Commun. Math. Anal. Appl.
doi: 10.4208/cmaa.2021-0005

Vol. **1**, No. 1, pp. 213-240
February 2022

On the Brachistochrone Problem

Philippe G. Ciarlet and Cristinel Mardare*

Department of Mathematics, City University of Hong Kong, Hong Kong, SAR, China.

Received 11 December 2021; Accepted 21 December 2021

Abstract. As a problem that dates back to the end of the seventieth century, the brachistochrone problem is one of the oldest problems in the calculus of variations and as such, has generated a myriad of publications. However, in most classical texts and in most papers, the favored way to solve the problem is to make two a priori assumptions, viz., that the brachistochrone lies in a vertical plane, and that it can be represented as the graph of a function in this plane; besides, with few exceptions, the existence of a solution is not rigorously established: instead it is sometimes even taken for granted that the solution is that of the associated Euler-Lagrange equations, even though these are well known to be only necessary conditions for the existence of a minimizer. The objective of this article is to show how all these shortcomings can be very simply, and rigorously, overcome, by means of arguments that do not need any a priori assumptions and that otherwise require only a modicum of basic notions from calculus, so as to rigorously establish the existence and uniqueness of the brachistochrone in full generality. One originality of our approach is that from the outset we seek the brachistochrone as a parameterized curve in the three-dimensional space, i.e., that can be represented by means of three parametric equations, instead of by means of a single graph in a vertical plane. Contrary to expectations, this increase of generality renders the ongoing analysis much simpler. Our objective is thus to show how the direct method of the calculus of variations based on the Euler-Lagrange equations can be used to solve the brachistochrone problem. Otherwise, there are other ways to solve this problem, for instance by means of convex optimization or optimal control; such methods are briefly described at the end of the paper.

AMS subject classifications: 49-01, 49K15

---

*Corresponding author. *Email addresses:* `mapgc@cityu.edu.hk` (P. G. Ciarlet), `cmardare@cityu.edu.hk` (C. Mardare)

Key words: Calculus of variations, Euler-Lagrange equations, existence of a minimizer, uniqueness of a minimizer, brachistochrone.

1 Introduction

Consider the following mechanical problem: A material point with mass m is subjected to uniform gravity and slides without friction along a curve joining a point $\boldsymbol{P}_0$ in the "horizontal plane" to a different point $\boldsymbol{P}_1$ situated below, or in, the horizontal plane.

Under the assumption that the velocity at $\boldsymbol{P}_0$ at the initial time vanishes, the brachistochrone problem consists in seeking whether there exists a smooth curve along which the time for the material point to go from $\boldsymbol{P}_0$ to $\boldsymbol{P}_1$ is the shortest (Fig. 1). If such a curve exists, it is called a brachistochrone ("brákhistos" means "shortest" and "khrónos" means "time" in ancient Greek).

This minimization problem was first proposed as an open one to his contemporaries by Johann Bernoulli in 1696 (cf. [4]). Answers were then quickly proposed by several outstanding mathematicians: Isaac Newton, Gottfried Wilhelm Leibniz, Johann Bernoulli, Jacob Bernoulli (brother of Johann), and Guillaume de l'Hôpital. An account of the history of the brachistochrone can be found in, e.g., Shafer [10], or Sussman & Willems [11].

This problem constitutes one of the oldest ones of the calculus of variations and has of course generated a myriad of publications. So, why write an additional one?

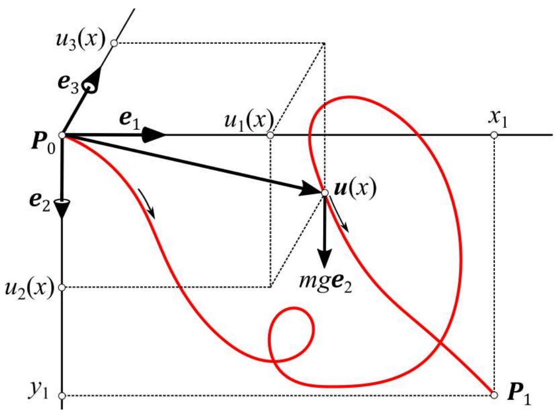

Figure 1: Given a point $\boldsymbol{P}_0$ chosen as the origin of the "horizontal plane" and a point $\boldsymbol{P}_1 = (x_1, y_1, 0)$ below $(y_1 > 0)$, or in $(y_1 = 0)$, the horizontal plane, the brachistochrone, if it exists, is a smooth curve along which the time for a material point with mass m sliding without friction along this curve under the influence of gravity, and with a zero velocity at $\boldsymbol{P}_0$, is the shortest for joining $\boldsymbol{P}_0$ to $\boldsymbol{P}_1$.

The most compelling reason for doing so is that the various "solutions" for solving this problem found in the existing literature by means of the Euler-Lagrange equations are either incomplete, or not satisfactory. This is all the more surprising, since solving this problem does not require any advanced techniques commonly used in the calculus of variations, such as convexity and coercivity, weak sequential lower semi-continuity, weak compactness in reflexive Banach spaces, etc. Instead it only requires a good knowledge of basic notions from calculus, such as ordinary differential equations, the fundamental formula of integral calculus, or how to compute the derivative of an inverse function of one variable.

Here is a list of commonly encountered shortcomings:

Perhaps the most serious one is that it is most of the time blithely taken for granted that the solution is that of the associated Euler-Lagrange equations, even though these are well known to be only necessary conditions for the existence of a minimizer. Incidentally, the issue of uniqueness of the solution is also seldom addressed. Likewise, it is often assumed from the outset, i.e., without a proof, that the brachistochrone lies in the vertical plane; that the Euler-Lagrange equations can be applied, even though the minimization problem is posed over a set that is not a vector space; and finally, that the brachistochrone can be represented as the graph of a function.

Our purpose in this article is to see how all these shortcomings can be very simply overcome. More specifically, we begin by carefully describing in Section 2 a natural mathematical model for this problem, which essentially relies on the well-known law of conservation of energy. One originality of our approach is that we seek the brachistochrone as a three-dimensional parameterized curve, i.e., that can be represented by means of three parametric equations, instead of by means of a graph in a vertical plane. Contrary to expectations, this increase of generality renders the ongoing analysis much simpler. More specifically, it is found that, if it exists, the brachistochrone should minimize a functional that is not quadratic over a set of "admissible parameterized curves" that is not a vector space, which indicates that the resulting Euler-Lagrange equations should be nonlinear. At this stage, there are thus three unknowns, the three components of the unknown parameterized curve.

For convenience, our proof is broken into a series of seven lemmas, which constitute Section 3. We first show in Lemma 3.1 by means of a very simple argument that the brachistochrone, if it exists, must lie in a vertical plane. This observation thus reduces the number of unknowns to two.

We then show in Lemma 3.2 that any would-be minimizer must satisfy two Euler-Lagrange equations, even though the set of admissible curves is not a vector space; note that there are two such Euler-Lagrange equations because the minimizer is sought as a parameterized curve in the vertical plane, instead of as a graph, in which case there would be only one Euler-Lagrange equation.

It is then established in Lemma 3.3 that, interestingly, the first Euler-Lagrange equation implies that the brachistochrone is a graph, in that the vertical component u is a function of the horizontal component. This observation, which seems to be new (it results from our considering parameterized curves from the outset), reduces the number of unknowns to one.

It is further established in Lemma 3.3 that the single unknown function should satisfy a two-point boundary value problem of a specific form, where the nonlinear ordinary differential equation is of the first order only, but fortunately the right-hand side is an unknown constant.

After showing in Lemma 3.4 that the second Euler-Lagrange equation provides some essential information regarding the sign of the derivative of the unknown function u (this information is usually overlooked, which results in an incomplete proof), we then solve in Lemma 3.5 this two-point boundary value problem by means of a suitable change of the unknown function, introduced here in a natural way, instead of being "handed from the sky" as is often the case. Because we show that the solution to this boundary value problem is unique, the uniqueness of the brachistochrone, if it exists, is thereby established.

We then establish the existence of the brachistochrone, which is achieved in two stages, depending on whether the point $\boldsymbol{P}_1$ is in the same horizontal plane as $\boldsymbol{P}_0$ or not. First, if the point $\boldsymbol{P}_1$ lies in the horizontal plane containing $\boldsymbol{P}_0$, then the existence of the brachistochrone is established in Lemma 3.6 by using an elegant argument due to Benson [3], based on an ingenious use of Cauchy-Schwarz inequality. Second, if the point $\boldsymbol{P}_1$ lies under the horizontal plane containing $\boldsymbol{P}_0$, then the existence of the brachistochrone is established in Lemma 3.7 by using a very simple (and apparently new) argument based on the existence result of Lemma 3.6.

All the results above are then assembled in a single theorem, the statement of which concludes Section 3.

This article thus gives a complete and elementary proof for the existence and uniqueness of the brachistochrone in full generality, which means in particular that we do not assume that the brachistochrone must be a graph in a vertical plane as is almost invariably the case in the existing literature. The only exceptions to this rule that we know of are the articles of Brookfield [5] and Sussmann & Willems [11]. Indeed, Brookfield's proof is much shorter, but it implicitly relies on the parametric equations of the brachistochrone (which can only be found by using Euler-Lagrange equations as in this article). Sussmann & Willems's proof is not elementary (it relies on a maximum principle in optimal control theory;

besides, most of the proof is "left to the reader"), by contrast to that of the present article.

Otherwise, the existing literature provides several different proofs for the existence and uniqueness of the brachistochrone, i.e., that do not rely on the Euler-Lagrange equations, but under the a priori assumption that the brachistochrone is a graph in the vertical plane containing the points $\boldsymbol{P}_0$ and $\boldsymbol{P}_1$. Such proofs are achieved, e.g., by convex optimization as in Kosmol [9] and Troutman [12] (see also Ball [2]), thanks to a change of the unknown function defining the graph of functions representing the curves; or by using the theory of parametric integrals as in Cesari [6]; or by a clever choice of the unknown function combined with Cauchy-Schwarz inequality as in Hrusa & Troutman [8], but this proof only applies under an additional assumption ensuring that the brachistochrone is contained in the first half of the cycloid (the above list is by no means exhaustive)

For completeness, these different approaches are briefly reviewed in Section 4.

2 Mathematical formulation of the brachistochrone problem; cf. Fig. 1

Let $\boldsymbol{e}_i$, $1\leq i\leq 3$, denote three mutually orthogonal unit vectors of the Euclidean space $\mathbb{E}^3$, chosen in such a way that the "horizontal plane" is that spanned by the vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_3$, while the "downward vertical direction" is that of the vector $\boldsymbol{e}_2$. Without loss of generality, we will assume that the point $\boldsymbol{P}_0$ is the origin of $\mathbb{E}^3$ and that the point $\boldsymbol{P}_1$ is a point with positive abscissa in the "vertical plane" spanned by the vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, i.e., that

$$\boldsymbol{P}_0=(0,0,0) \text{ and } \boldsymbol{P}_1=(x_1,y_1,0) \text{ for some constants } x_1>0 \text{ and } y_1\geq 0.$$

In what follows, the notation $\mathbb{E}^3$, $\boldsymbol{a}\cdot\boldsymbol{b}$, and $|\boldsymbol{a}|:=\sqrt{\boldsymbol{a}\cdot\boldsymbol{a}}$, where $\boldsymbol{a},\boldsymbol{b}\in\mathbb{E}^3$, respectively designate the three-dimensional Euclidean space, the Euclidean inner product, and the Euclidean norm, in $\mathbb{E}^3$. The notation $'$ and $''$ respectively designate the derivative and the second derivative of functions or vector fields of one variable.

Consider a material point of mass m that slides without friction along a curve $\boldsymbol{v}=(v_i)_{i=1}^3:[\theta_0,\theta_1]\to\mathbb{R}^3$ joining two points $\boldsymbol{P}_0$ to $\boldsymbol{P}_1$ in $\mathbb{E}^3$, parameterized by means of a parameter θ varying in some interval $[\theta_0,\theta_1]$ with $\theta_0<\theta_1$. We will assume that

$$\boldsymbol{v}\in\mathcal{C}^2\big([\theta_0,\theta_1];\mathbb{E}^3\big),\quad \boldsymbol{v}'(\theta_0)=\boldsymbol{0},\quad |\boldsymbol{v}'(\theta)|=\left(\sum_{i=1}^3|v_i'(\theta)|^2\right)^{\frac{1}{2}}>0,\quad \theta_0<\theta<\theta_1,$$

the last condition ensuring that a non-zero tangent vector is well-defined along the curve, except possibly at its end-points. It will be convenient to identify the curve with the image $\boldsymbol{v}([\theta_0,\theta_1])$ in $\mathbb{E}^3$ of the interval $[\theta_0,\theta_1]$ under the mapping $\boldsymbol{v}$, i.e., the set $\{\sum_{i=1}^3 v_i(\theta)\boldsymbol{e}_i \in \mathbb{E}^3; \theta_0 \le \theta \le \theta_1\}$.

During its movement along the curve, the position of the material point at any time $t\in[0,T]$, where $T>0$ is the time it takes for the material point to reach the point $\boldsymbol{P}_1$, is of the form

$$\boldsymbol{P}(t) = \big(P_i(t)\big)_{i=1}^3 = \boldsymbol{v}\big(\chi(t)\big), \quad 0\le t\le T,$$

where $\chi(t)$ designates the value of the parameter $\theta\in[\theta_0,\theta_1]$ at the time $t\in[0,T]$. Assume that

$$\chi\in\mathcal{C}^2[0,T], \quad \chi(0)=\theta_0, \quad \chi(T)=\theta_1, \quad \text{and} \quad \chi'(t)>0 \quad \text{for all} \quad 0<t<T.$$

Because the material point is assumed to slide without friction, the law of conservation of energy asserts that the sum of the kinetic energy $\frac{1}{2}m|\boldsymbol{P}'(t)|^2$ and the potential energy $-mgP_2(t)$, where g denotes the gravitational constant, is a constant function of the time $t\in[0,T]$. Since $\boldsymbol{P}'(0)=\boldsymbol{0}$ and $v_2(\theta_0)=0$ by assumption, it thus follows that

$$\frac{1}{2}m\big|\boldsymbol{P}'(t)\big|^2 = mgP_2(t), \quad 0\le t\le T.$$

Since by assumption $\chi'(t)>0$, $0<t<T$, the resulting relation

$$\big|\boldsymbol{P}'(t)\big| = \big|\boldsymbol{v}'(\chi(t))\chi'(t)\big| = \big|\boldsymbol{v}'(\chi(t))\big|\chi'(t) = \sqrt{2gv_2(\chi(t))}, \quad 0<t<T,$$

combined with the assumption $|\boldsymbol{v}'(\theta)|>0$, $\theta_0<\theta<\theta_1$, implies that, necessarily,

$$v_2(\theta)>0 \quad \text{at each} \quad \theta=\chi(t), \quad 0<t<T.$$

Noting that the function $\chi:[0,T]\to[\theta_0,\theta_1]$ is one-to-one and onto and differentiable with a strictly positive derivative on the open interval $(0,T)$, and that its inverse function χ^{-1} is thus differentiable on the open interval (θ_0,θ_1) with a derivative given by

$$(\chi^{-1})'(\theta) = \frac{1}{\chi'(\chi^{-1}(\theta))}, \quad \theta_0<\theta<\theta_1,$$

we infer that

$$\begin{aligned} T &= \chi^{-1}(\theta_1) - \chi^{-1}(\theta_0) = \int_{\theta_0}^{\theta_1} \big(\chi^{-1}\big)'(\theta)\,\mathrm{d}\theta \\ &= \int_{\theta_0}^{\theta_1} \frac{1}{\chi'(\chi^{-1}(\theta))}\,\mathrm{d}\theta = \int_{\theta_0}^{\theta_1} \frac{|\boldsymbol{v}'(\theta)|}{\sqrt{2gv_2(\theta)}}\,\mathrm{d}\theta. \end{aligned}$$

Observe that the time T does not depend on the function χ; as expected, T only depends on the parameterized curve $\boldsymbol{v}$.

The brachistochrone problem consists in seeking whether the integral

$$J(\boldsymbol{v}) := \int_{\theta_0}^{\theta_1} \frac{|\boldsymbol{v}'(\theta)|}{\sqrt{2gv_2(\theta)}} \mathrm{d}\theta = \int_{\theta_0}^{\theta_1} \sqrt{\frac{v_1'(\theta)^2 + v_2'(\theta)^2 + v_3'(\theta)^2}{2gv_2(\theta)}} \mathrm{d}\theta,$$

which thus measures the time it takes the material point to go from $\boldsymbol{P}_0$ to $\boldsymbol{P}_1$ along the parameterized curve $\boldsymbol{v}$, attains its infimum (this integral is obviously bounded below, by zero), and whether the minimizer is unique if it exists, when the vector fields $\boldsymbol{v} = (v_i)_{i=1}^3$ vary in a specific set $\boldsymbol{V}$ of parameterized curves, called the admissible curves, that satisfy the various conditions enumerated above, thus accordingly defined by

$$\begin{aligned} \boldsymbol{V} := \Big\{ & \boldsymbol{v} = (v_i)_{i=1}^3 \in \mathcal{C}^2\big([\theta_0,\theta_1];\mathbb{R}^3\big); \boldsymbol{v}(\theta_0) = (0,0,0), \boldsymbol{v}(\theta_1) = (x_1,y_1,0), \\ & \big|\boldsymbol{v}'(\theta)\big| > 0 \ \text{ and } \ v_2(\theta) > 0, \theta_0 < \theta < \theta_1 \Big\}. \end{aligned}$$

In other words, we look for a parameterized curve $\boldsymbol{u}$ such that

$$\boldsymbol{u} = (u_i)_{i=1}^3 \in \boldsymbol{V}, \quad J(\boldsymbol{u}) = \inf_{\boldsymbol{v} \in \boldsymbol{V}} J(\boldsymbol{v}).$$

If such a curve $\boldsymbol{u}$ exists and is unique, it is called the brachistochrone for joining the point $\boldsymbol{P}_0$ to the point $\boldsymbol{P}_1$.

Note that it may very well happen that $J(\boldsymbol{v}) = +\infty$ for some elements of $\boldsymbol{V}$. But of course, we are only interested in those $\boldsymbol{v} \in \boldsymbol{V}$ that satisfy $J(\boldsymbol{v}) < +\infty$, since the brachistochrone that we will eventually find fortunately fulfills this condition.

An important preliminary observation is that, when it is finite, the above integral $J(\boldsymbol{v})$ is invariant under changes of parameterizations "up to $\mathcal{C}^1$-diffeomorphisms": this means that, if the parameter $\theta \in [\theta_0,\theta_1]$ is replaced by another parameter $\lambda \in [\lambda_0,\lambda_1]$ in such a way that $\theta = \varphi(\lambda)$, where the function $\varphi : [\lambda_0,\lambda_1] \to [\theta_0,\theta_1]$ is one-to-one, onto, and continuously differentiable over $[\lambda_0,\lambda_1]$ with a derivative satisfying $\varphi'(\lambda) > 0$ for all $\lambda \in [\lambda_0,\lambda_1]$, then the functions w_i defined by $w_i(\lambda) := v_i(\varphi(\lambda))$, $\lambda_0 \le \lambda \le \lambda_1$, satisfy

$$\int_{\lambda_0}^{\lambda_1} \sqrt{\frac{v_1'(\lambda)^2 + v_2'(\lambda)^2 + v_3'(\lambda)^2}{2gw_2(\lambda)}} \mathrm{d}\lambda = \int_{\theta_0}^{\theta_1} \sqrt{\frac{v_1'(\theta)^2 + v_2'(\theta)^2 + v_3'(\theta)^2}{2gv_2(\theta)}} \mathrm{d}\theta$$

as is immediately verified.

In other words, the minimizing curve that we are seeking is in effect an equivalence class of all its parameterizations up to $\mathcal{C}^1$-diffeomorphisms between the intervals of definition of their parameters; consequently, we have the freedom to chose any representative in the equivalence class of a minimizer. This observation will be perfectly illustrated in the statement of the final existence and uniqueness theorem (Theorem 3.1), where the parameter $\theta \in [\theta_0, \theta_1]$ used for a "generic" parameterized curve $\boldsymbol{v}$ of the space $\boldsymbol{V}$ will be replaced by a specific parameter $\tau \in [0, \tau_1]$ used for the parameterized curve $\boldsymbol{u}$ that minimizes the functional J.

3 Solution to the brachistochrone problem

We now turn to the proof of the existence and uniqueness of the brachistochrone, which will be established in a series of lemmas (Lemma 3.1 to Lemma 3.7).

To begin with, we show that, if it exists, the brachistochrone necessarily lies in the "vertical" plane spanned by the vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ (Lemma 3.1). We then observe that, even though some assumptions classically made in the calculus of variations (see, e.g., the excellent textbook of Dacorogna [7]) are not satisfied by the above mathematical formulation of the brachistochrone problem (in particular, the set $\boldsymbol{V}$ is not a vector space), Euler-Lagrange equations can still be derived (Lemma 3.2), and that, again if it exists, the minimizer is necessarily the graph of a uniquely defined segment of cycloid (Lemmas 3.3 to 3.5). Following the ingenious approach due to Benson [3], we then show that, in the special case $y_1 = 0$ (i.e., when $\boldsymbol{P}_1$ is in the same horizontal plane as $\boldsymbol{P}_0$), an explicit lower bound for $\inf_{\boldsymbol{v} \in \boldsymbol{V}} J(\boldsymbol{v})$ can be found, and furthermore that this lower bound is attained, precisely when the admissible curve is the segment of the cycloid obtained by means of the Euler-Lagrange equations. Together, these two results establish the existence and uniqueness of a minimizer when $y_1 = 0$ (Lemma 3.6). Finally, we show how to extend the existence and uniqueness results to the case where $y_1 > 0$ (Lemma 3.7), and we state the final existence and uniqueness theorem (Theorem 3.1).

In Lemmas 3.1 to 3.5, the notation $\boldsymbol{u} = (u_i)_{i=1}^3 \in \boldsymbol{V}$ designates a solution, assumed to exist at this stage, to the minimization problem.

Lemma 3.1. *The curve $\boldsymbol{u}([\theta_0, \theta_1])$ lies in the plane spanned by the vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, i.e., $u_3 = 0$.*

Proof. Assume on the contrary that the function u_3 does not vanish on $[\theta_0, \theta_1]$. Since then $(u_1, u_2, \frac{1}{2} u_3)$ also belongs to the set $\boldsymbol{V}$ and (recall that $u_3(\theta_0) = u_3(\theta_1) = 0$)

$|u_3'|>0$ on an interval of length >0 contained in $[\theta_0,\theta_1]$, we have

$$\int_{\theta_0}^{\theta_1}\sqrt{\frac{u_1'(\theta)^2+u_2'(\theta)^2+(\frac{1}{2}u_3'(\theta))^2}{2gu_2(\theta)}}\mathrm{d}\theta<\int_{\theta_0}^{\theta_1}\sqrt{\frac{u_1'(\theta)^2+u_2'(\theta)^2+u_3'(\theta)^2}{2gu_2(\theta)}}\mathrm{d}\theta,$$

which contradicts that $\boldsymbol{u}$ is a solution to the minimization problem. Hence,

$$u_3(\theta)=0 \quad \text{for all} \quad \theta_0\leq\theta\leq\theta_1.$$

The proof is complete. $\square$

Taking into account that $u_3=0$, we will henceforth identify a solution $\boldsymbol{u}$ to the minimization problem with a vector field-still denoted by $\boldsymbol{u}$ for convenience-with only two components, u_1 and u_2.

Lemma 3.2. *The vector field $(u_i)_{i=1}^2$ satisfies the Euler-Lagrange equations associated with the functional J, which take the form*

$$\begin{cases} -\dfrac{\mathrm{d}}{\mathrm{d}\theta}\left\{\dfrac{u_1'(\theta)}{|\boldsymbol{u}'(\theta)|\sqrt{u_2(\theta)}}\right\}=0, & \theta_0<\theta<\theta_1, \\ -\dfrac{\mathrm{d}}{\mathrm{d}\theta}\left\{\dfrac{u_2'(\theta)}{|\boldsymbol{u}'(\theta)|\sqrt{u_2(\theta)}}\right\}-\dfrac{1}{2}\dfrac{|\boldsymbol{u}'(\theta)|}{(u_2(\theta))^{\frac{3}{2}}}=0, & \theta_0<\theta<\theta_1. \end{cases}$$

Proof. Define the Lagrangian $\mathcal{L}$ by

$$\mathcal{L}(\boldsymbol{v},\boldsymbol{\xi}):=\sqrt{\frac{\xi_1^2+\xi_2^2}{v_2}}$$

for each $\boldsymbol{v}=(v_i)_{i=1}^2\in\mathbb{R}^2$ with $v_2>0$ and $\boldsymbol{\xi}=(\xi_i)_{i=1}^2\in\mathbb{R}^2$.

Let a vector field $\boldsymbol{\varphi}=(\varphi_i)_{i=1}^2$ be given, where each function φ_i belongs to the space $\mathcal{C}^1[\theta_0,\theta_1]$ and has a compact support in the open interval (θ_0,θ_1). Then a simple compactness-continuity argument shows that there exists $\varepsilon_0(\boldsymbol{\varphi})>0$ such that $u_2(\theta)+\varepsilon\varphi_2(\theta)>0$ and $|(\boldsymbol{u}+\varepsilon\boldsymbol{\varphi})'(\theta)|>0$ for all $\theta_0<\theta<\theta_1$ and for all $|\varepsilon|\leq\varepsilon_0(\boldsymbol{\varphi})$, i.e., such that

$$(\boldsymbol{u}+\varepsilon\boldsymbol{\varphi})\in\boldsymbol{V} \quad \text{for all} \quad |\varepsilon|\leq\varepsilon_0(\boldsymbol{\varphi}).$$

By construction, the function

$$f:\varepsilon\in\big(-\varepsilon_0(\boldsymbol{\varphi}),\varepsilon_0(\boldsymbol{\varphi})\big)\to f(\varepsilon):=J(\boldsymbol{u}+\varepsilon\boldsymbol{\varphi}),$$

which is easily seen to be differentiable, has a minimum at $\varepsilon=0$. Therefore,

$$f'(0)=0.$$

This implies that the classical argument used when a functional is to be minimized over a vector space (in which case ε varies in $\mathbb{R}$ instead of only in an interval) can be re-used verbatim from this point on; for this reason, it is only briefly sketched below (for details see, e.g., Dacorogna [7, Chapter 2]). Noting that

$$f'(0)=\frac{1}{\sqrt{2g}}\int_{\theta_0}^{\theta_1}\sum_{i=1}^{2}\left\{\varphi_i'(\theta)\frac{\partial\mathcal{L}}{\partial\xi_i}\big(\boldsymbol{u}(\theta),\boldsymbol{u}'(\theta)\big)+\varphi_i(\theta)\frac{\partial\mathcal{L}}{\partial v_i}\big(\boldsymbol{u}(\theta),\boldsymbol{u}'(\theta)\big)\right\}\mathrm{d}\theta,$$

we thus infer by integrating by parts that

$$\int_{\theta_0}^{\theta_1}\sum_{i=1}^{2}\left\{-\frac{\mathrm{d}}{\mathrm{d}\theta}\left\{\frac{\partial\mathcal{L}}{\partial\xi_i}\big(\boldsymbol{u}(\theta),\boldsymbol{u}'(\theta)\big)\right\}+\frac{\partial\mathcal{L}}{\partial v_i}\big(\boldsymbol{u}(\theta),\boldsymbol{u}'(\theta)\big)\right\}\varphi_i(\theta)\,\mathrm{d}\theta=0$$

for all functions $\varphi_i\in\mathcal{C}^1[\theta_0,\theta_1]$, $i=1,2$, with compact support in (θ_0,θ_1).

The fundamental lemma of the calculus of variations therefore implies that the following Euler-Lagrange equations hold:

$$\begin{cases}-\dfrac{\mathrm{d}}{\mathrm{d}\theta}\left\{\dfrac{\partial\mathcal{L}}{\partial\xi_1}\big(\boldsymbol{u}(\theta),\boldsymbol{u}'(\theta)\big)\right\}+\dfrac{\partial\mathcal{L}}{\partial v_1}\big(\boldsymbol{u}(\theta),\boldsymbol{u}'(\theta)\big)=0, & \theta_0<\theta<\theta_1,\\ -\dfrac{\mathrm{d}}{\mathrm{d}\theta}\left\{\dfrac{\partial\mathcal{L}}{\partial\xi_2}\big(\boldsymbol{u}(\theta),\boldsymbol{u}'(\theta)\big)\right\}+\dfrac{\partial\mathcal{L}}{\partial v_2}\big(\boldsymbol{u}(\theta),\boldsymbol{u}'(\theta)\big)=0, & \theta_0<\theta<\theta_1.\end{cases}$$

The conclusion then follows by using the specific form of the Lagrangian $\mathcal{L}$. $\square$

Lemma 3.3. *The first Euler-Lagrange equation implies that the vector field $(u_i)_{i=1}^2$ is the graph of a function $u:[0,x_1]\to\mathbb{R}$ in the following sense: given any $x\in[0,x_1]$, there exists a unique $\theta\in[\theta_0,\theta_1]$ such that*

$$\big(u_1(\theta),u_2(\theta)\big)=\big(x,u(x)\big).$$

Besides, the function u is in the space $\mathcal{C}[0,x_1]\cap\mathcal{C}^2(0,x_1)$, and necessarily satisfies

$$\begin{aligned}&\big(1+u'(x)^2\big)u(x)=2R, \quad 0<x<x_1,\\ &u(0)=0, \quad u(x_1)=y_1\end{aligned}$$

for some constant R unknown at this stage that satisfies either $R\geq\frac{y_1}{2}$ if $y_1>0$ or $R>0$ if $y_1=0$.

Proof. The first Euler-Lagrange equation obviously implies that there exists a constant C such that

$$\frac{u_1'(\theta)}{|\boldsymbol{u}'(\theta)|\sqrt{u_2(\theta)}}=C,\quad \theta_0<\theta<\theta_1.$$

If $C\leq 0$, then $u_1'(\theta)\leq 0, \theta_0<\theta<\theta_1$, and thus

$$\int_{\theta_0}^{\theta_1} u_1'(\theta)\,\mathrm{d}\theta = u_1(\theta_1)-u_1(\theta_0)=x_1\leq 0,$$

in contradiction with the assumption $x_1>0$. Consequently, the only possibility is that $C>0$, which implies that

$$u_1'(\theta)>0,\quad \theta_0<\theta<\theta_1.$$

Besides,

$$C\leq \frac{1}{\sqrt{\sup_{\theta_0\leq\theta\leq\theta_1} u_2(\theta)}}\leq \frac{1}{\sqrt{y_1}}\quad \text{if}\quad y_1=u_2(\theta_1)>0.$$

The function

$$u_1:\theta\in[\theta_0,\theta_1]\to u_1(\theta)=\int_{\theta_0}^{\theta} u_1'(\psi)\,\mathrm{d}\psi,$$

which belongs to the space $\mathcal{C}^2[\theta_0,\theta_1]$ (since $\boldsymbol{u}\in\boldsymbol{V}$), is thus a bijection from $[\theta_0,\theta_1]$ onto $[0,x_1]$ since it is strictly increasing; besides, its inverse function u_1^{-1} belongs to the space $\mathcal{C}[0,x_1]\cap\mathcal{C}^2(0,x_1)$, with a derivative given by

$$\left(u_1^{-1}\right)'(x)=\frac{1}{u_1'(\theta)}\quad \text{at each}\quad x=u_1(\theta),\quad \theta_0<\theta<\theta_1.$$

Then the function $u:[0,x_1]\to\mathbb{R}$ defined by

$$u(x):=u_2\left(u_1^{-1}(x)\right)\quad \text{at each}\quad x\in[0,x_1],$$

which belongs to the space $\mathcal{C}[0,x_1]\cap\mathcal{C}^2(0,x_1)$, satisfies

$$\begin{aligned} u(x)\left(1+u'(x)^2\right)&=u_2(\theta)\left(1+\left(u_2'(\theta)\frac{1}{u_1'(\theta)}\right)^2\right)=\frac{u_2(\theta)\left(u_1'(\theta)^2+u_2'(\theta)^2\right)}{u_1'(\theta)^2}\\ &=2R:=\frac{1}{C^2}\quad \text{at each}\quad x=u_1(\theta),\quad \theta_0<\theta<\theta_1, \end{aligned}$$

and

$$u(0)=u_2(\theta_0)=0,\quad u(x_1)=u_2(\theta_1)=y_1.$$

The proof is complete. $\square$

Lemma 3.4. *The second Euler-Lagrange equation implies that the derivative u' of the function $u\in\mathcal{C}[0,x_1]\cap\mathcal{C}^2(0,x_1)$ found in Lemma 3.3 has at most one zero in the open interval $(0,x_1)$. Besides,*

- *If u' has no zero in $(0,x_1)$, then $y_1>0$ and $u'(x)>0$ at each $0<x<x_1$.*
- *If u' has a zero $\widetilde{x}\in(0,x_1)$, then $u'(x)>0$ at each $0<x<\widetilde{x}$ and $u'(x)<0$ at each $\widetilde{x}<x<x_1$.*

Proof. The second Euler-Lagrange equation, viz.,

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\left\{\frac{u_2'(\theta)}{|\boldsymbol{u}'(\theta)|\sqrt{u_2(\theta)}}\right\}=-\frac{1}{2}\frac{|\boldsymbol{u}'(\theta)|}{(u_2(\theta))^{\frac{3}{2}}},\quad \theta_0<\theta<\theta_1,$$

obviously implies that, as a strictly decreasing function, the function $\theta\in(\theta_0,\theta_1)\to\frac{u_2'(\theta)}{|\boldsymbol{u}'(\theta)|\sqrt{u_2(\theta)}}$ has at most one zero; consequently, the function $\theta\in(\theta_0,\theta_1)\to u_2'(\theta)$ has also at most one zero.

Since $u_1'(\theta)\neq 0$ and $u'(x)=\frac{u_2'(\theta)}{u_1'(\theta)}$ at each $x=u_1(\theta)$, $\theta_0<\theta<\theta_1$ (cf. Lemma 3.3), the function $x\in(0,x_1)\to u'(x)$ has likewise at most one zero.

If u' has no zero in $(0,x_1)$, then $y_1>0$ (to see this, use Rolle's theorem); therefore, $u'(x)>0$ at each $0<x<x_1$ since the function u is in this case necessarily increasing on the interval $[0,x_1]$.

If $u'(\widetilde{x})=0$ for some $\widetilde{x}\in(0,x_1)$, then $u'(x)$ is necessarily >0 for $x>0$ small enough since $u(0)=u_2(\theta_0)=0$ and $u(x)=u_2(\theta)>0$ at each $x=u_1(\theta)$, $\theta_0<\theta<\theta_1$ (by definition of the set $\boldsymbol{V}$); consequently, $u'(x)>0$ at each $0<x<\widetilde{x}$ and $u'(x)<0$ at each $\widetilde{x}<x<x_1$. $\square$

Lemma 3.5. *The function $u\in\mathcal{C}[0,x_1]\cap\mathcal{C}^2(0,x_1)$ found in Lemma 3.2 is uniquely defined, and is given in parametric form by*

$$\begin{aligned} x&=R(\tau-\sin\tau), && 0\le\tau\le\tau_1,\\ u(x)&=R(1-\cos\tau), && 0\le\tau\le\tau_1, \end{aligned}$$

where (R,τ_1) is the unique solution to the nonlinear system

$$\begin{aligned} R(\tau_1-\sin\tau_1)&=x_1,\\ R(1-\cos\tau_1)&=y_1. \end{aligned}$$

Proof. It was shown in Lemma 3.3 that the function $u \in \mathcal{C}[0,x_1] \cap \mathcal{C}^2(0,x_1)$ necessarily satisfies

$$\begin{aligned} &(1+u'(x)^2)u(x)=2R, \quad 0<x<x_1, \\ &u(0)=0, \quad u(x_1)=y_1, \end{aligned}$$

where the unknown constant R satisfies

$$\begin{cases} R \geq \dfrac{y_1}{2}, & \text{if} \quad y_1>0, \\ R>0, & \text{if} \quad y_1=0. \end{cases}$$

In particular then,

$$0<u(x) \leq 2R, \quad 0<x<x_1.$$

Assume first that the derivative u' does not vanish on the open interval $(0,x_1)$. Since $y_1>0$ and $u'(x)>0$ for all $0<x<x_1$ in this case (Lemma 3.4), the function $u \in \mathcal{C}[0,x_1] \cap \mathcal{C}^2(0,x_1)$ is a bijection from $[0,x_1]$ onto $[0,y_1]$, its inverse function $v:=u^{-1}$ belongs to the space $\mathcal{C}[0,y_1] \cap \mathcal{C}^2(0,y_1)$, and its derivative v' satisfies

$$v'(y)=\frac{1}{u'(x)}>0 \quad \text{at each} \quad y=u(x) \in (0,y_1).$$

Consequently,

$$\left(1+\frac{1}{v'(y)^2}\right)y=2R, \quad 0<y<y_1,$$

or equivalently,

$$(1+v'(y)^2)y=2R\big(v'(y)\big)^2, \quad 0<y<y_1.$$

That the new variable $y=u(x)$ varies in the interval $[0,y_1]$ and that $y_1 \leq 2R$ then suggest a change of variable, by means of the bijection

$$\tau \in [0,\tau_1] \to y=R(1-\cos\tau) \quad \text{at each} \quad 0 \leq y \leq y_1,$$

where, for any given constant $R \geq \frac{y_1}{2}$, the angle τ_1 is uniquely defined by the relations

$$0<\tau_1 \leq \pi, \quad R(1-\cos\tau_1)=y_1.$$

Note that τ_1 is a function of the unknown constant R at this stage. Together, the relations $v'(y)^2=\frac{y}{2R-y}$ and $v'(y)>0, 0<y<y_1$, imply that

$$v'(y)=\sqrt{\frac{y}{2R-y}}=\sqrt{\frac{1-\cos\tau}{1+\cos\tau}} \quad \text{at each} \quad y=R(1-\cos\tau), \quad 0<y<y_1.$$

Let the function $f:[0,\tau_1]\to\mathbb{R}$ be defined by

$$f(\tau):=v\big(R(1-\cos\tau)\big)=v(y)=x,\quad 0\le\tau\le\tau_1,$$

so that $f(0)=0$ and, for each $0<\tau<\tau_1$,

$$f'(\tau)=(R\sin\tau)v'\big(R(1-\cos\tau)\big)=R\sin\tau\sqrt{\frac{1-\cos\tau}{1+\cos\tau}}=R(1-\cos\tau).$$

Consequently,

$$x=R(\tau-\sin\tau),\quad 0\le\tau\le\tau_1.$$

In particular then, $R(\tau_1-\sin\tau_1)=x_1$, which, together with the relation $R(1-\cos\tau_1)=y_1$, uniquely determines the constants $R\ge\frac{y_1}{2}$ and $\tau_1\in(0,\pi]$ in this case; besides,

$$\begin{aligned}\tau_1<\pi \quad &\text{implies} \quad x_1<\pi R,\\ \tau_1=\pi \quad &\text{implies} \quad x_1=\pi R.\end{aligned}$$

Note that the case where $u'(x_1)=0$, and thus $u(x_1)=2R$, is covered by the above analysis: it corresponds to $2R=y_1$.

Assume next that the derivative u' vanishes on the open interval $(0,x_1)$. We already know that the restriction of the function u to the interval $[0,\pi R]$ is given in parametric form by

$$x=R(\tau-\sin\tau),\quad u(x)=R(1-\cos\tau),\quad 0\le\tau\le\pi$$

for some constant $R>0$ that satisfies $R\ge\frac{y_1}{2}$.

Since then $u'(\pi R)=0$, and the derivative u' vanishes at most once in $(0,x_1)$ (Lemma 3.4), it follows, again by Lemma 3.4, that $u'(x)<0$ for all $\pi R<x<x_1$. So we can re-use an argument similar to the one used in the first case for finding the parametric representation of the restriction of the function u to the interval $[\pi R,x_1]$, by means of the change of variable

$$\tau\in[\pi,\tau_1]\to y=R(1-\cos\tau)\quad \text{at each}\quad 2R\ge y\ge y_1,$$

where, for any given constant $R>0$ such that $R\ge\frac{y_1}{2}$, the angle τ_1 is this time uniquely defined by the relations

$$\pi<\tau_1\le 2\pi,\quad R(1-\cos\tau_1)=y_1.$$

Together, the relations $v'(y)^2 = \frac{y}{2R-y}$ and $v'(y) = \frac{1}{u'(x)} < 0$, $y_1 < y < 2R$, where $v \in \mathcal{C}[y_1,2R] \cap \mathcal{C}^2(y_1,2R)$ now denotes the inverse function of the function $u \in \mathcal{C}[\pi R,x_1] \cap \mathcal{C}^2(\pi R,x_1)$, imply that

$$v'(y) = -\sqrt{\frac{y}{2R-y}} = -\sqrt{\frac{1-\cos\tau}{1+\cos\tau}} \quad \text{at each} \quad y = R(1-\cos\tau), \quad y_1 < y < 2R.$$

Let the function $f : [\pi,\tau_1] \to \mathbb{R}$ be defined by

$$f(\tau) := v\big(R(1-\cos\tau)\big) = v(y) = x, \quad \pi \le \tau \le \tau_1,$$

so that

$$\begin{aligned} f'(\tau) &= (R\sin\tau)v'\big(R(1-\cos\tau)\big) = R\left(-\sqrt{\sin^2\tau}\right)\left(-\sqrt{\frac{1-\cos\tau}{1+\cos\tau}}\right) \\ &= R(1-\cos\tau), \quad \pi < \tau < \tau_1, \end{aligned}$$

and $f(\pi) = v(2R) = \pi R$. Consequently,

$$x = R(\tau - \sin\tau), \quad \pi \le \tau \le \tau_1.$$

In particular then, $R(\tau_1 - \sin\tau_1) = x_1$, which, together with the relation $R(1-\cos\tau_1) = y_1$, uniquely determines the constants $R > 0$ and $\tau_1 \in (\pi, 2\pi]$ in this case; besides,

$$\tau_1 > \pi \quad \text{implies} \quad x_1 > \pi R.$$

The conclusion follows by taking into account that, given any $x_1 > 0$ and any $y_1 \ge 0$, the nonlinear system

$$\begin{aligned} R(\tau_1 - \sin\tau_1) &= x_1, \\ R(1-\cos\tau_1) &= y_1 \end{aligned}$$

has a unique solution $(R,\tau_1) \in (0,\infty) \times (0,2\pi]$ (this is easily seen, as hinted at in Fig. 2). $\square$

Lemma 3.6. *The brachistochrone problem has one and only one solution when* $y_1 = 0$.

Proof. The uniqueness has been established in Lemma 3.5. The existence when $y_1 = 0$ is established below by means of a clever argument due to Benson [3], repeated here for the sake of completeness.

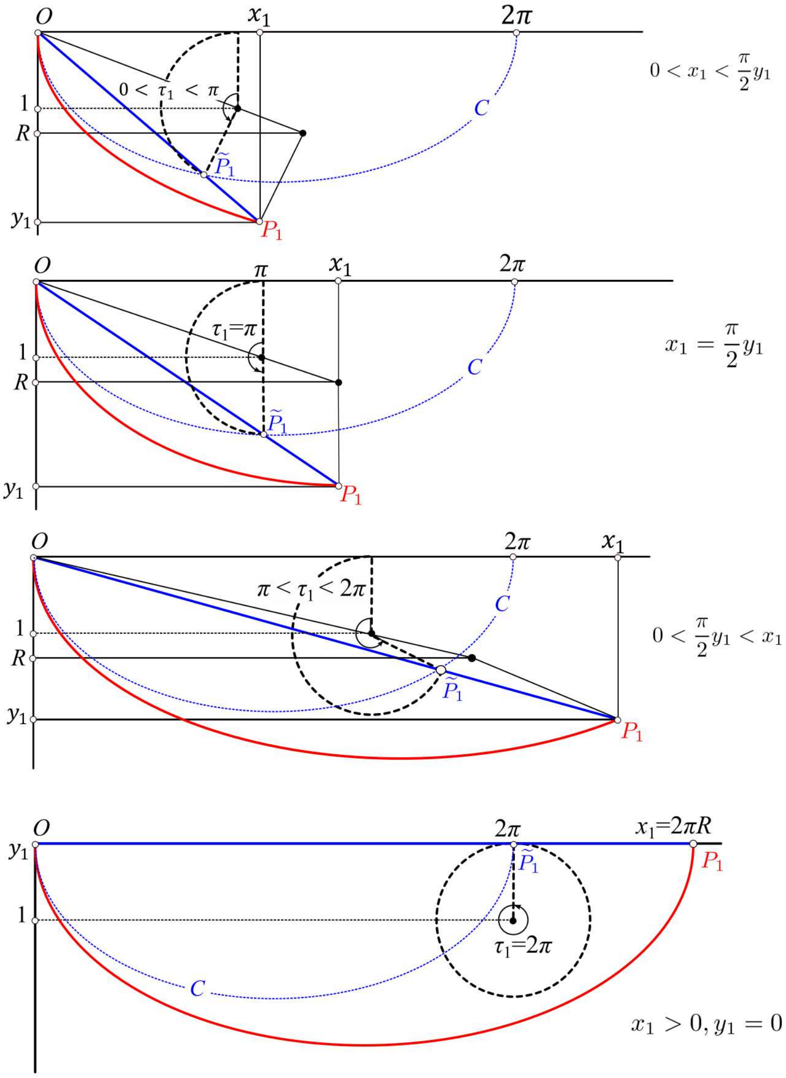

Figure 2: This figure suggests how to show that, given two constants $x_1 > 0$ and $y_1 \geq 0$, there exists one and only one solution (R,τ_1) to the nonlinear system

$$R(\tau_1 - \sin\tau_1) = x_1,$$
$$R(1-\cos\tau_1) = y_1,$$
$$R > 0, \quad 0 < \tau_1 \leq 2\pi.$$

More specifically, let C denote the "first arch" of the cycloid with $\mathbf{0}$ as a cusp and with one as the radius of the rolling circle, and let $\widetilde{\boldsymbol{P}}_1$ denote the intersection of C with the straight line passing through the points $\mathbf{0}$ and $\boldsymbol{P}_1 = (x_1, y_1)$ (such a point $\widetilde{\boldsymbol{P}}_1$ is uniquely defined, because the curve C is strictly concave on the interval $[0,2\pi]$; recall that the "y-axis" is pointing downward). Then the radius R and angle τ_1 are those corresponding to the segment of the cycloid obtained from C by means of a homothety with $\mathbf{0}$ as its center and $|\boldsymbol{P}_1|/|\widetilde{\boldsymbol{P}}_1|$ as its ratio. Note that there are three distinct cases if $y_1 > 0$ (depending on how x_1 compares with $\frac{\pi}{2} y_1$), and one case if $y_1 = 0$.

Given any vector field $\boldsymbol{v}=(v_i)_{i=1}^3\in\boldsymbol{V}$ such that $v_3=0$ on $[\theta_0,\theta_1]$ and $J(\boldsymbol{v})<\infty$, let

$$M:=\sup_{\theta_0\le\theta\le\theta_1} v_2(\theta)=v_2(\widetilde{\theta})>0 \quad \text{for some} \quad \theta_0<\widetilde{\theta}<\theta_1.$$

Cauchy-Schwarz' inequality implies that, for each $\theta_0<\theta<\theta_1$,

$$\sqrt{v_1'(\theta)^2+v_2'(\theta)^2}\sqrt{\frac{1}{v_2(\theta)}}\ge\left|v_1'(\theta)\right|\frac{1}{\sqrt{M}}+\left|v_2'(\theta)\right|\sqrt{\frac{1}{v_2(\theta)}-\frac{1}{M}}.$$

Consequently,

$$\int_{\theta_0}^{\theta_1}\sqrt{\frac{v_1'(\theta)^2+v_2'(\theta)^2}{v_2(\theta)}}\mathrm{d}\theta\ge\frac{1}{\sqrt{M}}\int_{\theta_0}^{\theta_1}\left|v_1'(\theta)\right|\mathrm{d}\theta+\int_{\theta_0}^{\theta_1}\left|v_2'(\theta)\right|\sqrt{\frac{1}{v_2(\theta)}-\frac{1}{M}}\mathrm{d}\theta.$$

Besides, the integrals in the right-hand side satisfy

$$\begin{aligned}
&\int_{\theta_0}^{\theta_1}\left|v_1'(\theta)\right|\mathrm{d}\theta\ge\left|\int_{\theta_0}^{\theta_1}v_1'(\theta)\mathrm{d}\theta\right|=|v_1(\theta_1)-v_1(\theta_0)|=x_1,\\
&\int_{\theta_0}^{\theta_1}\left|v_2'(\theta)\right|\sqrt{\frac{1}{v_2(\theta)}-\frac{1}{M}}\mathrm{d}\theta\\
\ge&\int_{\theta_0}^{\widetilde{\theta}}v_2'(\theta)\sqrt{\frac{1}{v_2(\theta)}-\frac{1}{M}}\mathrm{d}\theta-\int_{\widetilde{\theta}}^{\theta_1}v_2'(\theta)\sqrt{\frac{1}{v_2(\theta)}-\frac{1}{M}}\mathrm{d}\theta.
\end{aligned}$$

Let

$$g(v):=\sqrt{\frac{1}{v}-\frac{1}{M}},\quad G(v):=\int_M^v g(\rho)\mathrm{d}\rho \quad \text{at each} \quad v\in(0,M].$$

Then, given any $\widetilde{\theta}_0$ such that $\theta_0<\widetilde{\theta}_0<\widetilde{\theta}$, we have

$$\begin{aligned}
&\int_{\widetilde{\theta}_0}^{\widetilde{\theta}}v_2'(\theta)\sqrt{\frac{1}{v_2(\theta)}-\frac{1}{M}}\mathrm{d}\theta\\
=&\int_{\widetilde{\theta}_0}^{\widetilde{\theta}}g\big(v_2(\theta)\big)v_2'(\theta)\mathrm{d}\theta=\int_{\widetilde{\theta}_0}^{\widetilde{\theta}}\big(G(v_2(\theta))\big)'\mathrm{d}\theta,\\
=&G\big(v_2(\widetilde{\theta})\big)-G\big(v_2(\widetilde{\theta}_0)\big)=\int_{v_2(\widetilde{\theta}_0)}^{M}g(\rho)\mathrm{d}\rho
\end{aligned}$$

thanks to the fundamental formula of integral calculus applied to a function of class $\mathcal{C}^1$ on a compact interval.

Likewise, given any $\widetilde{\theta}_1$ such that $\widetilde{\theta}<\widetilde{\theta}_1<\theta_1$, we have

$$\begin{aligned}
&-\int_{\widetilde{\theta}}^{\widetilde{\theta}_1} v_2'(\theta)\sqrt{\frac{1}{v_2(\theta)}-\frac{1}{M}}\,\mathrm{d}\theta \\
=&-\int_{\widetilde{\theta}}^{\widetilde{\theta}_1}\big(G(v_2(\theta))\big)'\,\mathrm{d}\theta=G\big(v_2(\widetilde{\theta})\big)-G\big(v_2(\widetilde{\theta}_1)\big)=\int_{v_2(\widetilde{\theta}_1)}^{M} g(\rho)\,\mathrm{d}\rho.
\end{aligned}$$

Consequently, letting $\widetilde{\theta}_0$ approach θ_0 and $\widetilde{\theta}_1$ approach θ_1 gives

$$\begin{aligned}
&\int_{\theta_0}^{\widetilde{\theta}} v_2'(\theta)\sqrt{\frac{1}{v_2(\theta)}-\frac{1}{M}}\,\mathrm{d}\theta-\int_{\widetilde{\theta}}^{\theta_1} v_2'(\theta)\sqrt{\frac{1}{v_2(\theta)}-\frac{1}{M}}\,\mathrm{d}\theta \\
=&2\int_0^M\sqrt{\frac{1}{\rho}-\frac{1}{M}}\,\mathrm{d}\rho=\pi\sqrt{M},
\end{aligned}$$

where the last integral is computed as follows:

$$\int_0^1\sqrt{\frac{1-\sigma}{\sigma}}\,\mathrm{d}\sigma=-\int_0^\infty \tau\,\mathrm{d}\left(\frac{1}{1+\tau^2}\right)=\frac{\pi}{2}.$$

Finally then,

$$\begin{aligned}
&\int_{\theta_0}^{\theta_1}\sqrt{\frac{v_1'(\theta)^2+v_2'(\theta)^2}{v_2(\theta)}}\,\mathrm{d}\theta \\
\geq&\frac{1}{\sqrt{M}}x_1+\pi\sqrt{M}\geq\inf_{M>0}\left(\frac{1}{\sqrt{M}}x_1+\pi\sqrt{M}\right)=2\sqrt{\pi x_1}.
\end{aligned}$$

Let

$$u_1(\tau):=\frac{x_1}{2\pi}(\tau-\sin\tau),\quad u_2(\tau):=\frac{x_1}{2\pi}(1-\cos\tau),\quad \tau\in[0,2\pi]$$

be the parameterized curve given by Lemma 3.5 (where $\tau_1=2\pi$ and $R=\frac{x_1}{2\pi}$ since $y_1=0$). Then

$$\int_0^{2\pi}\sqrt{\frac{u_1'(\tau)^2+u_2'(\tau)^2}{u_2(\tau)}}\,\mathrm{d}\tau=2\sqrt{\pi x_1}=\inf_{v\in\boldsymbol{V}}\int_{\theta_0}^{\theta_1}\frac{\sqrt{v_1'(\theta)^2+v_2'(\theta)^2}}{\sqrt{v_2(\theta)}}\,\mathrm{d}\theta,$$

which establishes the existence of a minimizer as a segment of the cycloid with two cusps, one at $(0,0)$ and one at $(x_1,0)$ (as was noted in Section 2, we have the freedom to choose whichever parametrization we please for the curves $\boldsymbol{v}([\theta_0,\theta_1])$, $\boldsymbol{v}\in\boldsymbol{V}$). Thus the existence of the brachistochrone is established when $y_1=0$.

Note that this minimizer $\boldsymbol{u}=(u_i)_{i=1}^3$ indeed belongs to the space $\boldsymbol{V}$, since $u_3=0$, its components u_1 and u_2 are functions of class $\mathcal{C}^\infty$ on $[\theta_0,\theta_1]$,

$$\boldsymbol{u}(0)=(0,0,0), \quad \boldsymbol{u}(2\pi)=(x_1,0,0),$$

and

$$|\boldsymbol{u}'(\theta)|=\frac{x_1}{\sqrt{2}\pi}\sqrt{1-\cos\theta}>0, \quad u_2(\theta)=\frac{x_1}{2\pi}(1-\cos\theta)>0, \quad 0<\theta<2\pi.$$

The proof is complete. □

Lemma 3.7. *The brachistochrone problem has one and only one solution when $y_1>0$.*

Proof. The uniqueness has been established in Lemma 3.5.

Given a point $\boldsymbol{P}_1=(x_1,y_1,0)$ with $x_1>0$ and $y_1>0$, let $R>0$ be determined by solving the nonlinear system found in Lemma 3.5, and let C denote the segment of the cycloid joining the points $\boldsymbol{P}_0$ and $(2\pi R,0)$; cf. Fig. 3. Then we claim that the segment C_1 of the cycloid C that joins the point $\boldsymbol{P}_0$ to $\boldsymbol{P}_1$ is a solution to the

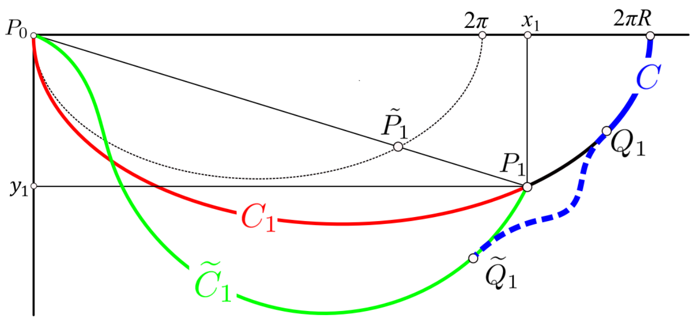

Figure 3: Existence of a minimizer when $y_1>0$: Given a point $\boldsymbol{P}_1=(x_1,y_1)$ with $y_1>0$, let $\widetilde{\boldsymbol{P}}_1$ denote the intersection of the line supporting the segment $[\boldsymbol{P}_0,\boldsymbol{P}_1]$ with the segment of the cycloid with one as the radius of the rolling circle and with only two cusps, at $x=0$ and at $x=2\pi$ (dashed curve passing by $\widetilde{\boldsymbol{P}}_1$), let C_1 denote the segment of the cycloid homothetic to the segment of this cycloid that joins $\boldsymbol{P}_0$ to $\widetilde{\boldsymbol{P}}_1$, with $\boldsymbol{P}_0$ as the center and $|\boldsymbol{P}_1|/|\widetilde{\boldsymbol{P}}_1|$ as the ratio of the homothety, and let R denote the radius of the rolling circle that generates the arc C_1. Then C_1 is the brachistochrone between the points $\boldsymbol{P}_0$ and $\boldsymbol{P}_1$ (the proof of this assertion is provided in Lemma 3.7).

minimization problem $\inf_{v\in V} J(v)$, thus establishing the existence of the brachistochrone when $y_1>0$.

Assume on the contrary that there exists a parameterized curve $\widetilde{C}_1$ of class $\mathcal{C}^2$ (cf. Fig. 3) joining P_0 and P_1 such that (with self-explanatory notations)

$$\int_{\theta_0}^{\theta_1}\sqrt{\frac{\widetilde{u}_1'(\theta)^2+\widetilde{u}_2'(\theta)^2}{\widetilde{u}_2(\theta)}}\mathrm{d}\theta<\int_{\theta_0}^{\theta_1}\sqrt{\frac{u_1'(\theta)^2+u_2'(\theta)^2}{u_2(\theta)}}\mathrm{d}\theta,$$

where the functions u_1 and u_2 appearing in the right-hand side are those defining the curve C_1. In other words, the time spent by the material point for going from P_0 to P_1 along $\widetilde{C}_1$ is strictly shorter than that along C_1. Besides, there is a nonzero tangent vector along $\widetilde{C}_1$ except perhaps at the points P_0 and P_1 (by definition of the set V).

By suitably modifying the curve $\widetilde{C}_1$ together with the segment $C-C_1$ of cycloid in a small enough neighborhood of the point P_1, so that the time for going from a point $\widetilde{Q}_1\in\widetilde{C}_1$ to a point $Q_1\in C$ in this neighborhood along the modified curve is small enough (see the dashed line in Fig. 3), a parameterized curve of class $\mathcal{C}^2$ could thus be constructed, along which the time spent for going from P_0 to $(2\pi R,0)$ would be strictly less than the time for going from P_0 to $(2\pi R,0)$ along C, thus contradicting that the time spent for going from P_0 to $(2\pi R,0)$ is the shortest (which has been established in Lemma 3.6). $\square$

Together, the results established in Lemmas 3.1 to 3.7 constitute the proof of the following theorem. Recall that $g>0$ denotes the gravitational constant.

Theorem 3.1 (Existence and uniqueness of the brachistochrone)**.** *Given constants $\theta_0<\theta_1$, $x_1>0$, and $y_1\geq 0$, define the set*

$$\begin{aligned}V:=\Big\{&v=(v_i)_{i=1}^3\in\mathcal{C}^2\big([\theta_0,\theta_1];\mathbb{R}^3\big);\, v(\theta_0)=(0,0,0),\, v(\theta_1)=(x_1,y_1,0),\\ &|v'(\theta)|>0 \text{ and } v_2(\theta)>0,\,\theta_0<\theta<\theta_1\Big\},\end{aligned}$$

and the functional

$$J(v):=\int_{\theta_0}^{\theta_1}\sqrt{\frac{v_1'(\theta)^2+v_2'(\theta)^2+v_3'(\theta)^2}{2gv_2(\theta)}}\mathrm{d}\theta.$$

Then the problem:

$$\textit{find } u=(u_i)_{i=1}^3\in V \textit{ such that } J(u)=\inf_{v\in V}J(v)$$

has one and only one solution $\boldsymbol{u}$ *up to* $\mathcal{C}^1$*-diffeomorphisms (cf. Section* 2*), called the brachistochrone between the points* $(0,0,0)$ *and* $(x_1,y_1,0)$*, which is the segment of the cycloid (Fig.* 4*) given (for instance) in parametric form by*

$$\left.\begin{aligned} u_1(\tau) &= R(\tau - \sin\tau) \\ u_2(\tau) &= R(1-\cos\tau) \\ u_3(\tau) &= 0 \end{aligned}\right\}, \quad 0 \le \tau \le \tau_1,$$

where the constants $R = R(x_1,y_1)$ *and* $\tau_1 = \tau_1(x_1,y_1)$ *are the unique solutions of the nonlinear system:*

$$\begin{aligned} &R(\tau_1 - \sin\tau_1) = x_1, \\ &R(1-\cos\tau_1) = y_1, \\ &R > 0, \quad 0 < \tau_1 \le 2\pi, \end{aligned}$$

(note that the change of parameters is given here by $\tau = \frac{\theta-\theta_0}{\theta_1-\theta_0}\tau_1$*).*

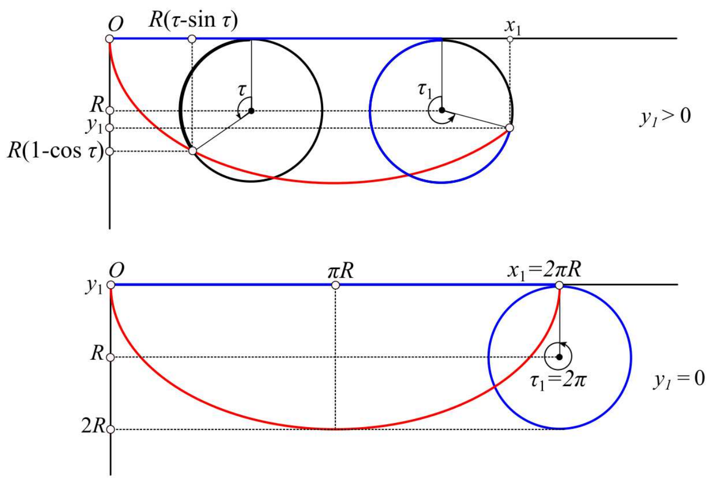

Figure 4: The brachistochrone is a segment of cycloid, with either one cusp if $y_1 < 0$ in which case $\tau_1 < 2\pi$, or with two cusps if $y_1 = 0$ in which case $\tau_1 = 2\pi$. The geometric interpretations of the parameter τ as an angle, of the constant R as the radius of a circle rolling without sliding along the axis supporting $\boldsymbol{e}_1$, and of the angle τ_1, are illustrated on the figure.

4 Final remarks

In this article, we showed that the classical method of the calculus of variations, based on the Euler-Lagrange equations, for finding the brachistochrone joining two points $\boldsymbol{P}_0$ and $\boldsymbol{P}_1$ in $\mathbb{R}^3$ can be rigorously justified in the general case, i.e., where the two points are joined by parameterized curves, by contrast with the particular case most often used in the literature whereby the points are a priori assumed to be joined by graphs of functions in the "vertical plane". Besides, our proof is deductive, in the sense that the solution is found by solving the Euler-Lagrange equations, and not just guessed to be the segment of cycloid joining the two given points $\boldsymbol{P}_0$ and $\boldsymbol{P}_1$ without explaining how the cycloid has been identified as a potential solution in the first place.

For the sake of completeness, we conclude by outlining alternative proofs to the brachistochrone problem found in the literature. These proofs fall into two categories, depending on whether they are given for the brachistochrone problem in the general case of parameterized curves in space (as in this paper), or to the brachistochrone problem in the special case where the admissible curves are a priori assumed to be graphs of functions in the vertical plane containing the two points $\boldsymbol{P}_0$ and $\boldsymbol{P}_1$. This means that the curves are assumed to be given by

$$\Big\{(x,y,z)\in\mathbb{R}^3;\ z=0 \text{ and } y=f(x),\, x\in[0,x_1]\Big\},$$

where $f:[0,x_1]\to[0,\infty)$ are sufficiently smooth functions such that $f(0)=0$ and $f(x_1)=y_1$.

For the brachistochrone problem in the general case, the authors are aware of two other proofs: one based on an ad-hoc change of variable (due to Brookfield [5]), and one based on optimal control theory and Hamiltonians (due to Sussmann & Willems [11]).

For the brachistochrone problem in the special case where the admissible curves are a priori assumed to be graphs of functions in the vertical plane containing the two points (which is substantially easier than the problem in the general case; it corresponds to Lemmas 3.6 and 3.7 in this paper), there are many alternative proofs: e.g., Kosmol [9] and Troutman [12], who independently proposed a change of unknown that transforms the brachistochrone problem into a convex optimisation problem for the new unknown; or Cesari [6] who solves the brachistochrone problem as a particular case of a general theorem about "parametric integrals" defined in ibid.; or Hrusa & Troutman [8], who solve the brachistochrone problem in a particular case where the solution is contained in the first half of the cycloid; or, finally, Balder [1], who solves the brachistochrone problem by using a method similar to the one used in Lemma 3.6 due to Benson (Balder's method applies for all $y_1\geq 0$, but under the assumption that the curve is a graph, while Benson's method applies for parameterized curves, but only for $y_1=0$).

4.1 Brookfield's proof

Brookfield's proof for the brachistochrone problem (cf. [5]) is based on an ingenious (although somewhat mysterious!) change of coordinates in the set $\{(x,y)\in\mathbb{R}^2; x>0, y\geq 0\}$, whereby the admissible curves joining the points $\boldsymbol{P}_0=(0,0,0)$ and $\boldsymbol{P}_1=(x_1,y_1,0)$ are defined for each $t\in[0,T]$ by

$$
\begin{aligned}
&x(t):=\rho(t)\sigma(t)-\rho(t)^2\sin\left(\frac{\sigma(t)}{\rho(t)}\right),\\
&y(t):=\rho(t)^2-\rho(t)^2\cos\left(\frac{\sigma(t)}{\rho(t)}\right),\\
&z(t)=0,
\end{aligned}
$$

where $T>0$ is a real parameter and the functions $\rho:[0,T]\to\mathbb{R}$ and $\sigma:[0,T]\to\mathbb{R}$ satisfy the following assumptions:

$$
\begin{aligned}
&\rho\in\mathcal{C}^1[0,T], \quad \rho(t)>0 && \text{for all} \quad 0\leq t\leq T,\\
&\sigma\in\mathcal{C}^1[0,T], \quad 0<\sigma(t)<2\pi\rho(t) && \text{for all} \quad 0<t<T, \quad \sigma(0)=0,\\
&\big(x(0),y(0)\big)=(0,0), \quad \big(x(T),y(T)\big)=(x_1,y_1).
\end{aligned}
$$

Then the law of conservation of energy implies that T equals the time needed for a material point to go from $\boldsymbol{P}_0$ to $\boldsymbol{P}_1$ if, at each $t\in[0,T]$,

$$
\big(g-\sigma'(t)^2\big)y(t)=16\rho^2(t)\left(\sin\frac{\sigma(t)}{2\rho(t)}-\frac{\sigma(t)}{2\rho(t)}\cos\frac{\sigma(t)}{2\rho(t)}\right)^2\big(\rho'(t)\big)^2.
$$

Since the right-hand side is obviously non-negative, it follows that

$$
T\geq\frac{\sigma(T)}{\sqrt{g}}
$$

with equality if and only if

$$
\rho'(t)=0, \quad \sigma'(t)=\sqrt{g} \quad \text{for all} \quad 0\leq t\leq T.
$$

This also implies that the shortest time T needed for joining the points $\boldsymbol{P}_0$ to $\boldsymbol{P}_1$ is given by $T=\tau_1\sqrt{R/g}$, where $(R,\tau_1)\in(0,\infty)\times(0,2\pi]$ is the unique solution of the nonlinear system

$$
\begin{aligned}
&R(\tau_1-\sin\tau_1)=x_1,\\
&R(1-\cos\tau_1)=y_1,
\end{aligned}
$$

and that the corresponding brachistochrone is uniquely defined as a segment of the cycloid given in parametric form at each $0 \le t \le T$ by

$$x(t) = R\left(t\sqrt{\frac{g}{R}} - \sin\left(t\sqrt{\frac{g}{R}}\right)\right),$$
$$y(t) = R\left(1 - \cos\left(t\sqrt{\frac{g}{R}}\right)\right).$$

4.2 Sussmann & Willems' proof

Sussmann & Willems' proof for the brachistochrone problem (cf. [11]) is based on a maximum principle in optimal control theory. In this approach, the brachistochrone problem is recast into the problem of minimising the functional $J(T,\boldsymbol{u},\boldsymbol{v}) := T$ over the set of all triples $(T,\boldsymbol{u},\boldsymbol{v})$, where the real number $T>0$ and the vector fields $\boldsymbol{u} = (u_1,u_2):[0,T] \to \mathbb{R}^2$ and $\boldsymbol{v} = (v_1,v_2):[0,T] \to \mathbb{R}^2$ are subject to the constrains

$$u_1(t)^2 + u_2(t)^2 = 1, \quad \boldsymbol{v}'(t) = \boldsymbol{u}(t)\sqrt{2g|v_2(t)|} \quad \text{for all} \quad t \in [0,T],$$
$$\boldsymbol{v}(0) = (0,0), \quad \boldsymbol{v}(T) = (x_1,y_1).$$

Then the maximum principle alluded to above states that any solution $(\hat{T},\hat{\boldsymbol{u}},\hat{\boldsymbol{v}})$ to the above minimisation problem necessarily satisfies the following conditions:

- First, there exists a vector field $\hat{\boldsymbol{p}} = (\hat{p}_1,\hat{p}_2) : [0,\hat{T}] \to \mathbb{R}^2$ and a constant $\hat{C} \ge 0$ such that
$$\big(\hat{\boldsymbol{p}}(t),\hat{C}\big) \ne ((0,0),0) \quad \text{for all} \quad t \in [0,\hat{T}].$$

- Second,
$$\hat{\boldsymbol{v}}'(t) = \frac{\partial H}{\partial \boldsymbol{p}}\big(\hat{\boldsymbol{v}}(t),\hat{\boldsymbol{u}}(t),\hat{\boldsymbol{p}}(t),\hat{C}\big), \quad t \in [0,\hat{T}],$$
$$-\hat{\boldsymbol{p}}'(t) = \frac{\partial H}{\partial \boldsymbol{v}}\big(\hat{\boldsymbol{v}}(t),\hat{\boldsymbol{u}}(t),\hat{\boldsymbol{p}}(t),\hat{C}\big), \quad t \in [0,\hat{T}],$$
where
$$H\big((\boldsymbol{v},\boldsymbol{u},\boldsymbol{p},C)\big) := (\boldsymbol{p}\cdot\boldsymbol{u})\sqrt{2g|v_2|} - C.$$

- Third, for each $t \in [0,\hat{T}]$,
$$H\big(\hat{\boldsymbol{v}}(t),\hat{\boldsymbol{u}}(t),\hat{\boldsymbol{p}}(t),\hat{C}\big) = \max_{a^2+b^2=1} H\big(\hat{\boldsymbol{v}}(t),(a,b),\hat{\boldsymbol{p}}(t),\hat{C}\big).$$

- Fourth, for each $t \in [0,\hat{T}]$,

$$H\big(\hat{\boldsymbol{v}}(t),\hat{\boldsymbol{u}}(t),\hat{\boldsymbol{p}}(t),\hat{C}\big)=0.$$

One then deduces from the above necessary conditions that the curve $t\in[0,\hat{T}]\to \hat{\boldsymbol{v}}(t)\in\mathbb{R}^2$ is the graph of a function $f:[0,x_1]\to\mathbb{R}$, where f satisfies the differential equation

$$1+f'(t)^2+2f(t)f''(t)=0, \quad t\in[0,\hat{T}],$$

then that

$$\hat{v}_2(t)>0 \quad \text{for all} \quad t\in(0,\hat{T}).$$

Noting that the differential equation above coincides with the Euler-Lagrange equation associated to the brachistochrone problem in the special case where the curves are assumed to be graphs of functions, one concludes that the segment of cycloid joining the points $\boldsymbol{P}_0$ and $\boldsymbol{P}_1$ is a solution to the brachistochrone problem.

4.3 Kosmol & Troutman's proof

Kosmol [9] and Troutman [12] independently solved the brachistochrone problem in the special case of curves that can be represented as graphs of functions by using a change of unknown that transforms the brachistochrone problem into a convex optimisation problem for the new unknown.

More specifically, assume that the curves are defined by

$$\Gamma_f:=\Big\{(x,y,z)\in\mathbb{R}^3;\, z=0 \text{ and } y=f(x),\, x\in[0,x_1]\Big\},$$

where $f:[0,x_1]\to[0,\infty)$ are sufficiently smooth functions such that $f(0)=0$ and $f(x_1)=y_1$. Then the time needed for a material point to go from $\boldsymbol{P}_0$ to $\boldsymbol{P}_1$ along Γ_f is given by

$$T(f):=\frac{1}{\sqrt{2g}}\int_0^{x_1}\sqrt{\frac{1+f'(x)^2}{f(x)}}\,\mathrm{d}x\in[0,+\infty].$$

Performing the change of unknown

$$f(x)=\sqrt{\phi(x)} \quad \text{for all} \quad x\in[0,x_1]$$

in the above integral then shows that

$$T(f)=I(\phi):=\frac{1}{\sqrt{2g}}\int_0^{x_1}\sqrt{4\phi'(x)^2+\frac{1}{\phi(x)^2}}\,\mathrm{d}x.$$

Since the function $\phi \to I(\phi)$ is strictly convex, the brachistochrone problem in the special case where the curves are graphs of functions admits at most one solution.

That this problem does have a solution can be established in at least three ways, each one having its advantages and disadvantages.

The most natural way is to use the theory of convex optimisation, in which case the function ϕ must be chosen to belong to the space of absolutely continuous functions on the closed interval $[0,x_1]$. Then one has to prove that the problem of minimising the functional $f \to T(f)$ under the assumption that f is an absolutely continuous function over $[0,x_1]$ is equivalent to the problem of minimising the functional $\phi \to I(\phi)$ under the assumption that ϕ is an absolutely continuous function over $[0,x_1]$, in spite of the fact the change of unknown does not preserve the absolute continuity of the functions; see e.g. Kosmol [9] or Ball [2].

Another way is to derive the Euler-Lagrange equation associated with the above minimisation problem, then to prove that the segment of cycloid joining the points $(0,0)$ and (x_1,y_1) satisfies this Euler-Lagrange equation.

Yet another way, due to Balder [1], is to directly prove that the segment of cycloid joining the points $(0,0)$ and (x_1,y_1) minimises the functional I, by noting that, for all sufficiently regular functions $\phi,\psi:[0,x_1]\to\mathbb{R}$,

$$I(\phi) \geq I(\psi)+K(\phi,\psi),$$

where

$$K(\phi,\psi) := \frac{1}{\sqrt{2g}} \int_0^{x_1} \frac{\psi(x)^{-2}-4\psi'(x)^2+4\psi'(x)\phi(x)-\psi(x)^{-3}\phi(x)}{\sqrt{4\psi'(x)^2+\psi(x)^{-2}}}\,\mathrm{d}x,$$

and that, for all sufficiently regular functions $\phi:[0,x_1]\to\mathbb{R}$,

$$K(\phi,\hat{\psi})=0,$$

where $\hat{\psi}:[0,x_1]\to\mathbb{R}$ is the function with the property that the curve

$$\left\{ \left(x,\hat{\psi}(x)^2\right) \in \mathbb{R}^2;\ x\in[0,x_1]\right\}$$

is the segment of cycloid joining $(0,0)$ to (x_1,y_1).

The key to proving the inequality $I(\phi)\geq I(\psi)+K(\phi,\psi)$ (mentioned above) is the following variant of the Cauchy-Schwarz inequality: For all $a>0$, $b\in\mathbb{R}$, $\alpha>0$ and $\beta\in\mathbb{R}$,

$$\sqrt{a^2+b^2}\sqrt{\alpha^2+\beta^2} \geq a\alpha+b\beta-\frac{a}{\alpha}(a-\alpha)^2,$$

or equivalently,

$$\sqrt{\alpha^2+\beta^2} \geq \sqrt{a^2+b^2} + \frac{a^2-b^2+b\beta-\frac{a^3}{\alpha}}{\sqrt{a^2+b^2}}$$

with the inequality becoming an equality if and only if $(a,b)=(\alpha,\beta)$.

4.4 Hrusa & Troutman's proof

Hrusa & Troutman solved the brachistochrone problem (cf. [8]) in the special case where $0<x_1<\frac{2y_1}{\pi}$, in which case the segment of cycloid joining the points $(0,0)$ and (x_1,y_1) can be represented by the graph of a function $y\in[0,y_1]\to x=\hat{f}(y)\in[0,x_1]$, and for curves that can also be represented as graph of a functions

$$y\in[0,y_1]\to x=f(y)\in[0,x_1].$$

In this case, the time needed for a material point to go from $(0,0)$ and (x_1,y_1) along the graph of the function $f:[0,y_1]\to[0,x_1]$ is given by

$$T(f):=\frac{1}{\sqrt{2g}}\int_0^{y_1}\sqrt{\frac{1+f'(y)^2}{y}}\mathrm{d}y.$$

Since the function $f\to T(f)$ is strictly convex, the brachistochrone problem in this special case can be solved by using convex optimisation theory, as in the proof of Kosmol [9] and Troutman [12] described above.

Acknowledgments

The authors warmly thank Dr. Amgad Abdrabrou, who expertly converted the original hand-drawn figures into electronic ones.

References

[1] E. Balder, *The brachistochrone problem made elementary*, Utrecht University, (2002), `https://webspace.science.uu.nl/~balde101/talks/vanbeek.pdf`.

[2] J. M. Ball, *The calculus of variations and material science*, Quart. Appl. Math., LVI (1998), 719–740.

[3] D. C. Benson, *An elementary solution of the brachistochrone problem*, Amer. Math. Monthly 76 (1969), 890–894.

[4] J. Bernoulli, *Problema novum ad cujus solutionem mathematice invitantur*, Acta Eruditorum Lipsiae, 18 (1696), p. 269.

[5] G. Brookfield, *Yet another elementary solution of the brachistochrone problem*, Math. Mag. 83 (2010), 59–63.

[6] L. Cesari, *Optimization-Theory and Applications*, Springer-Verlag, 1983.

[7] B. Dacorogna, *Introduction to the Calculus of Variations*, Imperial College Press, 2009.

[8] W. Hrusa and J. L. Troutman, *Elementary characterization of classical minima*, Amer. Math. Monthly 88 (1981), 321–327.

[9] P. Kosmol, *Bemerkungen zur brachistochrone*, Abh. Math. Sem. Univ. Hamburg, 54 (1984), 91–94.

[10] D. S. Shafer, *The brachistochrone: historical gateway to the calculus of variations*, Materials Mathematics, 5 (2007), 1–14.

[11] H. J. Sussmann and J. C. Willems, *300 years of optimal control: from the brachystochrone to the maximum principle*, IEEE Control Systems, LVI (1997), 32–44.

[12] J. L. Troutman, *Variational Calculus and Optimal Control, Optimization with Elementary Convexity*, Springer-Verlag, 1983.