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:

$$\frac{dx}{dt} = x_t.$$

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

$$\mathcal{F}\big(x(\cdot)\big) = \int_{t_1}^{t_2} F(x(t), x_t(t))dt \tag{2}$$

on the set of smooth functions $x = (y^T, z^T)^T : [t_1, t_2] \to \Omega$ with boundary conditions

$$z(t_1) = z_1, \quad z(t_2) = z_2, \quad y(t_1) = y_1, \tag{3}$$

and constraints

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

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

Assume that

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

and equation (4) can equivalently be written as

$$y_t = \Phi(y, z, z_t). \tag{6}$$

Definition 1. *Let a smooth function*

$$\tilde{x} : [t_1, t_2] \to \Omega, \quad \tilde{x}(t) = (\tilde{y}^T, \tilde{z}^T)^T(t)$$

be such that

$$a(\tilde{x}(t), \tilde{x}_t(t)) = 0, \quad \tilde{x}(t_1) = x_1 = (y_1^T, z_1^T)^T, \quad \tilde{z}(t_2) = z_2.$$

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

$$X : [t_1, t_2] \times (-\varepsilon_0, \varepsilon_0) \to \mathbb{R}^m, \quad X(t, \varepsilon) = (Y^T, Z^T)^T(t, \varepsilon), \quad \varepsilon_0 > 0$$

such that

1) $X\big([t_1, t_2] \times (-\varepsilon_0, \varepsilon_0)\big) \subset \Omega$;

2) $X(t, 0) = \tilde{x}(t), \quad t \in [t_1, t_2]$;

3) $X(t_1,\varepsilon)=x_1,\quad Z(t_2,\varepsilon)=z_2,\quad \varepsilon\in(-\varepsilon_0,\varepsilon_0)$ *;*
4) $a(X(t,\varepsilon),X_t(t,\varepsilon))=0,\quad (t,\varepsilon)\in[t_1,t_2]\times(-\varepsilon_0,\varepsilon_0)$
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 $\lambda(t)=(\lambda_1,\dots,\lambda_n)(t)$ such that $\tilde x$ satisfies the equations*

$$\frac{d}{dt}\frac{\partial F^*}{\partial x_t}-\frac{\partial F^*}{\partial x}=0,\quad F^*(t,x,x_t)=F(x,x_t)+\lambda(t)a(x,x_t), \tag{7}$$

and

$$\frac{\partial F}{\partial y_t}(\tilde x(t_2),\tilde x_t(t_2))+\lambda(t_2)\frac{\partial a}{\partial y_t}(\tilde x(t_2),\tilde x_t(t_2))=0. \tag{8}$$

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

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

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

where

$$B(x)=\begin{pmatrix} b^1_1(x) & b^1_2(x) & \cdots & b^1_m(x)\\ b^2_1(x) & b^2_2(x) & \cdots & b^2_m(x)\\ \vdots & \vdots & \ddots & \vdots\\ b^n_1(x) & b^n_2(x) & \cdots & b^n_m(x)\end{pmatrix}$$

is a matrix such that

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

Constraints (4) take the form

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

Equations (5), (6) imply that

$$B(x)x_t=A(x)y_t+C(x)z_t,\quad \det A\neq 0.$$

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

$$x_t(t)\in\mathcal{T}_{x(t)},\quad \mathcal{T}_x=\ker B(x)\subset T_xM,\quad \dim\mathcal{T}_x=m-n.$$

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

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

such that

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

and $x(t_2)\in\Sigma$.

Condition (8) takes the form:

$$\frac{\partial F^*}{\partial x_t}(t_2,\tilde x(t_2),\tilde x_t(t_2))v=0,\quad \forall v\in T_{\tilde x(t_2)}\Sigma.$$

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

$$x(t_1)=x_1,\quad x(t_2)\in\Sigma,\quad x_t(t)\in\mathcal{T}_{x(t)}.$$

Remark 1. *In general case equations (7) contain the derivatives λ_t and do not match with ones from classical mechanics with ideal constraints:*

$$\frac{d}{dt}\frac{\partial F}{\partial x_t}-\frac{\partial F}{\partial x}=\lambda B.$$

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

$$[F]_y=-\frac{d}{dt}\frac{\partial F}{\partial y_t}+\frac{\partial F}{\partial y},\quad [F]_z=-\frac{d}{dt}\frac{\partial F}{\partial z_t}+\frac{\partial F}{\partial z}$$

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

Let us put $Z(t,\varepsilon)=\tilde z(t)+\varepsilon\delta z(t)$,

$$\operatorname{supp}\delta z\subset[t_1,t_2]. \tag{10}$$

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

$$Y_t(t,\varepsilon)=\Phi(Y(t,\varepsilon),Z(t,\varepsilon),Z_t(t,\varepsilon)),\quad Y(t_1,\varepsilon)=y_1. \tag{11}$$

Remark 2. *That is why we can not impose condition $x(t_2)=x_2$. The value $Y(t_2,\varepsilon)$ has already been uniquely defined by other boundary conditions and the constraints. In other words if we add the condition $Y(t_2,\varepsilon)=y_2$ then the set of variations $\{X(t,\varepsilon)\}$ 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 t_2-t_1 small. Observe also that

$$Y_\varepsilon(t_1,\varepsilon)=0. \tag{12}$$

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

$$\begin{aligned}\frac{d}{d\varepsilon}\Big|_{\varepsilon=0}&\mathcal F\big(X(\cdot,\varepsilon)\big)\\ &=\int_{t_1}^{t_2}\Big([F]_z\delta z+[F]_yY_\varepsilon\Big)dt+\frac{\partial F}{\partial y_t}(\tilde x(t_2),\tilde x_t(t_2))Y_\varepsilon(t_2,0)=0.\end{aligned} \tag{13}$$

The function $\lambda(t)$ is still undefined but due to condition (5) the value $\lambda(t_2)$ is determined uniquely from (8).

From condition 4) of definition 1 it follows that

$$A(\varepsilon)=\int_{t_1}^{t_2}\lambda(t)a(X(t,\varepsilon),X_t(t,\varepsilon))dt=0.$$

By the same argument as above we have

$$\begin{aligned}\frac{d}{d\varepsilon}\Big|_{\varepsilon=0}A&=\int_{t_1}^{t_2}\Big([\lambda a]_z\delta z+[\lambda a]_yY_\varepsilon\Big)dt\\ &+\lambda(t_2)\frac{\partial a}{\partial y_t}(\tilde x(t_2),\tilde x_t(t_2))Y_\varepsilon(t_2,0)=0.\end{aligned} \tag{14}$$

Summing formulas (14) and (13) we yield

$$\int_{t_1}^{t_2}\Big([F^*]_z\delta z+[F^*]_yY_\varepsilon\Big)dt=0. \tag{15}$$

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

$$\lambda_t = \Lambda(t, \lambda).$$

Since we know $\lambda(t_2)$, by the existence and uniqueness theorem we obtain $\lambda(t)$ as a solution to the IVP for (16).

Equation (15) takes the form

$$\int_{t_1}^{t_2} [F^*]_z \delta z dt = 0.$$

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:*

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

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

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

that is $H(\tilde{x}(t), \tilde{x}_t(t)) = \mathrm{const}$.

Proof of Proposition 1. Consider a function $X(t, \varepsilon) = \tilde{x}(t + \varepsilon\varphi(t))$ with a smooth function φ such that $\operatorname{supp} \varphi \subset [t_1 + t', t_2 - t'], \quad t' > 0$.

Take $\varepsilon > 0$ small such that $\varepsilon|\varphi(t)| \leq t'$. Then the function X satisfies all the conditions of Definition 1. To check this use (17).

Furthermore we have

$$X = \tilde{x}(t) + \varepsilon\varphi(t)\tilde{x}_t(t) + O(\varepsilon^2),$$
$$X_t = \tilde{x}_t(t) + \varepsilon\big(\varphi_t(t)\tilde{x}_t(t) + \varphi(t)\tilde{x}_{tt}(t)\big) + O(\varepsilon^2)$$

and

$$\begin{aligned} &\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \mathcal{F}(X(\cdot, \varepsilon)) \\ &= \int_{t_1}^{t_2} \Big(\varphi(t)\frac{d}{dt}F(\tilde{x}, \tilde{x}_t) + \varphi_t(t)\frac{\partial F(\tilde{x}, \tilde{x}_t)}{\partial x_t}\tilde{x}_t(t)\Big)dt \\ &= \int_{t_1}^{t_2} H(\tilde{x}(t), \tilde{x}_t(t))\varphi_t(t)dt = 0. \end{aligned}$$

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.