3 One Dimensional Calculus of Variations

In this section we only state results of the Calculus of Variations. For derivations, see the book by Gelfand and Fomin[1], for example.

The standard problem in Calculus of Variations is to find functions $\boldsymbol{y}(\boldsymbol{s}) : [\mathbf{0}, \mathbf{1}] \to \mathbb{R}^{\boldsymbol{m}}$ which minimize the functional

$$\boldsymbol{I}[\boldsymbol{y}] = \int_{\mathbf{0}}^{\mathbf{1}} \boldsymbol{F}(\boldsymbol{s}, \boldsymbol{y}(\boldsymbol{s}), \boldsymbol{y}'(\boldsymbol{s}))\, \boldsymbol{ds}, \tag{3.1}$$

where $\boldsymbol{F}(\boldsymbol{s}, \boldsymbol{y}, \boldsymbol{p}) : \mathbb{R} \times \mathbb{R}^{\boldsymbol{m}} \times \mathbb{R}^{\boldsymbol{m}} \to \mathbb{R}$ is a given smooth function. Various conditions can be imposed on $\boldsymbol{y}$.

3.1 First and Second Variations

To find necessary conditions for $\boldsymbol{y}$ to be a minimizer, the idea is to consider the Taylor expansion

$$\boldsymbol{I}[\boldsymbol{y} + \boldsymbol{\epsilon h}] = \boldsymbol{I}[\boldsymbol{y}] + \boldsymbol{\epsilon}\, \boldsymbol{\delta I}[\boldsymbol{y}, \boldsymbol{h}] + \boldsymbol{\epsilon}^{\mathbf{2}}\, \boldsymbol{\delta}^{\mathbf{2}}\boldsymbol{I}[\boldsymbol{y}, \boldsymbol{h}] + \boldsymbol{O}(\boldsymbol{\epsilon}^{\mathbf{3}}), \tag{3.2}$$

where $\boldsymbol{\epsilon}$ is a small parameter and the function $\boldsymbol{h}(\boldsymbol{s}) : [\mathbf{0}, \mathbf{1}] \to \mathbb{R}^{\boldsymbol{m}}$ is an admissible (i. e. allowed by the imposed conditions on $\boldsymbol{y}$) variation. The functionals $\boldsymbol{\delta I}[\boldsymbol{y}, \boldsymbol{h}]$ and $\boldsymbol{\delta}^{\mathbf{2}}\boldsymbol{I}[\boldsymbol{y}, \boldsymbol{h}]$ are called the first and the second variations of $\boldsymbol{I}$. The first order necessary condition is

$$\boldsymbol{\delta I}[\boldsymbol{y}, \boldsymbol{h}] = \mathbf{0}, \quad \forall\, \boldsymbol{h} \text{ admissible.} \tag{3.3}$$

The second order necessary condition is

$$\boldsymbol{\delta}^{\mathbf{2}}\boldsymbol{I}[\boldsymbol{y}, \boldsymbol{h}] \geqslant \mathbf{0}, \quad \forall\, \boldsymbol{h} \text{ admissible.} \tag{3.4}$$

3.2 Euler-Lagrange Equations

In the case of imposed boundary conditions of the form

$$\boldsymbol{y}(\mathbf{0}) = \boldsymbol{\alpha}, \quad \boldsymbol{y}(\mathbf{1}) = \boldsymbol{\beta}, \tag{3.5}$$

$\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^{\boldsymbol{m}}$ given, the first order condition implies

$$-\frac{\boldsymbol{d}}{\boldsymbol{ds}}\left[\boldsymbol{F}_{\mathsf{p}}(\boldsymbol{s}, \boldsymbol{y}(\boldsymbol{s}), \boldsymbol{y}'(\boldsymbol{s}))\right] + \boldsymbol{F}_{\mathsf{y}}(\boldsymbol{s}, \boldsymbol{y}(\boldsymbol{s}), \boldsymbol{y}'(\boldsymbol{s})) = \mathbf{0}. \tag{3.6}$$

These $\boldsymbol{m}$ second-order ODEs are called the Euler-Lagrange equations.

[1] I. M. Gelfand, S. V. Fomin, 1963, *Calculus of Variations*, Prentice Hall

3.3 Natural Boundary Conditions

Suppose that there is an imposed boundary condition only at one end-point, e. g.

$$y(0) = \alpha. \tag{3.7}$$

Then, in addition to the Euler-Lagrange equations, the the first order condition implies a natural boundary condition at the other end-point:

$$F_{\mathsf{p}}(1, y(1), y'(1)) = 0. \tag{3.8}$$

3.4 Corner Conditions

The Euler-Lagrange equations were derived under the assumption that $y(s)$ and $y'(s)$ were continuous. Assume now that $y(s)$ is continuous but can have "corners", i. e. $y'(s)$ is only piecewise continuous. For example suppose that $y'(s)$ has a discontinuity at $\hat{s} \in (0, 1)$, i. e. $y'(\hat{s}^-) \neq y'(\hat{s}^+)$, where $y'(\hat{s}^-)$ [respectively $y'(\hat{s}^+)$] denotes the limit of $y'(s)$ as s tends to $\hat{s}$ from the left [respectively right]. Then the first order condition implies the corner condition

$$F_{\mathsf{p}}(\hat{s}, y(\hat{s}), y'(\hat{s}^-)) = F_{\mathsf{p}}(\hat{s}, y(\hat{s}), y'(\hat{s}^+)). \tag{3.9}$$

In the case where $F(s, y, p)$ is a convex function of p, i. e.

$$F_{\mathsf{pp}}(s, y, p) > 0, \quad \forall s, y, p, \tag{3.10}$$

the corner condition implies that $y'(\hat{s}^-) = y'(\hat{s}^+)$. So in this case the minimizers cannot have corners.

3.5 Isoperimetric Constraints

Consider the problem to minimize

$$I[y] = \int_0^1 F(s, y(s), y'(s))\, ds, \tag{3.11}$$

subject to the isoperimetric constraints

$$\int_0^1 G^{(i)}(s, y(s), y'(s))\, ds = 0, \quad i = 1, \ldots, k, \tag{3.12}$$

where $G^{(i)}(s, y, p) : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ are given smooth functions.

The first order necessary condition for y to be solution to this problem is that there are constants $\mu_i, \ i = 1, \dots, k,$ called Lagrange multipliers, such that the functional

$$J[y] = \int_0^1 \left(F(s, y(s), y'(s)) - \sum_{i=1}^{k} \mu_i G^{(i)}(s, y(s), y'(s)) \right) ds \quad (3.13)$$

is stationary.

The corresponding Euler-Lagrange equation is

$$-\frac{d}{ds}\left[F_{\mathsf{p}}(s, y(s), y'(s)) - \sum_{i=1}^{k} \mu_i G_{\mathsf{p}}^{(i)}(s, y(s), y'(s))\right] + F_{\mathsf{y}}(s, y(s), y'(s)) - \sum_{i=1}^{k} \mu_i G_{\mathsf{y}}^{(i)}(s, y(s), y'(s)) = 0. \quad (3.14)$$

3.6 Pointwise Constraints

Consider the problem to minimize

$$I[y] = \int_0^1 F(s, y(s), y'(s))\, ds, \quad (3.15)$$

subject to the pointwise constraints

$$g^{(i)}(s, y(s), y'(s)) = 0, \ \ \forall s \in [0, 1], \ \ i = 1, \dots, l, \quad (3.16)$$

where $g^{(i)}(s, y, p) : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ are given smooth functions.

The first order necessary condition for y to be a solution to this problem is that there are multipliers $\gamma_i(s) : [0, 1] \to \mathbb{R}$, $i = 1, \dots, l$, such that

$$-\frac{d}{ds}\left[F_{\mathsf{p}}(s, y(s), y'(s)) - \sum_{i=1}^{l} \gamma_i(s) g_{\mathsf{p}}^{(i)}(s, y(s), y'(s))\right] + F_{\mathsf{y}}(s, y(s), y'(s)) - \sum_{i=1}^{l} \gamma_i(s) g_{\mathsf{y}}^{(i)}(s, y(s), y'(s)) = 0. \quad (3.17)$$

3.7 Conjugate Point Theory

Suppose we have found a solution $\hat{y}$ of the first order necessary conditions (3.3). Suppose moreover that there is an imposed boundary condition at at least one end-point, say $\hat{y}(1) = \beta$. Then the conjugate point theory can be used to test the second order necessary condition (3.4) for $\hat{y}$.

Consider the Jacobi accessory equation

$$-\frac{d}{ds}\left[\hat{F}_{\mathsf{pp}}\, h'(s) + \hat{F}_{\mathsf{yp}}\, h(s)\right] + \hat{F}_{\mathsf{py}}\, h'(s) + \hat{F}_{\mathsf{yy}}\, h(s) = 0, \tag{3.18}$$

where a hat on a functional means that it is evaluated at $(s, \hat{y}(s), \hat{y}'(s))$. This system of m second order linear equations with known coefficients for $h(s)$ corresponds to the Euler-Lagrange equations for $\delta^2 I(\hat{y}, h)$. This system must be solved as an initial value problem with the m (imposed or natural) linearized boundary conditions at $s = 0$. For example, if we have the imposed boundary condition $\hat{y}(0) = \alpha$, then the corresponding linearized boundary condition is $h(0) = 0$. There is a m-parameter family of solutions to this initial value problem. A conjugate point arises at $s = \hat{s}$, $0 < \hat{s} < 1$, when there is a solution $\bar{h} \not\equiv 0$ in this family satisfying $\bar{h}(\hat{s}) = 0$. In this case $\delta^2 I(\hat{y}, h) \not\geqslant 0$ for any admissible h and $\hat{y}$ is not a minimizer. In the case where there are no conjugate points, $\delta^2 I(\hat{y}, h) \geqslant 0$ for any admissible h and $\hat{y}$ is a minimizer.

3.8 Hamiltonian Formulation

Consider the problem to find a minimizer y to the functional

$$I[y] = \int_0^1 F(s, y(s), y'(s))\, ds, \tag{3.19}$$

where $F(s, y, p)$ is a convex function of p, i. e.

$$F_{\mathsf{pp}}(s, y, p) > 0, \quad \forall s, y, p, \tag{3.20}$$

The first order necessary condition is the Euler-Lagrange equation

$$-\frac{d}{ds}\left[F_{\mathsf{p}}(s, y(s), y'(s))\right] + F_{\mathsf{y}}(s, y(s), y'(s)) = 0, \tag{3.21}$$

which is is a system of m second order ODEs. A Hamiltonian system is an equivalent system of $2m$ first order ODEs.

To construct the Hamiltonian system we start by defining the conjugate variable $z(s) : [0, 1] \to \mathbb{R}^m$ as

$$z \equiv F_{\mathsf{p}}(s, y, y'). \tag{3.22}$$

Suppose that this equation can be inverted into the form

$$y' = \phi(s, y, z), \tag{3.23}$$

where $\phi : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$. Then the Euler-Lagrange equations are equivalent to the system of $2m$ first order ODEs

$$\begin{cases} y' = \phi(s, y, z) \\ z' = F_{\mathsf{y}}(s, y, \phi(s, y, z)). \end{cases} \tag{3.24}$$

The Hamiltonian $H(s, y, z) : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is defined as the Legendre transform of the lagrangian F :

$$H(s, y, z) \equiv z \cdot \phi(s, y, z) - F(s, y, \phi(s, y, z)). \tag{3.25}$$

Then

$$\begin{aligned} H_{\mathsf{y}} &= z \cdot \phi_{\mathsf{y}} - F_{\mathsf{y}} - F_{\mathsf{p}} \cdot \phi_{\mathsf{y}} \\ &= z \cdot \phi_{\mathsf{y}} - F_{\mathsf{y}} - z \cdot \phi_{\mathsf{y}} \\ &= -F_{\mathsf{y}}, \end{aligned} \tag{3.26}$$

and

$$\begin{aligned} H_{\mathsf{z}} &= \phi + z \cdot \phi_{\mathsf{z}} - F_{\mathsf{p}} \cdot \phi_{\mathsf{z}} \\ &= \phi + z \cdot \phi_{\mathsf{z}} - z \cdot \phi_{\mathsf{z}} \\ &= \phi. \end{aligned} \tag{3.27}$$

Combining the equations (3.24), (3.26) and (3.27), we obtain the Hamiltonian equations

$$\begin{cases} y' = H_{\mathsf{z}} \\ z' = -H_{\mathsf{y}}. \end{cases} \tag{3.28}$$

In fact the Legendre transform defining the Hamiltonian corresponds to a Fenchel transform.

Definition. The *Fenchel transform* of the function $\psi(x) : \mathbb{R}^m \to \mathbb{R}$ is the function $\psi^*(v) : \mathbb{R}^m \to \mathbb{R}$ defined as

$$\psi^*(v) \equiv \max_{\mathsf{x}}(v \cdot x - \psi(x)). \tag{3.29}$$

Suppose that this maximum exists. Then, at the maximum, the first order condition is

$$v = \psi_{\mathsf{x}}(x). \tag{3.30}$$

Suppose this equation can be inverted as

$$x = \phi(v). \tag{3.31}$$

Then the Fenchel transform can be written as

$$\psi^*(v) = v \cdot \phi(v) - \psi(\phi(v)). \tag{3.32}$$

Comparing (3.25) to (3.32), we see that the Legendre transform of $F(s, y, y')$ is just the Fenchel transform between y' and z, holding s and y fixed.

Lemma. The Fenchel transform of $\bar{\psi}(x) \equiv \psi(x - a)$, where a is a constant shift, is

$$\bar{\psi}^*(v) = \psi^*(v) + v \cdot a. \tag{3.33}$$

Proof.

$$\begin{aligned}
\bar{\psi}^*(v) &= \max_{\mathsf{x}}(v \cdot x - \psi(x - a)) \\
&= \max_{\bar{\mathsf{x}}}(v \cdot (\bar{x} + a) - \psi(\bar{x})) \\
&= \max_{\bar{\mathsf{x}}}(v \cdot \bar{x} - \psi(\bar{x})) + v \cdot a \\
&= \psi^*(v) + v \cdot a.
\end{aligned} \tag{3.34}$$

Lemma. The Fenchel transform of

$$\psi(x) \equiv \tfrac{1}{2} x \cdot A x, \tag{3.35}$$

where A is a $m \times m$ symmetric, positive definite, constant matrix, is

$$\psi^*(v) = \tfrac{1}{2} v \cdot A^{-1} v. \tag{3.36}$$

Proof. By definition

$$\psi^*(v) = \max_{\mathsf{x}} \left(v \cdot x - \tfrac{1}{2} x \cdot A x \right). \tag{3.37}$$

The maximum occurs when $\boldsymbol{v} = \boldsymbol{A}\boldsymbol{x}$, or $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{v}$. Therefore

$$\begin{aligned} \psi^*(\boldsymbol{v}) &= \boldsymbol{v} \cdot \boldsymbol{A}^{-1}\boldsymbol{v} - \tfrac{1}{2}\boldsymbol{A}^{-1}\boldsymbol{v} \cdot \boldsymbol{A}\boldsymbol{A}^{-1}\boldsymbol{v} \\ &= \boldsymbol{v} \cdot \boldsymbol{A}^{-1}\boldsymbol{v} - \tfrac{1}{2}\boldsymbol{A}^{-1}\boldsymbol{v} \cdot \boldsymbol{v} \\ &= \boldsymbol{v} \cdot \boldsymbol{A}^{-1}\boldsymbol{v} - \tfrac{1}{2}\boldsymbol{v} \cdot \boldsymbol{A}^{-1}\boldsymbol{v} \\ &= \tfrac{1}{2}\boldsymbol{v} \cdot \boldsymbol{A}^{-1}\boldsymbol{v}. \end{aligned} \tag{3.38}$$

The last two lemmas imply the following corollary:

Corollary. The Fenchel transform of

$$\psi(\boldsymbol{x}) \equiv \tfrac{1}{2}(\boldsymbol{x} - \boldsymbol{a}) \cdot \boldsymbol{A}(\boldsymbol{x} - \boldsymbol{a}) \tag{3.39}$$

is

$$\psi^*(\boldsymbol{v}) = \tfrac{1}{2}\boldsymbol{v} \cdot \boldsymbol{A}^{-1}\boldsymbol{v} + \boldsymbol{v} \cdot \boldsymbol{a}. \tag{3.40}$$