (a), (b) as necessary to achieve a certain prescribed accuracy.
2. Compute the constant γ from equation (22).

We implemented this procedure in MATLAB. We show in Figure (2) the corresponding brachistochrone curves for $a = 6$, $b = -1$ and the cases $v_0 = 0.01, 2, 4$ with units consistent with those of a, b. The corresponding initial angles turned out to be -1.5694, -1.2585, and -0.96004. The times of descent are respectively 1.5499, 1.3195, and 1.0708.

7 The Compatibility Condition

The first equation in (18b) when evaluated at $\tau = \tau_1$ yields the relation:

$$v_0 = -\frac{1}{c_2} \cos \theta_1, \tag{30}$$

between the initial angle of inclination and the initial speed. We call this the *compatibility condition.* In this section we address the following questions: when the compatibility

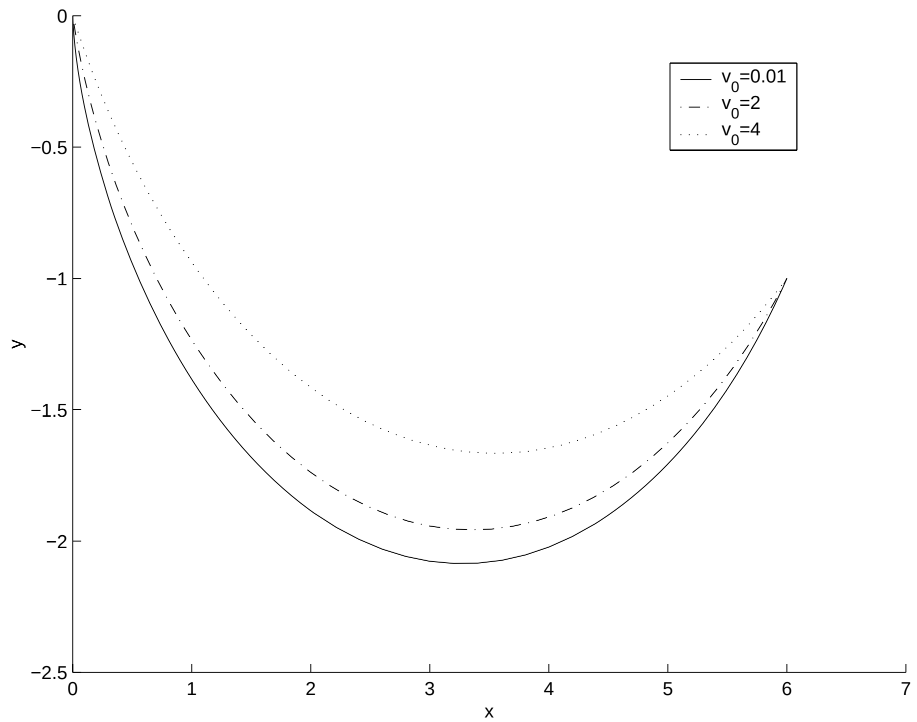

Figure 2: Graphs of the brachistochrones (cycloids) for initial speeds $v_0 = 0.01, 2, 4$ and final point $a = 6$, $b = -1$.

condition is not satisfied, is there a minimum time of descent?, and if so, does a minimizer exist?

To answer these questions we are going to consider the simpler scenario in which the descent curves are assumed to be functions of x from the start of the analysis. In this case the time integral (5) is given by (see e.g. [2]):

$$I(y, v_0) = \frac{1}{\sqrt{2g}} \int_0^a \sqrt{\frac{1 + y'(x)^2}{\alpha - y(x)}}\, \mathrm{d}x, \tag{31}$$

where $\alpha = v_0^2/2g$. The set of admissible functions $\mathcal{A}$ consists of continuous functions joining the initial and final point, with a piecewise continuous derivative, and such that $y(x) < \alpha$ for $x \in [0, a]$. Let

$$I^* = \inf_{y \in \mathcal{A}} I(y, v_0),$$

which exists and is attained by the cycloid joining the initial and final points ([2]).

If the initial angle θ_1 is specified, let

$$\mathcal{A}_{\theta_1} = \{y \in \mathcal{A} \,:\, y'(0) = \tan\theta_1\},$$

and

$$I_{\theta_1} = \inf_{y \in \mathcal{A}_{\theta_1}} I(y, v_0).$$

Since $\mathcal{A}_{\theta_1} \subset \mathcal{A}$ we have that

$$I_{\theta_1} \geq I^*. \tag{32}$$

We further show now that $I_{\theta_1} = I^*$! We do this by constructing a sequence in $\mathcal{A}_{\theta_1}$ with time integrals converging to I^*.

Let $y_c(x)$ be the cycloid yielding the minimum time of descent I^*. For any $\varepsilon > 0$, let $H_\varepsilon(x)$ be the cubic Hermite polynomial such that

$$H_\varepsilon(0) = y_c(0), \quad H'_\varepsilon(0) = \tan\theta_1, \quad H_\varepsilon(\varepsilon) = y_c(\varepsilon), \quad H'_\varepsilon(\varepsilon) = y'_c(\varepsilon),$$

and define

$$y_\varepsilon(x) = \begin{cases} H_\varepsilon(x) & , \quad 0 \leq x \leq \varepsilon, \\ y_c(x) & , \quad \varepsilon < x \leq a. \end{cases} \tag{33}$$

By construction $y_\varepsilon \in C^1[0,a]$ and joins the initial and final given points. One can easily show now by looking at the formula defining $H_\varepsilon(x)$, that for ε sufficiently small $y_\varepsilon \in \mathcal{A}$, i.e., $y_\varepsilon(x) < \alpha$ for all $x \in [0,a]$. Hence $y_\varepsilon \in \mathcal{A}_{\theta_1}$. Again, working with the expression for H_ε, one can show that there exists a constant $M > 0$ independent of ε such that

$$\max_{x\in[0,a]} |y_\varepsilon(x)| \leq M, \quad \max_{x\in[0,a]} |y'_\varepsilon(x)| \leq M. \tag{34}$$

Thus we have now that

$$\begin{aligned} I(y_\varepsilon, v_0) &= \frac{1}{\sqrt{2g}} \int_0^a \sqrt{\frac{1 + y'_\varepsilon(x)^2}{\alpha - y_\varepsilon(x)}}\, \mathrm{d}x, \\ &= \frac{1}{\sqrt{2g}} \int_0^\varepsilon \sqrt{\frac{1 + H'_\varepsilon(x)^2}{\alpha - H_\varepsilon(x)}}\, \mathrm{d}x + \frac{1}{\sqrt{2g}} \int_\varepsilon^a \sqrt{\frac{1 + y'_c(x)^2}{\alpha - y_c(x)}}\, \mathrm{d}x. \end{aligned}$$

It follows from (34) that

$$\lim_{\varepsilon \to 0} \int_0^\varepsilon \sqrt{\frac{1 + H'_\varepsilon(x)^2}{\alpha - H_\varepsilon(x)}}\, \mathrm{d}x = 0.$$

Using this in the previous identity we get that

$$\begin{aligned} \lim_{\varepsilon \to 0} I(y_\varepsilon, v_0) &= \lim_{\varepsilon \to 0} \frac{1}{\sqrt{2g}} \int_\varepsilon^a \sqrt{\frac{1 + y'_c(x)^2}{\alpha - y_c(x)}}\, \mathrm{d}x, \\ &= \frac{1}{\sqrt{2g}} \int_0^a \sqrt{\frac{1 + y'_c(x)^2}{\alpha - y_c(x)}}\, \mathrm{d}x = I^*. \end{aligned}$$

Combining this with (32) we get that $I_{\theta_1} = I^*$.

We can answer now both of the questions posed at the beginning of this section. Even when the compatibility condition (30) is not satisfied, there is a minimum time of descent which equals the minimum time of descent for the problem with initial speed v_0 disregarding the initial angle θ_1. However the minimum time of descent for the problem with both v_0 and θ_1 specified, is not attained, i.e., there is no minimizer. This is because a minimizer in $\mathcal{A}_{\theta_1}$ must satisfy the corresponding Euler–Lagrange equations of (31), the solutions of which are cycloids *satisfying the compatibility condition.*

8 Comments and Conclusions

If the compatibility condition (30) is not satisfied, then one can not specify in an arbitrary fashion the initial velocity of the particle. Since the extremals for our problem are given by (20), the candidates for the curve of minimum descent are always sections of a cycloid. When the initial velocity of the particle is such that the compatibility condition is not satisfied, then there is no extremal in the family (20) that connects the initial point to the final one. Thus the cycloids parametrized by γ in (20) either *overshoot* or *undershoot* the desired target at (a, b). There still a minimum time of descent (which equals the minimum time for the problem with the given initial speed disregarding the given initial angle of inclination) but there is no minimizer. Our analysis shows that if the initial angle is specified, then the initial speed can be adjusted to reach the required target in minimum time; or if the initial speed is given, then the initial angle can be adjusted to reach the required target in minimum time as well.

The problem in which the initial velocity of the bead does not necessarily lie in the vertical plane containing the initial and final points is another example of a problem of the calculus of variations in which the infimum exists but there is no minimizer. In this case the competing functions are three dimensional curves. However their projections onto the vertical plane containing the initial and final points, and the acceleration of gravity, are competing functions for the planar problem. Consequently the time of descent along the three dimensional curve must be greater than or equal than the optimal time of descent for the planar problem. One can now construct a sequence of three dimensional curves satisfying the initial velocity condition and whose times of descent approach the optimal one for the planar problem. (This last construction is similar to the one for the infimizing sequence (33), in which the three dimensional curves are essentially planar but with a little bump out of the plane at the beginning to satisfy the initial velocity condition.) However no planar curve can satisfy the condition with the initial velocity out of the vertical plane containing the initial and final points and consequently there is no minimizer.

In 1744, Euler solved a variation of the brachistochrone problem in which friction is included as a nonlinear function of the square of the speed of the bead. Although his solution was not explicit, it showed that the curve of minimum descend is no longer a cycloid as in the problem without friction. Ashby, Brittin, Love, and Wyss [1] and Lipp [4] have considered variations of the problem in which friction is either a linear function or the absolute value, of the component of the force acting on the particle that is normal to the curve (including the term corresponding to the acceleration in the direction of the normal). The question of whether or not there is some kind of compatibility condition for the initial velocity when friction is present shall be pursued elsewhere.

Acknowledgement: This research was sponsored in part by the National Security Agency (NSA) under grant number H98230-04-C-0486. Informal discussions with Errol Montes–Pizarro were very helpful.

References

[1] N. Ashby, W. E. Brittin, W. F. Love, and W. Wyss. Brachistochrome with Coulomb Friction. *American Journal of Physics*, 43(10):902–906, 1975.

[2] G. A. Bliss. Calculus of Variations. *The Mathematical Association of America*, The Open Court Publishing Co., Chicago, Illinois, 1925.

[3] L. Hawn and T. Kiser. Exploring the Brachistochrome Problem. *The American Mathematical Monthly*, 102(4):328–336, 1995.

[4] S. C. Lipp. Brachistochrome with Coulomb Friction. *SIAM J. Control Optim.*, 35(2):562–584, 1997.

[5] A. S. Parnovsky. Some Generalizations of Brachistochrone Problem. *Acta Physica Polonica*, A93 Supplement, S-55, 1998.

[6] H. Sagan. *Introduction to the Calculus of Variations.* Dover Publications, Inc., New York, 1992.

[7] J. L. Troutman. *Variational Calculus with Elementary Convexity.* Springer–Verlag, New York, 1983.

[8] G. J. Tee. Isochrones and Brachistochrones. *Neural, Parallel and Scientific Computations, Dynamic Publishers Inc.*, 7:311-342, 1999.