NECESSARY CONDITIONS OF THE CALCULUS OF VARIATIONS FOR A PROBLEM OF BOLZA-MAYER TYPE

(NEOBKHODIMYE USLOVIIA VARIATSIONNOGO ISCHISLENIIA DLIA ODNOI ZADACHI TIPA BOL'TSA-MAIERA)

PMM Vol.29, № 2, 1965, pp. 368-372

V.A.KOSMODEM'IANSKII
(Moscow)

(Received April 9, 1964)

The general problem of the optimization of certain processes of control is considered. It is supposed that the control functions depend parametrically on time and on the coordinates of the points of discontinuity of the first kind. The position of these points is to be determined from the equations satisfied by the extremals of a certain functional.

The applicability of the necessary conditions of the calculus of variations: the multiplier rule (Section 2), the conditions of Weierstrass (Section 3), of Clebsch (Section 4), and Jacobi (Section 5), is analyzed, for the problems of the type under consideration.

The theory is illustrated with the elementary example of the rectilinear motion of a two-stage rocket in a homogeneous gravity field, the resistence of the medium being neglected (Section 6).

1. Suppose that the process takes place in a dynamical system whose motion is governed by n ordinary differential equations of first order

$$g_s = x_s^{\cdot} - f_s(x_1, \ldots, x_n, u_1, \ldots, u_m, t) = 0 \qquad (s = 1, \ldots, n) \tag{1.1}$$

and the system of r finite equations

$$\Psi_k = \Psi_k \ (u_1, \ldots, u_m, t) = 0 \ (k = 1, \ldots, r < m) \tag{1.2}$$

In Equations (1.1) and (1.2), the $x_s(t)$ are the coordinates which fix the position of the system, and the $u_i(t, t_i)$ are the control functions, which have discontinuities of the first kind in the instants t_i.

It is supposed that the explicit dependence of the $m-r$ control functions $u_1, \ldots, u_{s-r}$ on the variables t and t_i is known.

The coordinates of the system satisfy p conditions at the ends (here t_0 and T are not prescribed)

$$\varphi_l = \varphi_l\,[t_0, T, x(t_0), x(T)] = 0 \qquad (l = 1, \ldots, p < 2n + 1) \tag{1.3}$$

The problem is to determine the instants of time t_i for which the functional of the form

$$J = g\,[t_0, x(t_0), T, x(T)] + \int_{t_0}^{T} f^0(t, x, u)\, dt \tag{1.4}$$

possess an extremum.

2. Consider the conditions (multiplier rule) for such an integral to be stationary. Suppose that the coordinates of an admissible family (Bliss) $x(t,b)$ are introduced; and that these functions are continuous and possess almost everywhere, with the exception of a finite number of values t_i, continuous derivatives with respect to t; that these functions also have continuous partial derivatives with respect to the parameter b, everywhere in the domain of the variables t and b under consideration. The variation of the family along a curve E (corresponding to the value $b=0$ of the parameter b) is defined by the Equations

$$\delta t_i(0)=\frac{\partial t_i}{\partial b_\sigma}db_\sigma=\xi_{i,\sigma}db_\sigma,\qquad \delta x_s=\frac{\partial x_s(t,0)}{\partial b_\sigma}db_\sigma=\eta_{s,\sigma}db_\sigma \tag{2.1}$$

$$\delta u_j=\frac{\partial u_j(t,t_i(0))}{\partial b_\sigma}db_\sigma=\zeta_{j,\sigma}db_\sigma$$

In these equations, b_σ denotes the totality of the parameters $b_1,\ldots,b_s$ while $\xi_{i,\sigma}$, $\eta_{s,\sigma}$, $\zeta_{j,\sigma}$ denote the variations corresponding to the parameter b_σ (here, and in what follows, summation is understood whenever a subscript is repeated). The variations $\eta_{s,\sigma}(t)$, $\zeta_{j,\sigma}(t)$ satisfy, along the curve E, the variational equations

(2.2)

$$\delta g_s=\dot\eta_{s,\sigma}-\frac{\partial f_s}{\partial x_i}\eta_{i,\sigma}-\frac{\partial f_s}{\partial u_j}\zeta_{j,\sigma}=0,\qquad \delta\psi_k=\frac{\partial\psi_k}{\partial u_j}\zeta_{j,\sigma}=0\quad\begin{pmatrix}i=1,\ldots,n;\ s=1,\ldots,n\\ k=1,\ldots,r;\ j=1,\ldots,m\end{pmatrix}$$

and also the variational equations at the end of the interval

$$\delta\varphi_l=(\varphi_{l,0}+\dot x_{s0}\varphi_{l,s0})\,\xi_{0,\sigma}+\varphi_{l,s0}\eta_{s,\sigma}(t_0)+(\varphi_{l,n}+\dot x_{sn}\varphi_{l,sn})\,\xi_{n,\sigma}+\varphi_{l,sn}\eta_{s,\sigma}(T) \tag{2.3}$$

where the subscript n refers to the number of parameters under consideration at the instant T.

Let us show that the problem posed in Section 1 is not trivial, by showing that a given curve E, satisfying Equations (1.1), may be imbedded in a family of curves which satisfy similar conditions as E.

The following imbedding lemma is valid (Bliss).

Let an admissible curve E satisfy Equations (1.1), (1.2) and $\xi_{i,\sigma}$, $\eta_{s,\sigma}(t)$ and $\zeta_{j,\sigma}(t)$ be an admissible set of variations which satisfy Equations (2.2) on E. Then there exists an s-parameter family of admissible curves, containing the given curve E when $b_\sigma=0$ $(\sigma=1,\ldots,s)$ which satisfy Equations (1.1) and (1.2), and are such that, for each $\sigma=1,\ldots,s$, the functions $\eta_{s,\sigma}(t)$ and $\zeta_{j,\sigma}(t)$ are the variations on E with respect to the parameters b_σ.

Cosider the function $t_i(b)$ defined by Equation

$$t_i(b)=t_i(0)+\frac{\partial t_i}{\partial b_\sigma}b_\sigma\qquad(|b_\sigma|<\varepsilon)$$

If $b_\sigma=0$, then $t_i(0)=t_i$. Suppose that, along the curve E the corresponding matrix has the same rank as the number of equations representing the control functions in the form $u_j=u_j(t,b)$.

Expanding these functions in the neighborhood of $b=0$ (i.e. on the curve E) with respect to the parameter b_σ, we obtain

$$u_j[t,t_i(b)]=u_j(t,t_i)+\zeta_{j,\sigma}b_\sigma$$

Then the system of differential equations becomes

$$\dot x_s-f_s[x,u+\zeta b,t]=0$$

Suppose that the curve E has a corner at the instant of time t_i.

From the existence theorem for systems of ordinary differential equations, it follows that, in the neighborgood of $(x_s, t, b=0)$ there is a solution of the normal system of ordinary differential equations (at least on the interval $[t_0, t_1]$) which may be representd in the form $x_s = X_s[t, t_0, x(t_0), b]$ with initial point t_0, $x_s(t_0)$. The functions

$$x_s = X_s[t, t_1, x(t_1) + b\eta(t_1), b] = x_s(t, b) \tag{2.4}$$

constitute an elementary family, whose curves satisfy Equations $\mathcal{G}_s = 0$ on the interval $[t_0, t_1]$. When $t = t_1$, the function $x_s(t)$ becomes

$$x_s(t_1, b) = X_s[t_1, x(t_1) + b\eta(t_1), b] = x_s(t_1, 0) + b_\sigma \eta_{s,\sigma}(t_1)$$

that is to say, the variations of the functions $x_{s,b}(t, 0)$ along the curve E have the initial values $\eta_{s,\sigma}(t_1)$.

Since the functions (2.4) satisfy the equations $\mathcal{G}_s = 0$, it follows that the functions $x_{s,b}(t, 0)$ satisfy the equations $\delta\mathcal{G}_s = 0$, and hence coincide with the variations $\eta_{s,\sigma}(t)$ inasmuch as these variations constitute the unique solution of the equations $\delta\mathcal{G}_s = 0$ which assume the initial values $\eta_{s,\sigma}(t)_1$.

In the similar way one can construct, on the interval (t_1, t_2), a new elementary family, adjoining the preceding one, etc.

Let us suppose that, on the interval $[t_0, T]$ under consideration, the control function u_j has a single point of discontinuity t_1

Let us distinguish the various functions, when considered on the interval (t_0, t_1), by a superscript minus sign, (for example $x_s^-(t)$, $u_j^-(t)$); similarly, when they are considered on the interval (t_1, T) a superscript plus sign will be used (for example $x_s^+(t)$, $u_j^+(t)$, etc.).

The explicit form of the functions $u_j^\pm(t, t_i, T)$, defined on their respective intervals for each concrete problem, does not play any role in the course of the proof, so that we shall simply write

$$u_j^- = u_j^-(t, t_0), \qquad u_j^+ = u_j^+(t, t_1)$$

Let us introduce the functions L and H

$$L = f^\circ + \lambda_s g_s - \mu_k \Psi_k = \lambda_s x_s^{\boldsymbol{\cdot}} - H$$

$$H = H_\lambda + H_\mu = \lambda_\alpha f_\alpha - \mu_k \psi_k \qquad (\lambda_0 = -1)$$

where $\lambda_s(t)$ and $\mu_k(t)$ are the Lagrange multipliers.

In constructing the expressions for the first variations, let us choose the multipliers $\lambda_s^\pm(t)$, $\mu_k^\pm(t)$ in such a way that the coefficients of $\eta_{s,\sigma}(t)$ $(s = 1, \ldots, n)$; $\zeta_{j,\sigma}(t)$ $(j = m, -r+1, \ldots, m)$; vanish; the remaining coefficients of the independent variations must then equal zero. We obtain then that the extremals must satisfy the differential equations

$$\frac{\partial L}{\partial \lambda_s^\pm} = 0 \quad (s = 1, \ldots, n), \qquad \frac{\partial L}{\partial \mu_k^\pm} = 0 \quad (k = 1, \ldots, r) \tag{2.5}$$

Further, the following differential equations must be satisfied by the functions $\lambda_s^\pm(t)$:

$$\frac{d}{dt}\frac{\partial L}{\partial x_s^{\boldsymbol{\cdot}\pm}} - \frac{\partial L}{\partial x_s^\pm} = 0 \qquad (s = 1, \ldots, n) \tag{2.6}$$

while the following, finite, equations determine the $\mu_k^\pm(t)$:

$$\frac{\partial L}{\partial u_j^\pm} = 0 \qquad (j = m - r + 1, \ldots, m) \tag{2.7}$$

Simultaneously, the following boundary condition hold

$$\frac{\partial \Phi}{\partial x_s^+(T)} + \lambda_s(T) = 0, \qquad -\frac{\partial \Phi}{\partial x_s(t_0)} + \lambda_s(t_0) = 0 \tag{2.8}$$

$$\frac{\partial \Phi}{\partial T} + \frac{\partial \Phi}{\partial x_s(T)} x_s^{\bullet}(T) + f^{\circ}(T) = 0 \qquad (\Phi = g + \rho_l \varphi_l)$$
$$-\frac{\partial \Phi}{\partial t_0} - \frac{\partial \Phi}{\partial x_s(t_0)} x_s^{\cdot}(t_0) + f^{\circ}(T) + \int_{t_0}^{t_1} \frac{\partial H}{\partial t_0} dt = 0 \tag{2.9}$$

Finally, one has the analogue of the Weierstrass-Erdmann corner condition

$$\left(\frac{\partial L}{\partial x_s^{\cdot}}\right)_{t_1-0} = \left(\frac{\partial L}{\partial x_s^{\cdot}}\right)_{t_1+0}, \qquad (H_\lambda)_{t_1-0} - (H_\lambda)_{t_1+0} + \int_{t_1}^{T} \frac{\partial H}{\partial t_1} dt = 0 \tag{2.10}$$

Thus, we have obtained $2n$ first order differential equations (2.6) satisfied by the multipliers $\lambda_s^{\pm}(t)$ $(s = 1, \ldots, n)$; $2r$ relations (2.7) which are satisfied by the $\mu_k^{\pm}(t)$ $(k = 1, \ldots, r)$; and, further, $2n$ differential equations (2.5) which are satisfied by the $x_s^{\pm}(t)$ $(s = 1, \ldots n)$.

The unknown quantities, so far, are: $4n$ arbitrary constants, arising from the solution of the corrsponding first order equations (2.5) and (2.6), the quantities t_0, t_1, T, and also the p multipliers ρ_l; in all, $4n + p + 3$ quantities.

For determining these unknowns we have $2n$ boundary conditions (2.8); n conditions of the $\lambda_s(t)$ multipliers continuity in (2.10); n conditions of the coordinates continuity in the point t_1 $x_s^-(t_1) = x_s^+(t_1)$ and p relationships (1.3); three equations (2.9), (2.10), in all $4n + p + 3$ quantities.

Equations (2.5) to (2.10) express the fact that the functional J is to be rendered stationary in our variational problem.

3. For the type of problem under consideration, the necessary condition of Weierstrass holds; this condition may be formulated as follows [3]: an admissible curve E, satisfying the system of equations (1.1) and the stationary conditions with the multipliers $\lambda_0 = -1$, $\lambda_s(t)$ is said to satisfy the necessary condition of Weierstrass with these multipliers,provided that for every element $(t, x, x^{\cdot}, u, \lambda, \mu)$ of the curve E, the following inequality holds:

$$E = L(x, X^{\cdot}, U, \lambda, \mu, t) - L(x, x^{\cdot}, u, \lambda, \mu, t) - \frac{\partial L}{\partial x_s^{\cdot}}(X_s^{\bullet} - x_s^{\cdot}) \geqslant 0 \tag{3.1}$$

for every admissible $(x, u, \lambda, \mu) \neq (X^{\cdot}, U, \lambda, \mu)$ which satisfies the systems (1.1), (1.2).

Employing the normal fundamental system of differential equations, the necessary conditions of Weierstrass may be written thus

$$E = H(x, u, \lambda, \mu, t) - H(x, U, \lambda, \mu, t)$$

that is to say, the following inequality is a necessary condition for the existence of a strong minimum:

$$H(x, u, \lambda, \mu, t) \geqslant H(x, H, \lambda, \mu, t) \tag{3.2}$$

The proof of the theorem coincides with the proof of similar theorems which appear in [2 and 3].

4. Suppose that the control U_k and the derivatives of the coordinates $X_s^{\bullet}$ differ from u_k and $x_s^{\bullet}$ by small quantities

$$U_k = u_k + \delta u_k, \qquad X_s^{\bullet} = x_s^{\bullet} + \delta x_s^{\bullet} \tag{4.1}$$

where δu_k, $\delta x_s^{\bullet}$ satisfies the variational equations on E

$$\eta_{s,\sigma}^{\bullet} - \frac{\partial f_s}{\partial u_j} \zeta_{j,\sigma} = 0, \qquad \frac{\partial \Psi_k}{\partial u_j} \zeta_{j,\sigma} = 0 \tag{4.2}$$

Substituting from (4.1) into the inequality (3.2), and expanding E in powers of $\delta u_k, \delta x_s^{\cdot}$ we obtain

$$\frac{\partial^2 H}{\partial u_\alpha \partial u_\beta} \delta u_\alpha \delta u_\beta \leqslant 0 \tag{4.3}$$

Thus, the necessary condition of Clebsch may be formulated as follows: An admissible curve E, satisfying Equations $g_s = 0, \psi_k = 0$, and the multiplier rule, is said to satisfy the necessary condition of Clebsch with respect to these multipliers, provided that for each element

$$(t, x, x^{\cdot}, u, \lambda, \mu) \in E$$

the inequality (4.3) holds for arbitrary $\delta u_k, \delta u_\beta$ which satisfy the variational equation (4.2).

5. Let us suppose, as was done before, that the curve E is imbedded in an s-parameter family and that it satisfies the multiplier rule, that is, $x_s = x_s^{\cdot}(t, b)$, with $|b| < \epsilon$. Let us express the first differential of J in the form

$$\Delta I = \Delta\Phi + [L(t_1 - 0) - L(t_1 + 0)]\,\delta t_1 + L(T)\,\delta T - L(t_0)\,\delta t_0 +$$

$$+ \int_{t_0}^{t_1} \left[\frac{\partial L}{\partial x_s^-} \delta x_s^- + \frac{\partial L}{\partial x_s^{\cdot -}} \delta x_s^{\cdot -} + \frac{\partial L}{\partial u_\beta^-} \delta u_\beta^- \right] dt +$$

$$+ \int_{t_1}^{T} \left[\frac{\partial L}{\partial x_s^+} \delta x_s^+ + \frac{\partial L}{\partial x_s^{\cdot +}} \delta x_s^{\cdot +} + \frac{\partial L}{\delta u_\beta^+} \delta u_\beta^+ \right] dt$$

Regarding this as a function of the parameter b we obtain the second differential

$$\Delta^2 I = \Delta\varphi + \left[\left(\frac{\partial L}{\partial t} - x_\alpha^{\cdot} \frac{\partial L}{\partial x_\alpha^{\cdot}} \right) \delta t^2 + 2 \frac{\partial L}{\partial x_\alpha} \Delta x_\alpha \delta t + 2 \frac{\partial L}{\partial u_j} \delta u_j \delta t \right]_{t_0}^{T} +$$

$$+ \left[\left(\frac{\partial L}{\partial t} - x_\alpha^{\cdot} \frac{\partial L}{\partial x_\alpha^{\cdot}} \right) \delta t^2 + 2 \frac{\partial L}{\partial x_\alpha} \Delta x_\alpha \delta t + 2 \frac{\partial L}{\partial u_j} \delta u_j \delta t \right]_{t_1+0}^{t_1-0} +$$

$$+ \int_{t_0}^{t_1} \left[2\omega^- + \frac{\partial L}{\partial u_\beta^-} \frac{\partial^2 u_\beta^-}{\partial t_0^2} \delta t_0^2 \right] dt + \int_{t_1}^{T} \left[2\omega^+ + \frac{\partial L}{\partial u_\beta^+} \frac{\partial^2 u_\beta^+}{\partial t_1^2} \delta t_1^2 \right] dt \tag{5.1}$$

where

$$\Delta x_\alpha = x_\alpha^{\cdot} \delta t + \delta x_\alpha, \qquad \Delta\varphi = \frac{\partial^2 \Phi}{\partial x_\alpha \partial x_\beta} \Delta x_\alpha \Delta x_\beta + 2 \frac{\partial^2 \Phi}{\partial t\, \partial x_\alpha} \Delta x_\alpha \delta t + \frac{\partial^2 \Phi}{\partial t^2} \delta t^2$$

$$-2\omega = \frac{\partial^2 H}{\partial x_\alpha \partial x_\beta} \delta x_\alpha \delta x_\beta + 2 \frac{\partial^2 H}{\partial x_\alpha \partial u_\beta} \delta x_\alpha \delta u_\beta + \frac{\partial^2 H}{\partial u_k \partial u_\beta} \delta u_k \delta u_\beta$$

$$(t = t_0,\ T,\ t_1 - 0,\ t_1 + 0)$$

Further, (5.1) must be nonnegative for arbitrary admissible sets of variations which satisfy the variational equations (2.2) on E.

Thus, we have shown the applicability of the fourth necessary condition for a minimum for variational problems of the kind under consideration, namely:

"A curve E with multipliers $\lambda_0 = -1$, $\lambda_s(t)$ is said to satisfy the fourth necessary condition for a minimum, provided that the second variation $\Delta^2 I$ of (5.1) is nonnegative on E."

The nonnegativity of the second variation may be obtained by solving the related problem of the minimum of the second variation [2 and 3].

6. By way of illustration of the above mentioned theory, we shall consider the simple example of the rectilinear motion of a two-stage rocket in a homogeneous gravity field, without taking into account the resistence of the medium. The equation of motion is

$$v^{\cdot} = -V^r \frac{u^{\cdot}}{u} - g \tag{6.1}$$

where v is the velocity of the composite rocket, g is the acceleration of gravity, $u = m/m_0$ is the dimensionless mass of the composite rocket, where m is the linearly varying mass of the composite rocket, m_0 is the initial mass of the rocket, and V^r is the relative velocity of the combustion products.

The optimum instant t_1 for the separation of a two-stage composite rocket which will give the maximum v at the end of the combustion period (at burn out), can be found from Equations

$$\beta_1 V_1{}^r u_1 u_{2-} = \beta_2 V_2{}^r u_2 u_{1-}, \qquad (\beta_i \text{ fuel consumption per sec.})$$

$$u_1 = u\,(t_1 + 0), \qquad u_{2-} = u\,(T - 0), \quad u_2 = u\,(T + 0), \qquad u_{1-} = u\,(t_1 - 0)$$

Since $\lambda(t) = -1$, the function H is given by

$$H = -\left(\frac{\beta V^r}{u} - g\right)$$

For this particular variational problem, let us verify that all the necessary conditions mentioned above are satisfied by the extremal E_2 which is obtained upon solving Equation (6.1). Weierstrass' inequality (3.2) becomes

$$-\frac{\beta_1 V_1{}^r}{u_{1-}} \geqslant -\frac{\beta_2 V_2{}^r}{u_1}, \qquad v^{\cdot}(t_1 - 0) \leqslant v^{\cdot}(t_1 + 0) \tag{6.2}$$

that is, the motor of the two-stage rocket must, at each instant of time, impart to the rocket an acceleration which is not less than the acceleration which a single-stage rocket would have received from its motor at that same instant.

Comparing E_2 with the solution obtained for a three-stage rocket, one arrives at an inequality analogous to (6.2).

The Weierstrass inequality enables us to show that a homogeneous n-stage rocket can impart a maximum velocity which exceeds that of a rocket with $k < n$ stages. It is readily verified that the necessary conditions of Clebsch and Jacobi are not satisfied along E_2, which represents a two-stage homogeneous rocket.

BIBLIOGRAPHY

1. Kosmodem'ianskii, V.A., Ob odnom tipe variatsionnykh zadach (On a type of variational problems). *PMM* Vol.27, № 6, 1963.
2. Troitskii, V.A., O variatsionnykh optimizatsii protsessov upravleniia (On variational problems of optimization of control processes). *PMM* Vol.26, № 1, 1962.
3. Bliss, G.A., Lektsii po variatsionnomu ischisleniiu (Lectures on Calculus of Variations). Izd.inostr.Lit., 1950.

Translated by J.B.D.