1. The Lagrange Problem

Let $\Omega \subset \mathbb{R}^m$ be an open domain with standard coordinates

$$x = (x^1, \dots, x^m)^T.$$

To proceed with formulations we split the vector x in two parts:

$$x = \begin{pmatrix} y^1 \\ \vdots \\ y^n \\ z^1 \\ \vdots \\ z^{m-n} \end{pmatrix} = \begin{pmatrix} y \\ z \end{pmatrix}, \quad n < m. \tag{1}$$

Below to denote derivatives we use letters in the subscripts:

$$\frac{dx}{ds} = x_s.$$

Such a notation does not lead to a confusion with number subscripts such as x_1.

Let $F : \Omega \times \mathbb{R}^m \to \mathbb{R}$ stand for a smooth function.

We will study stationary points of a functional

$$\mathcal{F}\big(x(\cdot)\big) = \int_{s_1}^{s_2} F(x(s), x_s(s))ds \tag{2}$$

on the set of smooth functions $x = (y^T, z^T)^T : [s_1, s_2] \to \Omega$ with boundary conditions

$$z(s_1) = z_1, \quad z(s_2) = z_2, \quad y(s_1) = y_1, \tag{3}$$

and constraints

$$a(x, x_s) = 0. \tag{4}$$

Here $a = (a^1, \dots, a^n)^T$ is a vector of functions that are smooth in $\Omega \times \mathbb{R}^m$.

Assume that

$$\det \frac{\partial a}{\partial y_s}(x, x_s) \neq 0, \quad (x, x_s) \in \Omega \times \mathbb{R}^m \tag{5}$$

and equation (4) can equivalently be written as

$$y_s = \Phi(y, z, z_s). \tag{6}$$

Definition 1. *Let a smooth function*

$$\tilde{x} : [s_1, s_2] \to \Omega, \quad \tilde{x}(s) = (\tilde{y}^T, \tilde{z}^T)^T(s)$$

be such that

$$a(\tilde{x}(s), \tilde{x}_s(s)) = 0, \quad \tilde{x}(s_1) = x_1 = (y_1^T, z_1^T)^T, \quad \tilde{z}(s_2) = z_2.$$

We shall say that $\tilde{x}$ is a stationary point of functional (2) with constraints (4) and boundary conditions (3) if the following holds.

For any smooth function

$$X : [s_1, s_2] \times (-\varepsilon_0, \varepsilon_0) \to \mathbb{R}^m, \quad X(s, \varepsilon) = (Y^T, Z^T)^T(s, \varepsilon), \quad \varepsilon_0 > 0$$

such that

1) $X\big([s_1, s_2] \times (-\varepsilon_0, \varepsilon_0)\big) \subset \Omega$*;*

2) $X(s, 0) = \tilde{x}(s), \quad s \in [s_1, s_2]$*;*

3) $X(s_1, \varepsilon) = x_1, \quad Z(s_2, \varepsilon) = z_2, \quad \varepsilon \in (-\varepsilon_0, \varepsilon_0)$ *;*

4) $a(X(s, \varepsilon), X_s(s, \varepsilon)) = 0, \quad (s, \varepsilon) \in [s_1, s_2] \times (-\varepsilon_0, \varepsilon_0)$

we have

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \mathcal{F}\big(X(\cdot, \varepsilon)\big) = 0.$$

The functions X with properties 1)-4) are referred to as variations.

Theorem 1 ([1])**.** *If the function $\tilde{x}$ is a stationary point of functional (2) with constraints (4) and boundary conditions (3) then there is a smooth function $\lambda(s) = (\lambda_1, \dots, \lambda_n)(s)$ such that $\tilde{x}$ satisfies the equations*

$$\frac{d}{ds}\frac{\partial F^*}{\partial x_s} - \frac{\partial F^*}{\partial x} = 0, \quad F^*(s, x, x_s) = F(x, x_s) + \lambda(s) a(x, x_s), \tag{7}$$

and

$$\frac{\partial F}{\partial y_s}(\tilde{x}(s_2), \tilde{x}_s(s_2)) + \lambda(s_2)\frac{\partial a}{\partial y_s}(\tilde{x}(s_2), \tilde{x}_s(s_2)) = 0. \tag{8}$$

This theorem remains valid if the functions a, F depend on s.

1.1. **The Linear Constraints and Some Geometry.** Consider a case when

$$a(x, x_s) = B(x)x_s,$$

where

$$B(x) = \begin{pmatrix} b_1^1(x) & b_2^1(x) & \cdots & b_m^1(x) \\ b_1^2(x) & b_2^2(x) & \cdots & b_m^2(x) \\ \vdots & \vdots & \ddots & \vdots \\ b_1^n(x) & b_2^n(x) & \cdots & b_m^n(x) \end{pmatrix}$$

is a matrix such that

$$\operatorname{rang} B(x) = n < m, \quad \forall x \in \Omega.$$

Constraints (4) take the form

$$B(x)x_s = 0. \tag{9}$$

Equations (5), (6) imply that

$$B(x)x_s = A(x)y_s + C(x)z_s, \quad \det A \neq 0.$$

We can consider the domain Ω as a coordinate patch in some smooth manifold M. Then equation (9) defines a differential system in M

$$x_s(s) \in \mathcal{T}_{x(s)}, \quad \mathcal{T}_x = \ker B(x) \subset T_x M, \quad \dim \mathcal{T}_x = m - n.$$

The boundary conditions (3) imply that there is an $n-$dimensional submanifold

$$\Sigma \subset M, \quad \Sigma = \{z = z_2\}$$

such that

$$T_x M = T_x \Sigma \oplus \mathcal{T}_x, \quad x \in \Sigma$$

and $x(s_2) \in \Sigma$.

Condition (8) takes the form:

$$\frac{\partial F^*}{\partial x_s}(s_2, \tilde{x}(s_2), \tilde{x}_s(s_2))v = 0, \quad \forall v \in T_{\tilde{x}(s_2)}\Sigma.$$

Summing up we look for a stationary point of the functional $\mathcal{F}$ in the following class of functions

$$x(s_1) = x_1, \quad x(s_2) \in \Sigma, \quad x_s(s) \in \mathcal{T}_{x(s)}.$$

Note also that in general case equations (7) contain the derivatives λ_s and do not match ones from classical mechanics with ideal constraints:

$$\frac{d}{ds}\frac{\partial F}{\partial x_s} - \frac{\partial F}{\partial x} = \lambda B.$$

1.2. **Proof of Theorem 1.** Introduce a notation

$$[F]_y = -\frac{d}{ds}\frac{\partial F}{\partial y_s} + \frac{\partial F}{\partial y}, \quad [F]_z = -\frac{d}{ds}\frac{\partial F}{\partial z_s} + \frac{\partial F}{\partial z}$$

and correspondingly $[F]_x = ([F]_y, [F]_z)$.

Let us put $Z(s, \varepsilon) = \tilde{z}(s) + \varepsilon \delta z(s)$,

$$\operatorname{supp} \delta z \subset [s_1, s_2]. \tag{10}$$

Then the function Y is uniquely determined from the following Cauchy problem

$$Y_s(s, \varepsilon) = \Phi(Y(s, \varepsilon), Z(s, \varepsilon), Z_s(s, \varepsilon)), \quad Y(s_1, \varepsilon) = y_1. \tag{11}$$

Remark 1. *That is why we can not impose condition $x(s_2) = x_2$. The value $Y(s_2, \varepsilon)$ has already been uniquely defined by other boundary conditions and the constraints. In other words if we add the condition $Y(s_2, \varepsilon) = y_2$ then the set of variations $\{X(s, \varepsilon)\}$ may turn up to be too thin to justify the Lagrange multipliers method.*

Cauchy problem (11) has the suitable solution at least for $|\varepsilon|$ and $s_2 - s_1$ small. Observe also that

$$Y_\varepsilon(s_1, \varepsilon) = 0. \tag{12}$$

Using the standard integration by parts technique and from formulas (12), (10) we obtain

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \mathcal{F}\big(X(\cdot,\varepsilon)\big) = \int_{s_1}^{s_2} \Big([F]_z \delta z + [F]_y Y_\varepsilon\Big) ds + \frac{\partial F}{\partial y_s}(\tilde{x}(s_2), \tilde{x}_s(s_2)) Y_\varepsilon(s_2, 0) = 0. \tag{13}$$

The function $\lambda(s)$ is still undefined but due to condition (5) the value $\lambda(s_2)$ is determined uniquely from (8).

From condition 4) of definition 1 it follows that

$$A(\varepsilon) = \int_{s_1}^{s_2} \lambda(s) a(X(s,\varepsilon), X_s(s,\varepsilon)) ds = 0.$$

By the same argument as above we have

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} A = \int_{s_1}^{s_2} \Big([\lambda a]_z \delta z + [\lambda a]_y Y_\varepsilon\Big) ds + \lambda(s_2) \frac{\partial a}{\partial y_s}(\tilde{x}(s_2), \tilde{x}_s(s_2)) Y_\varepsilon(s_2, 0) = 0. \tag{14}$$

Summing formulas (14) and (13) we yield

$$\int_{s_1}^{s_2} \Big([F^*]_z \delta z + [F^*]_y Y_\varepsilon\Big) ds = 0. \tag{15}$$

To construct the function λ consider an equation

$$[F^*]_y = 0. \tag{16}$$

This is a system of linear ordinary differential equations for λ. Due to assumption (5) this system can be presented in the normal form that is

$$\lambda_s = \Lambda(s, \lambda).$$

Since we know $\lambda(s_2)$, by the existence and uniqueness theorem we obtain $\lambda(s)$ as a solution to the IVP for (16).

Equation (15) takes the form

$$\int_{s_1}^{s_2} [F^*]_z \delta z ds = 0.$$

Since δz is an arbitrary function we get $[F^*]_z = 0$. Together with (16) this proves the theorem.

2. The Energy Integral

In this section assume the second argument of the function a to be defined on a conic domain $K \subset \mathbb{R}^m$. All the formulated above results and the argument of section 1.2 remain valid under such an assumption.

Recall that by definition the domain K is a conic domain iff

$$x \in K \Longrightarrow \alpha x \in K, \quad \forall \alpha > 0.$$

Proposition 1. *Assume that a is a homogeneous function of degree r in the second argument:*

$$a(x, \alpha x_s) = \alpha^r a(a, x_s), \quad \forall \alpha > 0, \quad \forall (x, x_s) \in \Omega \times K. \tag{17}$$

Then the stationary point $\tilde{x}$ preserves the "energy":

$$H(x, x_s) = \frac{\partial F}{\partial x_s} x_s - F$$

that is $H(\tilde{x}(s), \tilde{x}_s(s)) = \text{const}$.

Proof of Proposition 1. Consider a function $X(s, \varepsilon) = \tilde{x}(s + \varepsilon\varphi(s))$ with a smooth function φ such that $\operatorname{supp}\varphi \subset [s_1 + s', s_2 - s']$ and

$$|\varepsilon|, s' > 0$$

are small enough.

The function X satisfies all the conditions of Definition 1. To check this use (17).

Furthermore we have

$$X = \tilde{x}(s) + \varepsilon\varphi(s)\tilde{x}_s(s) + O(\varepsilon^2),$$
$$X_s = \tilde{x}_s(s) + \varepsilon\big(\varphi_s(s)\tilde{x}_s(s) + \varphi(s)\tilde{x}_{ss}(s)\big) + O(\varepsilon^2)$$

and

$$\begin{aligned} &\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \mathcal{F}(X(\cdot, \varepsilon)) \\ &\quad = \int_{s_1}^{s_2} \Big(\varphi(s)\frac{d}{ds}F(\tilde{x}, \tilde{x}_s) + \varphi_s(s)\frac{\partial F(\tilde{x}, \tilde{x}_s)}{\partial x_s}\tilde{x}_s(s)\Big) ds \\ &\quad = \int_{s_1}^{s_2} H(\tilde{x}(s), \tilde{x}_s(s))\varphi_s(s) ds = 0. \end{aligned}$$

Here we use integration by parts.

Since φ is an arbitrary function the proposition is proved.

References

[1] N. I. Akhiezer: The Calculus of Variations. Blaisdell, New York, 1962.