ISSN 0081-5438, Proceedings of the Steklov Institute of Mathematics, 2021, Vol. 315, pp. 27–55. © Pleiades Publishing, Ltd., 2021.
Russian Text © The Author(s), 2021, published in Trudy Matematicheskogo Instituta imeni V.A. Steklova, 2021, Vol. 315, pp. 34–63.

Refined Euler–Lagrange Inclusion for an Optimal Control Problem with Discontinuous Integrand

S. M. Aseev [a,b]

Received August 30, 2021; revised September 26, 2021; accepted October 1, 2021

Abstract—We study a free-time optimal control problem for a differential inclusion with mixed-type functional in which the integral term contains the characteristic function of a given open set of "undesirable" states of the system. The statement of this problem can be viewed as a weakening of the statement of the classical optimal control problem with state constraints. Using the approximation method, we obtain first-order necessary optimality conditions in the form of the refined Euler–Lagrange inclusion. We also present sufficient conditions for their nondegeneracy and pointwise nontriviality and give an illustrative example.

DOI: 10.1134/S0081543821050047

1. STATEMENT OF THE PROBLEM AND AUXILIARY RESULTS

We study the following optimal control problem (P):

$$J(T,x(\cdot)) = \varphi(T,x(0),x(T)) + \int_0^T \lambda(x(t))\,\delta_M(x(t))\,dt \to \min, \tag{1.1}$$

$$\dot{x}(t) \in F(x(t)), \tag{1.2}$$

$$x(0) \in M_0, \qquad x(T) \in M_1. \tag{1.3}$$

Here $x(t) \in \mathbb{R}^n$ is the state vector at time $t \geq 0$; $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a set-valued mapping with nonempty convex compact values that is locally Lipschitz continuous in the sense of the Hausdorff metric (hereinafter, the local Lipschitz continuity of mappings/functions is understood as their Lipschitz continuity on any bounded subset of their domain of definition); M_0 and M_1 are nonempty closed sets in $\mathbb{R}^n$; $\varphi\colon [0,\infty) \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^1$ is a locally Lipschitz continuous function; $\lambda\colon \mathbb{R}^n \to (0,\infty)$ is a continuously differentiable function ($C^1(\mathbb{R}^n)$); and $\delta_M(\cdot)$ is the characteristic function of a given open set M in $\mathbb{R}^n$, i.e.,

$$\delta_M(x) = \begin{cases} 1, & x \in M, \\ 0, & x \notin M. \end{cases} \tag{1.4}$$

Concerning the set M and its complement $G = \mathbb{R}^n \setminus M$, we assume that each of these sets is nonempty and for any $x \in G$ the Clarke tangent cone $T_G(x)$ (see [15]) has a nonempty interior ($\operatorname{int} T_G(x) \neq \varnothing$). The terminal time $T > 0$ of the control process is assumed to be free.

By trajectories of the system, we will mean all absolutely continuous solutions $x\colon [0,T] \to \mathbb{R}^n$ of the differential inclusion (1.2) that are defined on arbitrary time intervals $[0,T]$, $T > 0$. A trajectory $x(\cdot)$ defined on $[0,T]$ will be said to be admissible in problem (P) if it satisfies the boundary conditions (1.3). An admissible trajectory $x_*(\cdot)$ defined on $[0,T_*]$, $T_* > 0$, is optimal in problem (P)

[a] Steklov Mathematical Institute of Russian Academy of Sciences, ul. Gubkina 8, Moscow, 119991 Russia.
[b] Lomonosov Moscow State University, Moscow, 119991 Russia.
E-mail address: aseev@mi-ras.ru

if the value of the functional (1.1) on $(T_*, x_*(\cdot))$ is equal to its minimum value on the set of such pairs.

Optimal control problems for differential inclusions have been studied by many authors (see [15, 16, 21, 24, 30, 31]). Interest in these problems stems, on the one hand, from the fact that specifying a differential relation as a differential inclusion allows one to uniformly cover a large number of extremal problems for various dynamic systems, including feedback systems, control systems with regular mixed constraints, as well as systems defined by a family of differential equalities and inequalities. On the other hand, an optimal control problem for a differential inclusion arises naturally in relation to the question of whether the necessary optimality conditions obtained are invariant with respect to the method of specifying the differential relation.

The main difference between problem (P) and various "standard" optimal control problems (see [15, 19, 25, 31]) is that the integral term in the functional (1.1) contains a discontinuous function $\delta_M(\cdot)$ of the state variable. This integral term penalizes the appearance of the trajectories of the system in a given set M of "undesirable" states of the system. Such undesirable sets M ("risk zones") can appear in many applied problems. For example, the set M may correspond to the overload or unstable operation of a technical system. The positive function $\lambda(\cdot)$ in the functional (1.1) determines which states $x \in M$ are still more preferable. In optimal control theory, one usually models the presence of a set M of undesirable states by a state constraint (see [16; 25, Ch. 6; 31]) of the form

$$x(t) \in G = \mathbb{R}^n \setminus M, \qquad t \in [0, T].$$

In this case, the set G ("safety zone") is assumed to be closed (i.e., the risk zone M is an open set). Note that in contrast to problems with state constraints, an admissible trajectory $x(\cdot)$ in problem (P) may penetrate the set M (sometimes this is even inevitable), but such a situation is undesirable. Optimal control problems with state constraints can be regarded as the limit case of problem (P) with a constant function $\lambda(x) \equiv \lambda > 0$, $x \in \mathbb{R}^n$, as $\lambda \to \infty$. In this sense the statement of problem (P) weakens the statement of the classical optimal control problem with state constraints.

Optimal control problems involving a closed set M of undesirable system states have been previously studied in [11, 12, 26, 27, 29]. In [26, 27] the case of a linear control system and a convex compact set $M \subset \mathbb{R}^n$ was considered, and it was assumed that an optimal trajectory $x_*(\cdot)$ intersects the boundary of M at most at finitely many time points and only in a regular way (see [26, 27] for details). In the papers [11, 12, 29] based on an approximation method, a more general case was studied where the control system is affine in control and the set $M \subset \mathbb{R}^n$ is nonempty and closed, without any a priori assumptions about the behavior of the optimal trajectory $x_*(\cdot)$. Note that the constructions used in [11, 12, 29] cannot be directly applied to the case of an open set M. At the same time, the case of an open set M is of greatest interest. In this case, there is a natural connection between problem (P) and the classical optimal control problem with state constraints (see [10] for details). Moreover, in this case the integral term in (1.1), which contains the characteristic function $\delta_M(\cdot)$, is lower semicontinuous, which entails (under natural additional assumptions) the existence of a solution in problem (P) (see [8, Theorem 1]).

The case of an open set M was considered in [7–9], where necessary optimality conditions for problem (P) were obtained in Clarke's Hamiltonian form using an approximation method. The aim of this paper is to further develop the approximation method and use it to obtain necessary optimality conditions for problem (P) in the form of a refined Euler–Lagrange inclusion; in this form, the specific features of the differential relation given by the differential inclusion (1.2) are taken into account in the most complete way. As is well known, the refined Euler–Lagrange inclusion implies Clarke's Hamiltonian inclusion [15, 16], the Euler–Lagrange inclusion [14, 24], and the maximum condition (for a detailed discussion of various variants of necessary optimality conditions for optimal control problems with differential inclusions, see the review [22]). The results obtained here strengthen the necessary optimality conditions established earlier in [7–9].

The main difficulties in obtaining necessary optimality conditions for problem (P) are associated with the fact that the integrand in (1.1) is discontinuous in the state variable, which does not allow us to directly use any infinitesimal methods; moreover, the right-hand side of (1.2) is set-valued and nonsmooth, which also complicates the analysis. To overcome these difficulties, we apply the approximation method (see [5]). The original optimal control problem (P) with discontinuous integrand is approximated by a sequence of optimal control problems for smooth control systems with smooth integrands and nonsmooth terminal constraints, for which necessary optimality conditions are available (see [20, 21]). For the original problem (P), we then obtain necessary optimality conditions by passing to the limit in the relations of the Pontryagin maximum principle for the approximating problems. Previously, such an approach to deriving necessary optimality conditions in optimal control problems for differential inclusions with state constraints was used in [6] (see also [5]).

In what follows, by $N_A(a) = T_A^*(a)$ and $\widehat{N}_A(a)$ we denote, respectively, the Clarke normal cone [15] and the cone of generalized normals [21, 22] to a closed set $A \subset \mathbb{R}^n$ at a point $a \in A$; ∂A stands for the boundary of A. Next, by $H(A,\psi) = \sup_{a\in A}\langle a,\psi\rangle$ we denote the support function of a closed set $A \subset \mathbb{R}^n$. Then the function $H(F(\cdot),\cdot)\colon\ H(F(x),\psi) = \max_{f\in F(x)}\langle f,\psi\rangle$, $x \in \mathbb{R}^n$, $\psi \in \mathbb{R}^n$, is the Hamiltonian of the differential inclusion (1.2). By $\operatorname{graph} F(\cdot) = \{(x,y)\colon y \in F(x)\}$ we denote the graph of the set-valued mapping $F(\cdot)$, and by $\widehat{\partial}\varphi(T,x_1,x_2)$, the generalized gradient of the locally Lipschitz continuous function $\varphi(\cdot,\cdot,\cdot)$ at a point $(T,x_1,x_2) \in [0,\infty) \times \mathbb{R}^n \times \mathbb{R}^n$ (see [21, 22]). Let $A + B = \{a + b\colon a \in A,\, b \in B\}$ and $\alpha A = \{\alpha a\colon a \in A\}$ be the algebraic sum of sets A and B in $\mathbb{R}^n$ and the product of a number $\alpha \in \mathbb{R}^1$ and a set $A \in \mathbb{R}^n$, respectively.

For an arbitrary $x \in \mathbb{R}^n$ and any $i \in \mathbb{N} = \{1,2,\ldots\}$, we set

$$\widetilde{\delta}_i(x) = \min\{i\rho(x,G),\delta_M(x)\},$$

where $\rho(x,G) = \min\{\|x-\xi\|\colon \xi \in G\}$ is the distance from a point x to the nonempty closed set $G = \mathbb{R}^n \setminus M$, and the function $\delta_M(\cdot)$ is defined in (1.4).

For $i \in \mathbb{N}$ introduce a function $\delta_i\colon \mathbb{R}^n \to \mathbb{R}^1$ as

$$\delta_i(x) = \int\limits_{\mathbb{R}^n} \widetilde{\delta}_i(x+z)\omega_i(z)\,dz, \tag{1.5}$$

where $\omega_i(\cdot)$ is a smooth ($C^\infty(\mathbb{R}^n)$) centrally symmetric probability density with $\operatorname{supp}\omega_i(\cdot) \subset 3^{-i}\mathbb{B}^n$. Here $\mathbb{B}^n$ is the closed unit ball in $\mathbb{R}^n$ centered at zero. Then, for any i, the function $\delta_i(\cdot)$ is smooth, since it is the convolution with the smooth function $\omega_i(\cdot)$.

The following two results are direct consequences of the definition of the characteristic function $\delta_M(\cdot)$, the continuity of the positive function $\lambda(\cdot)$, and the Fatou lemma (see the proofs of similar Lemmas 1 and 2 in [7]).

Lemma 1. *For any $x \in \mathbb{R}^n$ one has the inequality*

$$\delta_i(x) \le \delta_M(x) + \frac{i}{3^i}, \qquad i \in \mathbb{N}.$$

Lemma 2. *If a sequence $\{x_i(\cdot)\}_{i=1}^\infty$ of continuous functions $x_i\colon [0,T] \to \mathbb{R}^n$, $T > 0$, converges uniformly to a continuous function $\overline{x}\colon [0,T] \to \mathbb{R}^n$, then*

$$\liminf_{i\to\infty} \int\limits_0^T \lambda(x_i(t))\,\delta_i(x_i(t))\,dt \ge \int\limits_0^T \lambda(\overline{x}(t))\,\delta_M(\overline{x}(t))\,dt.$$

Lemmas 1 and 2 imply (see [8, Theorem 1]) that the integral functional $J_M\colon C([0,T],\mathbb{R}^n)\to\mathbb{R}^1$, $T>0$, defined by the equality

$$J_M(x(\cdot))=\int\limits_0^T\lambda(x(t))\,\delta_M(x(t))\,dt \tag{1.6}$$

is lower semicontinuous.

The following result allows us to construct a sequence of smooth control systems that approximate the differential inclusion (1.2) (see [5, 6]).

Lemma 3. *Let $Z\subset\mathbb{R}^n$ be an open bounded set, L the Lipschitz constant of the set-valued mapping $F(\cdot)$ on $Z+\mathbb{B}^n$, and $S=\sup_{\{\xi\in F(x),\,x\in Z+\mathbb{B}^n\}}\|\xi\|$. Define a sequence of functions $\{H_i(\cdot,\cdot)\}_{i=1}^{\infty}$, $H_i\colon Z\times\mathbb{R}^n\to\mathbb{R}^1$, by the equality*

$$H_i(x,\psi)=\int\limits_{\mathbb{R}^{2n}}H(F(x+z),\psi+\|\psi\|v)\omega_i(z)\omega_i(v)\,dz\,dv+\frac{2(L+S)}{3^i}\|\psi\|, \tag{1.7}$$

where $\omega_i(\cdot)$ is a centrally symmetric smooth probability density with $\operatorname{supp}\omega_i(\cdot)\subset 3^{-i}\mathbb{B}^n$. Then for any i the function $H_i(\cdot,\cdot)$ is infinitely differentiable on $Z\times(\mathbb{R}^n\setminus\{0\})$. Moreover, for any $x\in Z$ the function $\psi\mapsto H_i(x,\psi)$ is positively homogeneous and subadditive on $\mathbb{R}^n$.

Lemma 3 follows directly from the properties of the convolution and [28, Proposition 2.5]. Indeed, being the convolution with a smooth function, the function $H_i(\cdot,\cdot)$ defined in (1.7) is infinitely differentiable on $Z\times(\mathbb{R}^n\setminus\{0\})$. By [28, Proposition 2.5], for any $x\in Z$, $z\in\mathbb{B}^n$ and an arbitrary $v\in\mathbb{R}^n$, the function

$$\psi\mapsto H(F(x+z),\psi-\|\psi\|v)+H(F(x+z),\psi+\|\psi\|v)$$

is positively homogeneous and subadditive on $\mathbb{R}^n$. Since the density $\omega_i(\cdot)$ is centrally symmetric, this implies that for any $x\in Z$ the function $\psi\mapsto H_i(x,\psi)$ is also positively homogeneous and subadditive on $\mathbb{R}^n$.

By Lemma 3, for any $i\in\mathbb{N}$ and all $x\in Z$ formula (1.7) defines the support function

$$H(F_i(x),\psi)=H_i(x,\psi),\qquad \psi\in\mathbb{R}^n, \tag{1.8}$$

of a convex compact set $F_i(x)\subset\mathbb{R}^n$, i.e., the value of a set-valued mapping $F_i\colon Z\rightrightarrows\mathbb{R}^n$. Moreover, in view of (1.7) we have

$$F(x)\subset F_i(x)\subset F(x)+\frac{2(L+S)}{3^{i-1}}\mathbb{B}^n \tag{1.9}$$

for any $i\in\mathbb{N}$ and all $x\in Z$. Indeed,

$$\begin{aligned}
&H(F_i(x),\psi)-H(F(x),\psi)\\
&\qquad=\int\limits_{\mathbb{R}^{2n}}\big[H(F(x+z),\psi+\|\psi\|v)-H(F(x),\psi)\big]\omega_i(z)\omega_i(v)\,dz\,dv+\frac{2(L+S)}{3^i}\|\psi\|\\
&\qquad\le\int\limits_{\mathbb{R}^{2n}}\big|H(F(x+z),\psi+\|\psi\|v)-H(F(x),\psi+\|\psi\|v)\big|\omega_i(z)\omega_i(v)\,dz\,dv\\
&\qquad\quad+\int\limits_{\mathbb{R}^{2n}}\big|H(F(x),\psi+\|\psi\|v)-H(F(x),\psi)\big|\omega_i(z)\omega_i(v)\,dz\,dv+\frac{2(L+S)}{3^i}\|\psi\|\\
&\qquad\le\left(\frac{1}{3^{2i}}+\frac{1}{3^i}\right)L\|\psi\|+\frac{1}{3^i}S\|\psi\|+\frac{2(L+S)}{3^i}\|\psi\|\le\frac{2(L+S)}{3^{i-1}}\|\psi\|,
\end{aligned}$$

which implies the second inclusion in (1.9). In a similar way, we obtain

$$H(F(x),\psi) - H(F_i(x),\psi) \le \left(\frac{1}{3^{2i}} + \frac{1}{3^i}\right)L\|\psi\| + \frac{1}{3^i}S\|\psi\| - \frac{2(L+S)}{3^i}\|\psi\| \le 0,$$

which implies the first inclusion in (1.9). Thus, inclusions (1.9) are proved.

For $i \in \mathbb{N}$ consider the differential inclusion

$$\dot{x}(t) \in F_i(x(t)), \tag{1.10}$$

where the set-valued mapping $F_i(\cdot)$ is defined on the set Z in terms of its support function given by (1.7) and (1.8). Then the set-valued mapping $F_i(\cdot)$ is Lipschitz continuous on Z, and all sets $F_i(x)$, $x \in Z$, are convex and compact. Therefore, for any compact set $D \subset \mathbb{R}^{n+1}$, the set $\mathbb{X}_i$ of all trajectories $x(\cdot)$ of the differential inclusion (1.10) that are defined on a fixed time interval $[0,T]$, $T > 0$, and satisfy the inclusion $(t, x(t)) \in D$ for all $t \in [0,T]$ is compact in the space $C([0,T],\mathbb{R}^n)$. Moreover, by virtue of (1.9) the set $\mathbb{X} = \bigcap_{i\in\mathbb{N}} \mathbb{X}_i$ coincides with the set of all trajectories of the differential inclusion (1.2) on $[0,T]$ that satisfy the inclusion $(t, x(t)) \in D$ for all $t \in [0,T]$.

Note that since the support function of the set-valued mapping $F_i(\cdot)$ is smooth on $Z \times (\mathbb{R}^n \setminus \{0\})$, the differential inclusion (1.10) can be explicitly represented as a smooth control system (see [5, 6] for details).

2. CONSTRUCTION OF A SEQUENCE OF APPROXIMATING PROBLEMS

Let $x_*(\cdot)$ be an optimal admissible trajectory in problem (P) and $T_* > 0$ be the corresponding optimal time. Let us choose a sequence of functions $\{z_i(\cdot)\}_{i=1}^{\infty}$, $z_i(\cdot) \in C^2([0,T_*],\mathbb{R}^n)$, such that the sequence $\{\|\dot{z}_i(t)\|\}_{i=1}^{\infty}$ is uniformly bounded on $[0,T_*]$ and $\dot{z}_i(\cdot) \to \dot{x}_*(\cdot)$ in $L^1([0,T_*],\mathbb{R}^n)$ as $i \to \infty$. Since $\operatorname{ess\,sup}_{t\in[0,T_*]}\|\dot{x}_*(t)\| < \infty$, it is clear that such a sequence $\{z_i(\cdot)\}_{i=1}^{\infty}$ exists. In what follows, we will assume that the trajectory $x_*(\cdot)$ and all functions $z_i(\cdot)$ are continuously extended to the infinite interval $[T_*,\infty)$ by constants: $x_*(t) \equiv x_*(T_*)$ and $z_i(t) \equiv z_i(T_*)$, $t \ge T_*$.

Define sets $\widetilde{M}_0$ and $\widetilde{M}_1$ as follows:

$$\widetilde{M}_0 = \begin{cases} M_0 & \text{if } x_*(0) \in M, \\ M_0 \cap G & \text{if } x_*(0) \in G \end{cases} \qquad \text{and} \qquad \widetilde{M}_1 = \begin{cases} M_1 & \text{if } x_*(T_*) \in M, \\ M_1 \cap G & \text{if } x_*(T_*) \in G. \end{cases}$$

For $i \in \mathbb{N}$ let $\widetilde{F}_i\colon [0,\infty) \times \mathbb{R}^n \rightrightarrows \mathbb{R}^{n+1}$ be a set-valued mapping defined by

$$\widetilde{F}_i(t,x) = \big\{(u,v)\colon\ u \in F(x),\ v = \|u - \dot{z}_i(t)\|\big\}, \qquad t \ge 0, \quad x \in \mathbb{R}^n. \tag{2.1}$$

It is easy to see that the mapping $\widetilde{F}_i(\cdot\,,\cdot)$ is locally Lipschitz continuous. Namely, if Z is a bounded subset of $\mathbb{R}^n$, L is the Lipschitz constant of $F(\cdot)$ on Z, and $K_i = \max_{t\in[0,T_*]}\|\ddot{z}_i(t)\|$, then $\widetilde{F}_i(\cdot\,,\cdot)$ satisfies the Lipschitz condition on $[0,\infty) \times Z$ with the constant $2L + K_i$. Moreover, the set-valued mapping $\widetilde{F}_i(t,\cdot)$ is Lipschitz continuous in x on Z uniformly in $t \in [0,\infty)$ with the Lipschitz constant $2L$ independent of i, and for any bounded subset Z in $\mathbb{R}^n$ there exists an $S > 0$ such that $\|\widetilde{F}_i(t,x)\| \le S$ for all $(t,x) \in [0,\infty) \times Z$. For the further reasoning, we choose a bounded open set Z such that $x_*(t) + \mathbb{B}^n \subset Z$ for all $t \in [0,T_*]$. Accordingly, we thereby fix the constants L and S.

Let $\Phi_i\colon [0,\infty) \times \mathbb{R}^n \rightrightarrows \mathbb{R}^{n+1}$, $i \in \mathbb{N}$, be a set-valued mapping defined as the convex hull of $\widetilde{F}_i(\cdot\,,\cdot)$:

$$\Phi_i(t,x) = \operatorname{conv}\widetilde{F}_i(t,x), \qquad t \ge 0, \quad x \in \mathbb{R}^n. \tag{2.2}$$

Then the set-valued mapping $\Phi_i(\cdot\,,\cdot)$ satisfies the same Lipschitz and boundedness conditions (with the same constants $2L$, K_i, and S) as $\widetilde{F}_i(\cdot\,,\cdot)$.

Let us introduce a new state variable $\widetilde{x} = (x, y)$, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^1$, and choose a sequence $\{\gamma_i\}_{i=1}^{\infty}$, $\gamma_i > 0$, such that $\gamma_i(1 + K_i) \to 0$ as $i \to \infty$.

Consider the following sequence of optimal control problems $\{(\widetilde{P}_i)\}_{i=1}^{\infty}$ with free terminal time $T > 0$:

$$\widetilde{J}_i(T, \widetilde{x}(\cdot)) = \varphi(T, x(0), x(T)) + \|x(0) - x_*(0)\|^2 + (T - T_*)^2 + \gamma_i y(T) + \int_0^T \lambda(x(t))\,\delta_M(x(t))\,dt \to \min,$$

$$\dot{\widetilde{x}}(t) = (\dot{x}(t), \dot{y}(t)) \in \Phi_i(t, x(t)), \tag{2.3}$$

$$|T - T_*| \le \frac{T_*}{2}, \qquad \|x(t) - x_*(t)\| \le \frac{1}{2}, \quad t \in [0, T],$$

$$x(0) \in \widetilde{M}_0, \qquad y(0) = 0, \qquad x(T) \in \widetilde{M}_1. \tag{2.4}$$

Note that for any $i \in \mathbb{N}$ the function $\widetilde{x}_i^*(\cdot) = (x_*(\cdot), y_i^*(\cdot))$ with $y_i^*(t) = \int_0^t \|\dot{x}_*(s) - \dot{z}_i(s)\|\,ds$, $t \in [0, T_*]$, is an admissible trajectory in problem $(\widetilde{P}_i)$. Moreover, the sequence $\{y_i^*(T_*)\}_{i=1}^{\infty}$ is bounded. Just as above, every admissible trajectory $\widetilde{x}(\cdot) = (x(\cdot), y(\cdot))$ in $(\widetilde{P}_i)$ defined on $[0, T_i]$, $T_i > 0$, is assumed to be continuously extended to the interval $[T_i, \infty)$ by a constant. Since the functional (1.6) is lower semicontinuous (see [8]) and the set of admissible trajectories $\widetilde{x}(\cdot)$ in $(\widetilde{P}_i)$ is compact in $C([0, 3T_*/2], \mathbb{R}^{n+1})$, for any i there exists a solution $\widetilde{x}_i(\cdot) = (x_i(\cdot), y_i(\cdot))$ of problem $(\widetilde{P}_i)$. Let T_i be the corresponding optimal terminal time.

Lemma 4. *The following conditions hold*:

$$\lim_{i \to \infty} T_i = T_*, \tag{2.5}$$

$$x_i(\cdot) \to x_*(\cdot) \qquad \text{in } C([0, T_*], \mathbb{R}^n) \quad \text{as } i \to \infty, \tag{2.6}$$

$$\dot{x}_i(\cdot) \to \dot{x}_*(\cdot), \quad \dot{y}_i(\cdot) \to 0 \qquad \text{in } L^1([0, T_*], \mathbb{R}^n) \quad \text{as } i \to \infty, \tag{2.7}$$

$$\lim_{i \to \infty} \int_0^{T_i} \lambda(x_i(t))\,\delta_M(x_i(t))\,dt = \int_0^{T_*} \lambda(x_*(t))\,\delta_M(x_*(t))\,dt. \tag{2.8}$$

Proof. Since $x_*(\cdot)$ is an optimal trajectory in problem (P), $x_i(\cdot)$ is an admissible trajectory in problem (P), $\widetilde{x}_i(\cdot)$ is an optimal trajectory in problem $(\widetilde{P}_i)$, and $\widetilde{x}_*(\cdot) = (x_*(\cdot), y_i^*(\cdot))$, $y_i^*(t) = \int_0^t \|\dot{x}_*(s) - \dot{z}_i(s)\|\,ds$, $t \in [0, T_*]$, is an admissible trajectory in problem $(\widetilde{P}_i)$, it follows that

$$\begin{aligned} J(T_*, x_*(\cdot)) &\le J(T_i, x_i(\cdot)) = \varphi(T_i, x_i(0), x_i(T_i)) + \int_0^{T_i} \lambda(x_i(t))\,\delta_M(x_i(t))\,dt \\ &\le \varphi(T_i, x_i(0), x_i(T_i)) + \|x_i(0) - x_*(0)\|^2 + (T_i - T_*)^2 + \gamma_i y_i(T_i) + \int_0^{T_i} \lambda(x_i(t))\,\delta_M(x_i(t))\,dt \\ &\le \varphi(T_*, x_*(0), x_*(T_*)) + \gamma_i y_i^*(T_*) + \int_0^{T_*} \lambda(x_*(t))\,\delta_M(x_*(t))\,dt = J(T_*, x_*(\cdot)) + \gamma_i y_i^*(T_*). \end{aligned} \tag{2.9}$$

Therefore,

$$\|x_i(0) - x_*(0)\|^2 + (T_i - T_*)^2 + \gamma_i y_i(T_i) \le J(T_*, x_*(\cdot)) - J(T_i, x_i(\cdot)) + \gamma_i y_i^*(T_*) \le \gamma_i y_i^*(T_*).$$

Since $y_i^*(T_*) = \int_0^{T_*} \|\dot{x}_*(s) - \dot{z}_i(s)\|\,ds \to 0$ as $i \to \infty$, we have $x_i(0) \to x_*(0)$ and $T_i \to T_*$ as $i \to \infty$ (see (2.5)) and $\gamma_i y_i(T_i) \le \gamma_i y_i^*(T_*)$ for all $i \in \mathbb{N}$. This last inequality implies

$$y_i(T_i) = \int\limits_0^{T_i} \|\dot{x}_i(s) - \dot{z}_i(s)\|\,ds \le y_i^*(T_*) \to 0, \qquad i \to \infty.$$

Therefore, $\dot{y}_i(\cdot) \to 0$ strongly in $L^1([0,T_*],\mathbb{R}^1)$ as $i \to \infty$ (see (2.7)).

Further, due to (2.1) and (2.2) we have

$$\dot{x}_i(t) = \sum_{j=1}^{n+2} \alpha_j^i(t) u_j^i(t), \qquad u_j^i(t) \in F(x_i(t)), \quad \alpha_j^i(t) \ge 0, \quad \sum_{j=1}^{n+2} \alpha_j^i(t) = 1,$$

$$y_i(t) = \sum_{j=1}^{n+2} \alpha_j^i(t) \|u_j^i(t) - \dot{z}_i(t)\|.$$

Hence, since $\dot{z}_i(\cdot)$ converges strongly to $\dot{x}_*(\cdot)$ in $L^1([0,T_*],\mathbb{R}^n)$, we find that $\dot{x}_i(\cdot)$ converges strongly to $\dot{x}_*(\cdot)$ in $L^1([0,T_*],\mathbb{R}^n)$ as $i \to \infty$ (see (2.7)). Recalling that $x_i(0) \to x_*(0)$ as $i \to \infty$, we conclude that $x_i(\cdot) \to x_*(\cdot)$ in $C([0,T_*],\mathbb{R}^n)$ (see (2.6)).

In view of (2.5)–(2.7) and the continuity of $\varphi(\cdot,\cdot,\cdot)$, it follows from (2.9) that

$$\limsup_{i\to\infty} \int\limits_0^{T_i} \lambda(x_i(t))\,\delta_M(x_i(t))\,dt \le \int\limits_0^{T_*} \lambda(x_*(t))\,\delta_M(x_*(t))\,dt.$$

On the other hand, by Lemmas 1 and 2 we have

$$\liminf_{i\to\infty} \int\limits_0^{T_i} \lambda(x_i(t))\,\delta_M(x_i(t))\,dt \ge \int\limits_0^{T_*} \lambda(x_*(t))\,\delta_M(x_*(t))\,dt.$$

This proves equality (2.8). $\square$

For every fixed $i \in \mathbb{N}$, using formula (1.7), we construct a smooth approximation

$$\dot{\widetilde{x}}(t) = (\dot{x}(t), \dot{y}(t)) \in \Phi_{i,j}(t, x(t)), \qquad j = 1, 2, \dots, \tag{2.10}$$

of the differential inclusion (2.3). Namely, for $i, j \in \mathbb{N}$ we set

$$H(\Phi_{i,j}(t,x),\widetilde{\psi}) = \int\limits_{\mathbb{R}^{2n+1}} H\big(\widetilde{F}_i(t,x+z),\widetilde{\psi} + \|\widetilde{\psi}\|\widetilde{v}\big)\,\omega_j(z)\widetilde{\omega}_j(\widetilde{v})\,dz\,d\widetilde{v} + \frac{2(L+S)}{3^j}\|\widetilde{\psi}\|. \tag{2.11}$$

Here $t \ge 0$, $x \in Z$, $\widetilde{\psi} = (\psi, \psi^{n+1})$ with $\psi \in \mathbb{R}^n$ and $\psi^{n+1} \in \mathbb{R}^1$, and the functions $\omega_j(\cdot)$ and $\widetilde{\omega}_j(\cdot)$ are smooth centrally symmetric probability densities with $\operatorname{supp}\omega_j(\cdot) \subset 3^{-j}\mathbb{B}^n$ and $\operatorname{supp}\widetilde{\omega}_j(\cdot) \subset 3^{-j}\mathbb{B}^{n+1}$, respectively.

By Lemma 3, formula (2.11) defines the support function for the right-hand side of the differential inclusion (2.10). If a trajectory $\widetilde{x}(\cdot)$ of (2.10) is defined on $[0,T]$, $T > 0$, then we will always assume that it is continuously extended to the infinite interval $[T,\infty)$ by a constant.

For a fixed $i \in \mathbb{N}$ and $j \to \infty$, consider the following sequence of optimal control problems $\{(P_{i,j})\}_{j=1}^{\infty}$ with free terminal time $T > 0$:

$$\begin{aligned}\widetilde{J}_{i,j}(T,\widetilde{x}(\cdot)) = \varphi(T, x(0), x(T)) + \|x(0) - x_*(0)\|^2 + (T - T_*)^2 + \gamma_i y(T)\\ + \int\limits_0^T \lambda(x(t))\,\delta_j(x(t))\,dt \to \min,\end{aligned} \tag{2.12}$$

$$\dot{\widetilde{x}}(t) = (\dot{x}(t), \dot{y}(t)) \in \Phi_{i,j}(t, x(t)), \tag{2.13}$$

$$|T - T_*| \le \frac{T_*}{2}, \qquad \|x(t) - x_*(t)\| \le \frac{1}{2}, \quad t \in [0, T], \tag{2.14}$$

$$x(0) \in \widetilde{M}_0, \qquad y(0) = 0, \qquad x(T) \in \widetilde{M}_1. \tag{2.15}$$

Lemma 5. *For any $i, j \in \mathbb{N}$, problem $(P_{i,j})$ has a solution $\widetilde{x}_{i,j}(\cdot) = (x_{i,j}(\cdot), y_{i,j}(\cdot))$ defined on $[0, T_{i,j}]$, where $T_{i,j} > 0$ is the corresponding optimal time. For every $i \in \mathbb{N}$, passing to a subsequence if necessary, we have*

$$\lim_{j\to\infty} T_{i,j} = T_i, \tag{2.16}$$

$$(x_{i,j}(\cdot), y_{i,j}(\cdot)) \to (x_i(\cdot), y_i(\cdot)) \qquad \text{in } C([0, T_i], \mathbb{R}^n) \quad \text{as } j \to \infty, \tag{2.17}$$

$$(\dot{x}_{i,j}(\cdot), \dot{y}_{i,j}(\cdot)) \to (\dot{x}_i(\cdot), \dot{y}_i(\cdot)) \qquad \text{weakly in } L^1([0, T_i], \mathbb{R}^{n+1}) \quad \text{as } j \to \infty, \tag{2.18}$$

$$\lim_{j\to\infty} \int_0^{T_{i,j}} \lambda(x_{i,j}(t))\delta_j(x_{i,j}(t))\, dt = \int_0^{T_i} \lambda(x_i(t))\delta_M(x_i(t))\, dt, \tag{2.19}$$

where $\widetilde{x}_i(\cdot) = (x_i(\cdot), y_i(\cdot))$ is a solution of problem $(\widetilde{P}_i)$ and $T_i > 0$ is the optimal time in $(\widetilde{P}_i)$ corresponding to $\widetilde{x}_i(\cdot)$.

Proof. The function $\widetilde{x}_i^*(\cdot) = (x_*(\cdot), y_i^*(\cdot))$ with $y_i^*(t) = \int_0^t \|\dot{x}_*(s) - \dot{z}_i(s)\|\, ds$, $t \in [0, T_*]$, is an admissible trajectory in problem $(P_{i,j})$. The functional (1.6) is lower semicontinuous and the set of admissible trajectories $\widetilde{x}(\cdot)$ in $(P_{i,j})$ is compact in $C([0, 3T_*/2], \mathbb{R}^{n+1})$, so for any $i, j \in \mathbb{N}$ problem $(P_{i,j})$ has a solution $\widetilde{x}_{i,j}(\cdot) = (x_{i,j}(\cdot), y_{i,j}(\cdot))$ defined on an interval $[0, T_{i,j}]$, $T_*/2 \le T_{i,j} \le 3T_*/2$.

For any fixed i the set of trajectories of system (2.13) that satisfy the constraints (2.14) and (2.15) and are continuously extended by constants to the interval $[T_{i,j}, \infty)$ is compact in the space $C([0, 3T_*/2], \mathbb{R}^{n+1})$. The sequence $\{T_{i,j}\}_{j=1}^\infty$ is bounded. Therefore, passing to a subsequence if necessary, we can assume that $T_{i,j} \to T_i > 0$ as $j \to \infty$ (see (2.16)) and the sequence $\{\widetilde{x}_{i,j}(\cdot)\}$ converges uniformly to some trajectory $\widetilde{x}_i(\cdot) = (x_i(\cdot), y_i(\cdot))$ of the inclusion (2.3) defined on the interval $[0, T_i]$; moreover, the trajectory $\widetilde{x}_i(\cdot)$ satisfies the constraints (2.14) and (2.15) for $T = T_i$. Thus, $\widetilde{x}_i(\cdot)$ is an admissible trajectory in problem $(\widetilde{P}_i)$ and condition (2.17) holds. Since the sequences $\{\dot{x}_{i,j}(t)\}_{j=1}^\infty$ and $\{\dot{y}_{i,j}(t)\}_{j=1}^\infty$ are uniformly bounded, condition (2.18) follows.

Let us prove that the admissible trajectory $\widetilde{x}_i(\cdot)$ is optimal in problem $(\widetilde{P}_i)$. Let $\widehat{x}_i(\cdot) = (\xi_i(\cdot), \zeta_i(\cdot))$ be an arbitrary admissible trajectory in $(\widetilde{P}_i)$ defined on some interval $[0, \widehat{T}_i]$, $\widehat{T}_i > 0$. Since $\widetilde{x}_{i,j}(\cdot)$ is an optimal admissible trajectory in problem $(P_{i,j})$, $j \in \mathbb{N}$, and $\widehat{x}_i(\cdot)$ is an admissible trajectory in $(P_{i,j})$, it follows from Lemma 1 that

$$\begin{aligned}
&\varphi\big(T_{i,j}, x_{i,j}(0), x_{i,j}(T_{i,j})\big) + (T_{i,j} - T_*)^2 + \gamma_i y_{i,j}(T_{i,j}) + \int_0^{T_{i,j}} \lambda(x_{i,j}(t))\delta_j(x_{i,j}(t))\, dt \\
&\le \varphi(\widehat{T}_i, \xi_i(0), \xi_i(\widehat{T}_i)) + (\widehat{T}_i - T_*)^2 + \gamma_i \zeta_i(\widehat{T}_i) + \int_0^{\widehat{T}_i} \lambda(\xi_i(t))\delta_j(\xi_i(t))\, dt \\
&\le \varphi(\widehat{T}_i, \xi_i(0), \xi_i(T_i)) + (\widehat{T}_i - T_*)^2 + \gamma_i \zeta_i(\widehat{T}_i) + \int_0^{\widehat{T}_i} \lambda(\xi_i(t))\delta_M(\xi_i(t))\, dt + \frac{j}{3^j} \int_0^{\widehat{T}_i} \lambda(\xi_i(t))\, dt \\
&\le \widetilde{J}_i(\widehat{x}_i(\cdot), \widehat{T}_i) + \frac{j}{3^{j-1}} T_* \max_{t\in[0,3T_*/2]} \lambda(\xi_i(t)).
\end{aligned} \tag{2.20}$$

On the other hand, Lemma 2 implies the inequality

$$\liminf_{j\to\infty} \int_0^{T_{i,j}} \lambda(x_{i,j}(t))\,\delta_j(x_{i,j}(t))\,dt \geq \int_0^{T_i} \lambda(x_i(t))\,\delta_M(x_i(t))\,dt.$$

Therefore, passing to the limit in (2.20) as $j\to\infty$, we obtain

$$\widetilde{J}_i(\widetilde{x}_i(\cdot),T_i) = \varphi(T_i,x_i(0),x_i(T_i)) + (T_i-T_*)^2 + \gamma_i y_i(T_i) + \int_0^{T_i} \lambda(x_i(t))\,\delta_M(x_i(t))\,dt \leq \widetilde{J}_i(\widehat{x}_i(\cdot),\widehat{T}_i).$$

Thus, the inequality $\widetilde{J}_i(\widetilde{x}_i(\cdot),T_i) \leq \widetilde{J}_i(\widehat{x}_i(\cdot),\widehat{T}_i)$ holds. Since $\widehat{x}_i(\cdot)$ is an arbitrary admissible trajectory in $(\widetilde{P}_i)$, we see that $\widetilde{x}_i(\cdot)$ is an optimal admissible trajectory in $(\widetilde{P}_i)$ and T_i is the corresponding optimal time in this problem.

Equality (2.19) follows from (2.16), (2.17), and Lemmas 1 and 2. $\square$

Theorem 1. *For all $i\in\mathbb{N}$ one can choose numbers $j(i)$ in such a way that* $\lim_{i\to\infty} j(i)=\infty$, $\lim_{i\to\infty} 3^{-j(i)}K_i = 0$, $(x_{i,j(i)}(\cdot),y_{i,j(i)}(\cdot)) \to (x_*(\cdot),y_*(\cdot))$ *in* $C([0,T_*],\mathbb{R}^{n+1})$, $(\dot{x}_{i,j(i)}(\cdot),\dot{y}_{i,j(i)}(\cdot)) \to (\dot{x}_*(\cdot),0)$ *in* $L^1([0,T_*],\mathbb{R}^{n+1})$ *as* $i\to\infty$, *and, in addition,*

$$\lim_{i\to\infty} \int_0^{T_{i,j(i)}} \lambda(x_{i,j(i)}(t))\,\delta_{j(i)}(x_{i,j(i)}(t))\,dt = \int_0^{T_*} \lambda(x_*(t))\,\delta_M(x_*(t))\,dt. \tag{2.21}$$

Proof. Since $\{x_{i,j}(\cdot)\}_{j=1}^{\infty}$ and $\{y_{i,j}(\cdot)\}_{j=1}^{\infty}$ converge uniformly to $x_i(\cdot)$ and $y_i(\cdot)$, respectively, on $[0,T_i]$ as $j\to\infty$, $T_i\to T_*$ as $i\to\infty$, and since $\{x_i(\cdot)\}_{i=1}^{\infty}$ and $\{y_i(\cdot)\}_{i=1}^{\infty}$ converge uniformly to $x_*(\cdot)$ and zero, respectively, on $[0,T_*]$ as $i\to\infty$, we can choose the sequence $\{j(i)\}_{i=1}^{\infty}$ in such a way that the sequences $\{x_{i,j(i)}(\cdot)\}_{i=1}^{\infty}$ and $\{y_{i,j(i)}(\cdot)\}_{i=1}^{\infty}$ converge uniformly to $x_*(\cdot)$ and zero, respectively, on $[0,T_*]$ as $i\to\infty$. Since the sequences $\{\dot{x}_{i,j(i)}(\cdot)\}_{i=1}^{\infty}$ and $\{\dot{y}_{i,j(i)}(\cdot)\}_{i=1}^{\infty}$ are uniformly bounded on $[0,T_*]$, we have $\dot{x}_{i,j(i)}(\cdot)\to\dot{x}_*(\cdot)$ weakly in $L^1([0,T_*],\mathbb{R}^n)$ and $\dot{y}_{i,j(i)}\to 0$ weakly in $L^1([0,T_*],\mathbb{R}^1)$ as $i\to\infty$. Due to the inequality $\dot{y}_{i,j(i)}(t)\geq 0$ on $[0,T_*]$, the weak convergence of $\{\dot{y}_{i,j(i)}(\cdot)\}$ to zero in $L^1([0,T_*],\mathbb{R}^1)$ implies its strong convergence to zero in $L^1([0,T_*],\mathbb{R}^1)$ as $i\to\infty$.

Next, using (2.13) we have

$$\dot{x}_{i,j(i)}(t) \overset{\text{a.e.}}{=} \sum_{k=1}^{n+2} \alpha_k^{i,j(i)}(t)u_k^{i,j(i)}(t) + l_{i,j(i)}(t),$$

$$\dot{y}_{i,j(i)}(t) \overset{\text{a.e.}}{=} \sum_{k=1}^{n+2} \alpha_k^{i,j(i)}(t)\big\|u_k^{i,j(i)}(t)-\dot{z}_i(t)\big\| + m_{i,j(i)}(t),$$

where

$$u_k^{i,j(i)}(t)\in\widetilde{F}_i(t,x_{i,j(i)}(t)) \qquad \text{and} \qquad \big(l_{i,j(i)}(t),m_{i,j(i)}(t)\big)\in\frac{L+S}{3^{j-1}}\mathbb{B}^{n+1}.$$

Therefore,

$$\begin{aligned}\|\dot{x}_{i,j(i)}(t)-\dot{z}_i(t)\| &\overset{\text{a.e.}}{\leq} \sum_{k=1}^{n+2} \alpha_k^{i,j(i)}(t)\big\|u_k^{i,j(i)}(t)-\dot{z}_i(t)\big\| + \|l_{i,j(i)}(t)\| \\ &\overset{\text{a.e.}}{\leq} \dot{y}_{i,j(i)}(t) + \|l_{i,j(i)}(t)\| + |m_{i,j(i)}(t)|.\end{aligned}$$

Since $\dot{z}_i(t) \to \dot{x}_*(t)$ strongly in $L^1([0,T_*],\mathbb{R}^n)$ as $i \to \infty$, $\dot{y}_{i,j(i)}(t) \to 0$ strongly in $L^1([0,T_*],\mathbb{R}^1)$ as $i \to \infty$, and $T_{i,j(i)} \to T_*$ as $i \to \infty$, it follows that the sequence $\{x_{i,j(i)}(\cdot)\}_{i=1}^{\infty}$ converges strongly to $x_*(\cdot)$ in $L^1([0,T_*],\mathbb{R}^n)$.

Increasing the number $j(i)$ if necessary, we can assume that $\lim_{i\to\infty} 3^{-j(i)}K_i = 0$ (without loss of generality).

The proof of condition (2.21) is completely similar to that of condition (2.19) in Lemma 5. $\square$

In what follows, for brevity, for each $i \in \mathbb{N}$ we will denote problem $(P_{i,j(i)})$ by (P_i), the optimal trajectory $\widetilde{x}_{i,j(i)}(\cdot)$ in it by $\widetilde{x}_i(\cdot) = (x_i(\cdot), y_i(\cdot))$, and the corresponding optimal time $T_{i,j(i)}$ by T_i. Then by Theorem 1 we have $\lim_{i\to\infty} 3^{-j(i)}K_i = 0$, $(x_i(\cdot), y_i(\cdot)) \to (x_*(\cdot), y_*(\cdot))$ in $C([0,T_*],\mathbb{R}^{n+1})$, $(\dot{x}_i(\cdot), \dot{y}_i(\cdot)) \to (\dot{x}_*(\cdot), 0)$ in $L^1([0,T_*],\mathbb{R}^{n+1})$ as $i \to \infty$, and

$$\lim_{i\to\infty} \int_0^{T_i} \lambda(x_i(t))\,\delta_{j(i)}(x_i(t))\,dt = \int_0^{T_*} \lambda(x_*(t))\,\delta_M(x_*(t))\,dt. \tag{2.22}$$

Corollary 1. *Under the hypotheses of Theorem* 1, *passing to a subsequence if necessary, we can assume that*

$$\lim_{i\to\infty} \delta_{j(i)}(x_i(t)) = \delta_M(x_*(t)) \qquad \textit{for a.e.} \quad t \in [0,T_*]. \tag{2.23}$$

Proof. Since the set M is open, it follows from the definition of the functions $\delta_M(\cdot)$ and $\delta_i(\cdot)$, $i \in \mathbb{N}$ (see (1.4) and (1.5)), and the uniform convergence of $\{x_i(\cdot)\}_{i=1}^{\infty}$ to $x_*(\cdot)$ on $[0,T_*]$ that $\lim_{i\to\infty} \delta_{j(i)}(x_i(t)) = \delta_M(x_*(t)) = 1$ for all $t \in [0,T_*]$ such that $x_*(t) \in M$. Consider now the set of those $t \in [0,T_*]$ for which $x_*(t) \in G$. In this case $\delta_M(x_*(t)) = 0$, and in view of (2.22) we have the equality

$$\lim_{i\to\infty} \int_{\{t\in[0,T_*]\colon\, x_*(t)\in G\}} \lambda(x_i(t))\,\delta_{j(i)}(x_i(t))\,dt = 0.$$

Since the functions $\lambda(x_i(\cdot))\,\delta_{j(i)}(x_i(\cdot))$, $i \in \mathbb{N}$, are nonnegative, we obtain

$$\lim_{i\to\infty} \operatorname{meas}\bigl\{t \in [0,T_*]\colon\; x_*(t) \in G,\; \lambda(x_i(t))\,\delta_{j(i)}(x_i(t)) > \varepsilon\bigr\} = 0$$

for any $\varepsilon > 0$; i.e., the sequence $\{\lambda(x_i(\cdot))\,\delta_{j(i)}(x_i(\cdot))\}_{i=1}^{\infty}$ converges on the set $\{t \in [0,T_*]\colon x_*(t) \in G\}$ to zero in measure. Therefore, passing to a subsequence if necessary, we can assume that the sequence $\{\lambda(x_i(t))\,\delta_{j(i)}(x_i(\cdot))\}_{i=1}^{\infty}$ converges to zero for a.e. $t \in [0,T_*]$ such that $x_*(t) \in G$, and hence for a.e. $t \in [0,T_*]$. Since $\lim_{i\to\infty} \lambda(x_i(t)) = \lambda(x_*(t))$, $t \in [0,T_*]$, and the function $\lambda(\cdot)$ is positive on M, this yields condition (2.23). $\square$

By Corollary 1 and Lebesgue's dominated convergence theorem (see [23, Ch. VI, § 3]), we can assume without loss of generality that the equality

$$\lim_{i\to\infty} \int_0^{t} \xi(x_i(s))\,\delta_{j(i)}(x_i(s))\,ds = \int_0^{t} \xi(x_*(s))\,\delta_M(x_*(s))\,ds \tag{2.24}$$

holds for every continuous function $\xi\colon \mathbb{R}^n \to \mathbb{R}^n$ and every $t \in [0,T_*]$.

3. MAIN RESULT

The following theorem is the main result of the paper.

Theorem 2. *Let $x_*(\cdot)$ be an optimal admissible trajectory in problem (P) and $T_* > 0$ be the corresponding optimal time. Then there exists a constant $\psi^0 \geq 0$, an absolutely continuous function $\psi\colon [0,T_*] \to \mathbb{R}^n$, and a bounded regular n-dimensional Borel measure η on $[0,T_*]$ such that the* *following conditions hold*:

(1) *the measure* η *is concentrated on the set* $\mathfrak{M} = \{t \in [0,T_*]\colon x_*(t) \in \partial G\}$ *and is nonpositive on the set of continuous functions* $y\colon \mathfrak{M} \to \mathbb{R}^n$ *with values* $y(t) \in T_G(x_*(t))$, $t \in \mathfrak{M}$, *i.e.*,

$$\int\limits_{\mathfrak{M}} y(t)\,d\eta \le 0;$$

(2) *the refined Euler–Lagrange inclusion holds for a.e.* $t \in [0,T_*]$:

$$\dot\psi(t) \in \operatorname{conv}\left\{u\colon \left(u,\, \psi(t) + \int\limits_0^t \lambda(x_*(s))\,d\eta + \psi^0 \int\limits_0^t \delta_M(x_*(s))\frac{\partial\lambda(x_*(s))}{\partial x}\,ds\right) \in \widehat{N}_{\operatorname{graph} F(\cdot)}(x_*(t),\dot x_*(t))\right\};$$

(3) *for* $t = T_*$ *as well as for any point* $t \in [0,T_*)$ *of right approximate continuity*[1] *of* $\delta_M(x_*(\cdot))$, *one has the equality*

$$H\left(F(x_*(t)),\, \psi(t) + \int\limits_0^t \lambda(x_*(s))\,d\eta + \psi^0 \int\limits_0^t \delta_M(x_*(s))\frac{\partial\lambda(x_*(s))}{\partial x}\,ds\right) - \psi^0\lambda(x_*(t))\delta_M(x_*(t)) = H(F(x_*(0)),\psi(0)) - \psi^0\lambda(x_*(0))\delta_M(x_*(0));$$

(4) *the transversality condition holds*:

$$\left(H\left(F(x_*(T_*)),\psi(T_*) + \int\limits_0^{T_*} \lambda(x_*(s))\,d\eta + \psi^0 \int\limits_0^{T_*} \delta_M(x_*(s))\frac{\partial\lambda(x_*(s))}{\partial x}\,ds\right),\, \psi(0),\right.$$
$$\left. -\psi(T_*) - \int\limits_0^{T_*} \lambda(x_*(s))\,d\eta - \psi^0 \int\limits_0^{T_*} \delta_M(x_*(s))\frac{\partial\lambda(x_*(s))}{\partial x}\,ds\right)$$
$$\in \psi^0\widehat{\partial}\varphi(T_*,x_*(0),x_*(T_*)) + \{0\}\times \widehat{N}_{\widetilde{M}_0}(x_*(0)) \times \widehat{N}_{\widetilde{M}_1}(x_*(T_*));$$

(5) *the nontriviality condition holds*:

$$\psi^0 + \|\psi(0)\| + \|\eta\| \ne 0.$$

Proof. Let $x_*(\cdot)$ be an optimal trajectory in problem (P), $T_* > 0$ an optimal time, and $\{(P_i)\}_{i=1}^\infty$ the sequence of approximating problems $(P_i) = (P_{i,j(i)})$, $i \in \mathbb{N}$ (see (2.12)–(2.15)). We will assume that the sequence $\{j(i)\}_{i=1}^\infty$ is chosen so that all the hypotheses of Theorem 1 are satisfied and equality (2.24) holds for every continuous function $\xi\colon \mathbb{R}^n \to \mathbb{R}^n$ and every $t \in [0,T_*]$.

Further, for any $i,j \in \mathbb{N}$ the differential inclusion (2.13) is equivalent (the sets of their trajectories coincide) to a smooth control system

$$\dot{\tilde x}(t) = (\dot x(t),\dot y(t)) = f_{i,j}(t,x(t),u(t)), \qquad u(t) \in U, \tag{3.1}$$

[1] Recall that a point $t \in [0,T)$, $T > 0$, is called a point of right approximate continuity of a function $\xi\colon [0,T] \to \mathbb{R}^1$ if there exists a Lebesgue measurable set $E \subset [t,T]$ such that t is a density point of E and the function $\xi(\cdot)$ is right-continuous at the point t along the set E (see [23, Ch. IX, § 6]).

where

$$U = \left\{ u = (\alpha_1, \dots, \alpha_{n+2}, \widetilde{v}_1, \dots, \widetilde{v}_{n+2})\colon\ \alpha_k \ge 0,\ \sum_{k=1}^{n+2} \alpha_k = 1,\ \widetilde{v}_k \in \mathbb{R}^{n+1},\ \|\widetilde{v}_k\| = 1 \right\}$$

and

$$f_{i,j}(t,x,u) = \sum_{k=1}^{n+2} \alpha_k \frac{\partial}{\partial \widetilde{v}} H(\Phi_{i,j}(t,x), \widetilde{v}_k).$$

Here the vector functions $f_{i,j}(\cdot,\cdot,\cdot)$ and $\partial f_{i,j}(\cdot,\cdot,\cdot)/\partial_x$ are locally Lipschitz continuous in their arguments (t,x,u) on $[0,\infty) \times Z \times U$ (see [6, Lemma 3]).

By Theorem 1, for all sufficiently large i, we in fact have strict inequalities in (2.14). Therefore, for all sufficiently large i, the optimal trajectory $\widetilde{x}_i\colon [0,T_i] \to \mathbb{R}^{n+1}$ in problem (P_i) satisfies the necessary optimality conditions given by the Pontryagin maximum principle for free-time optimal control problems without state constraints and with nonsmooth terminal constraints [20–22] (see also [6, Theorem 3, Lemma 2]). Namely, there exist numbers $\psi_i^0 \ge 0$ and absolutely continuous functions $\widetilde{\psi}_i\colon [0,T_i] \mapsto \mathbb{R}^{n+1}$, $\widetilde{\psi}_i(\cdot) = (\widehat{\psi}_i(\cdot), \widehat{\psi}_i^{n+1}(\cdot))$, such that

$$-\dot{\widehat{\psi}}_i(t) \overset{\text{a.e.}}{=} \frac{\partial}{\partial x} H_i(t, x_i(t), \widetilde{\psi}_i(t)) - \psi_i^0 \left(\lambda(x_i(t)) \frac{\partial \delta_{j(i)}(x_i(t))}{\partial x} + \delta_{j(i)}(x_i(t)) \frac{\partial \lambda(x_i(t))}{\partial x} \right), \tag{3.2}$$

$$\langle \widetilde{\psi}_i(t), \dot{\widetilde{x}}_i(t) \rangle \overset{\text{a.e.}}{=} H_i(t, x_i(t), \widetilde{\psi}_i(t)), \tag{3.3}$$

$$\begin{aligned} \big(h_i(T_i), \widehat{\psi}_i(0), -\widehat{\psi}_i(T_i)\big) \in \psi_i^0 \Big(\widehat{\partial} \varphi(T_i, x_i(0), x_i(T_i)) + \big(2(T_i - T_*), 2(x_i(0) - x_*(0)), 0\big) \Big) \\ + \{0\} \times \widehat{N}_{\widetilde{M}_1}(x_i(0)) \times \widehat{N}_{\widetilde{M}_2}(x_i(T_i)), \qquad -\widehat{\psi}_i^{n+1}(T_i) = \psi_i^0 \gamma_i, \end{aligned} \tag{3.4}$$

$$\dot{h}_i(t) \overset{\text{a.e.}}{\in} \partial_t H_i(t, x_i(t), \widetilde{\psi}_i(t)), \tag{3.5}$$

$$\psi_i^0 + \|\widetilde{\psi}_i(0)\| \ne 0. \tag{3.6}$$

Here the function $H_i(\cdot,\cdot,\cdot)$ is defined for all $t \ge 0$, $x \in Z$, and $\widetilde{\psi} = (\widehat{\psi}, \widehat{\psi}^{n+1})$, $\widehat{\psi} \in \mathbb{R}^n$, $\widehat{\psi}^{n+1} \in \mathbb{R}^1$, by the equality (see (2.11))

$$H_i(t, x, \widetilde{\psi}) = H(\Phi_{i,j(i)}(t,x), \widetilde{\psi}),$$

the absolutely continuous function $h_i(\cdot)$ is defined on $[0,T_i]$ by the equality

$$h_i(t) = H_i(t, x_i(t), \widetilde{\psi}_i(t)) - \psi_i^0 \lambda(x_i(t)) \delta_{j(i)}(x_i(t)), \tag{3.7}$$

and $\partial_t H_i(t, x_i(t), \widetilde{\psi}(t))$ is the partial Clarke subdifferential [15] of the locally Lipschitz continuous function $H_i(\cdot,\cdot,\cdot)$ with respect to t. Since $\widetilde{x} = (x,y)$, $x \in Z$, $y \in \mathbb{R}^1$, and the right-hand side of (3.1) does not depend on y, it follows (see the second condition in (3.4)) that $\widehat{\psi}_i^{n+1}(t) \equiv \widehat{\psi}_i^{n+1} = -\psi_i^0 \gamma_i$ on $[0,T_i]$.

Multiplying the adjoint variables ψ_i^0 and $\widetilde{\psi}_i(\cdot)$ by a positive factor and recalling (3.6) and the second condition in (3.4), we can assume without loss of generality that

$$\psi_i^0 + \|\widehat{\psi}_i(0)\| + \psi_i^0 \int_0^{T_i} \left\| \frac{\partial \delta_{j(i)}(x_i(t))}{\partial x} \right\| dt = 1. \tag{3.8}$$

We define adjoint variables $\eta_i, \psi_i\colon [0,T_i] \to \mathbb{R}^n$ as

$$\eta_i(t) = \psi_i^0 \frac{\partial \delta_{j(i)}(x_i(t))}{\partial x}, \qquad \psi_i(t) = \widehat{\psi}_i(t) - \int_0^t \lambda(x_i(s)) \eta_i(s)\, ds - \psi_i^0 \int_0^t \delta_{j(i)}(x_i(s)) \frac{\partial \lambda(x_i(s))}{\partial x}\, ds.$$

In terms of the variables $\eta_i(\cdot)$ and $\psi_i(\cdot)$, system (3.2) takes the form (see (2.11))

$$-\dot{\psi}_i(t) \overset{\text{a.e.}}{=} \int\limits_{\mathbb{R}^{2n+1}} \frac{\partial}{\partial x} H\Biggl(\widetilde{F}_i(t, x_i(t)+z), \Biggl(\psi_i(t) + \int\limits_0^t \lambda(x_i(s))\eta_i(s)\,ds + \psi_i^0 \int\limits_0^t \delta_{j(i)}(x_i(s)) \frac{\partial\lambda(x_i(s))}{\partial x}\,ds, \widehat{\psi}_i^{n+1}\Biggr) + \Biggl\| \Biggl(\psi_i(t) + \int\limits_0^t \lambda(x_i(s))\eta_i(s)\,ds + \psi_i^0 \int\limits_0^t \delta_{j(i)}(x_i(s)) \frac{\partial\lambda(x_i(s))}{\partial x}\,ds, \widehat{\psi}_i^{n+1}\Biggr)\Biggr\| \widetilde{v}\Biggr) \times \omega_{j(i)}(z)\,\widetilde{\omega}_{j(i)}(\widetilde{v})\,dz\,d\widetilde{v}. \tag{3.9}$$

In view of (3.8), passing to a subsequence if necessary, we can assume that $\psi_i^0 \to \psi^0 \geq 0$ and $\psi_i(0) = \widehat{\psi}_i(0) \to \psi_0$, $\|\psi_0\| \leq 1$, as $i \to \infty$. Also, using the condition $\lim_{i\to\infty} T_i = T_*$ and Helly's theorem (see, for example, [13, Theorem 15.1.i]), we can assume that the sequence $\{\eta_i(\cdot)\}_{i=1}^{\infty}$ converges weakly as $i \to \infty$ to a regular n-dimensional Borel measure η, $\operatorname{supp}\eta \subset [0, T_*]$; namely, for every continuous function $\xi\colon [0,\infty) \to \mathbb{R}^n$ we have

$$\lim_{\varepsilon\to 0} \lim_{i\to\infty} \int\limits_0^{T_*+\varepsilon} \langle \xi(t), \eta_i(t)\rangle\,dt = \int\limits_0^{T_*} \xi(t)\,d\eta. \tag{3.10}$$

Let us prove that condition (1) holds. Let $\tau \in [0, T_*]$ and either $x_*(\tau) \in \operatorname{int} M$ or $x_*(\tau) \in \operatorname{int} G$. Then, according to the definition (1.5) of $\delta_i(\cdot)$, $i \in \mathbb{N}$, and the uniform convergence of the sequence $\{x_i(\cdot)\}_{i=1}^{\infty}$ to $x_*(\cdot)$ on $[0, T_*]$, there exist $\varepsilon > 0$ and $\delta > 0$ such that either $x_i(t) + \varepsilon\mathbb{B}^n \subset M$ or $x_i(t) + \varepsilon\mathbb{B}^n \subset G$ for all sufficiently large i and all t in the δ-neighborhood of τ in $[0, T_*]$. Therefore, for all sufficiently large i, either $\delta_{j(i)}(x_i(t)) \equiv 1$ or $\delta_{j(i)}(x_i(t)) \equiv 0$ for all t in the δ-neighborhood of τ in $[0, T_*]$. Hence, for all sufficiently large i, we have $\eta_i(t) \equiv 0$ in the δ-neighborhood of τ in $[0, T_*]$. Therefore, being the weak limit of the sequence $\{\eta_i(\cdot)\}_{i=1}^{\infty}$, the measure η is concentrated on the set $\mathfrak{M} = \{t \in [0, T_*] \colon x_*(t) \in \partial G\}$. Clearly, the set $\mathfrak{M}$ is closed and bounded; i.e., it is a compact set in $\mathbb{R}^n$.

If $\mathfrak{M} = \varnothing$, then condition (1) holds. Suppose that $\mathfrak{M} \neq \varnothing$ and $y\colon \mathfrak{M} \to \mathbb{R}^n$ is a continuous function such that $y(t) \in T_G(x_*(t))$, $t \in \mathfrak{M}$. Since $\operatorname{int} T_G(x) \neq \varnothing$, $x \in G$, and the Clarke normal cone is upper semicontinuous in this case (see [15]), to prove condition (1) we can assume without loss of generality that there exists a $\delta > 0$ such that $y(t) \in N_\delta^*(t)$ for all $t \in \mathfrak{M}$ (see [3, Sect. 3]). Here

$$N_\delta(t) = \bigl\{\alpha y\colon\ \|y - \xi\| \leq \delta,\ \xi \in N_G(x_*(t)),\ \|\xi\| = 1,\ \alpha \geq 0\bigr\}$$

is the conical δ-neighborhood of the Clarke normal cone $N_G(x_*(t))$ and $N_\delta^*(t)$ is the polar cone of $N_\delta(t)$.

Let us choose an arbitrary $\tau \in \mathfrak{M}$ and show that there exists an $\varepsilon(\tau) > 0$ such that

$$\frac{\partial\delta_{j(i)}(x_i(t))}{\partial x} \in N_{\delta/2}(\tau), \qquad t \in [\tau - \varepsilon(\tau), \tau + \varepsilon(\tau)] \cap [0, T_i], \tag{3.11}$$

for all sufficiently large i.

Suppose that condition (3.11) fails. Then there exists a sequence $\tau_i \to \tau$, $i \to \infty$, such that

$$\frac{\partial\delta_{j(i)}(x_i(\tau_i))}{\partial x} \notin N_{\delta/2}(\tau).$$

According to the definition of $\delta_i(\cdot)$, we have

$$\frac{\partial \delta_{j(i)}(x_i(\tau_i))}{\partial x} = \int\limits_{\mathbb{R}^n} \frac{\partial \widetilde{\delta}_{j(i)}(x_i(\tau_i)+z)}{\partial x}\,\omega_{j(i)}(z)\,dz = \int\limits_{\{z\colon j(i)\rho(x_i(t)+z,G)\le 1\}} \frac{\partial \widetilde{\delta}_{j(i)}(x_i(t)+z)}{\partial x}\,\omega_{j(i)}(z)\,dz$$

$$= j(i) \int\limits_{\{z\colon j(i)\rho(x_i(\tau_i)+z,G)\le 1\}} \frac{\partial \rho(x_i(\tau_i)+z,G)}{\partial x}\,\omega_{j(i)}(z)\,dz.$$

Therefore, there exists a sequence $z_i \to 0$, $i \to \infty$, such that

$$v_i = \frac{\partial \rho(x_i(\tau_i)+z_i,G)}{\partial x} \notin N_{\delta/2}(\tau). \tag{3.12}$$

Since the distance function $\rho(\cdot\,,G)$ is Lipschitz continuous with constant 1, we have $\|v_i\| = 1$. Note that $v_i \in N_G(\xi_i)$, where ξ_i is a nearest point to $x_i(\tau_i)+z_i$ in G, and $x_i(\tau_i) \to x_*(\tau)$, $z_i \to 0$; hence, $\xi_i \to x_*(\tau)$ as $i \to \infty$. Passing to a subsequence, we obtain $v_i \to v$ as $i \to \infty$, where $\|v\| = 1$. Since the Clarke normal cone is upper semicontinuous (in the case under study, $\operatorname{int} T_G(x) \neq \varnothing$ for $x \in G$), we obtain the inclusion $v \in N_G(x_*(\tau))$, which contradicts condition (3.12). This completes the proof of condition (3.11).

Reducing $\varepsilon(\tau) > 0$ if necessary and recalling that $y(t) \in N_\delta^*(t)$ for all $t \in \mathfrak{M}$, we can assume that $y(t) \in N_{\delta/2}^*(\tau)$ for all $t \in \mathfrak{M} \cap [\tau - \varepsilon(\tau), \tau + \varepsilon(\tau)]$. Therefore,

$$\left\langle y(t), \frac{\partial \delta_{j(i)}(x_i(t))}{\partial x} \right\rangle \le 0, \qquad t \in \mathfrak{M} \cap [\tau - \varepsilon(\tau), \tau + \varepsilon(\tau)].$$

According to the definition of η, this condition implies that for any point $\tau \in \mathfrak{M}$ there exists an $\varepsilon(\tau) > 0$ such that the inequality

$$\int\limits_{\mathfrak{M} \cap [\tau - \varepsilon_1, \tau + \varepsilon_2]} y(t)\,d\eta \le 0$$

holds for all $0 < \varepsilon_1, \varepsilon_2 \le \varepsilon(\tau)$. Since the set $\mathfrak{M}$ is compact, this implies the validity of condition (1).

Lemma 6. *For a.e.* $t \in [0, T_*]$ *we have the equality*

$$\lim_{i\to\infty} \int\limits_0^t \eta_i(s)\,ds = \int\limits_0^t d\eta.$$

Proof. Let $\tau \in [0, T_*]$ be a continuity point of η, i.e., $\eta(\tau) = 0$, where $\eta(\tau)$ is the atomic part of the measure η at the point τ. Let us prove that

$$\lim_{k\to\infty} \lim_{i\to\infty} \psi_i^0 \int\limits_\tau^{\tau+1/k} \left\| \frac{\partial \delta_{j(i)}(x_i(s))}{\partial x} \right\| ds = 0. \tag{3.13}$$

Clearly, if $\tau \notin \mathfrak{M}$, condition (3.13) holds. Suppose that $\tau \in \mathfrak{M}$ and condition (3.13) is violated. Then, for some $\alpha > 0$ and all sufficiently large k, we have

$$\lim_{i\to\infty} \psi_i^0 \int\limits_\tau^{\tau+1/k} \left\| \frac{\partial \delta_{j(i)}(x_i(s))}{\partial x} \right\| ds \ge \alpha.$$

Due to the assumption $\operatorname{int} T_G(x) \neq \varnothing$, $x \in G$, there exists a number $\varepsilon > 0$ and a vector $g \in \mathbb{R}^n$, $\|g\| = 1$, such that $g \in N_\varepsilon^*(\tau)$. According to the definition of $\delta_i(\cdot)$ (see (1.5)), since the sequence

$\{x_i(\cdot)\}_{i=1}^{\infty}$ converges uniformly to $x_*(\cdot)$ on $[0,T_*]$, for all sufficiently large i and k and every $t\in[\tau-1/k,\tau+2/k]$ we have

$$\frac{\partial\delta_{j(i)}(x_i(t))}{\partial x}=\int\limits_{\mathbb{R}^n}\frac{\partial\widetilde{\delta}_{j(i)}(x_i(t)+z)}{\partial x}\,\omega_{j(i)}(z)\,dz$$

$$=j(i)\int\limits_{\{z\colon\, j(i)\rho(x_i(t)+z,G)\le 1\}}\frac{\partial\rho(x_i(t)+z,G)}{\partial x}\,\omega_{j(i)}(z)\,dz\in N_{\varepsilon/2}(\tau).$$

Therefore, for any $u\in\mathbb{R}^n$, $\|u\|\le\varepsilon/2$, the inclusion

$$\frac{\partial\delta_{j(i)}(x_i(t))}{\partial x}+\left\|\frac{\partial\delta_{j(i)}(x_i(t))}{\partial x}\right\|u\in N_{\varepsilon}(\tau)$$

holds for all $t\in[\tau-1/k,\tau+2/k]$. Hence we obtain

$$\left\langle\frac{\partial\delta_{j(i)}(x_i(t))}{\partial x},g\right\rangle\le-\frac{\varepsilon}{2}\left\|\frac{\partial\delta_{j(i)}(x_i(t))}{\partial x}\right\|,\qquad t\in\left[\tau-\frac{1}{k},\tau+\frac{2}{k}\right].$$

Consider a sequence $\{h_k(\cdot)\}_{k=1}^{\infty}$ of continuous functions $h_k(\cdot)\colon[0,\infty)\to[0,1]$ such that $h_k(t)=1$ for $t\in[\tau,\tau+1/k]$ and $h_k(t)=0$ for $t\notin[\tau-1/k,\tau+2/k]$. Then, for all sufficiently large k, we have

$$\lim_{i\to\infty}\psi_i^0\int\limits_{\tau-1/k}^{\tau+2/k}\left\langle\frac{\partial\delta_{j(i)}(x_i(t))}{\partial x},g\right\rangle h_k(t)\,dt\le\limsup_{i\to\infty}\psi_i^0\int\limits_{\tau}^{\tau+1/k}\left\langle\frac{\partial\delta_{j(i)}(x_i(t))}{\partial x},g\right\rangle dt$$

$$\le-\frac{\varepsilon}{2}\limsup_{i\to\infty}\psi_i^0\int\limits_{\tau}^{\tau+1/k}\left\|\frac{\partial\delta_{j(i)}(x_i(t))}{\partial x}\right\|dt\le-\frac{\varepsilon\alpha}{2}.$$

According to the definition of η and the choice of τ, this yields

$$0=\langle g,\eta(\tau)\rangle=\lim_{k\to\infty}\lim_{i\to\infty}\psi_i^0\int\limits_{\tau-1/k}^{\tau+2/k}\left\langle\frac{\partial\delta_{j(i)}(x_i(t))}{\partial x},g\right\rangle h_k(t)\,dt\le-\frac{\varepsilon\alpha}{2}.$$

We have arrived at a contradiction. Therefore, condition (3.13) holds.

The assertion of the lemma now follows from the definition of η, condition (3.13), and the fact that almost every point of the interval $[0,T_*]$ is a continuity point of η. $\square$

Let us prove that condition (2) holds. To this end, consider the sequence $\{\psi_i(\cdot)\}_{i=1}^{\infty}$. Combining equality (3.9), condition (3.8), and the fact that the set-valued mapping $\Phi_{i,j(i)}(t,\cdot)$ is Lipschitz continuous with constant $2L$ independent of i, we have $\|\dot\psi_i(t)\|\le\kappa(\|\psi_i(t)\|+1)$, $t\in[0,T_i]$, where $\kappa\ge0$ is a constant. Therefore, by Gronwall's lemma (see [13, Lemma 18.1.i]), we can assume without loss of generality that $\psi_i(\cdot)\to\psi(\cdot)$ in $C([0,T_*],\mathbb{R}^n)$ and $\dot\psi_i(\cdot)\to\dot\psi(\cdot)$ weakly in $L^1([0,T_*],\mathbb{R}^n)$ as $i\to\infty$, where $\psi\colon[0,T_*]\to\mathbb{R}^n$ is a Lipschitz function, $\psi(0)=\psi_0$. Since $\psi_i(\cdot)\to\psi(\cdot)$ uniformly on $[0,T_*]$ as $i\to\infty$ and the sequence $\{\|\dot\psi_i(t)\|\}_{i=1}^{\infty}$ is uniformly bounded on $[0,T_*]$, it follows that the inclusion $\dot\psi(t)\in\operatorname{conv}\operatorname{Ls}_{i\to\infty}\dot\psi_i(t)$ holds for a.e. $t\in[0,T_*]$, where

$$\operatorname*{Ls}_{i\to\infty}\dot\psi_i(t)=\left\{q\in\mathbb{R}^n\colon\ \exists\{i_j\}_{j=1}^{\infty}\colon\dot\psi_{i_j}(t)\to q,\ j\to\infty\right\}$$

is the upper topological limit of the sequence $\{\dot\psi_i(t)\}_{i=1}^{\infty}$ as $i\to\infty$. Further, $\dot z_i(\cdot)\to\dot x_*(\cdot)$ and $\dot x_i(\cdot)\to\dot x_*(\cdot)$ strongly in $L^1([0,T_*],\mathbb{R}^n)$, and also $\dot y_i(\cdot)\to0$ strongly in $L^1([0,T_*],\mathbb{R}^1)$ as $i\to\infty$.

Therefore, passing to a subsequence if necessary, we can assume that $\dot{z}_i(t) \to \dot{x}_*(t)$, $\dot{x}_i(t) \to \dot{x}_*(t)$, and $\dot{y}_i(t) \to 0$ as $i \to \infty$ for a.e. $t \in [0, T_*]$. Moreover, by Lemma 6, the equality $\int_0^t d\eta = \lim_{i\to\infty} \int_0^t \eta_i(s)\, ds$ holds for a.e. $t \in [0, T_*]$. Now we fix an arbitrary $t \in [0, T_*]$ for which all the above conditions are satisfied.

Let $q \in \mathrm{Ls}_{i\to\infty} \dot{\psi}_i(t)$. Passing to a subsequence if necessary, we assume that $\dot{\psi}_i(t) \to q$ as $i \to \infty$ and $\|\dot{z}_i(t) - x_*(t)\| \le 1/i$, $i \in \mathbb{N}$. Let us prove the inclusion

$$q \in \operatorname{conv}\left\{u\colon \left(u, \psi(t) + \int_0^t \lambda(x_*(s))\, d\eta + \psi^0 \int_0^t \delta_M(x_*(s)) \frac{\partial \lambda(x_*(s))}{\partial x}\, ds\right) \in \widehat{N}_{\operatorname{graph} F(\cdot)}(x_*(t), \dot{x}_*(t))\right\}. \quad (3.14)$$

To this end, it suffices to consider the case where $\lim_{i\to\infty}\big(\widehat{\psi}_i(t), \widehat{\psi}_i^{n+1}\big) \ne 0$. Indeed, in view of (3.9) we have

$$\|\dot{\psi}_i(t)\| \le 2L \big\|\big(\widehat{\psi}_i(t), \widehat{\psi}_i^{n+1}\big)\big\| \left(1 + \frac{1}{3^i}\right).$$

Hence, if $\lim_{i\to\infty}\big(\widehat{\psi}_i(t), \widehat{\psi}_i^{n+1}\big) = 0$, in view of (2.24) we obtain

$$\begin{aligned}\left(q, \psi(t) + \int_0^t \lambda(x_*(s))\, d\eta + \psi^0 \int_0^t \delta_M(x_*(s)) \frac{\partial \lambda(x_*(s))}{\partial x}\, ds\right) &= \lim_{i\to\infty} (\dot{\psi}_i(t), \widehat{\psi}_i(t)) \\ &= 0 \in \widehat{N}_{\operatorname{graph} F(\cdot)}(x_*(t), \dot{x}_*(t)).\end{aligned}$$

Therefore, in the case of $\lim_{i\to\infty}\big(\widehat{\psi}_i(t), \widehat{\psi}_i^{n+1}\big) = 0$ condition (2) holds.

So, suppose that $\lim_{i\to\infty}\big(\widehat{\psi}_i(t), \widehat{\psi}_i^{n+1}\big) \ne 0$. Then $\big(\widehat{\psi}_i(t), \widehat{\psi}_i^{n+1}\big) \ne 0$ for all sufficiently large i. By [6, Lemma 4] in this case for a.e. $(z, \widetilde{v})$, $z \in \mathbb{R}^n$, $\widetilde{v} = (v, v^{n+1})$, $v \in \mathbb{R}^n$, $v^{n+1} \in \mathbb{R}^1$, we have the inclusion

$$\left(-\frac{\partial}{\partial x} H\big(\widetilde{F}_i(t, x_i(t) + z), \widetilde{\psi}_i(t) + \|\widetilde{\psi}_i(t)\|\widetilde{v}\big), \widetilde{\psi}_i(t) + \|\widetilde{\psi}_i(t)\|\widetilde{v}\right) \in \Gamma_{\operatorname{graph} \widetilde{F}_i(t,\cdot)}\big(x_i(t) + z, \widetilde{p}_i(z, \widetilde{v})\big),$$

where

$$\widetilde{p}_i(z, \widetilde{v}) = \big(p_i(z, \widetilde{v}), p_i^{n+1}(z, \widetilde{v})\big) = \frac{\partial}{\partial(\psi, \psi^{n+1})} H\big(\widetilde{F}_i(t, x_i(t) + z), \widetilde{\psi}_i(t) + \|\widetilde{\psi}_i(t)\|\widetilde{v}\big)$$

is the unique supporting vector in the direction $\widetilde{\psi}_i(t) + \|\widetilde{\psi}_i(t)\|\widetilde{v}$ from the set $\Phi_i(t, x_i(t) + z) = \operatorname{conv} \widetilde{F}_i(t, x_i(t) + z)$, and $\Gamma_{\operatorname{graph} \widetilde{F}_i(t,\cdot)}(x_i(t) + z, \widetilde{p}_i(z, \widetilde{v}))$ is the subnormal cone (i.e., the polar cone of the contingent cone[2]) to the set $\operatorname{graph} \widetilde{F}_i(t, \cdot)$ at the point $(x_i(t) + z, \widetilde{p}_i(z, \widetilde{v}))$. Note that the uniqueness of $\widetilde{p}_i(z, \widetilde{v})$ implies the inclusion $\widetilde{p}_i(z, \widetilde{v}) \in \widetilde{F}_i(t, x_i(t) + z)$.

By the maximum condition (3.3), since the sets $\Phi_{i,j(i)}(t, x)$ are strictly convex, we have

$$\begin{aligned}(\dot{x}_i(t), \dot{y}_i(t)) = \int_{\mathbb{R}^{2n+1}} &\frac{\partial}{\partial(\psi, \psi^{n+1})} H\big(\widetilde{F}_i(t, x_i(t) + z), \widetilde{\psi}_i(t) + \|\widetilde{\psi}_i(t)\|\widetilde{v}\big) \omega_{j(i)}(z)\, \widetilde{\omega}_{j(i)}(\widetilde{v})\, dz\, d\widetilde{v} \\ &+ \frac{2(L+S)}{3^i} \frac{\widetilde{\psi}_i(t)}{\|\widetilde{\psi}_i(t)\|}.\end{aligned}$$

[2]Recall that the contingent cone to a set A at a point $\xi \in \overline{A}$ is the set $K_A(\xi) = \{v\colon \exists\, v_i \to v,\ \exists\, \alpha_i \to +0\colon \xi + \alpha_i v_i \in A\}$ (see [21]). Accordingly, $\Gamma_A(\xi) = K_A^*(\xi) = \{p\colon \langle p, v\rangle \le 0\ \forall\, v \in K_A(\xi)\}$.

This equality, together with the relation $\lim_{i\to\infty} \dot y_i(t) = 0$, implies

$$\lim_{i\to\infty} \int\limits_{\mathbb{R}^{2n+1}} \frac{\partial}{\partial\psi^{n+1}} H\big(\widetilde F_i(t, x_i(t) + z), \widetilde\psi_i(t) + \|\widetilde\psi_i(t)\|\widetilde v\big)\, \omega_{j(i)}(z)\, \widetilde\omega_{j(i)}(\widetilde v)\, dz\, d\widetilde v = 0.$$

By the definition of the set-valued mapping $\widetilde F_i(\cdot\,,\cdot)$ (see (2.1)), we have

$$\frac{\partial}{\partial\psi^{n+1}} H\big(\widetilde F_i(t, x_i(t) + z), \widetilde\psi_i(t) + \|\widetilde\psi_i(t)\|\widetilde v\big) = \bigg\| \frac{\partial}{\partial\psi} H\big(\widetilde F_i(t, x_i(t) + z), \widetilde\psi_i(t) + \|\widetilde\psi_i(t)\|\widetilde v\big) - \dot z_i(t) \bigg\|.$$

Therefore,

$$\lim_{i\to\infty} \int\limits_{\mathbb{R}^{2n+1}} \bigg\| \frac{\partial}{\partial\psi} H\big(\widetilde F_i(t, x_i(t) + z), \widetilde\psi_i(t) + \|\widetilde\psi_i(t)\|\widetilde v\big) - \dot z_i(t) \bigg\| \omega_{j(i)}(z)\, \widetilde\omega_{j(i)}(\widetilde v)\, dz\, d\widetilde v = 0.$$

Hence, we can choose the number $i(k) \ge k$ in such a way that

$$\int\limits_{\mathbb{R}^{2n+1}} \bigg\| \frac{\partial}{\partial\psi} H\big(\widetilde F_{i(k)}(t, x_{i(k)}(t) + z), \widetilde\psi_{i(k)}(t) + \|\widetilde\psi_{i(k)}(t)\|\widetilde v\big) - \dot z_{i(k)}(t) \bigg\| \omega_{j(i(k))}(z)\, \widetilde\omega_{j(i(k))}(\widetilde v)\, dz\, d\widetilde v \le \frac{1}{k^2}.$$

Setting

$$\Lambda_k = \bigg\{ (z, \widetilde v) \in \mathbb{R}^{2n+1}\colon\ \bigg\| \frac{\partial}{\partial\psi} H\big(\widetilde F_{i(k)}(t, x_{i(k)}(t) + z), \widetilde\psi_{i(k)}(t) + \|\widetilde\psi_{i(k)}(t)\|\widetilde v\big) - \dot z_{i(k)}(t) \bigg\| \ge \frac{1}{k} \bigg\},$$

we find that for any k

$$\int\limits_{\Lambda_k} \omega_{j(i(k))}(z)\, \widetilde\omega_{j(i(k))}(\widetilde v)\, dz\, d\widetilde v \le \frac{1}{k}. \tag{3.15}$$

Since the set-valued mapping $\widetilde F_i(t, \cdot)$ is Lipschitz continuous with the constant $2L$ independent of i and the sequence $\{\|\widetilde\psi_{i(k)}(t)\|\}_{i=1}^{\infty}$ is bounded, we can assume that

$$\int\limits_{\Lambda_k} \bigg\| \frac{\partial}{\partial x} H\big(\widetilde F_{i(k)}(t, x_{i(k)}(t) + z), \widetilde\psi_{i(k)}(t) + \|\widetilde\psi_{i(k)}(t)\|\widetilde v\big) \bigg\| \omega_{j(i(k))}(z)\, \widetilde\omega_{j(i(k))}(\widetilde v)\, dz\, d\widetilde v \le \frac{\kappa_1}{k}, \tag{3.16}$$

where $\kappa_1 \ge 0$ is a constant.

Consider the vector

$$\Big(-\frac{\partial}{\partial x} H\big(\widetilde F_{i(k)}(t, x_{i(k)}(t) + z), \widetilde\psi_{i(k)}(t) + \|\widetilde\psi_{i(k)}(t)\|\widetilde v\big), \widetilde\psi_{i(k)}(t) + \|\widetilde\psi_{i(k)}(t)\|\widetilde v \Big)$$
$$\in \Gamma_{\operatorname{graph} \widetilde F_{i(k)}(t,\cdot)}\big(x_{i(k)}(t) + z, \widetilde p_k(z, \widetilde v)\big),$$

where $(z, \widetilde v) \in \mathbb{R}^{2n+1} \setminus \Lambda_k$ and $\widetilde p_k(z, \widetilde v) = (p_k(z, \widetilde v), p_k^{n+1}(z, \widetilde v))$:

$$p_k(z, \widetilde v) = \frac{\partial}{\partial\psi} H\big(\widetilde F_{i(k)}(t, x_{i(k)}(t) + z), \widetilde\psi_{i(k)}(t) + \|\widetilde\psi_{i(k)}(t)\|\widetilde v\big),$$

$$p_k^{n+1}(z, \widetilde v) = \bigg\| \frac{\partial}{\partial\psi} H\big(\widetilde F_{i(k)}(t, x_{i(k)}(t) + z), \widetilde\psi_{i(k)}(t) + \|\widetilde\psi_{i(k)}(t)\|\widetilde v\big) - \dot z_{i(k)}(t) \bigg\|.$$

Since $\|\dot z_{i(k)}(t) - \dot x_*(t)\| \le 1/k$, we can assume (based on the definition of the sets Λ_k) that $\|p_k(z, \widetilde v) - \dot x_*(t)\| \le 2/k$ for a.e. $(z, \widetilde v) \in \mathbb{R}^{2n+1} \setminus \Lambda_k$.

Let

$$(u, w) \in K_{\operatorname{graph} F(\cdot)}\big(x_{i(k)}(t) + z, p_k(z, \widetilde v)\big), \qquad \|(u, w)\| = 1.$$

Then, by the definition of the contingent cone, there exists a number β, $|\beta|\le 1$, such that

$$(u,w,\beta)\in K_{\operatorname{graph}\widetilde{F}_{i(k)}(t,\cdot)}\bigl(x_{i(k)}(t)+z,\widetilde{p}_k(z,\widetilde{v})\bigr).$$

So, for a.e. $(z,\widetilde{v})\in\mathbb{R}^{2n+1}\setminus\Lambda_k$ such that $\widetilde{v}=(v,v^{n+1})\in\operatorname{supp}\widetilde{\omega}_{j(i(k))}(\cdot)$, we have

$$\Bigl\langle u,-\frac{\partial}{\partial x}H\bigl(\widetilde{F}_{i(k)}(t,x_{i(k)}(t)+z),\widetilde{\psi}_{i(k)}(t)+\|\widetilde{\psi}_{i(k)}(t)\|\widetilde{v}\bigr)\Bigr\rangle+\bigl\langle w,\widehat{\psi}_{i(k)}(t)+\|\widetilde{\psi}_{i(k)}(t)\|v\bigr\rangle$$
$$\le-\beta\bigl(\widehat{\psi}^{n+1}_{i(k)}+\|\widetilde{\psi}_{i(k)}(t)\|v^{n+1}\bigr)\le\bigl|\widehat{\psi}^{n+1}_{i(k)}\bigr|+\|\widetilde{\psi}_{i(k)}(t)\|\frac{1}{3^{j(i(k))}}.$$

Therefore, there exists a vector $\zeta_k(z,\widetilde{v})\in\mathbb{R}^{2n}$, $\|\zeta_k(z,\widetilde{v})\|\le\bigl|\widehat{\psi}^{n+1}_{i(k)}\bigr|+\|\widehat{\psi}_{i(k)}(t)\|\cdot 3^{-j(i(k))}$, such that

$$\Bigl(-\frac{\partial}{\partial x}H\bigl(F_{i(k)}(t,x_{i(k)}(t)+z),\widetilde{\psi}_{i(k)}(t)+\|\widetilde{\psi}_{i(k)}(t)\|\widetilde{v}\bigr),\widehat{\psi}_{i(k)}(t)+\|\widetilde{\psi}_{i(k)}(t)\|v\Bigr)$$
$$\in\Gamma_{\operatorname{graph}F(\cdot)}\bigl(x_{i(k)}(t)+z,\widetilde{p}_{i(k)}(z,\widetilde{v})\bigr)+\zeta_k(z,\widetilde{v}).$$

Fix an arbitrary $\varepsilon>0$. Since the equality

$$\widehat{N}_{\operatorname{graph}F(\cdot)}(x,y)=\operatorname*{Ls}_{(x_j,y_j)\xrightarrow{\operatorname{graph}F(\cdot)}(x,y)}\Gamma_{\operatorname{graph}F(\cdot)}(x_j,y_j)$$

holds (see [21]) and the graph of the set-valued mapping $\widehat{N}_{\operatorname{graph}F(\cdot)}(\cdot,\cdot)$ on the set $\operatorname{graph}F(\cdot)$ is closed, there exists a $\delta>0$ such that the inequality $\|(x,y)-(x_*(t),\dot{x}_*(t))\|\le\delta$ implies the inclusion

$$\widehat{N}^r_{\operatorname{graph}F(\cdot)}(x,y)\subset\widehat{N}^r_{\operatorname{graph}F(\cdot)}(x_*(t),\dot{x}_*(t))+\varepsilon\mathbb{B}^{2n},$$

where

$$\widehat{N}^r_{\operatorname{graph}F(\cdot)}(x,y)=r\mathbb{B}^{2n}\cap\widehat{N}_{\operatorname{graph}F(\cdot)}(x,y)$$

and r is a positive number such that for all k we have $2(2L+1)\|\widetilde{\psi}_{i(k)}(t)\|\le r$ for a.e. $(z,\widetilde{v})\in\mathbb{R}^{2n+1}\setminus\Lambda_k$ with $\widetilde{v}=(v,v^{n+1})\in\operatorname{supp}\widetilde{\omega}_{j(i(k))}(\cdot)$. Using the fact that the set-valued mapping $\widetilde{F}_i(t,\cdot)$ is Lipschitz continuous with the constant $2L$ independent of i, we obtain the inequality

$$\Bigl\|\Bigl(-\frac{\partial}{\partial x}H\bigl(F_{i(k)}(t,x_{i(k)}(t)+z),\widetilde{\psi}_{i(k)}(t)+\|\widetilde{\psi}_{i(k)}(t)\|\widetilde{v}\bigr),\widehat{\psi}_{i(k)}(t)+\|\widetilde{\psi}_{i(k)}(t)\|v\Bigr)\Bigr\|\le r.$$

Since the quantity $\widehat{\psi}^{n+1}_{i(k)}=-\psi^0_{i(k)}\gamma_{i(k)}$ is independent of z and $\widetilde{v}$ and $\lim_{i\to\infty}\widehat{\psi}^{n+1}_{i(k)}=0$, it follows that for all sufficiently large k the inclusion

$$\Bigl(-\frac{\partial}{\partial x}H\bigl(\widetilde{F}_{i(k)}(t,x_{i(k)}(t)+z),\widetilde{\psi}_{i(k)}(t)+\|\widetilde{\psi}_{i(k)}(t)\|\widetilde{v}\bigr),\widehat{\psi}_{i(k)}(t)+\|\widetilde{\psi}_{i(k)}(t)\|v\Bigr)$$
$$\in\widehat{N}_{\operatorname{graph}F(\cdot)}(x_*(t),\dot{x}_*(t))+\varepsilon\mathbb{B}^{2n}$$

holds on the set $\mathbb{R}^{2n+1}\setminus\Lambda_{i(k)}$. Therefore, on this set we have

$$-\frac{\partial}{\partial x}H\bigl(\widetilde{F}_{i(k)}(t,x_{i(k)}(t)+z),\widetilde{\psi}_{i(k)}(t)+\|\widetilde{\psi}_{i(k)}(t)\|\widetilde{v}\bigr)$$
$$\in\Bigl\{u\colon\ \bigl(u,\widehat{\psi}_{i(k)}(t)+\|\widetilde{\psi}_{i(k)}(t)\|v\bigr)\in\widehat{N}_{\operatorname{graph}F(\cdot)}(x_*(t),\dot{x}_*(t))+\varepsilon\mathbb{B}^{2n},\ \|u\|\le 4L\|\widetilde{\psi}_{i(k)}(t)\|\Bigr\}.$$

Using this inclusion, conditions (3.15) and (3.16), the condition

$$\lim_{i\to\infty}\widehat{\psi}_i(t)=\psi(t)+\int\limits_0^t\lambda(x_*(s))\,d\eta+\psi^0\int\limits_0^t\delta_M(x_*(s))\frac{\partial\lambda(x_*(s))}{\partial x}\,ds,$$

and the fact that the cone $\widehat{N}_{\operatorname{graph} F(\cdot)}(x_*(t),\dot{x}_*(t))$ is closed and $\varepsilon>0$ is arbitrary, from (3.9) we obtain inclusion (3.14). Condition (2) is proved.

Let us prove that condition (3) holds. For $i\in\mathbb{N}$ we define functions $p_i(\cdot)$ and $q_i(\cdot)$ as

$$p_i(t)=H_i\left(t,x_i(t),\left(\psi_i(t)+\int\limits_0^t\lambda(x_i(s))\eta_i(t)\,dt+\psi_i^0\int\limits_0^t\delta_{j(i)}(x_i(s))\frac{\partial\lambda(x_i(s))}{\partial x}\,ds,\widehat{\psi}_i^{n+1}\right)\right),$$

$$q_i(t)=\psi_i^0\lambda(x_i(t))\,\delta_{j(i)}(x_i(t)).$$

Then $h_i(t)=p_i(t)-q_i(t)$ for $t\in[0,T_i]$ (see (3.7)).

By Lemma 1, Theorem 1, the second condition in (3.4), and the constraint at time $t=0$ (see (2.4)), we have

$$\lim_{i\to\infty}p_i(0)=H(F(x_*(0)),\psi(0))\qquad\text{and}\qquad\lim_{i\to\infty}q_i(0)=\psi^0\lambda(x_*(0))\,\delta_M(x_*(0)),$$

which implies the boundedness of the sequence $\{|h_i(0)|\}_{i=1}^{\infty}$.

By condition (3.5) we have

$$\dot{h}_i(t)\overset{\text{a.e.}}{\in}\partial_t H_i\left(t,x_i(t),\left(\psi_i(t)+\int\limits_0^t\lambda(x_i(s))\eta_i(t)\,dt+\psi_i^0\int\limits_0^t\delta_{j(i)}(x_i(s))\frac{\partial\lambda(x_i(s))}{\partial x}\,ds,\widehat{\psi}_i^{n+1}\right)\right).$$

By virtue of (2.11) and [15, Theorems 2.7.2, 2.8.2], for a.e. t and arbitrary x and $\widetilde{\psi}=(\widehat{\psi},\widehat{\psi}^{n+1})$ we have the inclusions

$$\begin{aligned}\partial_t H_i(t,x,\widetilde{\psi})&\subset\int\limits_{\mathbb{R}^{n+1}}\partial_t\max_{u\in F(x+z)}\Big(\big\langle u,\widehat{\psi}+\|\widetilde{\psi}\|v\big\rangle+\|u-\dot{z}_i(t)\|\big(\widehat{\psi}^{n+1}+\|\widetilde{\psi}\|v^{n+1}\big)\Big)\\&\quad\times\omega_{j(i)}(z)\,\widetilde{\omega}_{j(i)}(\widetilde{v})\,dz\,d\widetilde{v}\\&\subset\int\limits_{\mathbb{R}^{2n+1}}\left\{\big(\widehat{\psi}^{n+1}+\|\widetilde{\psi}\|v^{n+1}\big)\bigcup_{u\in F(x+z)}\partial_t\|u-\dot{z}_i(t)\|\,\omega_{j(i)}(z)\,\widetilde{\omega}_{j(i)}(\widetilde{v})\right\}dz\,d\widetilde{v}.\end{aligned}$$

Hence, using the equality $\widehat{\psi}_i^{n+1}=-\psi_i^0\gamma_i$, we obtain

$$|\dot{h}_i(t)|\overset{\text{a.e.}}{\le}\int\limits_{\mathbb{R}^{2n+1}}K_i\left(\psi_i^0\gamma_i+\frac{1}{3^{j(i)}}\|\widetilde{\psi}_i(t)\|\right)\omega_{j(i)}(z)\,\widetilde{\omega}_{j(i)}(\widetilde{v})\,dz\,d\widetilde{v}.$$

In view of the conditions $\lim_{i\to\infty}K_i\gamma_i=0$ and $\lim_{i\to\infty}3^{-j(i)}K_i=0$ (see Theorem 1), this inequality implies that the sequence $\{\dot{h}_i(\cdot)\}_{i=1}^{\infty}$ converges to zero in $L^\infty([0,T_*],\mathbb{R}^1)$ as $i\to\infty$. Therefore, without loss of generality we can assume that

$$\lim_{i\to\infty}h_i(t)=h(t)\equiv H(F(x_*(0)),\psi(0))-\psi^0\lambda(x_*(0))\,\delta_M(x_*(0)),\qquad t\in[0,T_*].\tag{3.17}$$

According to the definition of η (see (3.10)) and the terminal constraint at time T_i (see (2.4)), we have

$$\begin{aligned} h(T_*) &= \lim_{i\to\infty} h_i(T_i) = \lim_{i\to\infty} \big(p_i(T_i) - \psi_i^0 \lambda(x_i(T_i))\,\delta_{j(i)}(x_i(T_i))\big) \\ &= H\left(F(x_*(T_*)), \psi(T_*) + \int_0^{T_*} \lambda(x_*(s))\,d\eta + \psi^0 \int_0^{T_*} \delta_M(x_*(s)) \frac{\partial \lambda(x_*(s))}{\partial x}\,ds\right) \\ &\quad - \psi^0 \lambda(x_*(T_*))\,\delta_M(x_*(T_*)). \end{aligned}$$

Thus, in view of (3.17), we have shown that condition (3) holds at time T_*.

Let $\tau \in [0, T_*)$ be a point of right approximate continuity of $\delta_M(x_*(\cdot))$. Then the function $\delta_M(x_*(\cdot))$ is right-continuous at the point τ along some measurable set E for which $\tau \in E$ is a right density point. By Corollary 1, we can assume without loss of generality that $\lim_{i\to\infty} \delta_{j(i)}(x_i(t)) = \delta_M(x_*(t))$ for a.e. $t \in E$. Therefore, there exists a sequence $\{\tau_k\}_{k=1}^{\infty}$, $\tau_k \in E$, such that $\tau_k \to \tau + 0$ and $\lim_{i\to\infty} \delta_{j(i)}(x_i(\tau_k)) = \delta_M(x_*(\tau_k))$. Hence, passing to a subsequence of $\{\tau_k\}_{k=1}^{\infty}$ if necessary, we can assume without loss of generality that $\lim_{i\to\infty} \delta_{j(i)}(x_i(\tau_i)) = \delta_M(x_*(\tau))$. Moreover, according to the definition of η (see (3.10)), we can assume that

$$\int_0^{\tau} \lambda(x_*(s))\,d\eta = \lim_{i\to\infty} \int_0^{\tau_i} \lambda(x_i(s))\eta_i(s)\,ds.$$

Therefore,

$$\begin{aligned} h(\tau) &= \lim_{i\to\infty} h_i(\tau_i) = \lim_{i\to\infty} \big(p_i(\tau_i) - \psi_i^0 \lambda(x_i(\tau_i))\,\delta_{j(i)}(x_i(\tau_i))\big) \\ &= H\left(F(x_*(\tau)), \psi(\tau) + \int_0^{\tau} \lambda(x_*(s))\,d\eta + \psi^0 \int_0^{\tau} \delta_M(x_*(s)) \frac{\partial \lambda(x_*(s))}{\partial x}\,ds\right) - \psi^0 \lambda(x_*(\tau))\,\delta_M(x_*(\tau)). \end{aligned}$$

In view of (3.17), this implies that condition (3) holds for every point $t \in [0, T_*)$ of right approximate continuity of $\delta_M(x_*(\cdot))$. Condition (3) is proved.

Since the cone of generalized normals $\widehat{N}_{\widetilde{M}_i}(\cdot)$ to the closed set $\widetilde{M}_i$, $i = 0, 1$, is upper semicontinuous and the generalized gradient $\widehat{\partial}\varphi(\cdot,\cdot,\cdot)$ of the locally Lipschitz continuous function $\varphi(\cdot,\cdot,\cdot)$ is also upper semicontinuous, condition (4) follows from Theorem 1 and the transversality condition (3.4).

Finally, let us prove that condition (5) holds. Suppose that $\psi^0 = 0$ and $\psi(0) = 0$. Let us show that then $\|\eta\| \neq 0$. Indeed, in this case condition (3.8) yields the equality

$$\lim_{i\to\infty} \psi_i^0 \int_0^{T_i} \left\| \frac{\partial \delta_{j(i)}(x_i(t))}{\partial x} \right\| dt = 1. \tag{3.18}$$

According to the definition of $\delta_i(\cdot)$ (see (1.5)), this means that

$$\lim_{i\to\infty} \psi_i^0 j(i) \int\limits_{\{t\in[0,T_i]\colon j(i)\rho(x_i(t),G)\le 1\}} \left\| \int_{\mathbb{R}^n} \frac{\partial \rho(x_i(t)+z, G)}{\partial x} \omega_{j(i)}(z)\,dz \right\| dt = 1. \tag{3.19}$$

Therefore, $\mathfrak{M} = \{t\colon x_*(t) \in \partial G\} \neq \varnothing$; indeed, otherwise by Theorem 1 for all sufficiently large i we would have $\partial\rho(x_i(t)+z, G)/\partial x \equiv 0$ for all $t \in [0, T_i]$ and $z \in \operatorname{supp}\omega_{j(i)}(\cdot)$, which would contradict (3.19).

Since $\operatorname{int} T_G(x) \neq \varnothing$, $x \in G$, there exists a $\delta > 0$ such that for every $\tau \in \mathfrak{M}$ one can find a vector $g(\tau) \in \mathbb{R}^n$, $\|g(\tau)\| = 1$, for which the following inclusion holds:

$$\big\{y\colon \|y - g(\tau)\| \le 2\delta\big\} \subset T_G(x_*(\tau)).$$

Therefore,

$$\langle g(\tau), \xi\rangle \le -2\delta\|\xi\|, \qquad \xi \in N_G(x_*(\tau)). \tag{3.20}$$

Next, for any $\tau \in \mathfrak{M}$ there exists an $\varepsilon(\tau) > 0$ such that

$$\frac{\partial\delta_{j(i)}(x_i(t))}{\partial x} \in N_{\delta/2}(\tau), \qquad t \in [\tau - \varepsilon(\tau), \tau + \varepsilon(\tau)] \cap [0, T_i],$$

for all sufficiently large i (see (3.11)). In this case, in view of (3.20), for all sufficiently large i we have

$$\left\langle g(\tau), \frac{\partial\delta_{j(i)}(x_i(t))}{\partial x}\right\rangle \le -\delta\left\|\frac{\partial\delta_{j(i)}(x_i(t))}{\partial x}\right\|, \qquad t \in [\tau - \varepsilon(\tau), \tau + \varepsilon(\tau)] \cap [0, T_i].$$

By Lemma 6, choosing sufficiently small numbers $0 < \varepsilon_1(\tau) \le \varepsilon(\tau)$ and $0 < \varepsilon_2(\tau) \le \varepsilon(\tau)$, we can assume that the equality

$$\lim_{i\to\infty} \int\limits_{\tau-\varepsilon_1(\tau)}^{\tau+\varepsilon_2(\tau)} \langle g(\tau), \eta_i(t)\rangle\, dt = \int\limits_{\tau-\varepsilon_1(\tau)}^{\tau+\varepsilon_2(\tau)} g(\tau)\, d\eta \tag{3.21}$$

holds for all $\tau \in \mathfrak{M}$. Further, since the set $\mathfrak{M}$ is compact, there exists a finite set of points $\{\tau_k\}_{k=1}^N$, $\tau_k \in \mathfrak{M}$, such that

$$\mathfrak{M} \subset \bigcup_{k=1}^N \big[\tau_k - \varepsilon_1(\tau_k), \tau_k + \varepsilon_2(\tau_k)\big]. \tag{3.22}$$

Then, using the fact that $\langle g(\tau_k), \partial\delta_{j(i)}(x_i(t))/\partial x\rangle$ is negative for all sufficiently large i and all $t \in [\tau_k - \varepsilon_1(\tau_k), \tau_k + \varepsilon_2(\tau_k)] \cap [0, T_i]$, $k = 1, \dots, N$, we obtain

$$\begin{aligned}
\psi_i^0 \sum_{k=1}^N \int\limits_{[\tau_k-\varepsilon_1(\tau_k),\tau_k+\varepsilon_2(\tau_k)]\cap[0,T_i]} &\left\langle g(\tau_k), \frac{\partial\delta_{j(i)}(x_i(t))}{\partial x}\right\rangle dt \\
&\le \psi_i^0 \int\limits_{\bigcup_{k=1}^N\{[\tau_k-\varepsilon_1(\tau_k),\tau_k+\varepsilon_2(\tau_k)]\cap[0,T_i]\}} \left\langle g(\tau_k), \frac{\partial\delta_{j(i)}(x_i(t))}{\partial x}\right\rangle dt \\
&\le -\delta\psi_i^0 \int\limits_{\bigcup_{k=1}^N\{[\tau_k-\varepsilon_1(\tau_k),\tau_k+\varepsilon_2(\tau_k)]\cap[0,T_i]\}} \left\|\frac{\partial\delta_{j(i)}(x_i(t))}{\partial x}\right\| dt.
\end{aligned}$$

By conditions (3.18) and (3.22), passing to the limit on the left- and right-hand sides of the last inequality, we obtain

$$\sum_{k=1}^N \int\limits_{[\tau_k-\varepsilon_1(\tau_k),\tau_k+\varepsilon_2(\tau_k)]} g(\tau_k)\, d\eta \le -\delta < 0.$$

Therefore, $\|\eta\| \neq 0$, and condition (5) is established. This completes the proof of Theorem 2. □

4. NONDEGENERACY CONDITIONS

The result proved in the previous section (Theorem 2) is similar in form to the necessary optimality conditions obtained in [6] in the case of an optimal control problem for a differential inclusion with state constraints. Just as in [6, Theorem 1] (as well as in a number of other versions of the Pontryagin maximum principle for problems with state constraints), a Borel measure η appears in the relations of Theorem 2, which may lead to a situation where any admissible trajectory satisfies these relations. In this case Theorem 2 does not provide any meaningful information, i.e., it degenerates. So conditions that would guarantee the nondegeneracy of Theorem 2 are of interest. For more details on the degeneration of necessary optimality conditions for problems with state constraints, see [2, § 2.4] and [3, Sect. 6] (see also [1, 4, 17, 18]).

The following result shows that in the abnormal case (i.e., for $\psi^0 = 0$) condition (3) of Theorem 2 implies condition (a) of [6, Theorem 1] (condition (b) of [3, Theorem 1]). This allows us to investigate the question of nondegeneracy of Theorem 2 in the same way as it was done in [3, 6] for problems with state constraints.

Lemma 7. *Suppose that an admissible trajectory $x_*(\cdot)$ is defined on an interval $[0,T_*]$, $T_* > 0$, and satisfies the conditions of Theorem 2 together with adjoint variables $\psi^0 = 0$, $\psi(\cdot)$ and a measure η. Then, for every $\tau \in [0,T_*]$, we have*

$$H\left(F(x_*(\tau)), \psi(\tau) + \int_0^\tau \lambda(x_*(s))\,d\eta\right) = H\left(F(x_*(\tau)), \psi(\tau) + \int_0^\tau \lambda(x_*(s))\,d\eta - \lambda(x_*(\tau))\eta(\tau)\right). \quad (4.1)$$

Proof. Since $\psi^0 = 0$, condition (3) of Theorem 2 at $t = T_*$ as well as at all points $t \in [0,T_*)$ of right approximate continuity of $\delta_M(x_*(\cdot))$ takes the form

$$H\left(F(x_*(t)), \psi(t) + \int_0^t \lambda(x_*(s))\,d\eta\right) = H(F(x_*(0)), \psi(0)). \quad (4.2)$$

Since the function $H(F(\cdot),\cdot)$ is continuous and the function $t \mapsto \psi(t) + \int_0^t \lambda(x_*(s))\,d\eta$ is right-continuous, from condition (4.2) we find that for any $\tau \in (0,T_*]$

$$H\left(F(x_*(\tau)), \psi(\tau) + \int_0^\tau \lambda(x_*(s))\,d\eta - \lambda(x_*(\tau))\eta(\tau)\right)$$
$$= \lim_{t\to\tau-0} H\left(F(x_*(t)), \psi(t) + \int_0^t \lambda(x_*(s))\,d\eta\right) = H(F(x_*(0)), \psi(0)).$$

Similarly, since the function $t \mapsto \psi(t) + \int_0^t \lambda(x_*(s))\,d\eta$ is right-continuous, it also follows from condition (4.2) that for any $\tau \in (0,T_*]$

$$H\left(F(x_*(\tau)), \psi(\tau) + \int_0^\tau \lambda(x_*(s))\,d\eta\right) = H(F(x_*(0)), \psi(0)).$$

Thus, (4.1) holds for every $\tau \in (0,T_*]$.

Further, since (4.1) holds for a.e. $t \in (0, T_*)$ and the function $H(F(\cdot), \cdot)$ is continuous, we have

$$H\big(F(x_*(0)), \psi(0) + \lambda(x_*(0))\eta(0)\big) = \lim_{t \to 0+0} H\left(F(x_*(t)), \psi(t) + \int\limits_0^t \lambda(x_*(s))\, d\eta \right)$$
$$= H(F(x_*(0)), \psi(0)).$$

Thus, condition (4.1) is satisfied at the point $\tau = 0$ as well. $\square$

Let us show that we can use this result to obtain nondegeneracy and pointwise nontriviality conditions for Theorem 2 similar to the well-known conditions for nondegeneracy and pointwise nontriviality of the maximum principle for problems with state constraints (see [3, 6]). The proofs are based on the use of condition (4.1) and are of technical character.

Following [3], we say that an admissible trajectory $x_*(\cdot)$ defined on $[0, T_*]$ is controllable at the endpoints $x_*(0)$ and $x_*(T_*)$ (with respect to the set G) if

$$H(F(x_*(0)), -g_0) > 0 \qquad \text{for any} \quad g_0 \in N_G(x_*(0)) \cap \big[-\widehat{N}_{\widetilde{M}_0}(x_*(0))\big], \quad g_0 \neq 0,$$

and

$$H(F(x_*(T_*)), g_1) > 0 \qquad \text{for any} \quad g_1 \in N_G(x_*(T_*)) \cap \big[-\widehat{N}_{\widetilde{M}_1}(x_*(T_*))\big], \quad g_1 \neq 0.$$

Theorem 3. *Let an admissible trajectory $x_*(\cdot)$ be controllable at the endpoints $x_*(0)$ and $x_*(T_*)$ and satisfy the conditions of Theorem 2. Then the following nondegeneracy condition holds:*

$$\psi^0 + \operatorname{meas}\left\{ t \in [0, T_*] \colon\ \psi(t) + \int\limits_0^t \lambda(x_*(s))d\eta + \psi^0 \int\limits_0^t \delta_M(x_*(s)) \frac{\partial \lambda(x_*(s))}{\partial x}\, ds \neq 0 \right\} > 0.$$

Proof. Suppose that the assertion of the theorem is false. Then

$$\psi^0 = 0 \qquad \text{and} \qquad \psi(t) + \int\limits_0^t \lambda(x_*(s))d\eta = 0 \quad \text{for a.e.} \quad t \in [0, T_*]. \tag{4.3}$$

Since $\psi^0 = 0$, it follows from Lemma 7 that condition (4.1) holds for any $t \in [0, T_*]$. This condition coincides with the measure-jump condition (b) in [3, Theorem 1], which allows us to follow the scheme of the proof of Theorem 2 in [3]. Indeed, by condition (2) of Theorem 2, we have

$$\|\dot{\psi}(t)\| \le \kappa \left\| \psi(t) + \int\limits_0^t \lambda(x_*(s))\, d\eta \right\|$$

for a.e. $t \in [0, T_*]$, where $\kappa \ge 0$ is a constant. Therefore, $\psi(t) \equiv \psi(0)$ and

$$h(t) = H\left(F(x_*(t)), \psi(t) + \int\limits_0^t d\eta \right) \equiv H(F(x_*(0)), \psi(0)) = 0, \qquad t \in [0, T_*].$$

Further, using (4.3) we obtain $\psi(0) + \lambda(x_*(0))\eta(0) = 0$. If $x_*(0) \in M$, then $\eta(0) = 0$, so $\psi(0) = 0$. Suppose that $x_*(0) \in G$. Then, by condition (1) of Theorem 2, the inclusion $\eta(0) \in N_G(x_*(0))$ holds, and hence $\psi(0) = -\lambda(x_*(0))\eta(0) \in -N_G(x_*(0))$. On the other hand, by condition (4) of Theorem 2 we have $\psi(0) \in \widehat{N}_{\widetilde{M}}(x_*(0))$. Therefore, due to the controllability of $x_*(\cdot)$ at the point $x_*(0)$, we again obtain the equalities $\psi(0) = \eta(0) = 0$. Using them together with the identity $\psi(t) \equiv \psi(0)$, $t \in [0, T_*]$, and the second equality in (4.3), which is valid for a.e. $t \in [0, T_*]$, we find that the measure η vanishes on $[0, T_*)$.

Consider the point $t = T_*$. Using equality (4.1) at time T_*, we derive the following equality: $H(F(x_*(T_*)), \lambda(x_*(T_*))\eta(T_*)) = 0$. Further, applying condition (1) of Theorem 2, we find that $\lambda(x_*(T_*))\eta(T_*) \in N_G(x_*(T_*))$, and by condition (4) we obtain $\lambda(x_*(T_*))\eta(T_*) \in -\widehat{N}_{\widetilde{M}_1}(x_*(T_*))$. Since the trajectory $x_*(\cdot)$ is controllable at $x_*(T_*)$, we therefore have $\eta(x_*(T_*)) = 0$. Thus, we have shown that $\psi(0) = 0$ and the measure η is zero, which contradicts the nontriviality condition (5) of Theorem 2. Hence, condition (4.3) leads to a contradiction. $\square$

Theorem 4. *Let an admissible trajectory $x_*(\cdot)$ be controllable at the endpoints $x_*(0)$ and $x_*(T_*)$ and satisfy the conditions of Theorem 2. Suppose in addition that*

$$H\big(F(x_*(t)), (-1)^i g\big) > 0 \qquad \forall g \in N_G(x_*(t)), \quad t \in (0, T_*), \quad i = 1, 2. \tag{4.4}$$

Then

$$\psi^0 + \left\| \psi(t) + \int\limits_0^t \lambda(x_*(s))\, d\eta + \psi^0 \int\limits_0^t \delta_M(x_*(s)) \frac{\partial \lambda(x_*(s))}{\partial x}\, ds \right\| > 0, \qquad t \in (0, T_*). \tag{4.5}$$

Proof. Consider the set

$$\Delta = \left\{ t \in (0, T_*)\colon\ \psi(t) + \int\limits_0^t \lambda(x_*(s))\, d\eta = 0 \right\}.$$

Suppose that condition (4.5) is violated. Then $\psi^0 = 0$ and $\Delta \neq \varnothing$. Therefore, by Lemma 7, condition (4.1) holds for any $t \in [0, T_*]$. Further, repeating the arguments used in the proof of Theorem 4 in [3] and employing condition (4.1), the controllability of $x_*(\cdot)$ at the endpoints $x_*(0)$ and $x_*(T_*)$, condition (4.4), and the definition of η, we can show that the set Δ is both open and closed with respect to the interval $(0, T_*)$. Therefore, $\Delta = (0, T_*)$. Thus, in this case, we have

$$\psi^0 + \left\| \psi(t) + \int\limits_0^t \lambda(x_*(s))\, d\eta + \psi^0 \int\limits_0^t \delta_M(x_*(s)) \frac{\partial \lambda(x_*(s))}{\partial x}\, ds \right\| = 0$$

for all $t \in (0, T_*)$. However, this contradicts the assertion of Theorem 3. $\square$

5. EXAMPLE

Consider the following optimal control problem (P_λ):

$$J_\lambda(T, x(\cdot)) = T + \lambda \int\limits_0^T \delta_M(x(t))\, dt \to \min, \tag{5.1}$$

$$\begin{cases} \dot{x}^1(t) = u^1(t), \\ \dot{x}^2(t) = u^2(t), \end{cases} \qquad u(t) = (u^1(t), u^2(t)) \in U = \big\{ u \in \mathbb{R}^2\colon\ \|u\| \le 1 \big\}, \tag{5.2}$$

$$x(0) = x_0 = (x^1(0), x^2(0)) = (-1, 0), \qquad x(T) = x_1 = (x^1(T), x^2(T)) = (1, 0).$$

Here $T > 0$ is the free terminal time of the control process, $\lambda > 0$, and

$$M = \big\{ x \in \mathbb{R}^2\colon\ \|x\| < 1 \big\}.$$

In this case, $G = \mathbb{R}^2 \setminus M = \{x \in \mathbb{R}^2\colon\ \|x\| \ge 1\}$. Clearly, for any $\lambda > 0$ problem (P_λ) is a particular case of problem (P). Since the integral part of the functional (5.1) is lower semicontinuous (see [8, Theorem 1]), there exists an optimal admissible trajectory $x_\lambda(\cdot)$ in problem (P_λ). Let $T_\lambda > 0$ be the corresponding optimal time. It is easy to see that $T_\lambda \le \pi$.

The optimal trajectory $x_\lambda(\cdot)$ satisfies the conditions of Theorem 2. Therefore, there exists a number $\psi_\lambda^0 \ge 0$, an absolutely continuous function $\psi_\lambda(\cdot) = (\psi_\lambda^1(\cdot), \psi_\lambda^2(\cdot))$, and a regular two-dimensional Borel measure $\eta_\lambda = (\eta_\lambda^1, \eta_\lambda^2)$ for which conditions (1)–(5) of Theorem 2 hold.

Since the right-hand side of (5.2) is independent of the state variable x and $F(x) \equiv U = \mathbb{B}^2$, $x \in \mathbb{R}^2$, the refined Euler–Lagrange inclusion (see condition (2) of Theorem 2) in this case reads

$$\dot\psi_\lambda(t) = 0, \qquad \psi_\lambda(t) + \lambda \int_0^t d\eta_\lambda \in N_U(\dot x_\lambda(t)).$$

Here $N_U(\dot x_\lambda(t))$ is the normal cone (in the sense of convex analysis) to the unit ball $U = \mathbb{B}^2$ at the point $\dot x_\lambda(t) \in U$. Therefore, for a.e. $t \in [0, T_\lambda]$ we have the equalities

$$\dot\psi_\lambda(t) = 0, \qquad \left\langle \psi_\lambda(t) + \lambda \int_0^t d\eta_\lambda, \dot x_\lambda(t) \right\rangle = \left\| \psi_\lambda(t) + \lambda \int_0^t d\eta_\lambda \right\|. \tag{5.3}$$

Further, the optimal trajectory $x_\lambda(\cdot)$ is controllable at the endpoints x_0 and x_1 (with respect to the set G) and condition (4.4) holds. Hence, by Theorem 4, the trajectory $x_\lambda(\cdot)$, the adjoint variables ψ_λ^0 and $\psi_\lambda(\cdot)$ ($\psi_\lambda(t) \equiv \psi_\lambda$), and the measure η_λ satisfy the pointwise nontriviality condition (4.5):

$$\psi_\lambda^0 + \left\| \psi_\lambda + \lambda \int_0^t d\eta_\lambda \right\| > 0, \qquad t \in (0, T_\lambda). \tag{5.4}$$

Let us show that the following stronger condition holds in problem (P_λ):

$$\psi_\lambda^0 > 0 \qquad \text{and} \qquad \left\| \psi_\lambda + \lambda \int_0^t d\eta_\lambda \right\| > 0, \qquad t \in [0, T_\lambda]. \tag{5.5}$$

We begin with showing that $\psi_\lambda^0 > 0$. Indeed, suppose that $\psi_\lambda^0 = 0$. Then, by conditions (3) and (4) of Theorem 2, for $t = T_\lambda$ we have

$$\|\psi_\lambda\| = H(F(x_0), \psi_\lambda) = H\left(F(x_1), \psi_\lambda + \lambda \int_0^{T_\lambda} d\eta_\lambda \right) = \left\| \psi_\lambda + \lambda \int_0^{T_\lambda} d\eta_\lambda \right\| = 0.$$

In addition, for $\psi_\lambda^0 = 0$, by Lemma 7, equality (4.1) holds for any $t \in [0, T_\lambda)$. Since the function $t \mapsto \int_0^t d\eta_\lambda$ is right-continuous on $[0, T_\lambda)$, we then obtain the equality

$$\left\| \psi_\lambda + \lambda \int_0^t d\eta_\lambda \right\| = 0$$

for any $t \in [0, T_\lambda]$, which, combined with the assumption $\psi_\lambda^0 = 0$, contradicts equality (5.4). Therefore, for any $\lambda > 0$, Theorem 2 holds for problem (P_λ) with $\psi_\lambda^0 > 0$. The first condition in (5.5) is proved.

Let us prove the second condition in (5.5). By conditions (3) and (4) of Theorem 2, we have

$$\left\| \psi_\lambda + \lambda \int_0^t d\eta_\lambda \right\| = \left\| \psi_\lambda + \lambda \int_0^{T_\lambda} d\eta_\lambda \right\| = \|\psi_\lambda\| = \psi_\lambda^0 > 0 \tag{5.6}$$

for a.e. $t \in [0, T_\lambda]$. Since the function $t \mapsto \int_0^t d\eta_\lambda$ is right-continuous on $[0, T_\lambda)$, this implies the second condition in (5.5).

Without loss of generality, we will assume that $\psi_\lambda^0 = 1$. Then, in view of (5.5) and (5.6), we have

$$\left\| \psi_\lambda + \lambda \int_0^t d\eta_\lambda \right\| \equiv 1, \qquad t \in [0, T_\lambda]. \tag{5.7}$$

According to the second equality in (5.3), this means that

$$\dot{x}_\lambda(t) \overset{\text{a.e.}}{=} \psi_\lambda + \lambda \int_0^t d\eta_\lambda, \qquad \|\dot{x}_\lambda(t)\| \overset{\text{a.e.}}{=} 1. \tag{5.8}$$

Thus, the optimal trajectory $x_\lambda(\cdot)$ always has the maximum (unit) velocity and may only lie on the boundary $\partial M = \{x \in \mathbb{R}^2 \colon \|x\| = 1\}$ of the open ball $M = \{x \in \mathbb{R}^2 \colon \|x\| < 1\}$ or on straight line segments connecting its boundary points. The set $\mathfrak{M}_\lambda = \{t \in [0, T_\lambda] \colon x_\lambda(t) \in \partial M\}$ is compact. Therefore, there exist at most countably many nonoverlapping open intervals $\{(\tau_{2i-1}, \tau_{2i})\}_{i=1}^\infty$, $(\tau_{2i-1}, \tau_{2i}) \subset [0, T_\lambda]$, such that $\|x_\lambda(t_{2i-1})\| = \|x_\lambda(t_{2i})\| = 1$ and on each interval $[\tau_{2i-1}, \tau_{2i}] \subset [0, T_\lambda]$ the trajectory $x_\lambda(\cdot)$ follows the straight line segment with vertices $x_\lambda(t_{2i-1})$ and $x_\lambda(t_{2i})$, $i \in \mathbb{N}$, in M.

Let $\Delta = T_\lambda - \sum_{i=1}^\infty (\tau_{2i} - \tau_{2i-1}) \ge 0$. Set

$$\widetilde{\tau}_1 = 0, \qquad \widetilde{\tau}_2 = \tau_2 - \tau_1, \qquad \widetilde{\tau}_{2i-1} = \widetilde{\tau}_{2i-2}, \qquad \widetilde{\tau}_{2i} = \widetilde{\tau}_{2i-1} + (\tau_{2i} - \tau_{2i-1}), \qquad i \ge 2.$$

We define another trajectory $\widetilde{x}_\lambda(\cdot)$ of system (5.2) on $[0, T_\lambda]$ as follows. For $t \in [\widetilde{\tau}_1, \widetilde{\tau}_2]$, the trajectory $\widetilde{x}_\lambda(\cdot)$ follows with unit velocity the straight line segment connecting the points $\widetilde{x}_\lambda(0) = x_0$ and $\widetilde{x}_\lambda(\widetilde{\tau}_2) \in \partial M$ in the same direction as $x_\lambda(\cdot)$ on the interval $[\tau_1, \tau_2]$. Then, on each of the adjacent intervals $[\widetilde{\tau}_{2i-1}, \widetilde{\tau}_{2i}]$, $i \ge 2$, the trajectory $\widetilde{x}_\lambda(\cdot)$ follows with unit velocity during time $\tau_{2i} - \tau_{2i-1}$ the straight line segment connecting the points $\widetilde{x}_\lambda(\tau_{2i-1}) \in \partial M$ and $\widetilde{x}_\lambda(\tau_{2i}) \in \partial M$ in the same direction as $x_\lambda(\cdot)$ on the interval $[\tau_{2i-1}, \tau_{2i}]$. On the final interval $[T_\lambda - \Delta, T_\lambda]$, the trajectory $\widetilde{x}_\lambda(\cdot)$ lies on the circle ∂M, where the motion occurs with unit velocity from the point $\widetilde{x}(T_\lambda - \Delta) = \lim_{i \to \infty} \widetilde{x}(\widetilde{\tau}_i) \in \partial M$ to the terminal point $\widetilde{x}(T_\lambda) = x_1$.

It is easy to see that the trajectory $\widetilde{x}_\lambda(\cdot)$ thus constructed is admissible in problem (P_λ). It transfers system (5.2) from the point x_0 to the point x_1 in the same time T_λ as $x_\lambda(\cdot)$. The time during which $\widetilde{x}_\lambda(\cdot)$ belongs to the set M coincides with that for $x_\lambda(\cdot)$ and is equal to $T_\lambda - \Delta$. This implies the optimality of the trajectory $\widetilde{x}_\lambda(\cdot)$. Clearly, since $\widetilde{x}_\lambda(\cdot)$ is optimal, the system of intervals $\{(\widetilde{\tau}_{2i-1}, \widetilde{\tau}_{2i})\}_{i=1}^\infty$ cannot contain more than one nonzero interval $(\widetilde{\tau}_1, \widetilde{\tau}_2)$ of motion in the set M. Indeed, if $\widetilde{x}_\lambda(\cdot)$ has two adjacent nonzero intervals $(\widetilde{\tau}_1, \widetilde{\tau}_2)$ and $(\widetilde{\tau}_2, \widetilde{\tau}_3)$ of motion in M, then we can replace the motion along the segments $[\widetilde{x}_\lambda(\widetilde{\tau}_1), \widetilde{x}_\lambda(\widetilde{\tau}_2)]$ and $[\widetilde{x}_\lambda(\widetilde{\tau}_2), \widetilde{x}_\lambda(\widetilde{\tau}_3)]$ with the motion with unit velocity along the segment $[\widetilde{x}_\lambda(\widetilde{\tau}_1), \widetilde{x}_\lambda(\widetilde{\tau}_3)]$, which yields a new admissible trajectory $\widehat{x}_\lambda(\cdot)$ transferring system (5.2) from x_0 to x_1 in time $\widehat{T}_\lambda < T_\lambda$ (by the triangle inequality), with the same time Δ of motion along the boundary of M as $\widetilde{x}_\lambda(\cdot)$; however, this contradicts the optimality of $\widetilde{x}_\lambda(\cdot)$.

Consider the optimal trajectory $\widetilde{x}_\lambda(\cdot)$ with the interval $[0, T_\lambda - \Delta]$, $\Delta \in [0, \pi]$, of motion with unit velocity along a straight line from the point $x_0 \in \partial M$ to the point $\widetilde{x}(T_\lambda - \Delta) \in \partial M$ and with the interval $[T_\lambda - \Delta, T_\lambda]$ of motion with unit velocity along an arc of the circle ∂M from the point $\widetilde{x}(T_\lambda - \Delta)$ to the point x_1. For definiteness, we will assume that the motion is clockwise. By a direct calculation, we find the corresponding value of the functional as a function of the parameter $\Delta \in [0, \pi]$:

$$J_\lambda(T_\lambda, \widetilde{x}_\lambda(\cdot)) = J(\Delta) = \Delta + 2(1 + \lambda) \cos\frac{\Delta}{2}.$$

Since for $\Delta \in (0,\pi)$ we have

$$\frac{d^2}{d\Delta^2}J(\Delta) = -\frac{(1+\lambda)}{2}\sin\frac{\Delta}{2} < 0,$$

the function $J(\cdot)$ is strictly concave on $[0,\pi]$. Therefore, it can take its minimum value only at the extreme points $\Delta = 0$ and $\Delta = \pi$. Hence,

$$J_\lambda(T_\lambda, \widetilde{x}_\lambda(\cdot)) = \min\{2(1+\lambda), \pi\}.$$

Thus, if $0 < \lambda < \pi/2 - 1$, we have $J_\lambda(T_\lambda, \widetilde{x}_\lambda(\cdot)) = 2(1+\lambda) < \pi$ and problem (P_λ) has a unique optimal trajectory

$$x_\lambda(t) = (-1+t, 0), \qquad t \in [0, T_\lambda^1], \qquad T_\lambda^1 = 2.$$

If $\lambda = \pi/2 - 1$, then $J_\lambda(T_\lambda, \widetilde{x}_\lambda(\cdot)) = \pi$ and there are three optimal trajectories in (P_λ):

$$x_{\lambda,1}(t) = (-1+t, 0), \quad t \in [0,2], \qquad x_{\lambda,2}(t) = (\cos(\pi - t), \sin(\pi - t)), \quad t \in [0,\pi],$$

and the trajectory

$$x_{\lambda,3}(t) = (\cos(\pi - t), -\sin(\pi - t)), \qquad t \in [0,\pi],$$

which is symmetric to $x_{\lambda,2}(\cdot)$ with respect to the x^1-axis. If $\lambda > \pi/2 - 1$, then $J_\lambda(T_\lambda, \widetilde{x}_\lambda(\cdot)) = \pi$ and there are two (symmetric with respect to the x^1-axis) optimal trajectories

$$x_{\lambda,2}(t) = (\cos(\pi - t), \sin(\pi - t)), \qquad x_{\lambda,3}(t) = (\cos(\pi - t), -\sin(\pi - t)), \qquad t \in [0,\pi].$$

Consider the case $\lambda = \pi/2 - 1$ and the first optimal trajectory in this case, $x_{\lambda,1}(t) = (-1+t, 0)$, $t \in [0, T_\lambda^1]$, $T_\lambda^1 = 2$. The trajectory $x_{\lambda,1}(\cdot)$ corresponds to the motion with unit velocity along the coordinate axis x^1 from the point $x_0 = (-1,0)$ to the point $x_1 = (1,0)$. In view of (5.6) and (5.8), we have $\|\psi_\lambda\| = 1$ and $\psi_\lambda + \lambda\int_0^t d\eta_\lambda = (1,0)$ for a.e. $t \in [0, T_\lambda^1]$. Condition (1) of Theorem 2 implies that the measure η_λ can take nonzero values only at time $t = 0$ and $t = T_\lambda$. Since the function $t \mapsto \int_0^t d\eta_\lambda$ is right-continuous at zero, it follows that $\psi_\lambda + \lambda\eta_\lambda(0) = (1,0)$. Further, condition (1) of Theorem 2 implies that the atom $\eta_\lambda(0)$ has the form $\eta_\lambda(0) = \alpha_0(1,0)$, $\alpha_0 \geq 0$. Therefore, $\psi_\lambda^2 = 0$. Since we have $\psi_\lambda^1 + (\pi/2 - 1)\alpha_0 = 1$ and $|\psi_\lambda^1| = \|\psi_\lambda\| = 1$, only the following two cases are possible:

(i) $\psi_\lambda = (1,0)$ and $\alpha_0 = 0$;

(ii) $\psi_\lambda = (-1,0)$ and $\alpha_0 = 4/(\pi - 2)$.

Similarly, condition (1) of Theorem 2 implies that the atom $\eta_\lambda(T_\lambda^1)$ has the form $\eta_\lambda(T_\lambda^1) = \alpha_1(-1,0)$, $\alpha_1 \geq 0$. Using condition (5.7) for $t = T_\lambda$, we obtain

$$\big\|\psi_\lambda + \lambda\eta_\lambda(0) + \lambda\eta_\lambda(T_\lambda)\big\| = \Big|1 - \Big(\frac{\pi}{2} - 1\Big)\alpha_1\Big| = 1.$$

Therefore, either $\alpha_1 = 0$ or $\alpha_1 = 4/(\pi - 2)$, regardless of which case (i) or (ii) takes place for $t = 0$. So, for the first optimal trajectory $x_{\lambda,1}(\cdot)$ in the case $\lambda = \pi/2 - 1$, Theorem 2 allows nonzero atoms $\eta_\lambda(0) = \alpha(1,0)$ at the initial moment $t = 0$ and $\eta_\lambda(T_\lambda) = -\alpha(1,0)$ at the final moment $T_\lambda = 2$, where $\alpha = 4/(\pi - 2)$. However, Theorems 3 and 4 show that this does not lead to the degeneration of the obtained necessary optimality conditions.

Consider now the second optimal trajectory $x_{\lambda,2}(t) = (\cos(\pi - t), \sin(\pi - t))$, $t \in [0, T_\lambda^2]$, $T_\lambda^2 = \pi$, in the case $\lambda = \pi/2 - 1$. The trajectory $x_{\lambda,2}(\cdot)$ corresponds to the motion with unit velocity along the unit circle (boundary of M) clockwise from the point $x_0 = (-1,0)$ to the point $x_1 = (1,0)$. In view of (5.8), we have $\psi_\lambda + \lambda\int_0^t d\eta_\lambda = (\sin(\pi - t), -\cos(\pi - t))$ for a.e. $t \in [0,\pi]$. Since the function $t \to \int_0^t d\eta_\lambda$ is right-continuous at zero, we then obtain $\psi_\lambda + \lambda\eta_\lambda(0) = (0,-1)$. Condition (1) of

Theorem 2 implies that the atom $\eta_\lambda(0)$ has the form $\eta_\lambda(0) = \alpha_0(1,0)$, $\alpha_0 \geq 0$. Therefore, $\psi_\lambda^2 = -1$. In view of (5.6) we have $\|\psi_\lambda\| = 1$. Hence, $\psi_\lambda^1 = 0$. Further, condition (5.7) yields the equality

$$\left\|\psi_\lambda + \frac{\pi-2}{2}(\alpha_0,0)\right\| = \left\|\left(\frac{\pi-2}{2}\alpha_0,-1\right)\right\| = 1.$$

This implies $\alpha_0 = 0$. Similarly, at time $T_\lambda^2 = \pi$ we have $\eta_\lambda(\pi) = \alpha_1(-1,0)$, $\alpha_1 \geq 0$. Then, recalling (5.7), we obtain

$$\left\|\psi_\lambda - \frac{\pi-2}{2}(\alpha_1,0)\right\| = \left\|\left(\frac{\pi-2}{2}\alpha_1,1\right)\right\| = 1.$$

Therefore, we have the equality $\alpha_1 = 0$. Thus, for the second optimal trajectory $x_{\lambda,2}(\cdot)$ in the case $\lambda = \pi/2 - 1$, Theorem 2 holds with a measure η_λ that has only a continuous part. By virtue of (5.8), the measure η_λ is absolutely continuous with respect to the Lebesgue measure μ on $[0,\pi]$, and

$$\frac{d\eta_\lambda^1}{d\mu} = -\cos(\pi - t), \qquad \frac{d\eta_\lambda^2}{d\mu} = \sin(\pi - t), \qquad t \in [0,\pi],$$

where $d\eta_\lambda^i/d\mu$ is the Radon–Nikodym derivative of the ith component η_λ^i of η_λ with respect to the Lebesgue measure μ, $i = 1,2$. Thus, for the second optimal trajectory $x_{\lambda,2}(\cdot)$ in the case $\lambda = \pi/2 - 1$, Theorem 2 holds with a nonzero measure η_λ that is absolutely continuous with respect to the Lebesgue measure.

The third optimal trajectory $x_{\lambda,3}(\cdot)$ in the case $\lambda = \pi/2 - 1$ is completely similar to the trajectory $x_{\lambda,2}(\cdot)$, so the same conclusions apply to it.

ACKNOWLEDGMENTS

The author is grateful to K. O. Besov for a number of useful comments.

FUNDING

This work is supported by the Russian Science Foundation under grant 19-11-00223.

REFERENCES

1. A. V. Arutyunov, "Perturbations of extremal problems with constraints and necessary optimality conditions," J. Sov. Math. **54** (6), 1342–1400 (1991) [transl. from Itogi Nauki Tekh., Ser. Mat. Anal. **27**, 147–235 (1989)].
2. A. V. Arutyunov, *Optimality Conditions: Abnormal and Degenerate Problems* (Faktorial, Moscow, 1997; Kluwer, Dordrecht, 2000).
3. A. V. Arutyunov and S. M. Aseev, "Investigation of the degeneracy phenomenon of the maximum principle for optimal control problems with state constraints," SIAM J. Control Optim. **35** (3), 930–952 (1997).
4. A. V. Arutyunov, D. Yu. Karamzin, and F. L. Pereira, "The maximum principle for optimal control problems with state constraints by R. V. Gamkrelidze: Revisited," J. Optim. Theory Appl. **149** (3), 474–493 (2011).
5. S. M. Aseev, "Methods of regularization in nonsmooth problems of dynamic optimization," J. Math. Sci. **94** (3), 1366–1393 (1999).
6. S. M. Aseev, "Extremal problems for differential inclusions with state constraints," Proc. Steklov Inst. Math. **233**, 1–63 (2001) [transl. from Tr. Mat. Inst. Steklova **233**, 5–70 (2001)].
7. S. M. Aseev, "Optimization of dynamics of a control system in the presence of risk factors," Tr. Inst. Mat. Mekh. (Ekaterinburg) **23** (1), 27–42 (2017).
8. S. M. Aseev, "On an optimal control problem with discontinuous integrand," Proc. Steklov Inst. Math. **304** (Suppl. 1), S3–S13 (2019) [transl. from Tr. Inst. Mat. Mekh. (Ekaterinburg) **24** (1), 15–26 (2018)].
9. S. M. Aseev, "An optimal control problem with a risk zone," in *Large-Scale Scientific Computing: 11th Int. Conf., LSSC 2017, Sozopol, 2017* (Springer, Cham, 2018), Lect. Notes Comput. Sci. **10665**, pp. 185–192.
10. S. M. Aseev, "A problem of dynamic optimization in the presence of dangerous factors," in *Stability, Control and Differential Games: Proc. Int. Conf. SCDG2019, Yekaterinburg, 2019* (Springer, Cham, 2020), Lect. Notes Control Inf. Sci. – Proc., pp. 273–281.

11. S. M. Aseev and A. I. Smirnov, "The Pontryagin maximum principle for the problem of optimally crossing a given domain," Dokl. Math. **69** (2), 243–245 (2004) [transl. from Dokl. Akad. Nauk **395** (5), 583–585 (2004)].
12. S. M. Aseev and A. I. Smirnov, "Necessary first-order conditions for optimal crossing of a given region," Comput. Math. Model. **18** (4), 397–419 (2007) [transl. from Nelinein. Din. Upr. **4**, 179–204 (2004)].
13. L. Cesari, *Optimization—Theory and Applications: Problems with Ordinary Differential Equations* (Springer, New York, 1983).
14. F. H. Clarke, "The Euler–Lagrange differential inclusion," J. Diff. Eqns. **19**, 80–90 (1975).
15. F. H. Clarke, *Optimization and Nonsmooth Analysis* (J. Wiley & Sons, New York, 1983).
16. F. Clarke, *Functional Analysis, Calculus of Variations and Optimal Control* (Springer, London, 2013), Grad. Texts Math. **264**.
17. M. M. A. Ferreira and R. B. Vinter, "When is the maximum principle for state constrained problems nondegenerate?," J. Math. Anal. Appl. **187** (2), 438–467 (1994).
18. F. A. C. C. Fontes and H. Frankowska, "Normality and nondegeneracy for optimal control problems with state constraints," J. Optim. Theory Appl. **166** (1), 115–136 (2015).
19. A. D. Ioffe and V. M. Tikhomirov, *Theory of Extremal Problems* (Nauka, Moscow, 1974; North-Holland, Amsterdam, 2009).
20. B. Sh. Mordukhovich, "Maximum principle in the problem of time optimal response with nonsmooth constraints," J. Appl. Math. Mech. **40**, 960–969 (1976) [transl. from Prikl. Mat. Mekh. **40** (6), 1014–1023 (1976)].
21. B. Sh. Mordukhovich, *Approximation Methods in Problems of Optimization and Control* (Nauka, Moscow, 1988) [in Russian].
22. B. Sh. Mordukhovich, "Optimal control of difference, differential, and differential–difference inclusions," J. Math. Sci. **100** (6), 2613–2632 (2000) [transl. from Itogi Nauki Tekh., Ser.: Sovrem. Mat. Prilozh., Temat. Obz. **61**, 33–65 (1999)].
23. I. P. Natanson, *Theory of Functions of a Real Variable* (Nauka, Moscow, 1974; Dover Publ., Mineola, NY, 2016).
24. E. S. Polovinkin and G. V. Smirnov, "Time-optimal problem for differential inclusions," Diff. Eqns. **22** (8), 940–952 (1986) [transl. from Diff. Uravn. **22** (8), 1351–1365 (1986)].
25. L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes* (Fizmatgiz, Moscow, 1961; Pergamon, Oxford, 1964).
26. B. N. Pshenichnyi and S. Ochilov, "On the problem of optimal passage through a given domain," Kibern. Vychisl. Tekh., Kiev **99**, 3–8 (1993).
27. B. N. Pshenichnyi and S. Ochilov, "On a special time-optimal problem," Kibern. Vychisl. Tekh., Kiev **101**, 11–15 (1994).
28. R. Schneider, "Equivariant endomorphisms of the space of convex bodies," Trans. Am. Math. Soc. **194**, 53–78 (1974).
29. A. I. Smirnov, "Necessary optimality conditions for a class of optimal control problems with discontinuous integrand," Proc. Steklov Inst. Math. **262**, 213–230 (2008) [transl. from Tr. Mat. Inst. Steklova **262**, 222–239 (2008)].
30. G. V. Smirnov, *Introduction to the Theory of Differential Inclusions* (Am. Math. Soc., Providence, RI, 2002), Grad. Stud. Math. **41**.
31. R. Vinter, *Optimal Control* (Birkhäuser, Boston, 2000).

Translated by K. Besov