I

Classical mathematics and physics

This part is concerned with variational theory prior to modern quantum mechanics. The exception, saved for Chapter 10, is electromagnetic theory as formulated by Maxwell, which was relativistic before Einstein, and remains as fundamental as it was a century ago, the first example of a Lorentz and gauge covariant field theory. Chapter 1 is a brief survey of the history of variational principles, from Greek philosophers and a religious faith in God as the perfect engineer to a set of mathematical principles that could solve practical problems of optimization and rationalize the laws of dynamics. Chapter 2 traces these ideas in classical mechanics, while Chapter 3 discusses selected topics in applied mathematics concerned with optimization and stationary principles.

1

History of variational theory

The principal references for this chapter are:

[5] Akhiezer, N.I. (1962). *The Calculus of Variations* (Blaisdell, New York).
[26] Blanchard, P. and Brüning, E. (1992). *Variational Methods in Mathematical Physics* (Springer-Verlag, Berlin).
[78] Dieudonné, J. (1981). *History of Functional Analysis* (North-Holland, Amsterdam).
[147] Goldstine, H.H. (1980). *A History of the Calculus of Variations from the 17th through the 19th Century* (Springer-Verlag, Berlin).
[210] Lanczos, C. (1966). *Variational Principles of Mechanics* (University of Toronto Press, Toronto).
[322] Pars, L.A. (1962). *An Introduction to the Calculus of Variations* (Wiley, New York).
[436] Yourgrau, W. and Mandelstam, S. (1968). *Variational Principles in Dynamics and Quantum Theory*, 3rd edition (Dover, New York).

The idea that laws of nature should satisfy a principle of simplicity goes back at least to the Greek philosophers [436]. The anthropomorphic concept that the engineering skill of a supreme creator should result in rules of least effort or of most efficient use of resources leads directly to principles characterized by mathematical extrema. For example, Aristotle (*De Caelo*) concluded that planetary orbits must be perfect circles, because geometrical perfection is embodied in these curves: "... of lines that return upon themselves the line which bounds the circle is the shortest. That movement is swiftest which follows the shortest line". Hero of Alexandria (*Catoptrics*) proved perhaps the first scientific minimum principle, showing that the path of a reflected ray of light is shortest if the angles of incidence and reflection are equal.

The superiority of circular planetary orbits became almost a religious dogma in the Christian era, intimately tied to the idea of the perfection of God and of His creations. It was replaced by modern celestial mechanics only after centuries in which the concept of esthetic perfection of the universe was gradually superseded by a concept of esthetic perfection of a mathematical theory that could account for the actual behavior of this universe as measured in astronomical observations. Aspects of value-oriented esthetics lay behind Occam's logical "razor" (avoid unnecessary hypotheses), anticipating the later development of observational science and the search for an explanatory theory that was both as general as possible and as simple as possible. The path from Aristotle to Copernicus, Brahe, Kepler, Galileo, and Newton retraces this shift from *a priori* purity of concepts to mathematical theory solidly based on empirical science. The resulting theory of classical mechanics retains extremal principles that are the basis of the variational theory presented here in Chapter 2.

Variational principles have turned out to be of great practical use in modern theory. They often provide a compact and general statement of theory, invariant or covariant under transformations of coordinates or functions, and can be used to formulate internally consistent computational algorithms. Symmetry properties are often most easily derived in a variational formalism.

1.1 The principle of least time

The law of geometrical optics anticipated by Hero of Alexandria was formulated by Fermat (1601–1655) as a principle of least time, consistent with Snell's law of refraction (1621). The time for phase transmission from point P to point Q along a path $\mathbf{x}(t)$ is given by

$$T = \int_P^Q \frac{ds}{v(s)}, \tag{1.1}$$

where ds is a path element, and v is the phase velocity. Fermat's principle is that the value of the integral T should be stationary with respect to any infinitesimal deviation of the path $\mathbf{x}(t)$ from its physical value. This is valid for geometrical optics as a limiting case of wave optics. The mathematical statement is that $\delta T = 0$ for all variations induced by displacements $\delta\mathbf{x}(t)$. In this and subsequent variational formulas, differentials defined by the notation $\delta \cdots$ are small increments evaluated in the limit that quadratic infinitesimals can be neglected. Thus for sufficiently small displacements $\delta\mathbf{x}(t)$, the integral T varies quadratically about its physical value. For planar reflection consider a ray path from $P : (-d, -h)$ to the observation point $Q : (-d, h)$ via an intermediate point $(0, y)$ in the reflection plane $x = 0$. Elapsed time in a uniform medium is

$$T(y) = \left\{\sqrt{d^2 + (h+y)^2} + \sqrt{d^2 + (h-y)^2}\right\} \Big/ v, \tag{1.2}$$

to be minimized with respect to displacements in the reflection plane parametrized by y. The angle of incidence θ_i is defined such that

$$\sin\theta_i = \frac{h+y}{\sqrt{d^2+(h+y)^2}}$$

and the angle of reflection θ_r is defined by

$$\sin\theta_r = \frac{h-y}{\sqrt{d^2+(h-y)^2}}.$$

The law of planar reflection, $\sin\theta_i = \sin\theta_r$, follows immediately from

$$\frac{\partial T}{\partial y} = (\sin\theta_i - \sin\theta_r)/v = 0.$$

To derive Snell's law of refraction, consider the ray path from point $P : (-d, -h)$ to $Q : (d, h)$ via point $(0, y)$ in a plane that separates media of phase velocity v_i $(x < 0)$ and v_r $(x > 0)$. The elapsed time is

$$T(y) = v_i^{-1}\sqrt{d^2+(h+y)^2} + v_r^{-1}\sqrt{d^2+(h-y)^2}. \tag{1.3}$$

The variational condition is

$$\frac{\partial T}{\partial y} = \sin\theta_i/v_i - \sin\theta_r/v_r = 0.$$

This determines parameter y such that

$$\frac{\sin\theta_i}{\sin\theta_r} = \frac{v_i}{v_r}, \tag{1.4}$$

giving Snell's law for uniform refractive media.

1.2 The variational calculus

Derivation of a ray path for the geometrical optics of an inhomogeneous medium, given $v(\mathbf{r})$ as a function of position, requires a development of mathematics beyond the calculus of Newton and Leibniz. The elapsed time becomes a functional $T[\mathbf{x}(t)]$ of the path $\mathbf{x}(t)$, which is to be determined so that $\delta T = 0$ for variations $\delta\mathbf{x}(t)$ with fixed end-points: $\delta\mathbf{x}_P = \delta\mathbf{x}_Q = 0$. Problems of this kind are considered in the calculus of variations [5, 322], proposed originally by Johann Bernoulli (1696), and extended to a full mathematical theory by Euler (1744). In its simplest form, the concept of the variation $\delta\mathbf{x}(t)$ reduces to consideration of a modified function $\mathbf{x}_\epsilon(t) = \mathbf{x}(t) + \epsilon\mathbf{w}(t)$ in the limit $\epsilon \to 0$. The function $\mathbf{w}(t)$ must satisfy conditions of continuity that are compatible with those of $\mathbf{x}(t)$. Then $\delta\mathbf{x}(t) = \mathbf{w}(t)\,d\epsilon$ and the variation of the derivative function is $\delta\mathbf{x}'(t) = \mathbf{w}'(t)\,d\epsilon$.

The problem posed by Bernoulli is that of the *brachistochrone*. If two points are connected by a wire whose shape is given by an unknown function $y(x)$ in a vertical plane, what shape function minimizes the time of descent of a bead sliding without friction from the higher to the lower point? The mass of a bead moving under gravity is not relevant. It can easily be verified by trial and error that a straight line does not give the minimum time of passage. Always in such problems, conditions appropriate to physically meaningful solution functions must be specified. Although this is a vital issue in any mathematically rigorous variational calculus, such conditions will be stated as simply as possible here, strongly dependent on each particular application of the theory. Clearly the assumed wire in the brachistochrone problem must have the physical properties of a wire. This requires $y(x)$ to be continuous, but does not exclude a vertical drop. Since no physical wire can have an exact discontinuity of slope, it is reasonable to require velocity of motion along the wire to be conserved at any such discontinuity, so that the hypothetical sliding bead does not come to an abrupt stop or bounce with undetermined loss of momentum. It can easily be verified that a vertical drop followed by a horizontal return to the smooth brachistochrone curve always increases the time of passage. Thus such deviations from continuity of the derivative function do not affect the optimal solution.

The calculus of variations [5, 322] is concerned with problems in which a function is determined by a stationary variational principle. In its simplest form, the problem is to find a function $y(x)$ with specified values at end-points x_0, x_1 such that the integral $J = \int_{x_0}^{x_1} f(x, y, y')dx$ is stationary. The variational solution is derived from

$$\delta J = \int \left\{ \delta y \frac{\partial f}{\partial y} + \delta y' \frac{\partial f}{\partial y'} \right\} dx = 0$$

after integrating by parts to eliminate $\delta y'(x)$. Because

$$\int \delta y' \frac{\partial f}{\partial y'} dx = \delta y \left. \frac{\partial f}{\partial y'} \right|_{x_0}^{x_1} - \int \delta y \frac{d}{dx} \frac{\partial f}{\partial y'} dx,$$

$\delta J = 0$ for fixed end-points $\delta y(x_0) = \delta y(x_1) = 0$ if

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} = 0. \tag{1.5}$$

This is a simple example of the general form of Euler's equation (1744), derived directly from a variational expression.

Blanchard and Brüning [26] bring the history of the calculus of variations into the twentieth century, as the source of contemporary developments in pure mathematics. A search for existence and uniqueness theorems for variational problems engendered deep studies of the continuity and compactness of mathematical entities that generalize the simple intuitive definitions assumed by Euler and Lagrange. The seemingly self-evident statement that, for free variations of the function $y(x)$,

$$\int \left(\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} \right) \delta y dx = 0$$

implies Euler's equation, was first proven rigorously by Du Bois-Reymond in 1879. With carefully stated conditions on the functions f and y, this made it possible to prove the fundamental theorem of the variational calculus [26], on the existence of extremal solutions of variational problems.

1.2.1 Elementary examples

A geodesic problem requires derivation of the shortest path connecting two points in some system for which distance is defined, subject to constraints that can be either geometrical or physical in nature. The shortest path between two points in a plane follows from this theory. The problem is to minimize

$$J = \int_{x_0}^{x_1} f(x, y, y')dx = \int_{x_0}^{x_1} dx \sqrt{1 + \left(\frac{dy}{dx}\right)^2},$$

where

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

In this example, Euler's equation takes the form of the geodesic equation

$$\frac{d}{dx} \frac{y'}{\sqrt{1 + y'^2}} = 0.$$

The solution is $y' = const$, or

$$y(x) = y_0 \frac{x_1 - x}{x_1 - x_0} + y_1 \frac{x - x_0}{x_1 - x_0},$$

a straight line through the points x_0, y_0 and x_1, y_1.

In Johann Bernoulli's problem, the brachistochrone, it is required to find the shape of a wire such that a bead slides from point $0, 0$ to x_1, y_1 in the shortest time T under the force of gravity. The energy equation $\frac{1}{2} m v^2 = -mgy$ implies $v = \sqrt{-2gy}$, so that

$$T = \int_0^{x_1} \frac{ds}{v} = \int_0^{x_1} f(y, y')\, dx,$$

where $f(y, y') = \sqrt{-(1 + y'^2)/2gy}$. Because $\partial f/\partial x = 0$, the identity

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

and the Euler equation imply an integral of motion,

$$y'\frac{\partial f}{\partial y'} - f = \frac{-1}{\sqrt{-2gy(1 + y'^2)}} = const.$$

On combining constants into the single parameter a this implies

$$1 + \left(\frac{dy}{dx}\right)^2 = \frac{-2a}{y}.$$

The solution for a bead starting from rest at the coordinate origin is a cycloid, determined by the parametric equations $x = a(\phi - \sin\phi)$ and $y = a(\cos\phi - 1)$. This curve is generated by a point on the perimeter of a circle of radius a that rolls below the x-axis without slipping. The lowest point occurs for $\phi = \pi$, with $x_1 = \pi a$ and $y_1 = -2a$. By adding a constant ϕ_0 to ϕ, a can be adjusted so that the curve passes through given points x_0, y_0 and x_1, y_1.

1.3 The principle of least action

Variational principles for classical mechanics originated in modern times with the principle of least action, formulated first imprecisely by Maupertuis and then as an example of the new calculus of variations by Euler (1744) [436]. Although not stated explicitly by either Maupertuis or Euler, stationary action is valid only for motion in which energy is conserved. With this proviso, in modern notation for generalized coordinates,

$$\delta \int_P^Q \mathbf{p} \cdot d\mathbf{q} = 0, \tag{1.6}$$

for a path from system point P to system point Q.

For a particle of mass m moving in the (x, y) plane with force per mass (X, Y), instantaneous motion is described by velocity v along the trajectory. An instantaneous radius of curvature ρ is defined by angular momentum $\ell = mv\rho$ such that the centrifugal force mv^2/ρ balances the true force normal to the trajectory. Hence, following Euler's derivation, Newtonian mechanics implies that

$$\frac{v^2}{\rho} = \frac{Y dx - X dy}{\sqrt{dx^2 + dy^2}}$$

along the trajectory. The principle of least action requires the action integral per unit mass

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

to be stationary. The variation of v along the trajectory is determined for fixed energy $E = T + V$ by

$$v\,dv = -\frac{1}{m}\left(\frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy\right) = X dx + Y dy.$$

Thus $v\frac{\partial v}{\partial x} = X$ and $v\frac{\partial v}{\partial y} = Y$. Euler's equation then takes the form

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

where $y' = dy/dx$. The local curvature of a trajectory is defined by

$$\frac{1}{\rho} = \frac{d}{dx}\left[y'/(1+y'^2)^{\frac{1}{2}}\right] = y''/(1+y'^2)^{\frac{3}{2}}.$$

Using this formula and $\frac{dv}{dx} = \frac{X+Yy'}{v}$, Euler's equation implies

$$\frac{v}{\rho} + \frac{(X+Yy')y'}{v\sqrt{1+y'^2}} - \frac{Y}{v}\sqrt{1+y'^2} = 0.$$

This reproduces the formula derived directly from Newtonian mechanics:

$$\frac{v^2}{\rho} = \frac{Y - Xy'}{\sqrt{1+y'^2}} = \frac{Y dx - X dy}{\sqrt{dx^2 + dy^2}}.$$

Euler's proof of the least action principle for a single particle (mass point in motion) was extended by Lagrange (*c.* 1760) to the general case of mutually interacting particles, appropriate to celestial mechanics. In Lagrange's derivation [436], action along a system path from initial coordinates P to final coordinates Q is defined by

$$A = \sum_a m_a \int_P^Q v_a\,ds_a = \sum_a m_a \int_P^Q \dot{\mathbf{x}}_a \cdot d\mathbf{x}_a. \tag{1.7}$$

Variations about a true dynamical path are defined by coordinate displacements $\delta\mathbf{x}_a$. Velocity displacements $\delta\dot{\mathbf{x}}_a$ are constrained so as to maintain invariant total energy. This implies modified time values at the displaced points [146]. The energy constraint condition is

$$\delta E = \sum_a \left(m_a\dot{\mathbf{x}}_a \cdot \delta\dot{\mathbf{x}}_a + \frac{\partial V}{\partial \mathbf{x}_a} \cdot \delta\mathbf{x}_a\right) = 0.$$

The induced variation of action is

$$\delta A = \sum_a m_a \int_P^Q (\dot{\mathbf{x}}_a \cdot d\delta\mathbf{x}_a + \delta\dot{\mathbf{x}}_a \cdot d\mathbf{x}_a)$$
$$= \sum_a m_a \dot{\mathbf{x}}_a \cdot \delta\mathbf{x}_a|_P^Q - \sum_a m_a \int_P^Q (d\dot{\mathbf{x}}_a \cdot \delta\mathbf{x}_a - \dot{\mathbf{x}}_a dt \cdot \delta\dot{\mathbf{x}}_a),$$

on integrating by parts and using $d\mathbf{x}_a = \dot{\mathbf{x}}_a dt$. The final term here can be replaced, using the energy constraint condition. Then, using $d\dot{\mathbf{x}}_a = \ddot{\mathbf{x}}_a dt$,

$$\delta A = \sum_a m_a \dot{\mathbf{x}}_a \cdot \delta\mathbf{x}_a|_P^Q - \sum_a \int_P^Q \left(m_a \ddot{\mathbf{x}}_a + \frac{\partial V}{\partial \mathbf{x}_a} \right) \cdot \delta\mathbf{x}_a dt.$$

If the end-points are fixed, the integrated term vanishes, and A is stationary if and only if the final integral vanishes. Since $\delta\mathbf{x}_a$ is arbitrary, the integrand must vanish, which is Newton's law of motion. Hence Lagrange's derivation proves that the principle of least action is equivalent to Newtonian mechanics if energy is conserved and end-point coordinates are specified.

2

Classical mechanics

The principal references for this chapter are:

[26] Blanchard, P. and Brüning, E. (1992). *Variational Methods in Mathematical Physics: A Unified Approach* (Springer-Verlag, Berlin).
[146] Goldstein, H. (1983). *Classical Mechanics*, 2nd edition (Wiley, New York).
[187] Jeffreys, H. and Jeffreys, B.S. (1956). *Methods of Mathematical Physics*, 3rd edition (Cambridge University Press, New York).
[208] Kuperschmidt, B.A. (1990). *The Variational Principles of Dynamics* (World Scientific, New York).
[210] Lanczos, C. (1966). *Variational Principles of Mechanics* (University of Toronto Press, Toronto).
[240] Mercier, A. (1959). *Analytical and Canonical Formalism in Physics* (Interscience, New York).
[323] Pauli, W. (1958). *Theory of Relativity*, tr. G. Field (Pergamon Press, New York).
[393] Synge, J.L. (1956). *Relativity: the Special Theory* (Interscience, New York).
[436] Yourgrau, W. and Mandelstam, S. (1968). *Variational Principles in Dynamics and Quantum Theory*, 3rd edition (Dover, New York).

2.1 Lagrangian formalism

Newton's equations of motion, stated as "*force equals mass times acceleration*", are strictly true only for mass points in Cartesian coordinates. Many problems of classical mechanics, such as the rotation of a solid, cannot easily be described in such terms. Lagrange extended Newtonian mechanics to an essentially complete nonrelativistic theory by introducing generalized coordinates q and generalized forces Q such that the work done in a dynamical process is $\sum_k Q_k dq_k$ [436]. Since this must be the same when expressed in Cartesian coordinates, it follows that $Q_k = \sum_a \mathbf{X}_a \cdot \frac{\partial \mathbf{X}_a}{\partial q_k}$, where the Newtonian force is $\mathbf{X}_a = -\frac{\partial V}{\partial \mathbf{X}_a}$. Equivalently, if the potential function is $V(\{q\})$ in generalized coordinates, then $Q_k = -\frac{\partial V}{\partial q_k}$. The Newtonian kinetic energy $T = \frac{1}{2}\sum_a m_a \dot{\mathbf{x}}_a^2$ defines momenta $\mathbf{p}_a = \frac{\partial T}{\partial \dot{\mathbf{x}}_a} = m_a \dot{\mathbf{x}}_a$, which becomes $p_k = \frac{\partial T}{\partial \dot{q}_k}$ when kinetic energy is expressed as $T(\{q, \dot{q}\})$. The equations of motion $\dot{\mathbf{p}}_a = \mathbf{X}$ transform into $\dot{p}_k = Q_k$. Although this can be shown by direct