ON A CERTAIN FORM OF THE OPTIMALITY CONDITIONS*

V. V. Dikusar*, Z. M. Wojtowicz†

*Computing Center RAS, dikussar@yandex.ru
†Polytechnical Institute, Radom, Poland, mwojt@hotmail.com

The use of the mathematical theory of optimal processes in economic investigations is hindered by its lack of suitability for the solution of large multidimensional problems and of problems which include constraints on the phase coordinates. In this paper we describe a certain form of the optimality conditions with the aim of bringing it up to such a degree of formalism that it would be convenient enough for machine analysis and for solving the multidimensional dynamical problems of mathematical economics. To prove the necessity of these conditions in the course of the development of a generalized calculus of variations there was worked out a special tool called the generalized Lagrange method which with great completeness permits us to write out the whole collection of optimality conditions, This tool, applied to the proof of the existence of the system of adjoint functions for some sufficiently general problem in the theory of optimal processes, allowed us, firstly, to note new peculiarities of these functions, and secondly, to formulate a form of the principle of optimality which by its own convenience can be useful for machine realization because it is written only in terms of inequality relations.

The sufficiency of the principle of optimality for the problems posed in the theory of optimal processes is proved by a method which, in spite of definite limitations, is suitable for nonlocalized problems and can be applied without special regard for unconditional boundaries arising at the expense of phase constraints and also without the restrictive consideration of time as a phase coordinate.

I. **Problem A.** In a space $\mathfrak{Z}$ of type B we define an element $\mathfrak{z}$ ensuring the condition

$$\max\{f(\mathfrak{z})\colon \mathfrak{r}(\mathfrak{z}) = 0, \quad \mathfrak{p}(\mathfrak{z}) \geq 0\},$$

in which $f(\mathfrak{z})$ is a functional, $\mathfrak{r}(\mathfrak{z})$ and $\mathfrak{p}(\mathfrak{z})$ are abstract functions belonging to the classes $(\mathfrak{z} \to \mathfrak{N})$ and $(\mathfrak{z} \to \mathfrak{Z})$, where $\mathfrak{N}$ and $\mathfrak{Z}$ are spaces of type B. The space $\mathfrak{Z}$ is semiordered with the aid of a certain convex cone $\mathfrak{C} \subset \mathfrak{Z}$. The notation $\mathfrak{p}' \geq \mathfrak{p}''$ is used for any $\mathfrak{p}'$, $\mathfrak{p}'' \in \mathfrak{Z}$ satisfying the condition $\mathfrak{p}' - \mathfrak{p}'' \in \mathfrak{C}$.

Variations. The difference $\mathfrak{z} - \mathfrak{z}_0 = \bar{\mathfrak{z}}$, where $\mathfrak{z}$, $\mathfrak{z}_0 \in \mathfrak{Z}$ and $\mathfrak{z}_0$ is an admissible element of the problem, is called a variation of the element $\mathfrak{z}_0$. For the variation $\bar{\mathfrak{z}}$ of the element $\mathfrak{z}_0$ we delineate the following sets: a) the set $\mathfrak{L}$ of variations admissible under equality constraints if $\mathfrak{r}(\mathfrak{z}_0 + \varepsilon\bar{\mathfrak{z}}) = o(\varepsilon)$, where ε is small; b) the set $\mathfrak{C}$ of variations admissible under inequality constraints if $\mathfrak{p}(\mathfrak{z} + \varepsilon(\hat{\mathfrak{z}} + \mathfrak{z})) > 0$ is valid for all sufficiently small $\hat{\mathfrak{z}} \in \mathfrak{Z}$ and for small $\varepsilon > 0$.

Assumptions. A_1. The solution of Problem A exists.

A_2. Gâteaux derivatives, being linear operators of the variations of the arguments exist in the abstract functions $f(\mathfrak{z})$, $\mathfrak{c}(\mathfrak{z})$ and $\mathfrak{p}(\mathfrak{z})$ on $\mathfrak{Z}$.

A_3. The set $\mathfrak{L}$ is a subspace of the space $\mathfrak{Z}$, while $\mathfrak{C}$ is a convex cone. Here, the intersection $\mathfrak{L} \cap \mathfrak{C}$ is not empty.

Note that the subspace $\mathfrak{L}$ is defined by the equality $\partial\mathfrak{c}(\mathfrak{z}_0)\bar{\mathfrak{z}}/\partial\mathfrak{z} = 0$, while sufficiently small variations $\bar{\mathfrak{z}}$, satisfying the inequality $\partial\mathfrak{p}(\mathfrak{z}_0)\bar{\mathfrak{z}}/\partial\mathfrak{z} + \mathfrak{p}(\mathfrak{z}_0) > 0$ belong to the cone $\mathfrak{C}$.

An arbitrary $\mathfrak{z} \in \mathfrak{Z}$, representable in the form $\mathfrak{z} = \mathfrak{z}_0 + \bar{\mathfrak{z}}$, where $\bar{\mathfrak{z}} \in \mathfrak{L} \cap \mathfrak{C}$, is called an admissible modified value of the element $\mathfrak{z}_0$.

Lemma 1 *A necessary condition for $\mathfrak{f}(\mathfrak{z})$ to have a maximum at an element $\mathfrak{z}_0$ among the admissible elements belonging to a neighborhood of $\mathfrak{z}_0$ is that $\mathfrak{f}(\mathfrak{z})$ have a maximum at the element $\mathfrak{z}_0$ among its admissible modified values in any sufficiently small neighborhood.*

Incompatibility. We use the notation $\mathfrak{p}' > \mathfrak{p}''$ if $\mathfrak{p}' \geq \mathfrak{p}''$ and $\mathfrak{p}' - \mathfrak{p}''$ does not coincide with the vertex of $\mathfrak{C}$. In the product $\sigma = E^1 \times \mathfrak{N} \times \mathfrak{B}$, where E^1 is a real one-dimensional space, we consider the following

*This paper was supported by Russian-Moldavian-Ukrainian Funds of fundamental investigation. Projects 08-01-90101, 08-01-90425.

convex sets for an admissible element $\mathfrak{z}_0$ and for some neighborhood $\mathfrak{Z}^*$ of the origin of $\mathfrak{Z}$:

$$\bar{\sigma} = \{\mathfrak{s}_0, \mathfrak{s}_1, \mathfrak{s}_2 \colon \mathfrak{s}_0 > 0,\ \mathfrak{s}_1 = 0, \mathfrak{s}_2 > 0\},$$

$$\sigma^*(\mathfrak{z}_0, \mathfrak{Z}^*) = \Big\{\mathfrak{s}_0, \mathfrak{s}_1, \mathfrak{s}_2 \colon \mathfrak{s}_0 \leq \frac{\partial}{\partial \mathfrak{z}}\, \mathfrak{f}(\mathfrak{z}_0)\bar{\mathfrak{z}},\ \mathfrak{s}_1 = \frac{\partial}{\partial \mathfrak{z}}\, \mathfrak{c}(\mathfrak{z}_0)\bar{\mathfrak{z}},$$

$$\mathfrak{s}_2 \leq \frac{\partial}{\partial \mathfrak{z}}\, \mathfrak{p}(\mathfrak{z}_0)\bar{\mathfrak{z}} + \mathfrak{p}(\mathfrak{z}_0),\ \bar{\mathfrak{z}} \in \mathfrak{Z}^*\Big\}.$$

Lemma 2 *A necessary condition for $\mathfrak{f}(\mathfrak{z})$ to have a maximum at an element $\mathfrak{z}_0$ among the admissible elements of the problem is the condition of incompatibility of the sets $\bar{\sigma}$ and $\sigma^*(\mathfrak{z}_0, \mathfrak{Z}^*)$, i.e., $\bar{\sigma} \cap \sigma^*(\mathfrak{z}_0, \mathfrak{Z}^*) = \varnothing$ for some $\mathfrak{Z}^* \subset \mathfrak{Z}$.*

Lemma 3 *There exists a functional $\lambda(\mathfrak{e})$, defined on σ, which is strictly positive on a and positive on $\sigma^*(\mathfrak{z}_0, \mathfrak{Z}^*)$. Sets of type B of linear continuous functional $\rho = \langle \rho, \mathfrak{c}\rangle$ and $\pi = \langle \pi, \mathfrak{p}\rangle$. defined, respectively, on $\mathfrak{R}$ and $\mathfrak{B}$, are denoted by R and Π.*

The functional π will be called nonnegative if it is nonnegative on $\mathfrak{C}$, and strictly positive if it is positive on the cone $\mathfrak{C}$ excepting its vertex.

Theorem 1 *For Problem A a necessary condition for the optimality of an element $\mathfrak{z}_0$ in some neighborhood is the existence of linear continuous functionals $\rho \in R$ and $\pi \in \Pi$, of which π is strictly positive, such that $\langle \pi, \mathfrak{p}(\mathfrak{z}_0)\rangle = 0$ is valid and*

$$\frac{\partial}{\partial \mathfrak{z}} = \mathfrak{F}(\mathfrak{z}_0, \rho, \pi)\bar{\mathfrak{z}} = 0,$$

where

$$\mathfrak{f}(\mathfrak{z}, \rho, \pi) = \mathfrak{f}(\mathfrak{z}) + \langle \rho, \mathfrak{r}(\mathfrak{z})\rangle + \langle \pi, \mathfrak{p}(\mathfrak{z})\rangle$$

is fulfilled for any $\mathfrak{z}$.

II. Let us consider a problem in the theory of optimal processes to which are reduced: the problem with parameters, by an augmentation of the phase coordinate vector, and the problem with moving end-points, by the substitution $t = t_1 + \tau(t_2 - t_1)$.

Problem B. On the segment $[t_1, t_2]$ there are defined the vector functions $x(t)$ and $u(t)$ which ensure $\max\{G : dx/dt = f, g = 0, h \geq 0\}$, where G is functional of $x(t_1)$, $x(t_2)$; f, g, h are vector functions of $x(t)$, $u(t)$, t, $x(t_1)$, $x(t_2)$. We shall treat (he regular case. It is characterized by the fact that the set $\{u \colon g = 0,\ h \geq 0\}$ is bounded for any finite $x(t)$, t, ξ, η.

We shall seek solutions of Problem B in the class of absolutely continuous $x(t)$ and bounded measurable $u(t)$.

We make the assumptions:

B_1. The optimal solution $x^0(t)$, $u^0(t)$ of Problem B exists.

B_2. The functional $G(\xi, \eta)$ is continuously differentiable in the neighborhood of $\xi^0 = x^0(t_1)$, $\eta^0 = x^0(t_2)$.

B_3. The functions f, g, h, considered as functions of x, u, t, ξ, η for any absolutely continuous $x(t)$, for a bounded measurable $u(t)$ and for any ξ, and η taken, respectively, from the neighborhoods of $x^0(t)$, $u^0(t)$, $x^0(t)$, $x^0(t_2)$, are summable with respect to t on $[t_1, t_2]$ together with their first order derivatives with respect to x, u, ξ, η.

B_4. The conditions $g = 0$ and $h \geq 0$ ensure a nonempty set of admissible solutions in the neighborhood of the optimal solution.

Theorem 2 *Under assumptions B_1–B_4 for Problem B there exist vector functions $\psi(t)$, $\omega(t)$ and $\varepsilon(t)$, defined on $[t_1, t_2]$ (of which $\psi(t)$) is absolutely continuous while $\omega(t)$ and $\varepsilon(t)$ are measurable and bounded almost everywhere), satisfying at the optimal values of $x(t)$, $u(t)$ the conditions*

$$\frac{d\psi}{dt} + \psi f'_x + \omega g'_x + \varepsilon h'_x = 0; \tag{1}$$

$$\psi f'_u + \omega g'_u + \varepsilon h'_u = 0; \tag{2}$$

$$F'_\xi + \psi(t_1) + \int_{t_1}^{t_2} (\psi f'_\xi + \omega g'_\xi + \varepsilon h'_\xi) dt = 0; \tag{3}$$

$$G'_n - \psi(t_2) + \int_{t_1}^{t_2} (\psi f'_\eta + \omega g'_\eta \varepsilon h'_\eta) dt = 0; \tag{4}$$

$$\varepsilon(t) h = 0; \tag{5}$$

$$\varepsilon(t) \geq 0, \tag{6}$$

in which the expressions (1), (2), (5) and (6) are fulfilled almost everywhere on $[t_1, t_2]$; $\xi = x(t_1)$, $\eta = x(t_2)$. The vector quantities x, u, t, g, h are written as columns while $\psi(t)$, $\omega(t)$, $\varepsilon(t)$, as rows. The rule for the differentation of a vector with respect to a vector is the usual one.