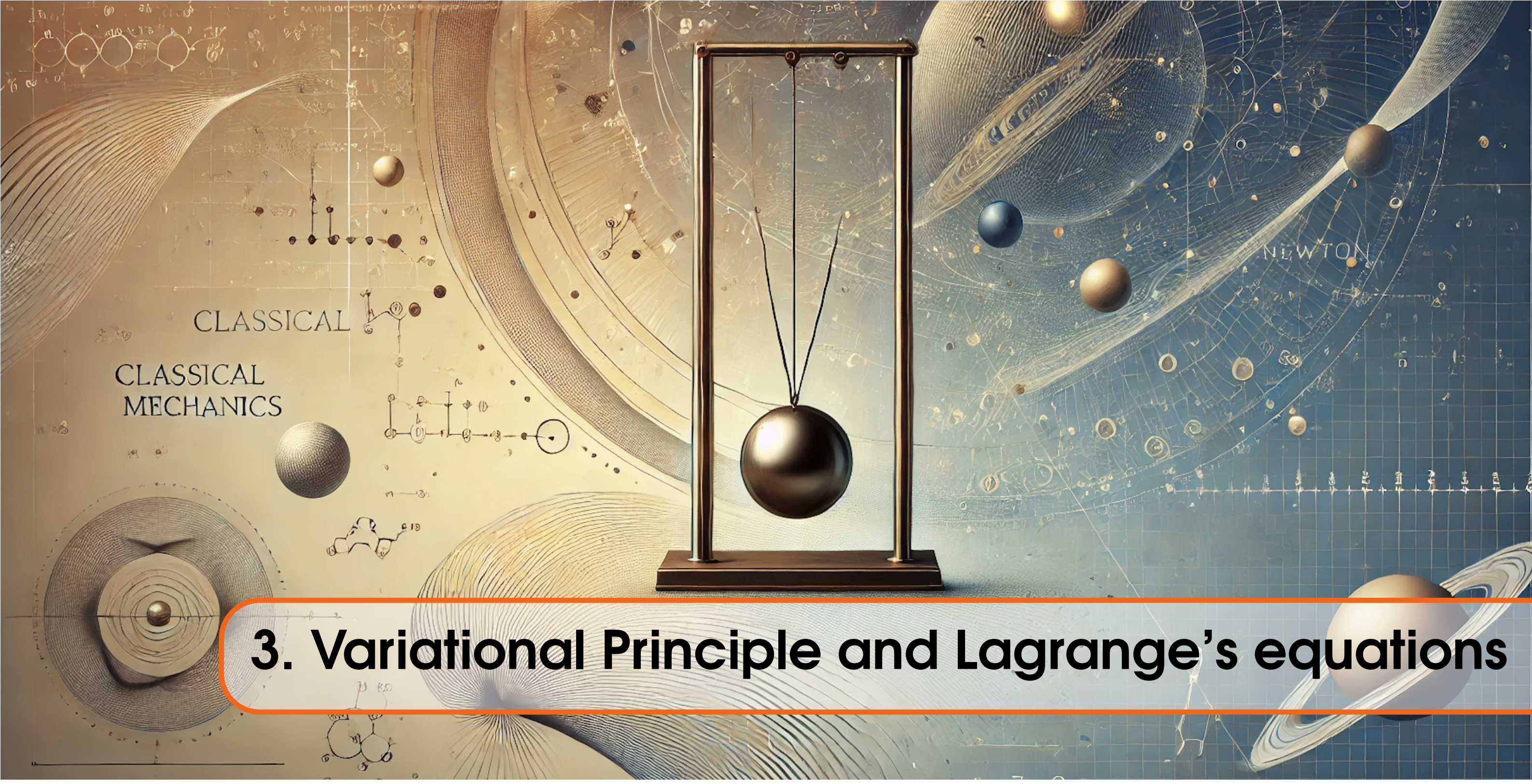
22PH102 CLASSICAL MECHANICS

I Semester

by Paulsamy MURUGANANADAM

Contents

3 Variational Principle and Lagrange's equations 5
3.1 Introduction 5
3.2 Variational principle 6
3.2.1 Hamilton's variational principle for monogenic systems 6
3.2.2 Calculus of variations 7
3.3 Derivation of Lagrange's equations from Hamilton's variational principle 13
3.4 Advantages of Variational Principle formulation 15
3.5 Conservation theorems 16
3.5.1 Conjugate momentum 16
3.5.2 Cyclic (or) Ignorable coordinates 17
3.5.3 Energy Function and the Conservation of Energy 18
3.6 Reference 19



3. Variational Principle and Lagrange's equations

3.1 Introduction

In the previous class, we derived Lagrange's equation by examining the system's instantaneous state and considering small virtual displacements around this state. In other words, we applied a ***differential principle***, such as D'Alembert's principle. Alternatively, the same Lagrange's equations can be obtained using a principle that accounts for the system's entire motion between times t_1 and t_2, along with small virtual variations of this motion from the actual trajectory. Such a principle is referred to as an ***integral principle***.

Before delving into the details, let us first clarify the meaning of the phrase "motion of the system between times t_1 and t_2." The system's instantaneous configuration is represented by the values of the n generalized coordinates $q_1, q_2, \ldots, q_n$, corresponding to a specific point in a Cartesian hyperspace where the q-coordinates serve as the n axes. The n-dimensional space defined by the

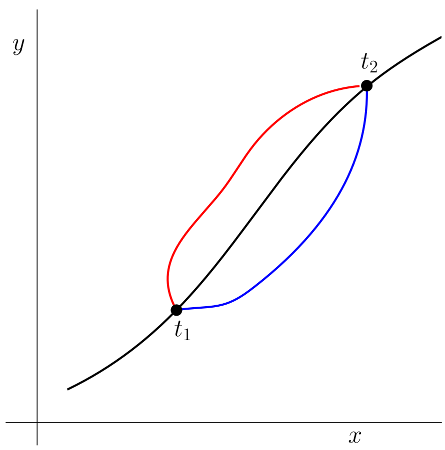


Figure 3.1: Path of the system in configuration space.

n generalized coordinates is known as the ***configuration space***. As time progresses, the system's state changes, causing the corresponding system point to traverse the configuration space and trace a curve called the *path of the motion of the system*. The phrase "motion of the system" refers to the movement of this system point along its path within the *configuration space*. Time serves as a parameter for this curve, with each point on the path corresponding to one or more specific time values.

3.2 Variational principle

The integral known as ***Hamilton's principle*** characterizes the motion of a mechanical system in which all forces, except constraint forces, can be derived from a generalized scalar potential. This potential may depend on the coordinates, velocities, and time. Such a system is referred to as ***monogenic***. If the scalar potential depends explicitly only on the coordinates, the monogenic system is also classified as ***conservative***. Therefore, conservative systems represent a specific subset of monogenic systems.

3.2.1 Hamilton's variational principle for monogenic systems

The motion of the system between times t_1 and t_2 is such that the line integral (known as the ***action*** or the ***action integral***)

$$I = \int_{t_1}^{t_2} L dt, \tag{3.1}$$

where $L = T - V$, has a ***stationary value*** for the ***actual path*** of the motion.

This means that among all possible paths, the system could take from its position at time t_1 to its position at time t_2, the system will follow the path for which the value of the integral (3.1) is ***stationary***. The value of the integral along the actual path is the same, up to first-order infinitesimal terms, as along all nearby paths. The concept of a stationary value for a line integral corresponds to the condition that the first derivative of the integral vanishes.

Hamilton's variational principle can be summarized by stating that the motion is such that the variation of the line integral I is zero for fixed t_1 and t_2:

$$\delta I = \delta \int_{t_1}^{t_2} L(q_1, q_2, \ldots, q_n, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n, t)\, dt = 0. \tag{3.2}$$

Mathematically, this means that the first variation of the action, δI, is zero at the extremum, analogous to the fact that the first derivative of a function is zero at a point of extremum (whether minimum or maximum). In other words, the trajectory is such that the variation of the action, δI, is zero when considering infinitesimal changes in the path, similar to how the first derivative of a function with respect to its variables is zero at a local extremum.

When the constraints are holonomic, ***Hamilton's principle*** (3.2) provides both a necessary and sufficient condition for Lagrange's equations. In other words, Hamilton's principle is a sufficient condition for deriving the equations of motion. This allows us to construct the mechanics of monogenic systems based on Hamilton's principle as the fundamental postulate, rather than relying on Newton's laws of motion. This formulation offers several advantages. For example, since the integral I is invariant under transformations of the generalized coordinates used to express L, the equations of motion will always take the Lagrangian form, regardless of how the generalized coordinates are transformed. Additionally, the variational principle provides a widely used approach to describing systems that appear non-mechanical.

3.2.2 Calculus of variations

We shall first explore the methods of the calculus of variations to determine the curve for which a given line integral attains a stationary value.

For simplicity, let us first consider a one-dimensional problem. We have a function $f(y, \dot{y}, x)$ defined along a path $y = y(x)$ between two points, x_1 and x_2, where $\dot{y}$ represents the derivative of y with respect to x, that is, $\dot{y} = \frac{dy}{dx}$. We wish to find a specific path such that the line integral J of the function between x_1 and x_2,

$$J = \int_{x_1}^{x_2} f(y, \dot{y}, x)\, dx, \tag{3.3}$$

has a stationary value relative to paths that differ infinitesimally from the true function $y(x)$. Here, the variable x serves as the analog of the parameter t in the context of variational principles. We restrict our consideration to paths for which the boundary conditions $y(x_1) = y_1$ and $y(x_2) = y_2$ are satisfied, as illustrated in Figure 3.2. Note that Figure 3.2 does not represent the configuration

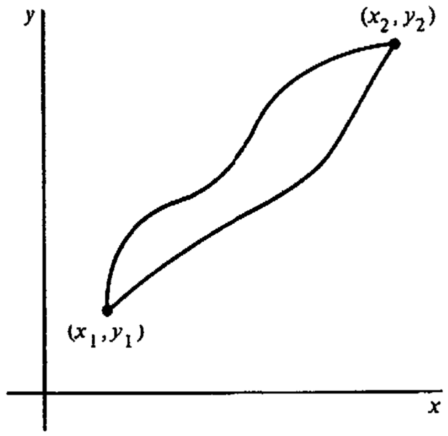


Figure 3.2: Varied paths of the function $y(x)$ in a one-dimensional problem of finding extrema (stationary points).

space. In this case, the configuration space is essentially one-dimensional. In the one-dimensional configuration space, both the correct and varied paths are segments of the same straight line connecting y_1 and y_2. The paths differ only in the functional relationship between y and x.

We put the problem in the familiar context of differential calculus, where stationary points of a function are determined. Since J must attain a stationary value for the correct path compared to any *neighbouring* path, its variation must be zero with respect to a particular set of neighbouring paths parameterized by an infinitesimal parameter α. These paths can be represented as $y(x, \alpha)$, where $y(x, 0)$ corresponds to the correct (stationary) path.

For example, if we select any function $\eta(x)$ that vanishes at $x = x_1$ and $x = x_2$, then a possible set of varied paths can be expressed as:

$$y(x, \alpha) = y(x, 0) + \alpha\eta(x). \tag{3.4}$$

Here, $y(x, 0)$ represents the original path, and $\eta(x)$ is an auxiliary function that satisfies the boundary conditions $\eta(x_1) = \eta(x_2) = 0$. It is assumed that both the correct path $y(x)$ and the auxiliary function $\eta(x)$ are well-behaved, meaning they are continuous and nonsingular in the interval $[x_1, x_2]$, with continuous first and second derivatives throughout the same interval.

For any such family of curves, the functional J depends on the parameter α:

$$J(\alpha) = \int_{x_1}^{x_2} f\left(y(x,\alpha), \dot{y}(x,\alpha), x\right) dx, \tag{3.5}$$

The condition for obtaining a stationary point of J is given by the requirement that the derivative of J with respect to α, evaluated at $\alpha = 0$, must vanish:

$$\left(\frac{dJ}{d\alpha}\right)_{\alpha=0} = 0. \tag{3.6}$$

By the usual method of differentiating under the integral sign, we find that

$$\frac{dJ}{d\alpha} = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial \dot{y}}\frac{\partial \dot{y}}{\partial \alpha}\right) dx. \tag{3.7}$$

Consider the second of these integrals

$$\int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}}\frac{\partial \dot{y}}{\partial \alpha}\, dx = \int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}}\frac{\partial^2 y}{\partial x \partial \alpha}\, dx.$$

Integrating by parts, the above integral becomes

$$\begin{aligned}\int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}}\frac{\partial^2 y}{\partial x \partial \alpha}\, dx &= \int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}}\frac{d}{dx}\left(\frac{\partial y}{\partial \alpha}\right) \\ &= \left.\frac{\partial f}{\partial \dot{y}}\frac{\partial y}{\partial \alpha}\right|_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right)\frac{\partial y}{\partial \alpha}\, dx.\end{aligned} \tag{3.8}$$

The condition on all the varied curves is that they pass through the points (x_1, y_1) and (x_2, y_2). Consequently, the partial derivatives with respect to α at x_1 and x_2 must vanish. Therefore, the first term in (3.8) becomes zero, and equation (3.7) reduces to:

$$\int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right)\right]\left(\frac{\partial y}{\partial \alpha}\right)_0 dx = 0. \tag{3.9}$$

The partial derivative of y with respect to α, as it appears in equation (3.9), is a function of x that is arbitrary but subject to continuity and endpoint conditions (i.e., it vanishes at the endpoints x_1 and x_2). For a parametric family of varied paths given by equation (3.4), this arbitrary function corresponds to $\eta(x)$.

$$\int_{x_1}^{x_2} M(x)\eta(x)\, dx = 0, \tag{3.10}$$

for all well-defined arbitrary functions $\eta(x)$ that are continuous and have continuous second derivatives, then $M(x)$ must identically vanish throughout the interval (x_1, x_2).

Thus, the integral J can have a stationary value only if the following condition is satisfied:

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right) = 0, \tag{3.11}$$

The differential quantity,

$$\left(\frac{\partial y}{\partial \alpha}\right)_0 d\alpha = \delta y, \tag{3.12}$$

represents the infinitesimal deviation of the varied path from the true path $y(x)$ at the point x, and thus corresponds to the virtual displacement. Similarly, the infinitesimal variation of J around the true path can be designated as

$$\left(\frac{dJ}{d\alpha}\right)_0 d\alpha = \delta J. \tag{3.13}$$

The condition that J is stationary for the correct path can be written as

$$\delta J = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right)\right] \delta y\, dx = 0, \tag{3.14}$$

which requires that $y(x)$ satisfies the differential equation (3.11).

■ **Example 3.1** **Shortest distance between two points in a plane:** One of the simplest applications of variational calculus is to prove that the shortest distance between two points is always a straight line. ■

For this purpose, we consider an element of length in a plane given by:

$$ds = \sqrt{dx^2 + dy^2} = dx\sqrt{1 + \left(\frac{dy}{dx}\right)^2},$$

and the total length of any curve connecting points 1 and 2 is:

$$I = \int_1^2 ds = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \equiv \int_{x_1}^{x_2} \sqrt{1+\dot{y}^2}\, dx, \quad \text{where} \quad \dot{y} = \frac{dy}{dx}.$$

The condition for the curve to be the shortest path is that I should be a minimum. This is an example of an extremum problem, expressed by equation (3.3), with $f = \sqrt{1+\dot{y}^2}$. Substituting into (3.11),

$$\frac{\partial f}{\partial y} = 0, \quad \frac{\partial f}{\partial \dot{y}} = \frac{\dot{y}}{\sqrt{1+\dot{y}^2}},$$

we have

$$0 - \frac{d}{dx}\left(\frac{\dot{y}}{\sqrt{1+\dot{y}^2}}\right) = 0,$$

or

$$\frac{\dot{y}}{\sqrt{1+\dot{y}^2}} = c,$$

where c is the constant of integration. This solution is valid only if

$$\dot{y} = a, \tag{3.15}$$

where a is a constant related to c by

$$a = \frac{c}{\sqrt{1+c^2}}.$$

Integrating (3.15), we obtain the equation of a straight line:

$$y = ax + b, \tag{3.16}$$

where b is an another constant.

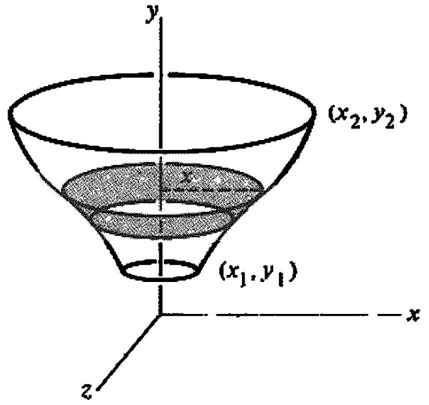


Figure 3.3: Minimum surface of revolution.

■ **Example 3.2 Minimum surface of revolution:** Another example is finding the curve that produces the minimum surface area of revolution. Consider a surface of revolution formed by rotating a curve, which passes between two fixed endpoints (x_1, y_1) and (x_2, y_2) in the xy-plane, about the y-axis as illustrated in Figure 3.3. ■

Here the problem is to find the curve for which the surface area is a minimum. The area of a strip (shaded region in Figure 3.3 is

$$\begin{aligned} dA &= 2\pi x ds \\ &= 2\pi x\sqrt{dx^2+dy^2} \\ &= 2\pi x\sqrt{1+\dot{y}^2}dx. \end{aligned}$$

The the total area is given by

$$A = 2\pi\int_1^2 x\sqrt{1+\dot{y}^2}dx.$$

The extremum of the above integral is given by [cf. equation (3.11)]

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right) = 0,$$

where $f = x\sqrt{1+\dot{y}^2}$, and

$$\frac{\partial f}{\partial y} = 0, \quad \frac{\partial f}{\partial \dot{y}} = \frac{x\dot{y}}{\sqrt{1+\dot{y}^2}}.$$

Substituting the above in the previous equation

$$0 - \frac{d}{dx}\left(\frac{x\dot{y}}{\sqrt{1+\dot{y}^2}}\right) = 0 \quad \Longrightarrow \quad \frac{x\dot{y}}{\sqrt{1+\dot{y}^2}} = a,$$

where a is a constant of integration. Squaring the above and factoring terms, we have

$$\dot{y}^2\left(x^2-a^2\right) = a^2 \quad \Longrightarrow \frac{dy}{dx} = \frac{a}{\sqrt{x^2-a^2}}.$$

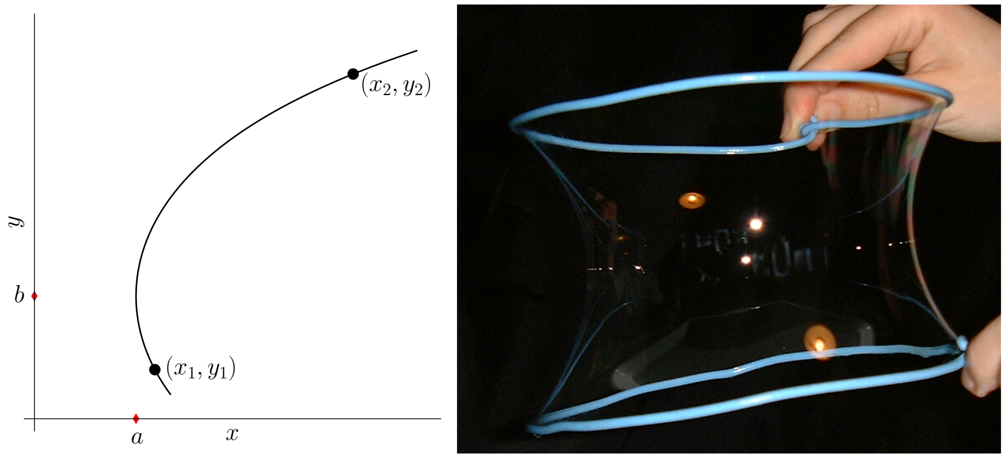


Figure 3.4: [Left panel] General catenary solution for a minimal surface of revolution: $(x_1,y_1) = (0.5927, 0.2)$, $(x_2,y_2) = (1.5685, 1.405)$, $a = 0.5$, and $b = 0.5$. [Right panel] Stretching a soap film between two parallel circular wire loops generates a catenoidal minimal surface of revolution.

The general solution to the above differential equation is

$$y = a\int \frac{dx}{\sqrt{x^2-a^2}} + b = a\cosh^{-1}\left(\frac{x}{a}\right) + b \quad\Longrightarrow\quad x = a\cosh\left(\frac{y-b}{a}\right),$$

which is the equation of a catenary. The two constants of integration, a and b, are determined by the requirement that the curve passes through two given points, as shown in Figure 3.4 (left panel). An example of this is when a soap film is stretched between two parallel circular wire loops, forming a catenoidal minimal surface of revolution, as shown in Figure 3.4 (right panel).

Exercise 3.1 Suppose that $F(y,\dot{y})$ does not depend explicitly on x, so that $\partial F/\partial x = 0$. Show that

$$\frac{d}{dx}\left(\dot{y}\frac{\partial F}{\partial \dot{y}} - F\right) = \dot{y}\left[\frac{d}{dx}\left(\frac{\partial F}{\partial \dot{y}}\right) - \frac{\partial F}{\partial y}\right].$$

Consider the left-hand-side of the above equation

$$\begin{aligned}\frac{d}{dx}\left(\dot{y}\frac{\partial F}{\partial \dot{y}} - F\right) &= \ddot{y}\frac{\partial F}{\partial \dot{y}} + \dot{y}\left(\frac{\partial^2 F}{\partial \dot{y}\partial y}\dot{y} + \frac{\partial^2 F}{\partial \dot{y}^2}\ddot{y}\right) - \left(\frac{\partial F}{\partial y}\dot{y} + \frac{\partial F}{\partial \dot{y}}\ddot{y}\right)\\ &= \dot{y}\left(\frac{\partial^2 F}{\partial \dot{y}\partial y}\dot{y} + \frac{\partial^2 F}{\partial \dot{y}^2}\ddot{y} - \frac{\partial F}{\partial y}\right)\\ &= \dot{y}\left[\frac{d}{dx}\left(\frac{\partial F}{\partial \dot{y}}\right) - \frac{\partial F}{\partial y}\right].\end{aligned}$$

Hence proved. Further, since F obeys the condition

$$\frac{d}{dx}\left(\frac{\partial F}{\partial \dot{y}}\right) - \frac{\partial F}{\partial y} = 0.$$

and if F is not an explicit function of x, one can extract a first integral as

$$\dot{y}\frac{\partial F}{\partial \dot{y}} - F = \text{constant}. \tag{3.17}$$

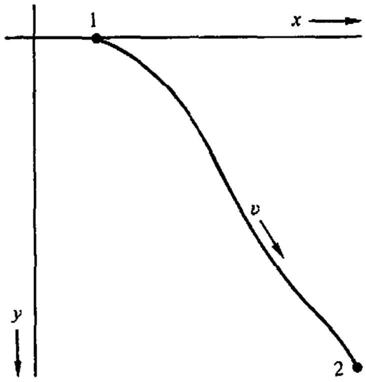


Figure 3.5: Illustration of the brachistochrone problem.

■ **Example 3.3** **The brachistochrone problem:** A classic problem involves finding the curve joining two points along which a particle, falling from rest under the influence of gravity, travels from the higher point to the lower point in the least time. The curve is called the *brachistochrone*, a name derived from the Greek *brachystos*, meaning *shortest*, and *chronos*, meaning *time*. ■

Let v be the speed along the curve. The time required to traverse an infinitesimal arc length ds is ds/v, and the problem is to find the minimum of the integral

$$t_{12} = \int_1^2 \frac{ds}{v}.$$

If the y axis is vertical and is measured down from the initial point of release, the conservation theorem for the energy of the particle can be written as

$$\frac{1}{2}mv^2 = mgy \quad \Longrightarrow \quad v = \sqrt{2gy},$$

and $ds = dx\sqrt{1+\dot{y}^2}$. Then the expression for t_{12} becomes

$$t_{12} = \int_{x_1}^{x_2} \frac{\sqrt{1+\dot{y}^2}}{\sqrt{2gy}} dx,$$

The extremum of the above integral is given by

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right) = 0,$$

with f is identified as

$$f = \frac{\sqrt{1+\dot{y}^2}}{\sqrt{2gy}}.$$

Since f is not an explicit function of x, we can use the first integral given in Eq. (3.17), that is,

$$\frac{\dot{y}^2}{\sqrt{2gy\,(1+\dot{y}^2)}} - \frac{\sqrt{1+\dot{y}^2}}{\sqrt{2gy}} = \text{constant}, \quad \left[\because\ \dot{y}\frac{\partial f}{\partial \dot{y}} - f \ \text{ is constant}\right]$$

By simplifying the above equation and multiplying both sides by -1, we obtain

$$\frac{1}{\sqrt{y\,(1+\dot{y}^2)}} = \frac{1}{c},$$

where c is a constant (with the factor $\sqrt{2g}$ absorbed into c). Squaring both sides and rearranging gives $y(1+\dot{y}^2) = c^2$, from which we obtain

$$\frac{dy}{dx} = \pm\sqrt{\frac{c^2}{y} - 1} \equiv \pm\sqrt{\frac{c^2 - y}{y}}.$$

The equation above is separable and can be solved. With the y-axis oriented vertically downward, one expects the solution $y(x)$ to increase as x increases. Therefore, it is appropriate to consider the $+$ sign. Separating the variables now gives

$$\int dx = \int \sqrt{\frac{y}{c^2 - y}}\, dy.$$

The above integral can be solved by substituting $y = c^2 \sin^2\phi$, so that $dy = 2c^2 \sin\phi \cos\phi\, d\phi$ as

$$\begin{aligned}
\int \sqrt{\frac{c^2 - y}{y}}\, dy &= 2c^2 \int \sin\phi \cos\phi \sqrt{\frac{c^2 \sin^2\phi}{c^2 - c^2 \sin^2\phi}}\, d\phi \\
&= 2c^2 \int \sin^2\phi\, d\phi \\
&= c^2 \int (1 - \cos 2\phi)\, d\phi \equiv \frac{1}{2}(2\phi - \sin 2\phi) + d,
\end{aligned}$$

where d is a constant.

In summary, we have

$$\begin{aligned}
x &= \frac{1}{2}(2\phi - \sin 2\phi) + d, \\
y &= c^2 \sin^2\phi \equiv \frac{1}{2}c^2(1 - \cos 2\phi)
\end{aligned}$$

When $\phi = 0$, these equations yield $x = d$ and $y = 0$. However, we have chosen our coordinate system such that $x = 0$ when $y = 0$; therefore, $d = 0$. The solutions, then, become

$$\begin{aligned}
x &= \frac{1}{2}(2\phi - \sin 2\phi), \\
y &= \frac{1}{2}c^2(1 - \cos 2\phi).
\end{aligned}$$

3.3 Derivation of Lagrange's equations from Hamilton's variational principle

The calculus of variations discussed above can be generalized to the case where the Lagrangian L is a function of many variables q_i, their derivatives $\dot{q}_i$, and time t. In this case, the variation of the integral I can be expressed as:

$$\delta I = \delta \int_{t_1}^{t_2} L(q_1(t), q_2(t), \ldots, q_n(t), \dot{q}_1(t), \dot{q}_2(t), \ldots, \dot{q}_n(t), t)\, dt. \tag{3.18}$$

We treat I as a function of the parameter α, which labels a family of possible curves $q_i(t,\alpha)$. The parameter α is introduced by defining the generalized coordinates as follows:

$$\begin{aligned}
q_1(t,\alpha) &= q_1(t,0) + \alpha\eta_1(t), \\
q_2(t,\alpha) &= q_2(t,0) + \alpha\eta_2(t), \\
&\vdots \\
q_n(t,\alpha) &= q_n(t,0) + \alpha\eta_n(t),
\end{aligned} \tag{3.19}$$

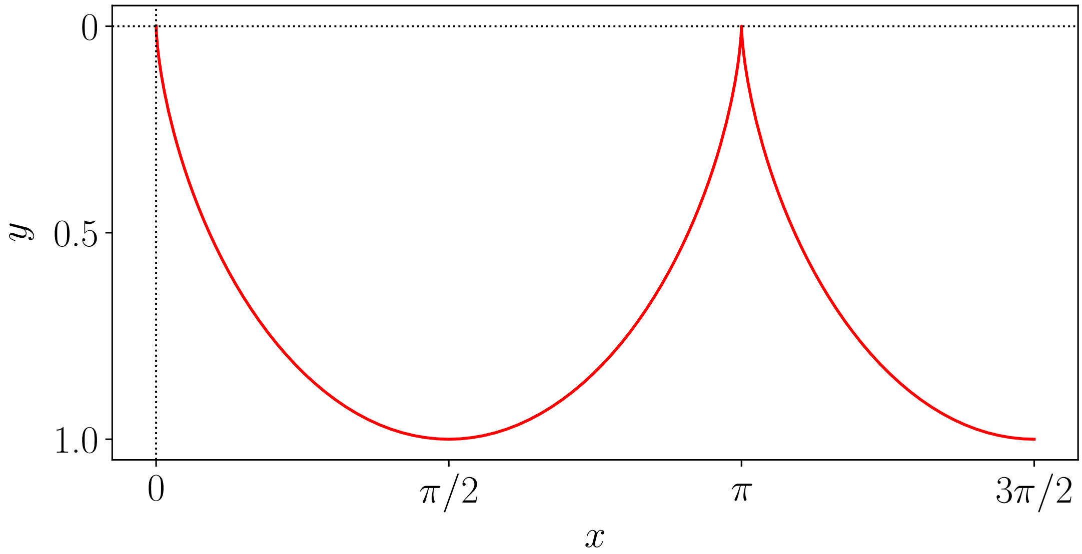


Figure 3.6: Solution to the brachistochrone problem.

where $q_1(t,0), q_2(t,0), \ldots, q_n(t,0)$ are the solutions to the extremum problem that are to be determined, and $\eta_1(t), \eta_2(t), \ldots, \eta_n(t)$ are independent functions of t that satisfy the following conditions: they vanish at the endpoints and are continuous through at least the second derivative. Apart from these conditions, the functions $\eta_1, \eta_2, \ldots, \eta_n$ are completely arbitrary.

The variation of I read as

$$\frac{\partial I}{\partial \alpha} d\alpha = \int_1^2 \sum_i \left(\frac{\partial L}{\partial q_i} \frac{\partial q_i}{\partial \alpha} d\alpha + \frac{\partial L}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial \alpha} d\alpha \right) dt. \tag{3.20}$$

The second sum in the above equation can integrated by parts as follows.

$$\begin{aligned} \int_1^2 \frac{\partial L}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial \alpha} dt = \int_1^2 \frac{\partial L}{\partial \dot{q}_i} \frac{\partial^2 q_i}{\partial t \partial \alpha} dt = \left. \frac{\partial L}{\partial \dot{q}_i} \frac{\partial q_i}{\partial \alpha} \right|_1^2 - \int_1^2 \frac{\partial q_i}{\partial \alpha} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) dt. \\ = -\int_1^2 \frac{\partial q_i}{\partial \alpha} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) dt, \end{aligned} \tag{3.21}$$

where the first term vanishes because, $\frac{\partial q_i}{\partial \alpha} = \eta_i(t)$, and $\eta_i(t_1) = 0$, $\eta(t_2) = 0$ (all curves pass through the end points). Substituting equation (3.21) into equation (3.19), we obtain

$$\begin{aligned} \frac{\partial I}{\partial \alpha} d\alpha &= \int_1^2 \sum_i \left[\frac{\partial L}{\partial q_i} \frac{\partial q_i}{\partial \alpha} d\alpha - \frac{\partial q_i}{\partial \alpha} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) d\alpha \right] dt, \\ &= \int_1^2 \sum_i \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \right] \frac{\partial q_i}{\partial \alpha} d\alpha \, dt. \end{aligned} \tag{3.22}$$

Or

$$\delta I = \int_1^2 \sum_i \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \right] \delta q_i \, dt, \tag{3.23}$$

where the variation δq_i is

$$\delta q_i = \left(\frac{\partial q_i}{\partial \alpha} \right)_{\alpha=0} d\alpha. \tag{3.24}$$

Since the q_i's are independent, the variations δq_i's are also independent (e.g., the functions $\eta_i(t)$ are independent of each other). By a straightforward application of the fundamental lemma (3.10), the condition that δI is zero requires that the coefficients of δq_i vanish separately:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0, \quad i = 1,2,\ldots,n. \tag{3.25}$$

Equations (3.25) represent the appropriate generalization of (3.11) for multiple variables and are known as the *Euler-Lagrange equations*.

3.4 Advantages of Variational Principle formulation

The variational principle formulation is particularly useful when a Lagrangian can be expressed in terms of the independent coordinates of the system under consideration. One of the key advantages of the variational principle is that it involves only physical quantities–specifically, kinetic and potential energies–that can be defined independently of any particular set of generalized coordinates. As a result, the formulation is automatically invariant under changes in the choice of coordinates for the system.

From the variational principle, it is also clear why the Lagrangian is always invariant under the addition of a total time derivative of any function of the coordinates and time. The time integral of such a total derivative between points 1 and 2 depends only on the values of the arbitrary function at the endpoints. Since the variation at the endpoints is zero, adding an arbitrary time derivative to the Lagrangian does not affect the variational behaviour of the integral.

Another advantage is that the Lagrangian formulation can be easily extended to describe systems not typically considered in classical dynamics, such as the elastic field, the electromagnetic field, and the field properties of elementary particles.

Consider a physical system consisting of a battery with voltage V, connected in series with an inductance L and a resistor with resistance R, where the electric charge q is chosen as the dynamical variable. The inductor contributes to the kinetic energy term, as its effect depends on the time rate of change of the charge. The resistor provides a dissipative term, while the potential energy is given by qV. The dynamic terms in the Lagrange equation with dissipation are:

$$T = \frac{1}{2}L\dot{q}^2, \quad \mathscr{F} = R\dot{q}^2, \tag{3.26}$$

and the potential energy is $U = qV$.

The Lagrangian is $\frac{1}{2}L\dot{q}^2 - qV$. The equation of motion is then given by

$$L\ddot{q} + R\dot{q} - V = 0, \qquad \Longrightarrow \qquad V = L\dot{I} + RI, \tag{3.27}$$

where $I = \dot{q}$ is the electric current. A solution for a battery connected to the circuit at time $t = 0$ is

$$I = I_0\left(1 - \mathrm{e}^{-Rt/L}\right), \tag{3.28}$$

where $I_0 = V/R$ is the final steady-state current.

The mechanical analog for this is a sphere of radius a and effective mass m' falling in a viscous fluid with constant density ρ and viscosity η under the force of gravity. The effective mass is the difference between the actual mass and the mass of the displaced fluid, and the direction of motion is along the y-axis. For this system,

$$T = \frac{1}{2}m'\dot{y}^2, \quad \mathscr{F} = 3\pi\eta a\dot{y}^2, \tag{3.29}$$

and the potential energy is $U = m'gy$, where the frictional drag force $F_f = 6\pi\eta a\dot{y}$ is described by Stokes' law. The equation of motion, given by Lagrange's equations, is

$$m'g = m'\ddot{y} + 6\pi\eta a\dot{y}. \tag{3.30}$$

Using $v = \dot{y}$, the solution (assuming the motion starts from rest at $t = 0$) is

$$v = v_0\left(1 - e^{-t/\tau}\right), \tag{3.31}$$

where $\tau = \dfrac{m'}{6\pi\eta a}$ is a measure of the time it takes for the sphere to reach $1/e$ of its terminal velocity, $v_0 = \frac{m'g}{6\pi\eta a}$.

3.5 Conservation theorems

So far, our primary concern has been to obtain the equations of motion. However, it is also important to understand how to solve these equations for a particular problem.

In general, a system with n degrees of freedom will have n second-order differential equations. The solution to each equation will require two integrations, resulting in $2n$ arbitrary constants of integration. In a specific problem, these constants will be determined by the initial conditions, that is, the initial values of the n generalized coordinates q_j and the n generalized velocities $\dot{q}_j$. Sometimes the equations of motion will be solvable in terms of known functions, but this is not always the case. In fact, the majority of problems are not completely integrable.

However, even when complete solutions cannot be obtained, it is often possible to extract a significant amount of information about the physical nature of the system's motion. Indeed, such information may be of greater interest to the physicist than the complete solution for the generalized coordinates as a function of time. It is important to determine how much can be understood about the motion of a given system without requiring a full integration of the problem.

In many problems, a number of first integrals of the equations of motion can be immediately identified. A first integral refers to relations of the form

$$f(q_1, q_2, \ldots, q_n, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n, t) = \text{constant}, \tag{3.32}$$

which are first-order differential equations. These first integrals are important because they provide physical insights into the system. They include the conservation laws that were identified earlier.

3.5.1 Conjugate momentum

Let us consider as an example a system of point masses under the influence of forces derived from potentials that depend only on position. Then, we have:

$$\begin{aligned}\frac{\partial L}{\partial \dot{x}_i} = \frac{\partial T}{\partial \dot{x}_i} - \frac{\partial V}{\partial \dot{x}_i} &= \frac{\partial}{\partial \dot{x}_i}\sum_i \frac{1}{2}m_i\left(\dot{x}_i^2 + \dot{z}_i^2 + \dot{z}_i^2\right)\\ &= m_i\dot{x}_i \equiv p_{ix},\end{aligned}$$

which is the x-component of the linear momentum associated with the i-th particle. This suggests an obvious extension of the concept of momentum. The ***generalized momentum*** associated with the coordinate q_j is defined as:

$$p_j = \frac{\partial L}{\partial \dot{q}_j}. \tag{3.33}$$

p_j is also known as the ***canonical momentum*** or ***conjugate momentum***.

Note: If q_j is not a Cartesian coordinate, then p_j does not necessarily have the dimension of linear momentum. Furthermore, if there is a velocity-dependent potential, even when q_j is a Cartesian coordinate, the associated generalized momentum will not be identical to the usual *mechanical* momentum.

For instance, in the case of a group of charged particles in an electromagnetic field, the Lagrangian is given by

$$L = \sum_i \frac{1}{2} m_i \dot{r}_i^2 - \sum_i q_i \phi(x_i) + \sum_i q_i \vec{A}(x_i) \cdot \dot{\vec{r}}_i, \tag{3.34}$$

where q_i denotes the charge of the i-th particle. The generalized momentum conjugate to $\vec{r}_i$ is

$$p_{i\alpha} = \frac{\partial L}{\partial \dot{r}_{i\alpha}} = m_i \dot{r}_{i\alpha} + q_i A_\alpha(x_i), \tag{3.35}$$

where $A_\alpha(x_i)$ is the α-th component of the vector potential $\vec{A}(x_i)$. Here, the conjugate momentum is equal to the mechanical momentum plus an additional term due to the electromagnetic interaction.

3.5.2 Cyclic (or) Ignorable coordinates

If the Lagrangian of a system does not depend on a given coordinate q_j (although it may depend on the corresponding velocity $\dot{q}_j$), then the coordinate is said to be ***cyclic*** or ***ignorable***.

The Lagrange's equations of motion

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0,$$

reduces, for a cyclic coordinate, to

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = 0,$$

Or

$$\frac{dp_j}{dt} = 0,$$

which means that

$$p_j = \text{constant}. \tag{3.36}$$

Hence, we can state the general conservation theorem: ***the generalized momentum conjugate to a cyclic coordinate is conserved***.

Note that the conditions for the conservation of generalized momenta are more general than the two momentum conservation theorems we discussed at the beginning. For instance, they provide a conservation theorem for cases where the law of action and reaction is not strictly obeyed, such as when electromagnetic forces are present.

Suppose we have a single particle in a field where neither ϕ nor $\vec{A}$ depends on x. Then x does not appear in the Lagrangian L and is, therefore, a cyclic coordinate. As a result, the corresponding canonical momentum p_x must be conserved. This momentum is given by

$$p_x = m\dot{x} + qA_x = \text{constant}. \tag{3.37}$$

In this case, it is not the mechanical linear momentum $m\dot{x}$ that is conserved, but rather the sum of $m\dot{x}$ and qA_x.

Note: It can be shown from classical electrodynamics that under these conditions, i.e., when neither $\vec{A}$ nor ϕ depends on x, qA_x is exactly the x-component of the electromagnetic linear momentum of the field associated with the charge q.

Exercise 3.2 The Lagrangian of a system is given by

$$L(\theta,\phi,\dot{\theta},\dot{\phi}) = \frac{1}{2}ml^2\left(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2\right) - mgl\cos\theta,$$

where m, l and g are constants. Which of the following is conserved?

(a) $\frac{\dot{\phi}}{\sin^2\theta}$ (b) $\dot{\phi}\sin^2\theta$ (c) $\frac{\phi}{\sin^2\theta}$ (d) $\phi\sin^2\theta$

3.5.3 Energy Function and the Conservation of Energy

Another conservation theorem that arises from the Lagrangian formulation is the conservation of total energy for systems where the forces are derivable from potentials that depend only on position.

Consider a general Lagrangian, which is a function of the coordinates q_j and the velocities $\dot{q}_j$, and may also depend explicitly on time. The total time derivative of L is then given by:

$$\frac{dL}{dt} = \sum_j \frac{\partial L}{\partial q_j}\frac{dq_j}{dt} + \sum_j \frac{\partial L}{\partial \dot{q}_j}\frac{d\dot{q}_j}{dt} + \frac{\partial L}{\partial t} \tag{3.38}$$

From Lagrange's equations

$$\frac{\partial L}{\partial q_j} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right),$$

and (3.38) can be rewritten as

$$\frac{dL}{dt} = \sum_j \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)\dot{q}_j + \sum_j \frac{\partial L}{\partial \dot{q}_j}\frac{d\dot{q}_j}{dt} + \frac{\partial L}{\partial t},$$

or

$$\frac{dL}{dt} = \sum_j \frac{d}{dt}\left(\dot{q}_j\frac{\partial L}{\partial \dot{q}_j}\right) + \frac{\partial L}{\partial t}.$$

It therefore follows that

$$\frac{d}{dt}\left(\sum_j \dot{q}_j\frac{\partial L}{\partial \dot{q}_j} - L\right) + \frac{\partial L}{\partial t} = 0. \tag{3.39}$$

The quantity in parenthesis is often called the ***energy function*** and will be denoted by h.

$$h(q_1,q_2,\ldots,q_n,\dot{q}_1,\dot{q}_2,\ldots,\dot{q}_n,t) = \sum_j \dot{q}_j\frac{\partial L}{\partial \dot{q}_j} - L, \tag{3.40}$$

and equation (3.39) can be seen as giving the total time derivative of h

$$\frac{dh}{dt} = -\frac{\partial L}{\partial t}. \tag{3.41}$$

If the Lagrangian is not an explicit function of time, meaning that t does not appear explicitly in L, but only implicitly through the time dependence of q and $\dot{q}$, then h is conserved according to equation (3.41). h is one of the first integrals of motion and is sometimes referred to as ***Jacobi's integral***.

Under certain circumstances, the function h represents the total energy of the system. To determine what these circumstances are, we note that the total kinetic energy of a system is often written as

$$T = T_0 + T_1 + T_2. \tag{3.42}$$

where $T_0 = T_0(q)$ is a function of the generalized coordinates only, $T_1 = T_1(q,\dot{q})$ is linear in the generalized velocities, and $T_2 = T_2(q,\dot{q})$ is a quadratic function of the generalized velocities $\dot{q}$. The Lagrangian can similarly be decomposed according to its functional dependence on the $\dot{q}$ variables:

$$L(q,\dot{q},t) = L_0(q,t) + L_1(q,\dot{q},t) + L_2(q,\dot{q},t). \tag{3.43}$$

Here, L_2 is a homogeneous function of the second degree in $\dot{q}$, while L_1 is homogeneous of the first degree in $\dot{q}$. The Lagrangian typically takes this form when the forces are derivable from a potential that does not involve the velocities. Even in the case of velocity-dependent potentials, we note that the Lagrangian for a charged particle in an electromagnetic field also exhibits this structure.

One may recall that Euler's theorem states that if f is a homogeneous function of degree n in the variables x_i, then

$$\sum_i x_i \frac{\partial f}{\partial x_i} = nf. \tag{3.44}$$

When applied to the function h, as given by equation (3.40), for Lagrangians of the form (3.43), this theorem implies that

$$h = 2L_2 + L_1 - L = L_2 - L_0. \tag{3.45}$$

If the transformation equations defining the generalized coordinates do not explicitly involve time, then $T = T_2$. Further, if the potential does not depend on the generalized velocities, then $L_2 = T$ and $L_0 = -V$, so that

$$h = T + V = E, \tag{3.46}$$

and the energy function is indeed the total energy. Under these conditions, if V does not explicitly depend on time, neither does L. Therefore, by equation (3.41), h (which, in this case, represents the total energy) will be conserved.

3.6 Reference

[1] H. Goldstein, J. L. Safko, and C. P. Poole. *Classical Mechanics, 3e*. Pearson Education, 2011.