Chapter 1

The Classical Calculus of Variations

For the convenience of the reader, we begin with some well-known definitions and facts from the classical calculus of variations. With the exception of Section 1.5, results are given without proofs. For proofs and detailed discussions, we refer the reader to one of the many books on the subject (e.g., Giaquinta and Hildebrandt, 1996; Troutman, 1996; van Brunt, 2004). Here we follow Chachuat, 2007, which gives all the necessary background for our purposes.

1.1 Problem Statement

We are concerned with the problem of finding minima (or maxima) of a functional $\mathcal{J} : \mathcal{D} \to \mathbb{R}$, where $\mathcal{D}$ is a subset of a (normed) linear space $\mathbf{D}$ of real-valued (or real-vector-valued) functions. The formulation of a problem of the calculus of variations requires two steps: the specification of a performance criterion, and the statement of physical constraints that should be satisfied. The performance criterion $\mathcal{J}$, also called cost functional (or objective), must be specified for evaluating quantitatively the performance of the system under study. The typical form of the cost is

$$\mathcal{J}(y) = \int_a^b L(t, y(t), y'(t))\, dt,$$

where $t \in [a, b]$ is the independent variable, usually called time; $y(t) \in \mathbb{R}^N$, $N \geq 1$, is a real vector variable, the functions $y(t)$, $a \leq t \leq b$, are generally called trajectories or curves; $y'(t) \in \mathbb{R}^N$ stands for the derivative of $y(t)$ with respect to time t; and $L : [a, b] \times \mathbb{R}^{2N} \to \mathbb{R}$ is a real-valued function, called the Lagrangian.

Enforcing constraints in the optimization problem reduces the set of candidate functions and leads to the following definition.

Definition 1.1. *A trajectory $y \in \mathbf{D}$ is said to be an admissible trajectory (or admissible function), provided it satisfies all the constraints of the problem along the interval $[a, b]$. The set of admissible trajectories is denoted by $\mathcal{D}$.*

A great variety of boundary conditions is of interest. The simplest one is to enforce both end-points fixed, e.g., $y(a) = y_a$ and $y(b) = y_b$, $y_a, y_b \in \mathbb{R}^N$. Alternatively, we may require that the trajectory $y \in \mathbf{D}$ joins a fixed point (a, y_a) to a specified curve $f(t)$, $a \leq t \leq T$. In this case, not only the optimal trajectory y shall be determined, but also the optimal value of b. Besides boundary constraints, another type of constraints is often required,

$$\mathcal{G}^j(y) = \int_a^b G^j(t, y(t), y'(t))dt = l_j, \quad j = 1, \ldots, r, \quad r \geq 1,$$

where $G^j : [a, b] \times \mathbb{R}^{2N} \to \mathbb{R}$, $j = 1, \ldots, r$. These constraints are often referred to as isoperimetric constraints. Similar constraints with $\leq$ sign can be considered. More generally, constraints of the form

$$G^j(t, y(t), y'(t))dt = 0, \quad j = 1, \ldots, r, \quad r \geq 1,$$

are called constraints of Lagrange form.

Having defined an objective functional $\mathcal{J}$ and constraints, one must then decide about the class of functions with respect to which the optimization shall be performed. The traditional choice in the calculus of variations is to consider the class of continuously differentiable functions, e.g., $C^1([a, b])$. We endow $C^1([a, b])$ with a norm. The most natural choice for a norm on $C^1([a, b])$ is

$$\|y\|_{1,\infty} := \max_{a \leq t \leq b} \|y(t)\| + \max_{a \leq t \leq b} \|y'(t)\|,$$

where $\| \cdot \|$ stands for the Euclidean norm in $\mathbb{R}^N$. The class of functions $C^1([a, b])$ endowed with $\| \cdot \|_{1,\infty}$ is a Banach space.

We now define what is meant by a minimum of $\mathcal{J}$ on $\mathcal{D}$.

Definition 1.2. *A trajectory $\bar{y} \in \mathcal{D}$ is said to be a local minimizer (resp. local maximizer) for $\mathcal{J}$ on $\mathcal{D}$, if there exists $\delta > 0$ such that $\mathcal{J}(\bar{y}) \leq \mathcal{J}(y)$ (resp. $\mathcal{J}(\bar{y}) \geq \mathcal{J}(y)$) for all $y \in \mathcal{D}$ with $\|y - \bar{y}\|_{1,\infty} < \delta$.*

The concept of variation of a functional is central to the solution of problems of the calculus of variations.

Definition 1.3. *The first variation of* $\mathcal{J}$ *at* $y \in \mathbf{D}$ *in the direction* $y \in \mathbf{D}$ *is defined as*

$$\delta \mathcal{J}(y; h) := \lim_{\varepsilon \to 0} \frac{\mathcal{J}(y + \varepsilon h) - \mathcal{J}(y)}{\varepsilon} = \frac{\partial}{\partial \varepsilon} \mathcal{J}(y + \varepsilon h) \bigg|_{\varepsilon = 0},$$

provided the limit exists.

Definition 1.4. *A direction* $h \in \mathbf{D}$, $h \neq 0$, *is said to be an admissible variation for* $\mathcal{J}$ *at* $y \in \mathcal{D}$ *if*

(i) $\delta \mathcal{J}(y; h)$ *exists; and*
(ii) $y + \varepsilon h \in \mathcal{D}$ *for all sufficiently small* ε.

The following well-known result offers a necessary optimality condition for the problems of the calculus of variations, based on the concept of variation.

Theorem 1.1. *Let* $\mathcal{J}$ *be a functional defined on* $\mathcal{D}$. *Suppose that* y *is a local minimizer (or local maximizer) for* $\mathcal{J}$ *on* $\mathcal{D}$. *Then,* $\delta \mathcal{J}(y; h) = 0$ *for each admissible variation* h *at* y.

1.2 The Euler–Lagrange Equations

In this section, we present a first-order necessary optimality condition for a problem which is known as the elementary (or basic or fundamental) problem of the calculus of variations.

The next lemma is an essential result upon which the calculus of variations depends. It is called the fundamental lemma of the calculus of variations, sometimes also called the DuBois–Reymond lemma.

Lemma 1.1 (The fundamental lemma of the calculus of variations). *If* $g(t)$ *is a continuous function of* t *for* $a \leq t \leq b$, *and if*

$$\int_a^b g(t) h(t)\, dt = 0$$

for all functions $h(t)$ *that are continuous for* $a \leq t \leq b$ *and are zero at* $t = a$ *and* $t = b$, *then* $g(t) = 0$ *for all* $a \leq t \leq b$.

We denote by $\partial_i K$, $i = 1, \ldots, M$ $(M \in \mathbb{N})$, the partial derivative of a function $K : \mathbb{R}^M \to \mathbb{R}$ with respect to its ith argument. The following theorem provides a necessary optimality condition for the elementary problem of the calculus of variations.

Theorem 1.2 (The Euler–Lagrange equations). *Consider the problem of minimizing (or maximizing) the functional*

$$\mathcal{J}(y) = \int_a^b L(t, y(t), y'(t))\, dt$$

on $\mathcal{D} = \{y \in \mathbf{D} : y(a) = y_a,\ y(b) = y_b\}$*, where* $L : [a,b] \times \mathbb{R}^{2N} \to \mathbb{R}$ *is a twice continuously differentiable function. Suppose that* y *gives a (local) minimum (or maximum) to* $\mathcal{J}$ *on* $\mathcal{D}$*. Then,*

$$\partial_i L(t, y(t), y'(t)) = \frac{d}{dt}\partial_{N+i} L(t, y(t), y'(t)), \quad i = 2, \ldots N+1, \tag{1.1}$$

for all $t \in [a,b]$.

Definition 1.5. *A function* y *that satisfies the system of Euler–Lagrange equations* (1.1) *on* $[a,b]$ *is called an extremal for the functional* $\mathcal{J}$.

If one of the boundary conditions $y(a) = y_a$ or $y(b) = y_b$ is not present in the problem (it is possible that all of them are not present), then in order to find the extremizers we must add another necessary condition, usually called the natural boundary condition (or transversality condition).

Theorem 1.3 (Natural boundary conditions). *If* y *is a local minimizer (or maximizer) to the functional*

$$\mathcal{J}(y) = \int_a^b L(t, y(t), y'(t))\, dt,$$

then y *satisfies the Euler–Lagrange equations* (1.1)*. Moreover,*
(i) if $y(a) = y_a$ *is free, then the natural boundary conditions*

$$\partial_{N+i} L(a, y(a), y'(a)) = 0, \quad i = 2, \ldots N+1, \tag{1.2}$$

hold;
(ii) if $y(b)$ *is free, then the natural boundary conditions*

$$\partial_{N+i} L(b, y(b), y'(b)) = 0, \quad i = 2, \ldots N+1, \tag{1.3}$$

hold.

1.3 Problems with Isoperimetric Constraints

An isoperimetric problem of the calculus of variations is a problem wherein one or more constraints involve the integral of a given function over part or all of the integration horizon $[a, b]$. Such isoperimetric constraints arise frequently in geometry problems, such as the determination of the curve (resp. surface) enclosing the largest surface (resp. volume) subject to a fixed perimeter (resp. area). The following theorems provide a characterization of the extremals for isoperimetric problems, based on the method of Lagrange multipliers.

Theorem 1.4. *Consider the problem of minimizing (or maximizing) the functional*

$$\mathcal{J}(y) = \int_a^b L(t, y(t), y'(t))\, dt$$

on $\mathcal{D}$ given by those $y \in \mathbf{D}$ such that $y(a) = y_a$, $y(b) = y_b$, and

$$\mathcal{G}(y) = \int_a^b G(t, y(t), y'(t))dt = l, \tag{1.4}$$

where $L, G : [a, b]\times\mathbb{R}^{2N} \to \mathbb{R}$ are twice continuously differentiable functions. Suppose that y gives a (local) minimum (or maximum) to this problem. Assume that $\delta\mathcal{G}(y; h)$ does not vanish for all $h \in \mathbf{D}$. Then there exists a constant $\lambda \in \mathbb{R}$ such that y is a solution of the Euler–Lagrange equations

$$\partial_i F(t, y(t), y'(t), \lambda) = \frac{d}{dt}\partial_{N+i} F(t, y(t), y'(t), \lambda), \quad i = 2, \dots N + 1,$$

where $F(t, y, y', \lambda) = L(t, y, y') - \lambda G(t, y, y')$.

Remark 1.1. The equality (1.4) is called an isoperimetric constraint. Observe that $\delta\mathcal{G}(y; h)$ does not vanish for all $h \in \mathbf{D}$ if, and only if, y is not an extremal for $\mathcal{G}$.

Theorem 1.4 can be generalized to r conditions of integral type (to r isoperimetric constraints).

Theorem 1.5. *Consider the problem of minimizing (or maximizing) the functional*

$$\mathcal{J}(y) = \int_a^b L(t, y(t), y'(t))\, dt$$

on $\mathcal{D}$ given by those $y \in \mathbf{D}$ such that $y(a) = y_a$, $y(b) = y_b$, and

$$\mathcal{G}^j(y) = \int_a^b G^j(t, y(t), y'(t))dt = l_j, \quad j = 1, \ldots, r,$$

where $L, G^j : [a, b] \times \mathbb{R}^{2N} \to \mathbb{R}$, $j = 1, \ldots, r$, are twice continuously differentiable functions. Suppose that y gives a (local) minimum (or maximum) to this problem. Assume that there are functions $h^1, \ldots, h^r \in \mathbf{D}$ such that the matrix

$$A = (a_{kl}), \quad a_{kl} := \delta\mathcal{G}^k(y; h^l), \tag{1.5}$$

has maximal rank r. Then there exist constants $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ such that y is a solution of the Euler–Lagrange equations

$$\partial_i F(t, y(t), y'(t), \lambda) = \frac{d}{dt}\partial_{N+i} F(t, y(t), y'(t), \lambda), \quad i = 2, \ldots N+1,$$

where
$F(t, y, y', \lambda) = F(t, y, y', \lambda_1, \ldots, \lambda_r) = L(t, y, y') - \sum_{j=1}^r \lambda_j G^j(t, y, y')$.

1.4 Sufficient Optimality Conditions via Joint Convexity

In this section we present a sufficient condition for an extremal to be a global extremizer (minimizer or maximizer).

Definition 1.6. *Given a function $f \in C^1([a, b] \times \mathbb{R}^{2N}; \mathbb{R})$, we say that $f(\underline{x}, y, v)$ is jointly convex (resp. jointly concave) in (y, v), if*

$$f(x, y + y^0, v + v^0) - f(x, y, v) \geq (\leq) \sum_{i=2}^{N+1} \partial_i f(x, y, v) y^0_{i-1} + \sum_{i=2}^{N+1} \partial_{N+i} f(x, y, v) v^0_{i-1}$$

for all (x, y, v),$(x, y + y^0, v + v^0) \in [a, b] \times \mathbb{R}^{2N}$.

Theorem 1.6. *Let $L(\underline{x}, y, v)$ be jointly convex (resp. jointly concave) in (y, v). If y satisfies the system of N Euler–Lagrange equations* (1.1) *with $y(a) = y_a$ and $y(b) = y_b$, then y is a global minimizer (resp. global maximizer) to*

$$\mathcal{J}(y) = \int_a^b L(t, y(t), y'(t))\, dt$$

on $\mathcal{D} = \{y \in \mathbf{D} : y(a) = y_a, y(b) = y_b\}$.

1.5 Noether's Theorem

We now review one of the most beautiful results of the calculus of variations: the classical theorem of Emmy Noether. This result explains all the conservation laws in mechanics (e.g., conservation of momentum or conservation of energy). There exist several ways to prove this result. In this section we recall one of those proofs. The proof is done in two steps: we begin by proving Noether's theorem without transformation of the time (without transformation of the independent variable); then, using a technique of time-reparameterization, we obtain Noether's theorem in its general form. This technique is not so popular while proving Noether's theorem (Torres, 2004c), but it turns out to be, as we shall see, very useful in the context of the fractional calculus of variations.

We begin by formulating the fundamental problem of the calculus of variations, using now the usual notation of physics:

$$I[q(\cdot)] = \int_a^b L\left(t, q(t), \dot{q}(t)\right) dt \longrightarrow \min \tag{1.6}$$

subject to the boundary conditions $q(a) = q_a$ and $q(b) = q_b$, and where $\dot{q} = \frac{dq}{dt}$. The Lagrangian $L : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is assumed here to be a C^2-function with respect to all its arguments.

Definition 1.7 (Invariance without transforming the time). *Functional* (1.6) *is said to be invariant under an ε-parameter group of infinitesimal transformations*

$$\bar{q}(t) = q(t) + \varepsilon\xi(t, q) + o(\varepsilon) \tag{1.7}$$

if

$$\int_{t_a}^{t_b} L\left(t, q(t), \dot{q}(t)\right) dt = \int_{t_a}^{t_b} L\left(t, \bar{q}(t), \dot{\bar{q}}(t)\right) dt \tag{1.8}$$

for any subinterval $[t_a, t_b] \subseteq [a, b]$.

We denote by $\partial_i L$ the partial derivative of L with respect to its ith argument, $i = 1, 2, 3$.

Theorem 1.7 (Necessary condition of invariance). *If* (1.6) *is invariant under transformations* (1.7), *then*

$$\partial_2 L\left(t, q, \dot{q}\right) \cdot \xi + \partial_3 L\left(t, q, \dot{q}\right) \cdot \dot{\xi} = 0\,. \tag{1.9}$$

Proof. Having in mind that condition (1.8) is valid for any subinterval $[t_a, t_b] \subseteq [a, b]$, we can get rid off the integral signs in (1.8): equation (1.8) is equivalent to

$$L(t, q, \dot{q}) = L(t, q + \varepsilon\xi + o(\varepsilon), \dot{q} + \varepsilon\dot{\xi} + o(\varepsilon)) . \tag{1.10}$$

Differentiating both sides of equation (1.10) with respect to ε, then substituting $\varepsilon = 0$, we obtain equality (1.9). □

Definition 1.8 (Conserved quantity). *A quantity $C(t, q(t), \dot{q}(t))$ is said to be conserved if $\frac{d}{dt}C(t, q(t), \dot{q}(t)) = 0$ along all the solutions of the Euler–Lagrange equations*

$$\frac{d}{dt}\partial_3 L(t, q, \dot{q}) = \partial_2 L(t, q, \dot{q}) . \tag{1.11}$$

Theorem 1.8 (Noether's theorem without transforming time). *If functional* (1.6) *is invariant under the one-parameter group of transformations* (1.7)*, then*

$$C(t, q, \dot{q}) = \partial_3 L(t, q, \dot{q}) \cdot \xi(t, q) \tag{1.12}$$

is conserved.

Proof. Using the Euler–Lagrange equations (1.11) and the necessary condition of invariance (1.9), we obtain:

$$\begin{aligned}
&\frac{d}{dt}\left(\partial_3 L(t, q, \dot{q}) \cdot \xi(t, q)\right) \\
&\quad = \frac{d}{dt}\partial_3 L(t, q, \dot{q}) \cdot \xi(t, q) + \partial_3 L(t, q, \dot{q}) \cdot \dot{\xi}(t, q) \\
&\quad = \partial_2 L(t, q, \dot{q}) \cdot \xi(t, q) + \partial_3 L(t, q, \dot{q}) \cdot \dot{\xi}(t, q) \\
&\quad = 0 .
\end{aligned}$$

□

Remark 1.2. In classical mechanics, $\partial_3 L(t, q, \dot{q})$ is interpreted as the generalized momentum.

Definition 1.9 (Invariance of (1.6)**).** *Functional* (1.6) *is said to be invariant under the one-parameter group of infinitesimal transformations*

$$\begin{cases} \bar{t} = t + \varepsilon\tau(t, q) + o(\varepsilon) , \\ \bar{q}(t) = q(t) + \varepsilon\xi(t, q) + o(\varepsilon) , \end{cases} \tag{1.13}$$

if

$$\int_{t_a}^{t_b} L(t, q(t), \dot{q}(t))\, dt = \int_{\bar{t}(t_a)}^{\bar{t}(t_b)} L(\bar{t}, \bar{q}(\bar{t}), \dot{\bar{q}}(\bar{t}))\, d\bar{t}$$

for any subinterval $[t_a, t_b] \subseteq [a, b]$.

Theorem 1.9 (Noether's theorem). *If functional* (1.6) *is invariant, in the sense of Definition 1.9, then*

$$C(t,q,\dot{q}) = \partial_3 L\left(t,q,\dot{q}\right)\cdot\xi(t,q) + \left(L(t,q,\dot{q}) - \partial_3 L\left(t,q,\dot{q}\right)\cdot\dot{q}\right)\tau(t,q) \quad (1.14)$$

is conserved.

Proof. Every non-autonomous problem (1.6) is equivalent to an autonomous one, considering t as a dependent variable. For that we consider a Lipschitzian one-to-one transformation $[a,b] \ni t \longmapsto \sigma \in [\sigma_a, \sigma_b]$ such that

$$\begin{aligned} I\left[q(\cdot)\right] &= \int_a^b L\left(t,q(t),\dot{q}(t)\right)dt = \int_{\sigma_a}^{\sigma_b} L\left(t(\sigma), q(t(\sigma)), \frac{\frac{dq(t(\sigma))}{d\sigma}}{\frac{dt(\sigma)}{d\sigma}}\right)\frac{dt(\sigma)}{d\sigma}d\sigma \\ &= \int_{\sigma_a}^{\sigma_b} L\left(t(\sigma), q(t(\sigma)), \frac{q_\sigma^{'}}{t_\sigma^{'}}\right)t_\sigma^{'} d\sigma \doteq \int_{\sigma_a}^{\sigma_b} \bar{L}\left(t(\sigma), q(t(\sigma)), t_\sigma^{'}, q_\sigma^{'}\right)d\sigma \\ &\doteq \bar{I}\left[t(\cdot), q(t(\cdot))\right] , \end{aligned}$$

where $t(\sigma_a) = a$, $t(\sigma_b) = b$, $t_\sigma^{'} = \frac{dt(\sigma)}{d\sigma}$, and $q_\sigma^{'} = \frac{dq(t(\sigma))}{d\sigma}$. If functional $I[q(\cdot)]$ is invariant in the sense of Definition 1.9, then functional $\bar{I}[t(\cdot), q(t(\cdot))]$ is invariant in the sense of Definition 1.7. Applying Theorem 1.8, we obtain that

$$C\left(t,q,t_\sigma^{'},q_\sigma^{'}\right) = \partial_4\bar{L}\cdot\xi + \partial_3\bar{L}\tau \quad (1.15)$$

is a conserved quantity. Since

$$\begin{aligned} \partial_4\bar{L} &= \partial_3 L\left(t,q,\dot{q}\right) , \\ \partial_3\bar{L} &= -\partial_3 L\left(t,q,\dot{q}\right)\cdot\frac{q_\sigma^{'}}{t_\sigma^{'}} + L(t,q,\dot{q}) \\ &= L(t,q,\dot{q}) - \partial_3 L\left(t,q,\dot{q}\right)\cdot\dot{q} \, , \end{aligned} \quad (1.16)$$

substituting (1.16) into (1.15), we arrive at (1.14). □