ISSN 0012-2661, Differential Equations, 2015, Vol. 51, No. 1, pp. 33–46. © Pleiades Publishing, Ltd., 2015.
Original Russian Text © A.V. Arutyunov, D.Yu. Karamzin, 2015, published in Differentsial'nye Uravneniya, 2015, Vol. 51, No. 1, pp. 34–47.

CONTROL THEORY

Maximum Principle in an Optimal Control Problem with Equality State Constraints

A. V. Arutyunov and D. Yu. Karamzin
Peoples' Friendship University of Russia, Moscow, Russia
Computing Centre of the Russian Academy of Sciences, Moscow, Russia
e-mail: arutun@orc.ru, dmitry_karamzin@mail.ru

Received December 25, 2013

Abstract—We consider an optimal control problem with equality state constraints. We prove nondegenerate necessary optimality conditions in the form of the Pontryagin maximum principle.

DOI: 10.1134/S0012266115010048

1. STATEMENT OF THE PROBLEM AND MAIN DEFINITIONS

Consider the optimal control problem

$$\begin{aligned} &K_0(p) + \int_{t_1}^{t_2} f_0(x,u,t)\,dt \to \min, \qquad \dot{x} = f(x,u,t), \qquad t \in [t_1,t_2], \qquad t_1 < t_2, \\ &G(x,t) = 0, \qquad K_1(p) \le 0, \qquad K_2(p) = 0, \qquad p = (x_1,x_2,t_1,t_2), \qquad u(t) \in U. \end{aligned} \tag{1}$$

The given vector functions G and K_i take values in arithmetic spaces of dimensions $d(G)$ and $d(K_i)$, respectively, K_0 and f_0 are scalar functions, $U \subseteq \mathbb{R}^m$ is a given closed set, $\dot{x} = dx/dt$, $t \in [t_1,t_2]$ is the time (the time instants t_1 and t_2 are not assumed to be fixed), x is a state variable taking values in the n-dimensional arithmetic space $\mathbb{R}^n$, and $u \in \mathbb{R}^m$ is a control parameter. The vector $p \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^1 \times \mathbb{R}^1$ is said to be terminal. For the class of admissible controls we take measurable essentially bounded functions $u(\cdot)$ ranging in the set U (see [1, p. 86; 2]).

Let $u(t)$, $t \in [t_1,t_2]$, be an admissible control, let $x(t)$, $t \in [t_1,t_2]$, be the corresponding trajectory, and let p be the corresponding terminal vector. The triple (p,x,u) is referred to as an *admissible process* if it satisfies the terminal constraints $K_1(p) = 0$ and $K_2(p) \le 0$ and the state constraints $G(x(t),t) = 0$ $\forall t \in [t_1,t_2]$.

In what follows, we assume that all functions occurring in the statement of the problem are continuously differentiable and the function G is twice continuously differentiable. Let us introduce notions needed below.

In the following, we assume that the state constraints are regular, i.e., that the matrix $\dfrac{\partial G}{\partial x}(x,t)$ has full rank for all $(x,t)\colon\ G(x,t) = 0$.

We say that the terminal constraints are regular at a point $p = (x_1,x_2,t_1,t_2)\colon\ K_1(p) \le 0$, $K_2(p) = 0$ if

$$\operatorname{rank} \frac{\partial K_2}{\partial p}(p) = d(K_2)$$

and there exists a $d \in \ker \dfrac{\partial K_2}{\partial p}(p)$ such that $\left\langle \dfrac{\partial K_1^j}{\partial p}(p), d \right\rangle > 0$ for all j such that $K_1^j(p) = 0$. (Here and throughout the following, superscripts stand for the coordinates of a vector or a vector function.)

We say that the state constraints at a point p^* are coordinated with the terminal constraints if there exists a number $\varepsilon > 0$ such that

$$\left\{p \in \mathbb{R}^{2n+2} : \ |p^* - p| \le \varepsilon, \ K_1(p) = 0, \ K_2(p) \le 0\right\} \subseteq \{p : \ G(x_1, t_1) = 0, \ G(x_2, t_2) = 0\}.$$

Let us introduce the notion of regularity of a trajectory. Let (p^*, x^*, u^*) be an admissible process in problem (1). Here $p^* = (x_1^*, x_2^*, t_1^*, t_2^*)$. Set

$$\Gamma(x, u, t) = \frac{\partial G}{\partial x}(x, t) f(x, u, t) + \frac{\partial G}{\partial t}(x, t),$$
$$U(x, t) := \{u \in U : \ \Gamma(x, u, t) = 0\}, \qquad T = [t_1^*, t_2^*].$$

Definition 1. A trajectory $x^*(\cdot)$ is said to be *regular* if the relations

$$\operatorname{rank} \frac{\partial \Gamma}{\partial u}(x^*(t), u, t) = d(G), \qquad \operatorname{im} \frac{\partial \Gamma^*}{\partial u}(x^*(t), u, t) \cap N_U(u) = \{0\}$$

hold for all $t \in T$ and $u \in U(x^*(t), t)$.

Here $N_U(u)$ is the limit normal cone of the set U at the point u in the Mordukhovich sense (see [3, p. 4]). For $u \in U$, this cone is defined by the relation

$$N_U(u) := \operatorname{Ls}_{y \to u} \operatorname{cone}(y - \Pi_U(y)).$$

Here $\Pi_U(y)$ is the projection of a vector y into the set U; i.e.,

$$\Pi_U(y) = \{u \in U : \ |y - u| = \operatorname{dist}(y, U)\},$$

where $\operatorname{dist}(y, U) = \inf_{u \in U} |y - u|$ is the distance of a point to a set, cone is the conical hull of a set, and Ls is the upper topological limit of a family of sets. [It consists of all possible limit points of sequences of vectors in $\operatorname{cone}(y_k - \Pi_U(y_k))$ as $y_k \to u$.]

2. MAXIMUM PRINCIPLE

Consider the extended Hamilton–Pontryagin function

$$\bar{H}(x, u, t, \psi, \mu, \lambda^0) = \langle \psi, f(x, u, t) \rangle - \langle \mu, \Gamma(x, u, t) \rangle - \lambda^0 f_0(x, u, t)$$

and the small Lagrangian

$$l(p, \lambda) = \lambda^0 K_0(p) + \langle \lambda^1, K_1(p) \rangle + \langle \lambda^2, K_2(p) \rangle,$$

where $\lambda = (\lambda^0, \lambda^1, \lambda^2)$, $\lambda^0 \in \mathbb{R}$, $\lambda^1 \in \mathbb{R}^{d(K_1)}$, and $\lambda^2 \in \mathbb{R}^{d(K_2)}$.

Definition 2. We say that an *admissible process* (p^*, x^*, u^*) *satisfies the Pontryagin maximum principle* if there exists a vector $\lambda = (\lambda^0, \lambda^1, \lambda^2)$, $\lambda^0 \ge 0$, $\lambda^1 \ge 0$, $\langle \lambda^1, K_1(p^*) \rangle = 0$, an absolutely continuous function $\psi : \ T \to \mathbb{R}^n$, and a measurable essentially bounded function $\mu : T \to \mathbb{R}^{d(G)}$ such that

$$\text{either} \quad \lambda^0 > 0, \quad \text{or} \quad \psi(t) \notin \operatorname{im} \frac{\partial G^*}{\partial x}(t) \quad \forall t \in T, \tag{2}$$

$$\dot{\psi}(t) = -\frac{\partial \bar{H}}{\partial x}(t) \quad \text{for a.a. } t \in T, \tag{3}$$

$$\psi(t_s^*) = (-1)^{s+1} \frac{\partial l}{\partial x_s}(p^*, \lambda), \quad s = 1, 2, \tag{4}$$

$$\max_{u \in U(t)} \bar{H}(u, t) = \bar{H}(t) \quad \text{for a.a. } t \in T, \tag{5}$$

$$\dot{h}(t) = \frac{\partial \bar{H}}{\partial t}(t) \quad \text{for a.a. } t \in T, \tag{6}$$

$$h(t_s^*) = (-1)^s \frac{\partial l}{\partial t_s}(p^*, \lambda), \qquad s = 1, 2, \tag{7}$$

$$\frac{\partial \bar{H}}{\partial u}(t) \in \operatorname{conv} N_U(u^*(t)) \quad \text{for a.a. } t \in T, \tag{8}$$

where $h(t) := \max_{u\in U(t)} \bar H(u,t)$; moreover, the above-mentioned maximum over $u \in U(t)$ is attained at each point $t \in T$.

A process (p^*, x^*, u^*) satisfying the maximum principle is referred to as an *extremal*, and the set (λ, ψ, μ, r) is referred to as the *Lagrange multipliers* corresponding to the process (p^*, x^*, u^*) in view of the maximum principle.

Here and throughout the following, we use the following agreement on notation. First, if some arguments of the mappings $\bar H$, G, Γ, f, U, and so on and their derivatives are omitted, then they are replaced by the optimal values $x^*(t)$ and $u^*(t)$ or the Lagrange multipliers $\psi(t)$, $\mu(t)$, and λ^0. Second, all Lagrange multipliers are treated as row vectors, while the vector functions or vectors f, x, and u are treated as column vectors. The gradients of functions are treated as rows, and the entries of the Jacobi matrix of a mapping $F(x):\ \mathbb{R}^n \to \mathbb{R}^k$ have the form $\dfrac{\partial F^i}{\partial x_j}(x)$; i.e., its rows are the gradients of the coordinate functions F^i.

Theorem 1. *Let a process* (p^*, x^*, u^*) *be optimal in problem* (1). *In addition, let* $x^*(t)$ *be an optimal trajectory, let the terminal constraints be regular and coordinated with the state constraints at the point* p^*, *and let the sets* $U(t)$ *be uniformly bounded for* $t \in T$. *Then the process* (p^*, x^*, u^*) *satisfies the maximum principle.*

Remark 1. By virtue of the formulas

$$\begin{aligned}
\frac{d}{dt}\left(\frac{\partial G}{\partial x}(t)\right) &= \frac{\partial^2 G}{\partial x^2}(t) f(t) + \frac{\partial^2 G}{\partial x\,\partial t}(t) = \frac{\partial \Gamma}{\partial x}(t) - \frac{\partial G}{\partial x}(t)\frac{\partial f}{\partial x}(t),\\
\frac{d}{dt}\left(\frac{\partial G}{\partial t}(t)\right) &= \frac{\partial^2 G}{\partial x\,\partial t}(t) f(t) + \frac{\partial^2 G}{\partial t^2}(t) = \frac{\partial \Gamma}{\partial t}(t) - \frac{\partial G}{\partial x}(t)\frac{\partial f}{\partial t}(t),
\end{aligned} \tag{9}$$

one can readily show that, in addition to the Lagrange multipliers (λ, ψ, μ, r), the set of Lagrange multipliers

$$\lambda, \quad \psi(t) + a\frac{\partial G}{\partial x}(t), \quad \mu(t) + a, \quad r(t),$$

where a is an arbitrary vector in $\mathbb{R}^{d(G)}$, satisfies the assumptions of the maximum principle as well. Indeed, under this transformation, the Hamiltonian $h(t)$ is replaced by $-a\dfrac{\partial G}{\partial t}(t)$, and the condition of maximum is preserved, because, by definition,

$$\frac{\partial G}{\partial x}(t) f(u,t) = -\frac{\partial G}{\partial t}(t) \quad \forall u \in U(t).$$

In this case, conditions (3) and (6) hold by virtue of (9).

Moreover, by virtue of the above argument, conditions (3)–(8) can always be satisfied trivially by choosing the following set of Lagrange multipliers:

$$\lambda = 0, \quad \psi(t) = a\frac{\partial G}{\partial x}(t), \quad \mu(t) = a, \quad r = 0,$$

where $a \in \mathbb{R}^{d(G)}$ is an arbitrary vector. However, obviously, such a set of Lagrange multipliers does not satisfy condition (2). This justifies the importance of condition (2) in the sense that it cannot be weakened by the replacement by the ordinary condition of nontriviality of the set of Lagrange multipliers.

The nontriviality condition (2) is also needed for a different reason. Let $\lambda^0 = 0$. Then, by virtue of condition (2), we have $\psi(t) \notin \operatorname{im}\dfrac{\partial G^*}{\partial x}(t)$ $\forall t \in T$. This condition is important, because if

$\psi(t) \in \operatorname{im} \dfrac{\partial G^*}{\partial x}(t)$ for some t, then the maximum condition (5) at the point t becomes less meaningful since

$$\bar{H}(u,t) = -\left\langle a(t), \frac{\partial G}{\partial t}(t)\right\rangle \quad \forall u \in U(t),$$

where $a(t) \in \mathbb{R}^{d(G)}$ is chosen from the condition $\psi(t) = a(t)\dfrac{\partial G}{\partial x}(t)$; therefore, $\bar{H}(u,t)$ treated as a function of u is constant on $U(t)$.

Remark 2. Consider the classical Hamilton–Pontryagin function [1, p. 11]

$$H(x,u,t,\psi,\lambda^0) = \langle \psi, f(x,u,t)\rangle - \lambda^0 f_0(x,u,t).$$

Since $\Gamma(x,u,t) = 0$ $\forall u \in U(x,t)$, it follows that the function $\bar{H}$ in the maximum condition (5) can be replaced by H.

Proof of Theorem 1. Without loss of generality, one can assume that $f_0 = 0$. This can always be achieved by the introduction of an additional state variable.

First, consider the auxiliary problem

$$\varphi(p) \to \min, \quad \dot{x} = f(x,u,t), \quad t \in T, \quad G(x,t) = 0, \quad p \in C, \quad p = (x_1, x_2), \quad u(t) \in U \tag{10}$$

on a fixed time interval $T = [t_1^*, t_2^*]$; for this problem, we prove the maximum principle without condition (6) and with condition (2) replaced by the condition

$$\psi(t_1^*) \in \ker \frac{\partial G}{\partial x}(t_1^*). \tag{11}$$

In the preceding, $\varphi(p) := K_0(p, t_1^*, t_2^*)$, and C is a given closed subset of $\mathbb{R}^{2n}$. In problem (10), the terminal constraints have a more general form than in problem (1). In this case, the coordination of terminal constraints with the state constraints implies that $C \subseteq \mathcal{G}$, where

$$\mathcal{G} = \{p = (x_1, x_2) : \ G(x_1, t_1^*) = 0, \ G(x_2, t_2^*) = 0\}.$$

Let (p^*, x^*, u^*) be an optimal process in problem (10). For problem (10), we prove the existence of Lagrange multipliers λ^0, ψ, and μ that do not vanish simultaneously and satisfy conditions (3), (5), (8), and (11) and the transversality condition

$$(\psi(t_1^*), -\psi(t_2^*)) \in \lambda^0 \frac{\partial \varphi}{\partial p}(p^*) + N_C(p^*). \tag{12}$$

For $\theta > 0$, set

$$\tilde{C} = \tilde{C}(\theta) := \bigcup_{p \in C:\, |p - p^*| \le \theta} (p + N_{\mathcal{G}}(p)).$$

By virtue of the regularity of the state constraints, the set $\mathcal{G}$ is a smooth manifold. Therefore, its normal cone $N_{\mathcal{G}}(p)$ is a subspace. However, by construction, $p^* + N_{\mathcal{G}}(p^*) \subseteq \tilde{C}$, and hence $N_{\tilde{C}}(p^*)$ lies in the orthogonal complement of $N_{\mathcal{G}}(p^*)$. Therefore,

$$N_{\tilde{C}}(p^*) \cap N_{\mathcal{G}}(p^*) = \{0\}. \tag{13}$$

By using the regularity of the state constraints, by performing the linearization of the mapping that defines the smooth manifold $\mathcal{G}$, and by taking into account the corresponding properties of the limit normal cone under local diffeomorphisms (see [4, Th. 5.2, formula (5.2)]), we find that

$$N_C(p^*) = N_{\tilde{C}}(p^*) + N_{\mathcal{G}}(p^*) \tag{14}$$

and, for the point $p^* \in \mathcal{G}$, there exists a neighborhood O such that

$$O \cap \mathcal{G} \cap (p + N_{\mathcal{G}}(p)) = p \quad \forall p \in O \cap \mathcal{G}.$$

By using the last relation and the definition of the set $\tilde{C}$, we take a number $\theta > 0$ small enough to ensure that $\mathcal{G} \cap \tilde{C} \subseteq C$; i.e.,

$$\text{if} \quad p \in \tilde{C} \cap \mathcal{G}, \quad \text{then} \quad p \in C. \tag{15}$$

Take positive numbers c and δ such that $|u| \le c$ for arbitrary $u \in U(x,t)$ and x satisfying the condition $|x - x^*(t)| \le \delta$ $\forall t \in T$. This is possible by virtue of the uniform boundedness of the sets $U(t)$ and the continuity of the mapping Γ.

Along with problem (10), consider the problem in which the state constraints are replaced by the mixed constraints

$$\begin{aligned} &\varphi(p) \to \min, \qquad \dot{x} = f(x,u,t), \\ &G(x_1, t_1^*) = 0, \qquad p \in \tilde{C}, \qquad p = (x_1, x_2), \qquad u(t) \in U, \\ &\Gamma(x,u,t) = 0 \quad \text{for a.a. } t \in T. \end{aligned} \tag{16}$$

By virtue of the obvious identity

$$G(x(t),t) = G(x_1, t_1^*) + \int_{t_1^*}^{t} \Gamma(x(s), u(s), s)\, ds \quad \forall t \in T,$$

and condition (15), the sets of admissible processes in problems (10) and (16) coincide. Therefore, the process (p^*, x^*, u^*) is also optimal in problem (16).

For convenience, we assume that $\varphi(p^*) = 0$. By $\mathcal{M} \subseteq \mathbb{R}^{2n} \times \mathbb{L}_1^m(T)$ we denote the set of pairs $(p, u(\cdot))$, $p = (x_1, x_2)$, such that $G(x_1, t_1^*) = 0$, $u(t) \in U(x(t), t)$ for almost all $t \in T$, $|x(t) - x^*(t)| \le \delta$ $\forall t \in T$, and $x_2 = x(t_2^*)$, where $x(\cdot)$ is the trajectory corresponding to the control $u(\cdot)$ and the initial condition $x(t_1^*) = x_1$. The set $\mathcal{M}$ is nonempty, because it contains the element $(p^*, u^*(\cdot))$. Moreover, one can readily see that $\mathcal{M}$ is closed; consequently, it is a complete metric space with metric induced by the norm $|p| + \|u\|_{L_1}$.

For $a \in \mathbb{R}$, set $a^+ = \max\{a, 0\}$. For each positive integer i, set $\varepsilon_i = i^{-1}$ and $\varphi_i(p) = (\varphi(p) + \varepsilon_i)^+$. On the set $\mathcal{M}$, we introduce the functional

$$F_i(p, u(\cdot)) = ((\varphi_i(p))^2 + (\operatorname{dist}(p, \tilde{C}))^2)^{1/2}.$$

The functional F_i is continuous and positive on $\mathcal{M}$. [The terms in the radicand do not vanish simultaneously, because $\varphi(p^*) = 0$.]

For fixed i, consider the problem

$$F_i(p, u(\cdot)) \to \min, \qquad (p, u(\cdot)) \in \mathcal{M}.$$

Obviously, $F_i(p^*, u^*(\cdot)) = \varepsilon_i$. Let us apply the Ekeland variational principle [5] to this problem. Then, for each i, there exists an element $(p_i, u_i(\cdot)) \in \mathcal{M}$, $p_i = (x_{1,i}, x_{2,i})$, such that

$$F_i(p_i, u_i(\cdot)) \le F_i(p^*, u^*(\cdot)) = \varepsilon_i, \tag{17}$$

$$|p_i - p^*| + \int_{t_1^*}^{t_2^*} |u_i(t) - u^*(t)|\, dt \le \sqrt{\varepsilon_i}\,, \tag{18}$$

and the pair $(p_i, u_i(\cdot))$ is a solution of the problem

$$\begin{aligned} &((\varphi_i(p))^2 + (\operatorname{dist}(p, \tilde{C}))^2)^{1/2} + \sqrt{\varepsilon_i}\left(|p - p_i| + \int_{t_1^*}^{t_2^*} |u - u_i(t)|\, dt \right) \to \min, \qquad \dot{x} = f(x,u,t), \\ &G(x_1, t_1^*) = 0, \qquad u(t) \in U(x(t), t) \quad \text{for a.a. } t, \qquad |x(t) - x^*(t)| \le \delta \quad \forall t \in T. \end{aligned} \tag{19}$$

We denote the optimal trajectory in this problem by $x_i(t)$.

It follows from inequality (18) that, after passage to a subsequence $p_i \to p^*$, we have $u_i(t) \to u^*(t)$ for almost all t, and hence $x_i(t) \rightrightarrows x^*(t)$ uniformly on T; therefore, in the derivation of optimality conditions in problem (19) for large values i, one can omit the constraint $|x(t) - x^*(t)| \le \delta\ \forall t \in T$. Therefore, problem (19) contains only terminal constraints at the left endpoint of the trajectory and mixed constraints $u \in U(x,t)$, which are regular for large i (see Definition 3 and Lemma 1 in Section 4) by virtue of the regularity of the optimal trajectory (see Definition 1).

To this problem, we apply the maximum principle for a problem without terminal constraints, which is to be proved in the appendix (Theorem 3 and Remarks 3 and 4); in this connection, we set $\lambda^0 = 1$. By virtue of this maximum principle, for each i, there exists a vector $\xi_i \in \mathbb{R}^{d(G)}$, an absolutely continuous function ψ_i, a measurable bounded function μ_i, and a number $\varkappa > 0$ independent of i such that

$$\dot{\psi}_i(t) = -\frac{\partial \bar{H}_i}{\partial x}(t) \quad \text{for a.a. } t \in T, \tag{20}$$

$$\begin{aligned} &\psi_i(t_1^*) \in \lambda_i^0 \frac{\partial \varphi_i}{\partial x_1}(p_i) + \varrho_i \partial_{x_1} \operatorname{dist}(p_i, \tilde{C}) + \xi_i \frac{\partial G_i}{\partial x}(t_1^*) + \sqrt{\varepsilon_i}\, B_{\mathbb{R}^n}, \\ &\psi_i(t_2^*) \in -\lambda_i^0 \frac{\partial \varphi_i}{\partial x_2}(p_i) - \varrho_i \partial_{x_2} \operatorname{dist}(p_i, \tilde{C}) + \sqrt{\varepsilon_i}\, B_{\mathbb{R}^n}, \end{aligned} \tag{21}$$

$$\max_{u \in U_i(t)} (\bar{H}_i(u,t) - \sqrt{\varepsilon_i}\, |u - u_i(t)|) = \bar{H}_i(t) \quad \text{for a.a. } t \in T, \tag{22}$$

$$\frac{\partial \bar{H}_i}{\partial u}(t) \in \operatorname{conv} N_U(u_i(t)) + \sqrt{\varepsilon_i}\, B_{\mathbb{R}^m} \quad \text{for a.a. } t \in T, \tag{23}$$

$$|\mu_i(t)| \le \varkappa(1 + |\psi_i(t)|) \quad \text{for a.a. } t \in T. \tag{24}$$

Here

$$\lambda_i^0 = \frac{\varphi_i(p)}{((\varphi_i(p))^2 + (\operatorname{dist}(p_i, \tilde{C}))^2)^{1/2}}, \qquad \varrho_i = \frac{\operatorname{dist}(p_i, \tilde{C})}{((\varphi_i(p))^2 + (\operatorname{dist}(p_i, \tilde{C}))^2)^{1/2}};$$

the subscript i at $\bar{H}$, G, and U implies that the variables x, u, ψ, and μ are replaced by their values $x_i(t)$, $u_i(t)$, $\psi_i(t)$, and $\mu_i(t)$; B_X is the closed unit ball in the space X, and $\partial_{x_s} h(\bar{p})$ is the limit subdifferential of the Lipschitz function $h(p)$ at the point $\bar{p}$ with respect to the variable x_s, $s = 1, 2$ (see [3, p. 82]). In condition (22), we have taken into account Remark 2.

The vector $\psi_i(t_1^*)$ can be represented in the form $\psi_i(t_1^*) = a_i + b_i$, where $a_i \in \operatorname{im} \dfrac{\partial G_i^*}{\partial x}(t_1^*)$ and $b_i \in \ker \dfrac{\partial G_i}{\partial x}(t_1^*)$. By virtue of Remark 1, not only the Lagrange multipliers ψ_i and μ_i but also the functions $\bar{\psi}_i$ and $\bar{\mu}_i$ satisfy conditions (20)–(23), where

$$\bar{\psi}_i(t) := \psi_i(t) - a_i \frac{\partial G_i}{\partial x}(t), \qquad \bar{\mu}_i(t) := \mu_i(t) - a_i.$$

Therefore, the inclusion

$$\bar{\psi}_i(t_1^*) \in \ker \frac{\partial G_i}{\partial x}(t_1^*) \tag{25}$$

holds for all i. It follows from conditions (20)–(23) that

$$\dot{\bar{\psi}}_i(t) = -\frac{\partial \bar{H}_i}{\partial x}(\bar{\psi}_i(t), \bar{\mu}_i(t), t), \qquad t \in T, \tag{26}$$

$$\begin{aligned} &\bar{\psi}_i(t_1^*) \in \lambda_i^0 \frac{\partial \varphi_i}{\partial x_1}(p_i) + \varrho_i \partial_{x_1} \operatorname{dist}(p_i, \tilde{C}) + (\xi_i - a_i) \frac{\partial G_i}{\partial x}(t_1^*) + \sqrt{\varepsilon_i}\, B_{\mathbb{R}^n}, \\ &\bar{\psi}_i(t_2^*) \in -\lambda_i^0 \frac{\partial \varphi_i}{\partial x_2}(p_i) - \varrho_i \partial_{x_2} \operatorname{dist}(p_i, \tilde{C}) - a_i \frac{\partial G_i}{\partial x}(t_2^*) + \sqrt{\varepsilon_i}\, B_{\mathbb{R}^n}, \end{aligned} \tag{27}$$

$$\max_{u\in U_i(t)}(\bar{H}_i(u,\bar{\psi}_i(t),\bar{\mu}_i(t),t)-\sqrt{\varepsilon_i}\,|u-u_i(t)|)=\bar{H}_i(\bar{\psi}_i(t),\bar{\mu}_i(t),t)\quad\text{for a.a. } t\in T, \tag{28}$$

$$\frac{\partial\bar{H}_i}{\partial u}(\bar{\psi}_i(t),\bar{\mu}_i(t),t)\in\operatorname{conv}N_U(u_i(t))+\sqrt{\varepsilon_i}\,B_{\mathbb{R}^m}\quad\text{for a.a. } t\in T. \tag{29}$$

Note that

$$(\lambda_i^0)^2+(\varrho_i)^2=1\quad\forall i. \tag{30}$$

Hence, by taking into account relations (21) and (24), by following the standard argument, and by using the Gronwall inequality, from (20) we find that the family of functions $\psi_i(t)$ is uniformly bounded. Since the $\psi_i(t)$ are uniformly bounded, it follows that the sequence of vectors a_i is bounded. Then the functions $\bar{\psi}_i(t)$ are uniformly bounded as well. In addition, by virtue of (24), $|\bar{\mu}_i(t)|\le\text{const}$ for almost all $t\in T$ and for all i, where const is independent of i and t. This implies the uniform boundedness of the derivatives $\dfrac{d\bar{\psi}_i}{dt}(t)$.

By using the above argument, by taking into account the compactness of the unit ball in the Euclidean space, the Arzelá–Ascoli theorem, and the weak sequential compactness of a unit ball in $\mathbb{L}_2$, and by passing to a subsequence, we obtain $\lambda_i^0\to\lambda^0$, $\varrho_i\to\varrho$, $\bar{\psi}_i(t)\rightrightarrows\psi(t)$, and $\bar{\mu}_i\to\mu$ weakly in $\mathbb{L}_2(T)$.

By passing to the limit in conditions (26)–(29) in a standard way (see [6, Subsec. 2.5]), we obtain relations (3), (5), (8), (12), and (11). Here, to derive the transversality conditions (12) from (27), one can use the properties of the subdifferential of the distance function (see [3, item 1.3.3, Ths. 1.97 and 1.105]), the upper semicontinuity of the limit normal cone, and formula (14). The inclusion (11) follows from (25).

It remains to note that λ^0, ψ, and μ cannot vanish simultaneously. Indeed, otherwise, by virtue of relation (30), we have

$$\varrho_i\to 1,\qquad \lambda_i^0\to 0,\qquad \bar{\psi}_i(t_1^*)\to 0,\qquad \bar{\psi}_i(t_2^*)\to 0.$$

Since $\varrho_i\to 1$, it follows that $p_i\notin\tilde{C}$ for large i. But, by Theorem 1.105 in [3], we have $|h|=1$ and $h\in\partial\operatorname{dist}(p,\tilde{C})$ for all $p\notin\tilde{C}$. Therefore, the transversality condition (27) with large i contradicts relation (13). Consequently, λ^0, ψ, and μ do not vanish simultaneously.

The maximum principle with conditions (3), (5), (8), (11), and (12) is thereby proved for the auxiliary problem (10) on a fixed time interval.

By using the approach suggested in [6, Subsec. 2.11] and by considering the so-called v-problem, we prove the existence of Lagrange multipliers λ^0, ψ, and μ that do not vanish simultaneously and satisfy conditions (3)–(8) and the condition

$$\frac{\partial G}{\partial x}(t_1^*)\psi(t_1^*)=h(t_1^*)\frac{\partial G}{\partial t}(t_1^*) \tag{31}$$

(for details, see [6]). In this case, the transversality conditions (4) and (7) are consequences of condition (12) and the regularity of the terminal conditions, and condition (31) follows from the inclusion (11).

It remains to prove the nontriviality condition (2). Suppose that condition (2) fails. Then $\lambda^0=0$, and there exists a point $\tau\in T$ and a vector $a\in\mathbb{R}^{d(G)}$, $a\ne 0$, such that

$$\psi(\tau)=a\frac{\partial G}{\partial x}(\tau).$$

By Remark 1, not only the Lagrange multipliers $\lambda=(0,\lambda_1,\lambda_2)$, ψ, and μ but also the triple λ, $\bar{\psi}(t)$, and $\bar{\mu}(t)$, where $\bar{\psi}(t)=\psi(t)-a\dfrac{\partial G}{\partial x}(t)$ and $\bar{\mu}(t)=\mu(t)-a$, satisfy conditions (3)–(8) and (12).

Obviously, $\bar{\psi}(\tau)=0$. On the other hand, it follows from the inclusion (8) and the regularity of the optimal trajectory $x^*(t)$ that $|\bar{\mu}(t)|\le\text{const}\times|\bar{\psi}(t)|$ for almost all $t\in T$, where the constant is independent of t. Hence, by using the Gronwall inequality and the fact that the function $\bar{\psi}$ is absolutely continuous and satisfies Eq. (3), we obtain the relations $\bar{\psi}(t) = 0\ \forall t \in T$, and $\bar{\mu}(t) = 0$ for almost all $t \in T$. Consequently, $\psi(t) \equiv a\dfrac{\partial G}{\partial x}(t)$ and $\mu(t) \equiv a$. Therefore, by the definition of Γ, from condition (5), we have $h(t) \equiv -\left\langle a, \dfrac{\partial G}{\partial t}(t)\right\rangle$. Thus,

$$\psi(t_1^*) = a\frac{\partial G}{\partial x}(t_1^*), \qquad h(t_1^*) = -\left\langle a, \frac{\partial G}{\partial t}(t_1^*)\right\rangle.$$

Now, by using condition (31), we obtain

$$\left(\frac{\partial G}{\partial x}(t_1^*)\frac{\partial G^*}{\partial x}(t_1^*) + \frac{\partial G}{\partial t}(t_1^*)\frac{\partial G^*}{\partial t}(t_1^*)\right) a = 0,$$

and consequently, $a = 0$, because the parenthesized expression is the sum of two symmetric positive definite matrices. If $a = 0$, then $\psi = 0$ and $\mu = 0$. Therefore, the Lagrange multipliers λ^0, ψ, and μ are zero simultaneously, which contradicts the above-proved assertions. Consequently, condition (2) is satisfied. The proof of the theorem is complete.

Note that, among all sets of Lagrange multipliers corresponding to a given optimal process, there does not necessarily exist a set for which μ is a function of bounded variation, and in general μ is a measurable essentially bounded function. As an example, we present the two-dimensional problem

$$\int_0^1 u_1 u_2\, dt \to \min, \qquad \dot{x} = u, \qquad u = (u_1, u_2) \in \mathbb{R}^2, \qquad x_1 = 0, \qquad x(0) = 0,$$

whose solution is given by the control $u(t) = (0, u_2(t))$, where $u_2(t) = \cos(1/t)$.

The problem with inequality state constraints was studied in [7]. A maximum principle in which the function $\bar{H}$ in the maximum condition is maximized on the entire set U rather than its subset $U(x^*(t), t)$ (like in Theorem 1), was obtained for it. In this connection, we encounter the problem as to whether it is possible to obtain a similar result for a problem with equality state constraints, i.e., replace the set $U(t)$ in the maximum condition (5) by a larger set U. However, the above-considered example gives the negative answer to this question. Indeed, by the maximum condition, $\max_{(u_1,u_2)\in\mathbb{R}^2}(\langle \psi, u\rangle - u_1\mu - \lambda^0 u_1 u_2) = 0$, which implies that $\lambda^0 = 0$ and $\psi = (\mu, 0)$. However, this contradicts the nontriviality condition (2) (and even the weaker nontriviality condition in [7]).

3. THE CASE OF UNBOUNDED SET U

Let us generalize the obtained maximum principle to the case in which the sets $U(t)$ are not bounded. This generalization is required to study a problem of variational calculus. In what follows, we additionally suppose that the set U can be represented in the form

$$U = \{u \in \mathbb{R}^m :\ q(u) \le 0\},$$

where $q:\ \mathbb{R}^m \to \mathbb{R}^{d(q)}$ is a given smooth mapping.

Set $I(u) = \{i :\ q^i(u) = 0\}$. We assume that the following regularity condition is satisfied.

Assumption R. The vectors $\dfrac{\partial q^i}{\partial u}(u)$ and $\dfrac{\partial \Gamma^j}{\partial u}(x^*(t), u, t)$ are linearly independent for $i \in I(u)$, $j = 1, \dots, d(G)$, for all $t \in T$, and for all $u \in U(x^*(t), t)$.

Obviously, this condition is stronger that the regularity of the trajectory $x^*(t)$ from Definition 1.

Theorem 2. *Let a process (p^*, x^*, u^*) be optimal in problem (1). In addition, let the mappings q and Γ be $(n+m+1)$ continuously differentiable, let Assumption R be satisfied, and let the terminal constraints be regular and be coordinated with the state constraints at the point p^*. Then the process (p^*, x^*, u^*) satisfies the maximum principle.*

Proof. Let $S \subseteq \{1, \dots, d(q)\}$ be some (possibly empty) set of indices. Consider the sets

$$M_S = \{(x, u, t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^1 : \ \Gamma(x, u, t) = 0, \ q^i(u) = 0 \ \forall i \in S, \ q^i(u) < 0 \ \forall i \notin S\}.$$

The set of various subsets of indices S is finite; therefore, the number of such sets M_S is also finite.

Take an arbitrary $c > \operatorname*{ess\,sup}_{t \in T} |u^*(t)| + 1$. By using Assumption R, we take a number $\varepsilon = \varepsilon(c) > 0$ such that each of the sets

$$M_S^* = M_S \cap \{(x, u, t) : \ |x - x^*(t)| < \varepsilon, \ |u| < c, \ t \in (t_1^* - \varepsilon, t_2^* + \varepsilon)\}$$

is a smooth manifold.

Fix an arbitrary S. By ξ_S we denote the restriction of the function $|u|^2$ to the manifold M_S^*. By virtue of the smoothness class of the functions q and Γ and by the Morse–Sard theorem [8, p. 95 of the Russian translation], there exists a number $\delta \in (0, 1)$ such that the number $c - \delta$ is a regular value of the function ξ_S for all index sets S. Hence it follows that, at all points (x, u, t) of the manifold M_S^* at which $|u|^2 = c - \delta$, the vector u does not belong to the linear span of the vectors $\dfrac{\partial q^i}{\partial u}(u)$ and $\dfrac{\partial \Gamma^j}{\partial u}(x, u, t)$, where $i \in S$ and $j = 1, \dots, d(G)$.

Therefore, in problem (1) with the additional constraints

$$|u|^2 \le c - \delta, \qquad |x - x^*(t)| < \varepsilon, \qquad |p - p^*| < \varepsilon,$$

the optimal trajectory $x^*(t)$ is regular in the sense of Definition 1. Taking into account the fact that the mapping $U(x, t)$ is already bounded in the new problem, we apply Theorem 1 to it. By passing in the resulting conditions of the maximum principle to the limit as $c \to \infty$ in a standard way and by using Assumption R, we complete the proof of the theorem.

From Theorem 2, we obtain the Euler–Lagrange equation and the Legendre condition for the problem of variational calculus on the smooth regular surface $M = \{x \in \mathbb{R}^n : \ G(x) = 0\}$. In particular, they imply the equation of geodesics on the surface M.

4. APPENDIX

On a fixed time interval $T = [t_1^*, t_2^*]$, consider the problem

$$\varphi(p) + \int_{t_1^*}^{t_2^*} f_0(x, u, t)\, dt \to \min, \qquad \dot{x} = f(x, u, t), \qquad t \in T, \qquad R(x, u, t) \in C. \tag{32}$$

It is a problem without terminal constraints but with so-called constrains of mixed type $R(x, u, t) \in C$, where C is a given closed subset of $\mathbb{R}^{d(R)}$. We assume that φ is a Lipschitz function.

Suppose that the process (p^*, x^*, u^*) is a solution of problem (32).

Consider the multimapping $U(x, t) := \{u \in \mathbb{R}^m : \ R(x, u, t) \in C\}$.

Definition 3. The mixed constraints in problem (32) are said to be regular if

$$N_C(R(u, t)) \cap \ker \frac{\partial R^*}{\partial u}(u, t) = \{0\}$$

for all $u \in U(t)$ and t.

The regularity of mixed constraints implies that the Robinson condition [3, p. 418] is satisfied at each admissible point $(x^*(t), u, t)$.

The simplest example of regular mixed constraints is given by geometric constraints $u \in C$ or, more generally, by the case in which the matrix $\dfrac{\partial R}{\partial u}(x, u, t)$ has full rank for all $u \in U(x, t)$ and for all x and t.

Theorem 3 (the maximum principle for problem without terminal constraints). *Let* (p^*, x^*, u^*) *be the optimal process in problem* (32). *In addition, let the mixed constraints be regular, and let the sets* $U(t)$ *be bounded uniformly with respect to* $t \in T$. *Then there exist a number* $\lambda^0 > 0$, *an absolutely continuous function* $\psi : T \to \mathbb{R}^n$, *and a measurable essentially bounded vector function* $r : T \to \mathbb{R}^{d(R)}$ *such that*

$$r(t) \in \operatorname{conv} N_C(R(t)) \quad \textit{for a.a. } t, \tag{33}$$

$$\dot{\psi}(t) = -\frac{\partial H}{\partial x}(t) + r(t)\frac{\partial R}{\partial x}(t) \quad \textit{for a.a. } t, \tag{34}$$

$$(\psi(t_1^*), -\psi(t_2^*)) \in \lambda^0 \partial\varphi(p^*), \tag{35}$$

$$\max_{u \in U(t)} H(u,t) = H(t) \quad \textit{for a.a. } t, \tag{36}$$

$$\frac{\partial H}{\partial u}(t) - r(t)\frac{\partial R}{\partial u}(t) = 0 \quad \textit{for a.a. } t. \tag{37}$$

In addition, there exists a constant $\varkappa > 0$ *such that*

$$|r(t)| \le \varkappa(\lambda^0 + |\psi(t)|) \quad \text{for a.a. } t. \tag{38}$$

Proof. Without loss of generality, we assume that $f_0 = 0$ and $\varphi(p^*) = 0$. Take positive numbers c and δ such that the inequality $|u| \le c$ holds for arbitrary $u \in U(x,t)$ and x satisfying the condition $|x - x^*(t)| \le \delta$ $\forall t \in T$. This is possible by virtue of the uniform boundedness of the sets $U(t)$ and the continuity of the mapping R.

By $\mathcal{M} \subseteq \mathbb{R}^{2n} \times \mathbb{L}_1^m(T)$ we denote the set of pairs $(p, u(\cdot))$, $p = (x_1, x_2)$, such that $|u(t)| \le c + 1$ and $|x(t) - x^*(t)| \le \delta$ for almost all $t \in T$, and $x(t_2^*) = x_2$, where $x(\cdot)$ is the trajectory corresponding to the control $u(\cdot)$ and the initial condition $x(t_1^*) = x_1$. The set $\mathcal{M}$ is nonempty because it contains $(p^*, u^*(\cdot))$. Moreover, $\mathcal{M}$ is closed and hence is a complete metric space with the metric induced by the norm $|p| + \|u\|_{L_1}$.

Take a positive integer $i > 1$ and set $\varepsilon_i = i^{-1}$ and $\varphi_i(p) = (\varphi(p) + \varepsilon_i)^+$. For nonnegative numbers $\alpha \ge 0$ and $\beta \ge 0$, set

$$\Delta(\alpha,\beta) = \begin{cases} \alpha\beta^{-2} & \text{for} \quad \beta > 0, \\ 1 & \text{for} \quad \alpha > 0,\ \beta = 0, \\ 0 & \text{for} \quad \alpha = \beta = 0. \end{cases}$$

Note that the function Δ is lower semicontinuous on the set $\alpha \ge 0$, $\beta \ge 0$. On the space $\mathcal{M}$, consider the functional

$$F_i(p, u(\cdot)) = \varphi_i(p) + \Delta\left(\int_T (\operatorname{dist}(R(x,u,t), C))^2 dt, \varphi_i(p)\right).$$

By construction, the functional F_i is continuous and positive on $\mathcal{M}$. [Both of its terms do not vanish simultaneously, because $\varphi(p^*) = 0$.]

Consider the problem

$$F_i(p, u(\cdot)) \to \min, \qquad (p, u(\cdot)) \in \mathcal{M}.$$

To this problem, we apply the Ekeland variational principle [5]. One can readily see that

$$F_i(p^*, u^*(\cdot)) = \varepsilon_i.$$

Therefore, there exists an element $(p_i, u_i(\cdot)) \in \mathcal{M}$, $p_i = (x_{1,i}, x_{2,i})$, satisfying conditions (17) and (18), and the pair $(p_i, u_i(\cdot))$ is a solution of the problem

$$F_i(p, u(\cdot)) + \sqrt{\varepsilon_i}\left(|p - p_i| + \int_{t_1^*}^{t_2^*} |u - u_i(t)|\, dt\right) \to \min, \qquad (p, u(\cdot)) \in \mathcal{M}.$$

Note that if $\varphi_i(p_i) = 0$, then the mixed constraints fail, and by the definition of Δ, we have $F_i(p_i, u_i(\cdot)) = 1$, which contradicts condition (17). Therefore, $\varphi_i(p_i) > 0$, and hence the pair $(p_i, u_i(\cdot))$ is a solution of the problem

$$\begin{gathered}\varphi_i(p) + \int\limits_{t_1^*}^{t_2^*} y^{-2}(\operatorname{dist}(R(x,u,t),C))^2\,dt + \sqrt{\varepsilon_i}\left(|p-p_i| + \int\limits_{t_1^*}^{t_2^*} |u-u_i(t)|\,dt\right) \to \min,\\ \dot{x} = f(x,u,t), \qquad \dot{y} = 0, \qquad y(t_1^*) = \varphi_i(p), \qquad G(x_1,t_1^*) = 0,\\ |u(t)| \le c+1 \quad \text{for a.a. } t, \quad |x(t)-x^*(t)| \le \delta, \quad y(t) > 0 \quad \forall t \in T.\end{gathered} \tag{39}$$

The optimal trajectory in this problem is denoted by x_i, y_i.

It follows from inequality (18) that $u_i(t) \to u^*(t)$ for almost all t after the passage to a subsequence $p_i \to p^*$, and then $x_i(t) \rightrightarrows x^*(t)$ uniformly on T. Therefore, in the derivation of necessary optimality conditions, the state constraints in problem (39) for large i can be omitted. Therefore, problem (39) contains only terminal constraints at the left endpoint.

To this problem, we apply the maximum principle in [3, Th.[1] 6.27]. For each i, there exists a number $\lambda_i^0 > 0$, absolutely continuous functions $\psi_i(t)$ and $\sigma_i(t)$, and a measurable function $\eta_i(t)$ ranging in the set $\operatorname{conv}\partial\operatorname{dist}(R_i(t),C)$ for almost all t such that

$$\begin{gathered}\dot{\psi}_i(t) = -\frac{\partial H_i}{\partial x}(t) + 2\lambda_i^0\,\frac{\operatorname{dist}(R_i(t),C)}{y_i^2}\,\eta_i(t)\frac{\partial R_i}{\partial x}(t) \quad \text{for a.a. } t \in T,\\ \sigma_i(t) = \sigma_i(t_1^*) - 2\lambda_i^0 \int\limits_{t_1^*}^{t} \frac{[\operatorname{dist}(R_i(\varsigma),C)]^2}{y_i^3}\,d\varsigma, \qquad \sigma_i(t_2^*) = 0,\end{gathered} \tag{40}$$

$$(\psi_i(t_1^*), -\psi_i(t_2^*)) \in (\lambda_i^0 - \sigma_i(t_1^*))\partial\varphi_i(p_i) + \lambda_i^0\sqrt{\varepsilon_i}B_{\mathbb{R}^{2n}}, \tag{41}$$

$$\begin{aligned}&\max_{u\colon |u|\le c+1}(H_i(u,t) - \lambda_i^0 y_i^{-2}(\operatorname{dist}(R_i(u,t),C))^2 - \lambda_i^0\sqrt{\varepsilon_i}|u-u_i(t)|)\\ &\qquad = H_i(t) - \lambda_i^0 y_i^{-2}(\operatorname{dist}(R_i(t),C))^2 \quad \text{for a.a. } t \in T,\end{aligned} \tag{42}$$

$$\begin{aligned}&\frac{\partial H_i}{\partial u}(t) \in 2\lambda_i^0\frac{\operatorname{dist}(R_i(t),C)}{y_i^2}\,\eta_i(t)\frac{\partial R_i}{\partial u}(t)\\ &\qquad + \lambda_i^0\sqrt{\varepsilon_i}B_{\mathbb{R}^m} \quad \text{for a.a. } t \in T, \quad \text{such that} \quad |u_i(t)| < c+1,\end{aligned} \tag{43}$$

$$\lambda_i^0 + \max_{t\in T}|\psi_i(t)| = 1. \tag{44}$$

Set $r_i(t) = 2\lambda_i^0 y_i^{-2}\operatorname{dist}(R_i(t),C)\eta_i(t)$. Then conditions (40) and (43) acquire the form

$$\dot{\psi}_i(t) = -\frac{\partial H_i}{\partial x}(t) + r_i(t)\frac{\partial R_i}{\partial x}(t) \quad \text{for a.a. } t \in T, \tag{45}$$

$$\frac{\partial H_i}{\partial u}(t) \in r_i(t)\frac{\partial R_i}{\partial u}(t) + \lambda_i^0\sqrt{\varepsilon_i}B_{\mathbb{R}^m} \quad \text{for a.a. } t \in T \quad \text{such that} \quad |u_i(t)| < c+1. \tag{46}$$

Suppose firstly that $\lambda_i^0 y_i^{-2} \to \infty$ as $i \to \infty$.

Let us show that there exists a constant $\varkappa > 0$ such that

$$|r_i(t)| \le \varkappa(\lambda_i^0 + |\psi_i(t)|) \quad \text{for a.a. } t \in T \quad \forall i. \tag{47}$$

The regularity of the mixed constraints and Theorem 4.37 in [3] imply the following property. There exists a $\gamma > 0$ such that, for any $t \in T$ and for each sequence of vectors $\xi_i \to \bar{x}$, where

[1] Theorem 6.27 was proved under quite general assumptions and for differential inclusions. The below-represented conditions are derived from that theorem in a standard way. The existence of a measurable function η_i follows from the theorem on a measurable selector of a measurable multimapping.

$\xi_i, \bar{x} \in x^*(t) + \gamma B_{\mathbb{R}^n}$, and each vector $\bar{u} \in U(\bar{x}, t)$, there exists a vector sequence $v_i \in U(\xi_i, t)$, $v_i \to \bar{u}$.

Since $x_i(t) \rightrightarrows x^*(t)$, it follows from the above-mentioned property that, starting from some sufficiently large index i, for each $t \in T$, there exists a vector $\tilde{u}_i(t)$ such that $R(x_i(t), \tilde{u}_i(t), t) \in C$.

By substituting the value $u = \tilde{u}_i(t)$ into the left-hand side of the maximum condition (42) for almost all t and by taking into account relation (44), we obtain

$$\lambda_i^0 y_i^{-2} (\mathrm{dist}(R_i(t), C))^2 \leq \mathrm{const} \times (\lambda_i^0 + |\psi_i(t)|) \leq \mathrm{const}.$$

Since $\lambda_i^0 y_i^{-2} \to \infty$, it follows that $\mathrm{dist}(R_i(t), C) \to 0$ for almost all $t \in T$ uniformly with respect to $t \in T$. Therefore, $|u_i(t)| < c + 1$ for almost all $t \in T$ by virtue of the choice of the constant c for all large i.

Using this fact, we apply the Lagrange principle (see Theorem 5.5 in [3]) to the nonsmooth maximization problem

$$H_i(u, t) - \lambda_i^0 y_i^{-2} (\mathrm{dist}(R_i(u, t), C))^2 - \lambda_i^0 \sqrt{\varepsilon_i}\, |u - u_i(t)| \to \max, \qquad |u| \leq c + 1,$$

for a given $t \in T$. By virtue of the maximum condition (42), the point $u = u_i(t)$ is a solution of this problem; therefore,

$$\frac{\partial H_i}{\partial u}(t) \in 2\lambda_i^0 y_i^{-2}\, \mathrm{dist}(R_i(t), C) \partial\, \mathrm{dist}(R_i(t), C) \frac{\partial R_i}{\partial u}(t) + \lambda_i^0 \sqrt{\varepsilon_i}\, B_{\mathbb{R}^m}.$$

This, together with the theorem on the existence of a measurable selector of a measurable multimapping, implies in a standard way that

$$\frac{\partial H_i}{\partial u}(t) \in \omega_i(t) \frac{\partial R_i}{\partial u}(t) + \lambda_i^0 \sqrt{\varepsilon_i}\, B_{\mathbb{R}^m} \quad \text{for a.a. } t \in T, \tag{48}$$

where $\omega_i(t) = 2\lambda_i^0 y_i^{-2}\, \mathrm{dist}(R_i(t), C) n_i(t)$, and $n_i(t)$ is some measurable mapping such that $n_i(t) \in \partial\, \mathrm{dist}(R_i(t), C)$ for almost all $t \in T$. From the properties of the subdifferential of the distance function, we have $|n_i(t)| \leq 1$ for almost all $t \in T$; moreover, by Theorem 1.105 in [3], $|n_i(t)| = 1$ for almost all t such that $R_i(t) \notin C$.

It follows from the definition of the function $\omega_i(t)$ that $\omega_i(t) = r_i(t) = 0$ for almost all t such that $R_i(t) \in C$. In addition, it is known that $|\eta_i(t)| \leq 1$ for almost all $t \in T$. This, together with the above-performed considerations, implies that $|r_i(t)| \leq |\omega_i(t)|$ for almost all $t \in T$.

It follows from the regularity of mixed constraints, the upper semicontinuity of the normal cone $N_C(y)$ with respect to y, and the compactness that there exist positive numbers ε and θ such that

$$\left| \frac{\partial R^*}{\partial u}(x, u, t) y \right| \geq \varepsilon |y| \quad \forall y \in N_C(\xi) \quad \forall\, \xi \in \Pi_C(R(x, u, t)), \tag{49}$$
$$\text{and} \quad \forall (x, u, t) \quad \text{such that} \quad |x - x^*(t)| \leq \delta, \quad |u| \leq c + 1, \quad \mathrm{dist}(R(x, u, t), C) \leq \theta.$$

The estimate (49) can readily be proved by contradiction.

By Theorem 1.105 in [3], we have

$$n_i(t) \in \bigcup_{y \in \Pi_C(R_i(t))} N_C(y) \quad \text{for a.a. } t \in T \quad \text{such that} \quad R_i(t) \notin C. \tag{50}$$

By taking into account the definition of ω_i and the fact that $\mathrm{dist}(R_i(t), C) \to 0$ for almost all $t \in T$ uniformly with respect to t and by using the inequality $|r_i(t)| \leq |\omega_i(t)|$ for almost all $t \in T$, one can derive the estimate (47) and the existence of the desired number $\varkappa$ directly from conditions (48)–(50).

It follows from (47) and (44) that the functions $r_i(t)$ are uniformly bounded. Let us show that $\sigma_i(t_1^*) \to 0$. To this end, by virtue of (40), it suffices to show that the function sequence

$\lambda_i^0 y_i^{-3}(\mathrm{dist}(R_i(t),C))^2$ converges to zero in the norm of the space $\mathbb{L}_1(T)$. It follows from the definition of $r_i(t)$ that

$$\int\limits_{t_1^*}^{t_2^*} 2\lambda_i^0 y_i^{-3}(\mathrm{dist}(R_i(t),C))^2\,dt = \int\limits_{t_1^*}^{t_2^*} y_i^{-1}\,\mathrm{dist}(R_i(t),C)|r_i(t)|\,dt.$$

Therefore, since $\|r_i\|_{L_\infty} \le \mathrm{const}$, it suffices to show that $\mathrm{dist}(R_i(t),C)y_i^{-1} \to 0$ in the norm of $\mathbb{L}_1(T)$. However, this sequence is convergent even in $\mathbb{L}_2(T)$ by virtue of condition (17). Thus, $\sigma_i(t_1^*) \to 0$.

Relations (40) and (44), together with the estimate $\|r_i\|_{L_\infty} \le \mathrm{const}$, imply that the function sequence $\{\psi_i\}$ is equicontinuous and uniformly bounded. Then, by using the Arzelá–Ascoli theorem and the weak sequential compactness of the unit ball in $\mathbb{L}_2$ and by passing to a subsequence, we find that there exist λ^0, ψ, and r such that $\lambda_i^0 \to \lambda^0$, $\psi_i \rightrightarrows \psi$, and $r_i \to r$ weakly in $\mathbb{L}_2^{d(R)}(T)$.

The inclusion (33) follows from the weak convergence, the definition of r_i, the upper semicontinuity of the limit subdifferential, and Theorem 1.97 in [3].

By passing in relation (45) to the limit as $i \to \infty$, we obtain condition (34).

By passing in (41) to the limit and by taking into account the convergence $\sigma_i(t_1^*) \to 0$, we obtain condition (35).

Let us prove condition (36). It was mentioned above that the function sequence $\mathrm{dist}(R_i(t),C)y_i^{-1}$ converges to zero in the norm of $\mathbb{L}_1(T)$. By passing to a subsequence, we find that

$$\mathrm{dist}(R_i(t),C)y_i^{-1} \to 0$$

for almost all $t \in T$. Take a point t at which the above-mentioned convergence takes place, the maximum condition (42) is satisfied for all i, and $u_i(t) \to u^*(t)$. Take $\bar u \in U(t)$. Consider a sequence $v_i \in U_i(t)$ such that $v_i \to \bar u$. By substituting the value $u = v_i$ into the maximum condition (42) and by passing to the limit, we obtain the inequality $H(\bar u,t) \le H(t)$, which, in view of the arbitrary choice of the vector $\bar u$, holds for all $\bar u \in U(t)$. This completes the proof of relation (36).

Condition (37) is obtained by the integration of the inclusion (46) over the closed interval $[t_1^*,t]$ and the subsequent passage to the limit as $i \to \infty$ (see [6, Subsec. 2.5]). The estimate (38) follows from inequality (47).

Consider the second case. Assume that the numerical sequence $\{\lambda_i^0 y_i^{-2}\}$ does not tend to infinity. In this case, since the function sequence $\mathrm{dist}(R_i(t),C)$ converges to zero in $\mathbb{L}_1$, by passing to a subsequence and by taking into account the definition of the function $r_i(t)$, we find that the sequence $r_i(t)$ is uniformly bounded on T and $r_i(t) \to 0$ for almost all $t \in T$. Since $\|r_i\|_{L_\infty} \le \mathrm{const}$, just as above, we show that $\sigma_i(t_1^*) \to 0$. By passing to the limit as $i \to \infty$ just as above, we obtain all conditions of the maximum principle. In this case, the estimate (47) holds for each $\varkappa > 0$ because $r = 0$.

Let us show that $\lambda^0 > 0$. Indeed, if $\lambda^0 = 0$, then, by (38), (34) and by taking into account the relation $\psi(t_1^*) = 0$ valid by virtue of (35) and by using the Gronwall inequality, we obtain the relation $\psi = 0$. However, this contradicts the fact that λ^0 and ψ do not vanish simultaneously. Therefore, $\lambda^0 > 0$. The proof of the theorem is complete.

Remark 3. It follows from the proof of Theorem 3 that the constant $\varkappa$ in the estimate (38) can be chosen to depend only on R, C, c, and δ. This follows from the fact that the number ε in condition (49) depends only on R, C, c, and δ.

Remark 4. Theorem 3 can be generalized to the case in which the function f_0 only satisfies the Lipschitz condition with respect to u. Then condition (37) acquires the form

$$0 \in \mathrm{conv}\,\partial_u H(t) - r(t)\frac{\partial R}{\partial u}(t) \quad \text{for a.a. } t.$$

Lemma 1. *Let the function R and the set C be defined in the form $R(x,u,t) = (\Gamma(x,u,t),u)$ and $C = \{0\} \times U$, where Γ and U are introduced in Section* 1. *Then the notions of regularity of mixed constraints and regularity of the trajectory $x^*(t)$ in Definitions* 1 *and* 3 *coincide.*

Proof. The proof is by a straightforward verification of the fact that the relations

$$\ker \frac{\partial R^*}{\partial u}(u,t) = \left\{ \xi = (\xi_1, \xi_2) \in \mathbb{R}^{d(G)} \times \mathbb{R}^m : \ \xi_1 \frac{\partial \Gamma}{\partial u}(u,t) = \xi_2 \right\},$$
$$N_C(y) = \{ \xi = (\xi_1, \xi_2) \in \mathbb{R}^{d(G)} \times \mathbb{R}^m : \ \xi_2 \in N_U(u) \}$$

hold for $y = (0, u)$.

ACKNOWLEDGMENTS

This research was supported by the Russian Foundation for Basic Research (projects nos. 13-01-00494 and 14-01-31185) and by the Ministry of Education and Science of the Russian Federation (project no. 1.333.2014/K).

REFERENCES

1. Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., and Mishchenko, E.F., *Matematicheskaya teoriya optimal'nykh protsessov* (The Mathematical Theory of Optimal Processes), Moscow: Nauka, 1983.
2. Gamkrelidze, R.V., Optimum-Rate Processes with Bounded State Coordinates, *Dokl. Akad. Nauk SSSR*, 1959, vol. 125, no. 3, pp. 475–478.
3. Mordukhovich, B.S., *Variational Analysis and Generalized Differentiation*, vols. 1, 2, New York, 2005.
4. Mordukhovich, B.S., Nguyen Mau Nam, and Bingwu Wang, Metric Regularity of Mappings and Generalized Normals to Set Images, *Set-Valued Anal.*, 2009, vol. 17, no. 4, pp. 359–387.
5. Ekeland, I., On the Variational Principle, *J. Math. Anal. Appl.*, 1974, vol. 47, pp. 324–353.
6. Arutyunov, A.V., *Usloviya ekstremuma. Anormal'nye i vyrozhdennye zadachi* (Extremum Conditions. Abnormal and Degenerate Problems), Moscow, 1997.
7. Arutyunov, A.V., Karamzin, D.Yu., and Pereira, F.L., The Maximum Principle for Optimal Control Problems with State Constraints by R.V. Gamkrelidze, *Revisited J. Optim. Theory Appl.*, 2011, vol. 149, pp. 474–493.
8. Hirsch, M.W., *Differential Topology*, New York–Heidelberg: Springer-Verlag, 1976. Translated under the title *Differentsial'naya topologiya*, Moscow: Mir, 1979.