# The brachistochrone problem made elementary

Erik Balder, Mathematical Institute, University of Utrecht, the Netherlands

February 27, 2002

Johann Bernoulli's brachistochrone problem can be solved completely by using only standard calculus and the Cauchy-Schwarz inequality.

Johann Bernoulli's famous 1696 brachistochrone problem asks for the optimal shape of a metal wire that connects two fixed points $A$ and $B$ in space. A bead of unit mass falls along this wire, without friction, under the sole influence of gravity. The shape of the wire is defined to be optimal if the bead falls from $A$ to $B$ in as short a time as possible. In a standard manner (see e.g. [3, p. 20] for details) this problem can be modeled in the inverted Euclidean plane, with the positive $y$-axis pointing *downward*, by taking $A = (0, 0)$ and $B = (x_1, y_1)$ with $x_1 > 0$ and $y_1 \geq 0$ (for obvious physical reasons the point $B$ cannot lie above the horizontal axis). If the function $f$ describes the shape of the wire connecting $A$ and $B$, then the associated falling time $T(f)$ follows by elementary mechanics:

$$T(f) = \int_0^{x_1} \frac{\sqrt{1 + f'^2(x)}}{\sqrt{2gf(x)}}\, dx.$$

Here $g$ is the gravitational constant. Formally, this expression must be minimized over the set $F$ of all continuously differentiable functions $f : (0, x_1) \to (0, \infty)$ that extend continuously to $[0, x_1]$, with $f(0) = 0$ and $f(x_1) = y_1$.

Because of their strictly local character, the classical necessary conditions for optimality are by themselves not adequate for a complete solution, as is known to every student of optimization theory. What is needed are either *sufficient* conditions for optimality (which bypass the necessary conditions altogether) or an additional result that ensures the existence of an optimal solution. Unfortunately, because the integrand of $T(f)$ has a singularity for $f(0) = 0$, application of both the necessary conditions and the required additional existence result is very difficult. This has led to the following remarkable situation in the literature: The brachistochrone problem, one of the oldest and most famous problems in continuous optimization, is solved in a mathematically rigorous way in very few textbooks (e.g., [2] is such an exception). The other route to a complete solution of the brachistochrone, via convexifying transformations and the concomitant possibility to apply sufficient conditions for optimality, seems to have gained attention only in more recent years [4, 5]. While such transformations and conditions are hard to find and apply in general, it is the purpose of this note to carry them to an extreme in the case of the brachistochrone: We present a new, completely *ad hoc* solution of the brachistochrone problem that is based *only* on standard calculus and the Cauchy-Schwarz inequality. This inequality is recorded here in the following form.

**Lemma 1** *For every* $a > 0, \xi > 0$ *and* $b, \eta \in \mathbb{R}$

$$\sqrt{a^2 + b^2}\sqrt{\xi^2 + \eta^2} \geq -\frac{a}{\xi}(a - \xi)^2 + a\xi + b\eta.$$

*This can be rewritten equivalently as*

$$\sqrt{\xi^2 + \eta^2} \geq \sqrt{a^2 + b^2} + \frac{a^2 - a^3\xi^{-1} + b\eta - b^2}{\sqrt{a^2 + b^2}}.$$

*and holds with equality if and only if* $(\xi, \eta) = (a, b)$.

Following ideas introduced in [4] and [5] (see Example 1, p. 257), let us rewrite the brachistochrone problem by means of the "convexifying" transformation $\phi := \sqrt{f}$. This gives falling time the following form: $S(\phi) := \sqrt{2g}T(\phi^2) = \int_0^{x_1} \sqrt{\phi^{-2}(x) + 4\phi'^2(x)}\, dx$

The corresponding image $\Phi$ of the class $F$ under $f \mapsto \sqrt{f}$ is the set of all continuously differentiable $\phi : (0, x_1) \to (0, \infty)$ that extend continuously to $[0, x_1]$, with $\phi(0) = 0$ and $\phi(x_1) = \sqrt{y_1}$.

We shall now prove *directly* the optimality of $\hat{\phi} : [0, x_1] \to \mathbb{R}$, given by the differential equation

$$4\hat{\phi}'^2(x)\hat{\phi}^2(x) = C\hat{\phi}^{-2}(x) - 1,\ 0 < x < x_1, \tag{1}$$

and the initial value $\hat{\phi}(0) = 0$. Here $C > 0$ is chosen such that $\hat{\phi}(x_1) = \sqrt{y_1}$. Such a continuously differentiable solution $\hat{\phi}$ exists and is well-known (see below). As it is easy to see that $\hat{\phi}(x) > 0$ for $0 < x < x_1$, we have $\hat{\phi} \in \Phi$. Unlike [4, 5], which use some convex analysis, we shall only invoke Lemma 1.

**Theorem 1** *The inequality $S(\phi) > S(\hat{\phi})$ holds for every $\phi \in \Phi$, $\phi \neq \hat{\phi}$.*

Proof. Let $\ell_\phi(x) := \sqrt{\phi^{-2}(x) + 4\phi'^2(x)}$. First we prove $S(\phi) \geq S(\hat{\phi})$. By Lemma 1 one has $\ell_\phi(x) \geq \ell_{\hat{\phi}}(x) + \rho(x)$ for every $x \in (0, x_1)$, where $\rho := (\hat{\phi}^{-2} - \phi\hat{\phi}^{-3} + 4\hat{\phi}'\phi' - 4\hat{\phi}'^2)/\ell_{\hat{\phi}}$. Since $S(\phi) = S(\hat{\phi}) + \int_0^{x_1} \rho$, it is enough to prove $\int_0^{x_1} \rho = 0$. By (1) we have on $(0, x_1)$ that $\ell_{\hat{\phi}} = \hat{\phi}^{-2}\sqrt{C}$, so $\rho = r/\sqrt{C}$, where $r := 1 - \hat{\phi}^{-1}\phi + 4\hat{\phi}^2\hat{\phi}'\phi' - 4\hat{\phi}^2\hat{\phi}'^2$. Let $p := 4\hat{\phi}^2\hat{\phi}'$. Using (1), a routine calculation shows that $p' = -\hat{\phi}^{-1}$ on $(0, x_1)$. Therefore, $r = -p'(\hat{\phi} - \phi) + p(\phi' - \hat{\phi}') = R'$ with $R := p(\phi - \hat{\phi})$. This implies $\int_0^{x_1} r = R(x_1) - R(0) = 0$. Hence, $\int_0^{x_1} \rho = 0$, which proves $S(\phi) \geq S(\hat{\phi})$.

Next, we complete the proof by demonstrating that $S(\phi) = S(\hat{\phi})$ implies $\phi = \hat{\phi}$. Since $\int_0^{x_1} \rho = 0$ was just shown, we get $\int_0^{x_1}(\ell_\phi - \ell_{\hat{\phi}} - \rho) = 0$. By Lemma 1, the integrand $\ell_\phi - \ell_{\hat{\phi}} - \rho$ here is nonnegative and continous, so we conclude that it is zero on $(0, x_1)$. By the last part of Lemma 1, this implies $(\phi, \phi') = (\hat{\phi}, \hat{\phi}')$ on $(0, x_1)$. So $\phi = \hat{\phi}$ on $[0, x_1]$. QED

From Theorem 1 it follows that $\hat{f} := \hat{\phi}^2$ is the unique solution to the original brachistochrone problem. This function describes a *cycloid* connecting the points $A$ and $B$, as can be seen by standard substitutions [2, 5]. Incidentally, it is possible to take instead of $\Phi$ a similar but broader class of functions, viz. absolutely continuous functions. The proof needs very few changes to achieve this.

*Open question:* It would be interesting to gain more insight in the nature of "convexifying" transformations of the kind used here. For instance, which variational problems allow such transformations?

*Postscriptum.* After this note had been written the author became aware of [1], where another proof based on the Cauchy-Schwarz inequality is given. However, that note only addresses the case where the points $A$ and $B$ lie in a horizontal plane (i.e., $y_1 = 0$).

# References


[1] Benson, D.C. An elementary solution of the brachistochrone problem. American Mathematical Monthly **76** (1969), 890-894.

[2] Cesari, L. *Optimization – Theory and Applications: problems with ordinary differential equations*, Springer-Verlag, Berlin, 1983.

[3] Ioffe, A. and V. Tichomirov. *Theorie der Extremalaufgaben*, Deutscher Verlag der Wissenschaften, 1979 (original version in Russian: Nauka, Moscow, 1974).

[4] Kosmol, P. Bemerkungen zur Brachistochrone. Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg **54** (1984), 91-94.

[5] Troutman, J.L. *Variational Calculus and Optimal Control: optimization with elementary convexity*, Springer-Verlag, Berlin, 1996.