The Brachistochrone

P C Deshmukh, Parth Rajauria, Abiya Rajans, B R Vyshakh, and Sudipta Dutta

The brachistochrone problem posed by Bernoulli and its solution highlights one of the most famous experiments in physics which illustrates the variational principle. This pedagogical study is designed to ignite a classroom discussion on the variational problem. We overview the Euler–Lagrange formalism of the variational principle and obtain the solution to the brachistochrone problem. We demonstrate the success of the variational method using brachistochrone models which were fabricated specially for this purpose.

1. Introduction

According to the first law of motion of classical mechanics, due to Galileo–Newton, we know that the mechanical states of an object, whether at rest or at constant velocity, are completely equivalent. The frame of reference in which this equivalence holds, termed as the ‘inertial frame’, is determined only by the initial conditions. The object is then said to be in ‘equilibrium’ in the inertial frame. Newtonian mechanics accounts for any change in equilibrium by invoking a stimulus (referred to as ‘force’). The object on which the force acts responds to it by undergoing an acceleration, which is linearly proportional to the force:

\[ \vec{F} = \frac{d\vec{p}}{dt} = m\vec{a} = m\frac{d^2\vec{x}}{dt^2} \] (in usual notations).

The proportionality \( m \) between the acceleration (response of the object) and the stimulus (force) is the inertia (or mass) of the object. This second-order differential equation, named Newton’s second law, embodies causality/determinism: the cause (force) determines the effect (acceleration). Solution to this second-order differential equation requires two constants of integration, which come from the initial conditions of position and velocity.
Figure 1. Alternative trajectories of a system with one degree of freedom. Newtonian mechanics explains the selection of the path taken by the object to reach the final position $q_f$ at a later time $t_f$ if it started out from position $q_i$ at initial time $t_i$ by natural laws described by the cause-effect relationship $F = m \frac{d^2q}{dt^2}$. Hamilton’s variational principle accounts for the selection of the path as one that makes ‘action’ (defined below by (1), Section 3) an extremum. Lagrange’s equation included in this figure is discussed in Section 3.

The ‘variational principle’ offers an alternative formulation of mechanics. It does not build on the principle of causality and determinism contained in Newton’s second law. This method is extensively discussed in several texts [1, 2]. There have been several recommendations [3–5] in recent years to propose that these methods be introduced rather early in college curriculum. These recommendations are inspired by the fact that the variational principle lends itself easily to be adapted in frontier research areas in physics, apart from providing some distinctive advantages over the Newtonian methods. Specifically, the variational principle, often referred to as Hamilton’s variational principle, is best adapted to provide a ‘backward integration’ for the development of quantum theory which supersedes classical mechanics. Besides, the formulation of the laws of physics on the basis of the variational principle also provides transparent relationships between symmetry and conservation principles which
are of fundamental importance in understanding the laws of nature [6].

It nearly feels like a deep conspiracy of nature that the principle of causality/determinism and the variational principle produce results that are completely equivalent, even as the former makes no use of the variational principle, and the latter makes no use of force. The essence of the difference between Newtonian formulation and the alternative based on the variation principle, due to Lagrange and Hamilton is illustrated in Figure 1.

The variational principle is also known as the ‘principle of extremum action’. The general mathematical framework for the development and application of this technique is the ‘calculus of variation’. Its beginnings can be traced to the solution provided by Isaac Newton, to the famous ‘brachistochrone problem’, which was posed by Johann (also known as Jean or John) Bernoulli in 1696. We hasten to add, nevertheless, that the formulation of the principle has a rich and intense history which dates back to periods even before the Bernoulli–Newton episode which we highlight in this paper.

2. The Challenge Posed by Johann Bernoulli

In essence, the brachistochrone problem posed by Johann Bernoulli is the following:

*Given two points A and B in a vertical plane (Figure 2), what is the curve traced out by a particle acted on only by gravity, which starts at A and reaches B in the shortest time?*

In Greek, ‘brachistos’ means ‘the shortest’ and ‘chronos’ means ‘time’, hence the name ‘brachistochrone’ of the curve along which the object traverses in the least time. The history of this problem is rather romantic, and is worth visiting.

Declared Bernoulli:
Figure 2. The solution to the brachistochrone problem is not along the shortest distance, which is of course the straight line, but along a path described by a cycloid. This is now best understood in the framework of the calculus of variation.

“I, Johann Bernoulli, address the most brilliant mathematicians in the world. Nothing is more attractive to intelligent people than an honest, challenging problem, whose possible solution will bestow fame and remain as a lasting monument. Following the example set by Pascal, Fermat, etc., I hope to gain the gratitude of the whole scientific community by placing before the finest mathematicians of our time a problem which will test their methods and the strength of their intellect. If someone communicates to me the solution of the proposed problem, I shall publicly declare him worthy of praise.”

Bernoulli knew the solution, but he challenged other mathematicians to tackle this problem. Five solutions to this problem are famous. Those are by (1) Isaac Newton (1642–1727), (2) Johann’s younger brother Jacob Bernoulli (also known as James or Jacques Bernoulli) (1655–1705), (3) G W Leibniz (1646–1716), (4) Guillaume-François Antoine, Marquis de l’Hôpital (1661–1704), and (5) Johann Bernoulli himself. Newton’s solution was published anonymously by the Royal Society (with the help of Charles Montague), but Johann Bernoulli immediately recognized that it was Newton’s. He said, “we know the lion by his claw.” Newton’s solution (Figure 3a) [7], was essentially the following:
From the given point A, let there be drawn an unlimited straight line APCZ parallel to the horizontal, and on it let there be described an arbitrary cycloid AQP meeting the straight line AB (assumed drawn, and produced if necessary) at the point Q, and further a second cycloid ADC whose base and height are to the base and height of the former as AB is to AQ respectively. This last cycloid will pass through the point B, and it will be that curve along which a weight, by the force of its gravity, shall descend most swiftly from the point A to the point B.

Jacob Bernoulli then solved it by what we now call as the ‘separation of variables’ method.

Leibniz sent Bernoulli his solution, just a week after the problem was posed. Leibniz provided a trajectory along which the body would move, and called it ‘tachystoptota’, which means ‘curve

The principle of extremum action is a very profound mathematical design which in its various incarnations explains why objects move the way they do!

Figure 3. (a) Newton’s response to Johann Bernoulli’s challenge. (b) Marquis de l’Hôpital’s graphical solution. (c) Johann Bernoulli’s solution. The cycloid path emerges as the limit of the strips.
of quickest descent’. Leibniz provided the correct trajectory, but did not recognize it as the cycloid. The graphical solution [7] by Marquis de l’Hôpital (Figure 3b) was published 300 years later by Jeanne Peiffer in 1988. Johann Bernoulli’s solution (Figure 3c) divided the vertical plane into strips [7]. Bernoulli proposed that the particle followed a piecewise linear path in each strip. The problem then reduced to the determination of the angle the straight line segment in each strip made, and to determine this Bernoulli invoked Fermat’s principle – that light always follows the shortest possible time of travel. In the limit, as the strips become infinitely thin, the line segments tend to a curve where at each point, the angle the line segment made with the vertical, becomes the angle the tangent to the curve makes with the vertical. At the end of the solution, Johann Bernoulli said:

“Before I end, I must voice once more the admiration I feel for the unexpected identity of Huygens’ tautochrone and my brachistochrone. I consider it especially remarkable that this coincidence can take place only under the hypothesis of Galileo, so that we even obtain from this a proof of its correctness. Nature always tends to act in the simplest way, and so it here lets one curve serve two different functions, while under any other hypothesis we should need two curves.”

The principle to which Johann Bernoulli alluded is known after Pierre Fermat (1601–1665, Figure 4). It is the precursor to the variational principle. Originally, it explained why a ray of light takes the path it does (including refraction) when it meets the boundary of an interface between two media. The principle of extremum action is a very profound mathematical design which in its various incarnations explains why objects move the way they do! For example, it explains why a stream of water running down a hill would take the path of steepest descent, and also the trajectories of any mechanical system subject to initial conditions and governed by an equation of motion that is derivable from the principle of extremum action.

Rich contributions following Fermat’s work by Maupertuis (1698–1759) and Euler (1707–1783) culminated in the works of La-
3. The Variational Principle

The variational principle [1–5] rests on the premise similar to Galileo–Newtonian mechanics that an isolated mechanical state of the system with one degree of freedom is represented by the pair \((q, \dot{q})\), with \(q\) representing its position, and \(\dot{q}\) its velocity. However, the position and velocity in this scheme have a broader sense which accommodates two distinctive features:

(i) \(q\) is the instantaneous coordinate of the object, called its ‘generalized coordinate’, in the sense that it provides the essential information about the position of the object under study after accounting for any constraints (if any) which provide partial information about the position.

(ii) \(\dot{q}\) provides the instantaneous velocity, called the ‘generalized velocity’. It is the time-derivative of the generalized coordinate.

\(q\) and \(\dot{q}\) need not necessarily have the dimensions \(L\) and \(LT^{-1}\). These can be, for example, an angle and angular velocity. Likewise, the ‘generalized momentum’ is defined as \(p = \frac{\partial L}{\partial \dot{q}}\). It is not necessarily the ‘mass times velocity’ of the Newtonian mechan-
ics, although it may well be just that as a possible special case.

The difference between a generalized coordinate and a physical coordinate in 3-dimensional space becomes important when we consider an N-particle system subjected to some constraints. If there are m constraints, the number of degrees of freedom reduces from $3N$ to $(3N - m)$, and these are then represented by $(3N - m)$ generalized coordinates. We refer the readers to some excellent text books [1, 2], on classical mechanics which detail the generalized coordinates, velocities, and momenta.

Hamilton’s variational principle states that a mechanical system evolves in such a way that a mathematical quantity called action is an extremum; action being defined as the definite integral $S$ (over time $t$) of the Lagrangian function $L(q(t), \dot{q}(t))$, which itself will be defined shortly:

$$S = \int_{t_1}^{t_2} L(q(t), \dot{q}(t)) dt.$$  \hspace{1cm} (1)

For action to be an extremum, any variation in it when determined along alternative trajectories, such as those shown in Figure 1, must be zero; i.e., just as the derivative of a function of one-variable is zero at an extremum (maximum or minimum). At these points, the function is ‘stationary’ and any variation in it is zero. We therefore ask the following question: what is the variation or change in action, if we imagine the system to evolve along alternative paths from the initial time $t_1$ to the end time $t_2$ of the time-interval under consideration? Hamilton’s principle tells us that the system evolves along a path for which the variation in action is zero. Nature selects that path for which the action is an extremum.

The change in action with respect to (a) the path along which the system evolves as per the laws of nature, and (b) an alternative imaginable path, is:

$$\delta S = \int_{t_1}^{t_2} L(q + \delta q, \dot{q} + \delta \dot{q}) dt - \int_{t_1}^{t_2} L(q, \dot{q}) dt = 0.$$  \hspace{1cm} (2)
From (2), it follows that:

$$0 = \delta S = \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right) dt. \quad (3)$$

Now, \(\delta q\) is an arbitrary change in the path; at each instant one could think of various alternative paths for each of which \(\delta q(t)\) would be different. The condition under which (3) would hold for arbitrary \(\delta q\) is, as shown below in Section 4, that:

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}} \right) = 0. \quad (4)$$

(4) is known as the Euler–Lagrange equation, or in the present context only as the Lagrange equation. It is an ‘equation of motion’, just like the one in Newton’s second law, but for the degree of freedom \(q\) rather than for each physical coordinate. If there are \(N\) generalized coordinates (one for each degree of freedom) \(\{q_i, i = 1, 2, 3, ..., N\}\), then each \(q_i\) satisfies the Lagrange equation corresponding to it, and \(q\) in (4) must be replaced by \(q_i\).

4. The Euler–Lagrange Equation

Leonhard Euler (1707–1783) is one of the most celebrated mathematician-physicist of all times, respected for his seminal contributions to a large number of disciplines, including mechanics, optics, fluid dynamics, and musicology apart from fundamental mathematics. In order to appreciate the solution to the brachistochrone problem, we visit the variational principle [8], which provides the basis for (4). Consider a function \(\phi_\gamma(\xi)\) of an independent variable \(\xi\). Just how the function \(\phi_\gamma(\xi)\) can be written in terms of \(\xi\), however depends on a continuous parameter \(\gamma\), which therefore appears as subscript. For example, we may write:

$$\phi_\gamma(\xi) = \phi_{\gamma=0}(\xi) + \gamma \chi(\xi). \quad (5)$$

We now consider another function \(\psi_\gamma\) of \(\xi\) which has a rather
complicated dependence on $\xi$, but it can be expressed in a simple manner by expressing it as a combination of dependence on three terms, (i) direct dependence of $\psi$ on $\xi$ as some simple function of $\xi$, (ii) dependence of $\psi$ on $\phi(\xi)$, which in turn of course depends on $\xi$ with details determined by the parameter $\gamma$, and (iii) also a dependence of $\psi$ on $\frac{d}{d\xi}\phi(\xi)$. We can then write $\psi$ in terms of its dependencies on 3 arguments:

$$\psi = \psi\left(\phi(\xi), \frac{d}{d\xi}\phi(\xi), \xi\right).$$  \hspace{1cm} (6)

The Euler equation is a mathematical condition that must be satisfied for the integral $I_\gamma$ to be an extremum with regard to the variations in its value for different $\gamma$.

The third argument of $\psi$ indicates the explicit dependence on $\xi$ of $\psi(\xi)$, normally expressible as some simple function of $\xi$. In addition to such an explicit dependence of $\psi$ on $\xi$, the first two arguments in (6) stand for an implicit dependence of $\psi$ on $\xi$, through its dependence on $\phi(\xi)$ and $\frac{d}{d\xi}\phi(\xi)$. All the three arguments are required to express the complete dependence of $\psi$ on $\xi$. Leaving out one or the other would limit our capacity to know $\psi$ in terms of $\xi$. However, not all of the 3 arguments may always be required. We now consider the following definite integral of $\psi(\xi)$ over the independent variable $\xi$:

$$I_\gamma = \int_{\xi_1}^{\xi_2} \psi\left(\phi(\xi), \frac{d}{d\xi}\phi(\xi), \xi\right) d\xi.$$  \hspace{1cm} (7)

The Euler equation is a mathematical condition that must be satisfied for the integral $I_\gamma$ to be an extremum with regard to the variations in its value for different $\gamma$. The parameter $\gamma$ influences how $\phi(\xi)$ depends on $\xi$. The extremum can be either a maximum or a minimum, or even a saddle point. Specifically, it is the stationary property of the extremum that is underscored in this analysis. Suppose the extremum value of $I_\gamma$ occur at $\gamma = 0$. Any change in $\gamma$ therefore results in increasing (or decreasing) the value of $I_\gamma$ if the extremum is a minimum (or maximum). As we change $\gamma$, the
rate at which the integral $I_\gamma$ changes is given by:

\[
\frac{\partial I_\gamma}{\partial \gamma} = \frac{\partial}{\partial \gamma} \int_{\xi_1}^{\xi_2} \psi_\gamma \left( \phi_\gamma(\xi), \frac{d}{d\xi} \phi_\gamma(\xi), \xi \right) d\xi, \tag{8a}
\]

i.e.,

\[
\frac{\partial I_\gamma}{\partial \gamma} = \int_{\xi_1}^{\xi_2} \left[ \frac{\partial}{\partial \gamma} \psi_\gamma \left( \phi_\gamma(\xi), \frac{d}{d\xi} \phi_\gamma(\xi), \xi \right) \right] d\xi, \tag{8b}
\]

hence,

\[
\frac{\partial I_\gamma}{\partial \gamma} = \int_{\xi_1}^{\xi_2} \left[ \frac{\partial \psi_\gamma}{\partial \phi_\gamma} \frac{\partial \phi_\gamma}{\partial \gamma} + \frac{\partial \psi_\gamma}{\partial \phi_{\gamma'}} \frac{\partial \phi_{\gamma'}}{\partial \gamma} \right] d\xi, \tag{8c}
\]

where, 

\[
\phi'_{\gamma} = \frac{\partial \phi_\gamma}{\partial \xi}. \tag{8d}
\]

From (5) we have:

\[
\frac{\partial \phi_\gamma}{\partial \xi} = \gamma \frac{\partial \chi}{\partial \xi} \quad \text{i.e.,} \quad \phi'_{\gamma} = \gamma \chi'. \tag{9a}
\]

Hence, \( \frac{\partial \phi_\gamma}{\partial \gamma} = \chi(\xi) \) and \( \frac{\partial}{\partial \gamma} \phi'_\gamma(\xi) = \chi'(\xi). \) \tag{9b}

Accordingly, using (9b) in (8c), we have:

\[
\frac{\partial I_\gamma}{\partial \gamma} = \int_{\xi_1}^{\xi_2} \left[ \frac{\partial \psi_\gamma}{\partial \phi_\gamma} \chi(\xi) + \frac{\partial \psi_\gamma}{\partial \phi_{\gamma'}} \chi'(\xi) \right] d\xi. \tag{10a}
\]

Now, integrating the second of the above terms by parts, we get:

\[
\frac{\partial I_\gamma}{\partial \gamma} = \int_{\xi_1}^{\xi_2} \left[ \frac{\partial \psi_\gamma}{\partial \phi_\gamma} \chi(\xi) \right] d\xi + \left. \left[ \frac{\partial \psi_\gamma}{\partial \phi_{\gamma'}} \chi(\xi) \right] \right|_{\xi_1}^{\xi_2} - \int_{\xi_1}^{\xi_2} \left[ \frac{d}{d\xi} \left( \frac{\partial \psi_\gamma}{\partial \phi_{\gamma'}} \right) \chi(\xi) \right] d\xi. \tag{10b}
\]

Now, \( \chi(\xi_2) = 0 = \chi(\xi_1) \), \( \xi_1 \) and \( \xi_2 \) being the end points of the range of the definite integral, and hence the term in the middle of (10b) is zero.

\[
\text{Hence,} \quad \frac{\partial I_\gamma}{\partial \gamma} = \int_{\xi_1}^{\xi_2} \left[ \frac{\partial \psi_\gamma}{\partial \phi_\gamma} - \frac{d}{d\xi} \left( \frac{\partial \psi_\gamma}{\partial \phi_{\gamma'}} \right) \right] \chi(\xi) d\xi. \tag{10c}
\]

Now, to get \( \frac{\partial I_\gamma}{\partial \gamma} = 0 \) when \( I_\gamma \) is an extremum for arbitrary \( \chi(\xi) \), the necessary condition would be:

\[
\left[ \frac{\partial \psi_\gamma}{\partial \phi_\gamma} - \frac{d}{d\xi} \left( \frac{\partial \psi_\gamma}{\partial \phi_{\gamma'}} \right) \right] = 0. \tag{11}
\]
This is a necessary condition for the integral $I_\gamma$ to have a stationary value and is known as the Euler equation. Identifying \{I, \psi, \phi, \xi\} respectively as \{S, L, q, t\}, we get (4) of Section 3.

5. The Brachistochrone

We now address the brachistochrone problem mentioned in Section 1 using the Euler–Lagrange formalism. Let us take a quick look at Figure 2. We assume a mass $m$ to go down under gravity along frictionless curves such as those sketched in Figure 2. Bernoulli’s question was to determine the curve that would enable the object to reach the point B, if $m$ is left under gravity from the point A at zero initial speed, in the least time. We use the method of variational calculus described above to determine this path. We take the zero of the energy scale to be given by the energy of the particle at rest at point A. As the particle falls under gravity, it would pick up kinetic energy, while its potential energy would become negative, the total energy remaining constant. The potential energy at a point where the particle’s coordinate is $y$ is $V(y) = -mg y$, and the kinetic energy would therefore be $T(y) = \frac{1}{2} m v^2 = -V(y) = mgy$. Accordingly, the particle’s speed at any instant is $v = \sqrt{2gy}$. The time interval $\delta t$ taken by the particle to traverse a tiny, (infinitesimal) distance $\delta s$ is then:

$$\delta t = \frac{\delta s}{v} = \frac{\delta x}{v} \sqrt{1 + \left(\frac{\delta y}{\delta x}\right)^2} \sqrt{2gy}.$$ (12)

From the above equation, we see that the shape of the curve along which the particle would traverse under gravity in least time can be parametrized in terms of the height function $y$ instead of the time parameter $t$. The total time taken for the particle to reach...
point B, at which the x-coordinate is \( x_B \), is therefore:

\[
\tau = \int_0^{x_B} \left( \frac{\sqrt{1 + \left( \frac{dy}{dx} \right)^2}}{\sqrt{2gy}} \right) dx = \int_0^{x_B} \left( \frac{\sqrt{1 + y'^2}}{\sqrt{2gy}} \right) dx,
\]

(13a)

\[
\left\{ \text{with } y' = \frac{dy}{dx} \right\}.
\]

hence, \( (\sqrt{2}) \tau = \int_0^{x_B} \left( \frac{\sqrt{1 + y'^2}}{\sqrt{2gy}} \right) dx = \int_0^{x_B} \Psi(y, y') dx,
\]

(13b)

where, \( \Psi(y, y') = \frac{\sqrt{1 + y'^2}}{\sqrt{2gy}} \).

(13c)

From the Euler equation (11), and equivalently from (4), the condition that the time \( \tau \) is an extremum is:

\[
0 = \left[ \frac{\partial \Psi}{\partial y} - \frac{d}{dx} \frac{\partial \Psi}{\partial y'} \right] y' = \left[ \frac{\partial \Psi}{\partial y} y' - y' \frac{d}{dx} \frac{\partial \Psi}{\partial y'} \right].
\]

(14)

Now,

\[
\frac{d}{dx} \left[ \Psi(y, y') \right] = \frac{\partial \Psi}{\partial y} dx + \frac{\partial \Psi}{\partial y'} dx = \frac{\partial \Psi}{\partial y} y' + \frac{\partial \Psi}{\partial y'} y''.
\]

(15a)

\[i.e., \quad \frac{\partial \Psi}{\partial y} y' = \frac{d}{dx} \left[ \Psi - y' \frac{\partial \Psi}{\partial y'} \right].\]

(15b)

Hence,

\[
0 = \frac{d}{dx} \left[ \Psi - y' \frac{\partial \Psi}{\partial y'} \right] - y' \frac{d}{dx} \frac{\partial \Psi}{\partial y'}
\]

(16a)

\[i.e., \quad 0 = \frac{d}{dx} \left[ \Psi - y' \frac{\partial \Psi}{\partial y'} \right].\]

(16b)

Essentially, this means that:

\[
\left[ \Psi - y' \frac{\partial \Psi}{\partial y'} \right] = \kappa, \; \text{a constant.}
\]

(17)

(17) provides a relationship between the coordinate \( y \) and its derivative \( y' \). This relation provides the equation to the trajectory for the fastest path for the particle to get to point B from point A under
gravity, as detailed below.

Now,

$$\frac{\partial \Psi}{\partial y'} = \frac{\partial}{\partial y'} \sqrt{1 + y'^2} = \frac{y'}{\sqrt{gy(1 + y'^2)}}.$$  \hfill (18a)

Hence we get,

$$\left[ \Psi - \frac{y'^2}{\sqrt{gy(1 + y'^2)}} \right] = \frac{C}{\sqrt{g}}, \quad \text{a constant.} \hfill (18b)$$

Separating the factor $\sqrt{g}$ and placing it in the denominator of (18b) enables us to write subsequent relations in a very convenient form.

i.e., $$\frac{1}{\sqrt{y(1 + y'^2)}} = C$$ or squaring $$y(1 + y'^2) = \frac{1}{C^2} = \alpha.$$ \hfill (18c)

i.e., $$y\left(\frac{dy}{dx}\right)^2 = \alpha - y.$$ \hfill (18d)

We thus get the following relation between $x$ and $y$ which essentially gives us the trajectory, i.e., the shape of the curve, along which the time taken for the mass $m$ to traverse under gravity from point A to B would be the least:

$$\sqrt{\frac{y}{(\alpha - y)}} (\delta y) = (\delta x),$$ \hfill (19a)

i.e., $$x = \int \sqrt{\frac{y}{(\alpha - y)}} dy.$$ \hfill (19b)

Using the coordinate system shown in Figure 5, a simple substitution:

$$y = \frac{\alpha}{2} (1 - \cos \theta) = \alpha \sin^2 \frac{\theta}{2}$$ \hfill (20)

renders the relationship in a familiar form:

$$x = \alpha \int \sin^2\left(\frac{\theta}{2}\right) d\theta = \frac{\alpha}{2} (\theta - \sin \theta) + \beta,$$ \hfill (21)
where $\beta$ is the constant of integration. From the initial condition $\theta = 0$ when $y = 0$, we get $\beta = 0$. The other constant to be determined is $\alpha$, and this is also easily determined since the path must pass through point B for which the coordinates are given, say, by $(x_A, y_B)$. Hence, $x_B = \frac{a}{2} (\theta_b - \sin \theta_b)$ and $y_B = \frac{a}{2} (1 - \cos \theta_b)$. From these two equations, one can determine the constant $\alpha$. The equation to the path of the brachistochrone (i.e., the path along which the time taken is the least) is therefore given by:

$$x = \frac{\alpha}{2} (\theta - \sin \theta); \quad y = \frac{\alpha}{2} (1 - \cos \theta),$$

which defines a ‘cycloid’ (Figure 5) with a radius of the circle given by $\frac{a}{2}$. A cycloid can be described as the locus of a point on the rim of a circle, such as bicycle wheel, while the center of the circle itself traverses along a straight line.

From Figure 5, we see that $\theta = 0$, when P is at the origin. The circle rolls down through a horizontal distance OS along the x-axis. This distance is exactly equal to the arc-length $PS = r\theta$, since the radius of the circle is $r$. The Cartesian coordinates $(x, y)$ of the center of the circle C, are obtainable in terms of the parameter $\theta$, since $x_C = OS = \text{arc-length } PS = r\theta$, and $y_C = CS = \text{radius of the circle}$. 

Figure 5. The solution to the brachistochrone problem turns out to be a cycloid, described in this figure in terms of angle $\theta$ in terms of which the $(x, y)$ coordinates are given by (23).
circle = \( r \). The Cartesian coordinates of point \( P \) are given by:

\[
\begin{align*}
    x_P &= OS - PQ = r\theta - r \sin \theta = r(\theta - \sin \theta), \\
    y_P &= CS - CQ = r - r \cos \theta = r(1 - \cos \theta).
\end{align*}
\]

The above equations describe a cycloid with \( r = \frac{\pi}{2} \).

In Figure 6a is shown a small brachistochrone (cycloid) model, along with three other slopes which we machined from an acrylic \((C_5O_2H_8)_n\) sheet of thickness 15 mm. To machine the brachistochrone, the curvature of an acrylic sheet was contoured to produce the cycloid from numerical data generated by a small computer program which gave us 400 \((x, y)\) points as per (23), with the angle \( 0 \leq \theta \leq 4\pi \) in steps of \( \delta \theta = \frac{\pi}{100} \). The resulting numerical data generated the cycloid curve, which was plotted using Mastercam X7 CAD/CAM software. We generated the ‘G-Codes’ to machine the acrylic sheet in CNC (Computerized Numerical Controller) – VMC (Vertical Milling Center) machine. A 6 mm carbide flat end-mill cutter was used to machine the acrylic sheet. Care was taken to use sufficient coolant at the machining area to avoid the melting of the acrylic sheet and get smooth curve edges. Since the cycloid curve represents the brachistochrone path, this model is labeled P\(_B\). The width of the track was about 1.45 cm, and brachistochrone was set in a frame of height 22.9 cm and breadth 25.6 cm.

For comparison, three additional models, with different curvatures (including a straight shortest path) were machined from acrylic sheets. For bookkeeping, these three additional slopes were labeled as P\(_L\), P\(_S\), and P\(_D\), to respectively represent the linear (straight) path between the points A and B, a path shallower than the brachistochrone, and a path deeper than it. The dimensions of the four models were machined to enable the release from rest a metallic object held by an electromagnet (Figure 6b), to come down under gravity alone, from identical start-points (A) of release, up to identical end-points (B) of their respective trajectories (Figures 6a and 6b). The start-time at the release of the metallic object at point A, and the end-time when the object crosses point B, were recorded using two photo-gates (Figure 6b) and the time
interval $\delta t$ the object took to traverse the four paths $P_L$, $P_S$, $P_B$, and $P_D$ from A to B were measured.

The experiment was done 50 times, and the results for the time intervals along the four paths are presented in the table below:

It is clear from Table 1 that the experimental results are completely in accordance with the conclusions arrived from the variational principle. The cycloid path $P_B$ is the brachistochrone. We note that in our experiment, the object that traversed the four paths mentioned above was a tiny metallic ball of diameter $\sim 13$ mm. Since the ball would roll down rather than skid, the rotational kinetic energy would need to be added to the translational kinetic energy in obtaining (12). Hence, the kinetic energy at

**Figure 6a.** The experimental setup (left panel) to perform the brachistochrone experiment. The time intervals taken by an object to come down under gravity alone from identical points ‘A’ to ‘B’ (right panel) were measured using two photogates, seen in the left panel (also seen in Figure 7, below). The experiment was done on four different paths – $P_L$, $P_S$, $P_B$, and $P_D$ as described in the text.

**Figure 6b.** The brachistochrone model placed for conducting the experiment to determine the time-interval $\delta t$ to measure the time taken by a mass $m$ to traverse under gravity from point ‘A’ to ‘B’ shown in Figure 6a. A metallic object is released from rest (zero initial speed) by the electromagnet at point A. The photogates A and B respectively record the start and end instants of time from whose difference the interval $\delta t$ is determined.
Table 1. The time-interval $\delta t$ (recorded by the photo-gates) taken by the metallic object released at rest from point ‘A’ to reach point ‘B’ as shown in Figure 6.

<table>
<thead>
<tr>
<th>Path</th>
<th>Time Difference $\delta t$</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_L$</td>
<td>0.345</td>
<td>0.0033</td>
</tr>
<tr>
<td>$P_S$</td>
<td>0.295</td>
<td>0.002</td>
</tr>
<tr>
<td>$P_B$ (cycloid)</td>
<td>0.286</td>
<td>0.0025</td>
</tr>
<tr>
<td>$P_D$</td>
<td>0.312</td>
<td>0.0023</td>
</tr>
</tbody>
</table>

a point where the particle’s coordinate is $y$ and potential energy is $V(y) = -mg y$, would be $-T(y) = \frac{1}{2} \left( m + \frac{I}{r^2} \right) v^2$, instead of $T(y) = \frac{1}{2} mv^2 = -V(y) = mg y$. The total energy would be conserved and hence the particle’s speed would be $v = \sqrt{2 \left( \frac{m}{m + I/r^2} \right) gy}$ instead of $v = \sqrt{2gy}$. Effectively, this would only scale the acceleration due to gravity $g$ by a dimensionless factor $\frac{m}{m + I/r^2}$, where $m$ is the mass of the ball, $I$ its moment of inertia, and $r$ its radius, without affecting the primary analysis.

There are other brachistochrone paths in nature; two of which were machined in our workshops which we call as the ‘brachistochrone carom boards’ and are shown in Figure 7. A striker hit from one of the two foci of the elliptic carom board passes through the other. On the parabolic carom board, if the striker strikes the parabolic reflector along the axis of the parabola, then after reflection, it passes through its focus, as shown in Figure 7. The parabolic reflector’s focus had a distance of 14.5 cm for its shortest distance with the parabolic curvature. The major axis of the elliptic carom board was 55.8 cm while its minor axis was 33.5 cm.

The results of the carom board experiments are akin to the path taken by a ray of light, as is well known in the laws of optical reflection, which are completely determined by Fermat’s principle, which is the precursor to the variational principle, as discussed above.
6. Conclusion

There are other situations in nature where the optimal path would be the one as determined by the variational principle. If a life-guard wishes to run and save a person drowning in a river, it is best that he approaches the bank through a path along which he would reach the accident victim fastest, and this must be optimized, considering the fact that the speed at which he would run on land to approach the bank would be different from the speed at which he would swim after jumping in the waters. Clearly, the solution would be given by the variational principle, and the most rapid path would be given as per Fermat’s principle. It is reported that when a bunch of ants approach a grain of food and need to crawl over surfaces on which their speeds are different, they follow Fermat’s principle, and also that when an eagle dives to catch a prey, it does so along a cycloid. These phenomena in nature are curious, but we are not aware of any rigorous analysis of these natural phenomena which seem to be consistent with the variational principle. We trust that the pedagogical analysis of the variational principle in this work, and the accompanying experiments shown in Figures 6 and 7 would aid the students in learning the founding principles of classical mechanics.

Suggested Reading


Figure 7. The brachistochrone carom boards. Carom board strikers’ incident along any line parallel to AQ in the parabolic reflector are reflected back by the parabolic surface through the focus F. In the elliptic reflector, a striker struck from the focus F1 always gets reflected through the second focus F2 regardless of which point it was aimed at.


Downloaded on 21st Nov’, 2016.