The Hybrid Maximum Principle is a consequence of Pontryagin Maximum Principle

A.V. Dmitruk, A.M. Kaganovich

Lomonosov Moscow State University,
Russia 119992, Moscow, Leninskie Gory, VMK MGU
e-mail: dmitruk@member.ams.org, sm99@yandex.ru

Abstract

We give a simple proof of the Maximum Principle for smooth hybrid control systems by reducing the hybrid problem to an optimal control problem of Pontryagin type and then by using the classical Pontryagin Maximum Principle.

1 Introduction

In a broad sense, hybrid control systems are control systems involving both continuous and discrete variables. In recent years, optimization problems for hybrid systems attracted a significant attention of specialists in control. One of the most important result in the study of such problems is Hybrid Maximum Principle proved in [4] and [5]. This proof is rather difficult, since it follows the standard line of the full procedure of direct proof of Maximum Principle (MP), based on the introduction of special class of control variations (e.g., needle-like ones), calculation of the increments of the cost and all constraints, etc. As is well known, this procedure is very heavy and cumbersome even in the case of classical optimal control problem without discrete variables. However, as will be shown later, one does not need to perform all this heavy procedure for obtaining the hybrid MP if one supposes known the Pontryagin MP for the standard optimal control problem with nonseparated terminal constraints. After some transformation of the hybrid problem, MP for it is an easy consequence of the classical Pontryagin MP.

The statement of hybrid optimal control problem supposes the presense of a finite number of control systems, each of which is defined on its own space of variables (possibly of different dimensions). A trajectory moving under one of these systems, at some moment of time can switch to any other system, and it can do so a finite number of times. The hybridity in such systems just means the presence of both continuous and discontinuous dynamics of state variables. One needs to choose a sequence of control systems, durations of motion under each system, and control variable for each system that minimize the given cost functional.

The sequence of control systems under which a trajectory moves is not defined a priori. However, when investigating a given trajectory for optimality, we suppose this sequence defined and fixed, like in other papers known for us. (The matter is that variations of this sequence generate trajectories which are "far" from the given one, and hence they cannot be compared by methods of analysis.)

2 Statement of the problem

Let $t_0 < t_1 < \ldots < t_\nu$ be real numbers. Denote by Δ_k the time interval $[t_{k-1}, t_k]$. For any collection of continuous functions $x_k : \Delta_k \to \mathbf{R}^{n_k}$, $k = 1, \ldots, \nu$, define a vector

$$p = \Big(t_0,\, (t_1, x_1(t_0), x_1(t_1)),\, (t_2, x_2(t_1), x_2(t_2)),\, \ldots,\, (t_\nu, x_\nu(t_{\nu-1}),\, x_\nu(t_\nu)) \Big)$$

of dimension $d = 1 + \nu + 2 \sum\limits_{k=1}^{\nu} n_k$.

On the time interval $\Delta = [t_0, t_\nu]$ consider the optimal control problem

$$\textbf{Problem A :} \qquad \begin{cases} \dot{x}_k = f_k(t, x_k, u_k), \quad u_k \in U_k, \\ \qquad\qquad \text{for } t \in \Delta_k, \quad k = 1, \ldots, \nu, \\ \eta_j(p) = 0, \quad j = 1, \ldots, q, \\ \varphi_i(p) \le 0, \quad i = 1, \ldots, m, \\ J = \varphi_0(p) \to \min, \end{cases}$$

where $x_k \in \mathbf{R}^{n_k}$, $u_k \in \mathbf{R}^{r_k}$, the functions $x_k(t)$ are absolutely continuous, $u_k(t)$ are measurable and essentially bounded on the corresponding Δ_k. The time instants $t_0, t_1, \ldots, t_\nu$ are not fixed, a priori they just satisfy the above equality and inequality constraints on the vector p.

Suppose the following assumptions to hold:

A1) every function f_k is defined and continuous on an open set $\mathcal{Q}_k \subset \mathbf{R}^{1+n_k+r_k}$ and takes values in $\mathbf{R}^{n_k}$; moreover, it has partial derivatives f_{kt}, f_{kx}, which are continuous on $\mathcal{Q}_k$ w.r.t. the triple of their arguments;

A2) the functions $\varphi_i(p)$ and $\eta_j(p)$ are defined on an open set $\mathcal{P} \subset \mathbf{R}^d$ and continuously differentiable there;

A3) U_k are arbitrary sets in $\mathbf{R}^{r_k}$.

The Problem A satisfying assumptions A1–A3 will be called smooth.

Definition 1. The tuple $w = (t_0;\ t_k, x_k(t), u_k(t),\ k = 1, \ldots, \nu)$ is called an admissible process in Problem A if it satisfies all the constraints, and for every $k = 1, \ldots, \nu$ there exists a compact set $\Omega_k \subset \mathcal{Q}_k$ such that $(t, x_k(t), u_k(t)) \in \Omega_k$ a.e. on $\Delta_k = [t_{k-1}, t_k]$.

The existence of compact sets Ω_k means here that the admissible process is not allowed to come arbitrarily close to the boundary of domain $\mathcal{Q}_k$ (otherwise, even uniformly small variations of the process can bring it out from the domain $\mathcal{Q}_k$, and therefore, one can not actually vary it).

Let w be an admissible process in Problem A. It is convenient to introduce the functions $x(t), u(t)$, that take values $x_k(t), u_k(t)$ respectively for $t \in (t_{k-1}, t_k)$, $k = 1, \ldots, \nu$. For the internal points t_k, $k = 1, \ldots, \nu - 1$, we admit that the state variable can have two values: $x(t_k - 0) = x_k(t_k)$ and $x(t_k + 0) = x_{k+1}(t_k)$. For the measurable function $u(t)$ this uncertainty is inessential. Note that for time moments t' and t'' belonging to different Δ_k, the values both of $x(t)$ and $u(t)$ may be vectors of different dimensions. An admissible process w can be now written as $w = (\theta, x(t), u(t))$, where $\theta = \{t_0, t_1, \ldots, t_\nu\}$.

Definition 2. An admissible process $w^0 = (\theta^0, x^0(t), u^0(t))$ is called optimal (globally minimal) in Problem A if $J(w^0) \le J(w)$ for any admissible process w.

Definition 3. We say that a process $w^0 = (\theta^0, x^0(t), u^0(t))$ defined on a time interval $\Delta^0 = [t_0^0, t_\nu^0]$ gives a *strong minimum* in Problem A if there exists an $\varepsilon > 0$ such that for any admissible process $w = (\theta, x(t), u(t))$ defined on a time interval $\Delta = [t_0, t_\nu]$ and satisfying the conditions

$$\|x_k^0 - x_k\|_C < \varepsilon \quad \text{for } k = 1, \ldots, \nu, \qquad |t_k^0 - t_k| < \varepsilon \quad \text{for } k = 0, \ldots, \nu,$$

there holds $J(w^0) \le J(w)$.

Definition 4. We say that a process $w^0 = (\theta^0, x^0(t), u^0(t))$ gives a *Pontryagin minimum* in Problem A if for any constant N there exists an $\varepsilon = \varepsilon(N) > 0$ such that for any admissible process $w = (\theta, x(t), u(t))$ satisfying the conditions

$$\|x_k^0 - x_k\|_C < \varepsilon \quad \text{for all } k = 1, \ldots, \nu, \qquad |t_k^0 - t_k| < \varepsilon \quad \text{for } k = 0, \ldots, \nu,$$
$$\|u_k^0 - u_k\|_1 < \varepsilon, \qquad \|u_k^0 - u_k\|_\infty \le N \quad \text{for } k = 1, \ldots, \nu,$$

there holds $J(w^0) \le J(w)$.

The functions $x_k(t)$ and $u_k(t)$ in Definitions 3, 4 are defined on a time interval Δ_k that differs from the time interval Δ_k^0, and so, all the norms should be considered on the common time interval, i.e., on $\Delta_k \cap \Delta_k^0$.

Note the following obvious relations between the introduced types of optimality: if a process w^0 gives the global minimum, then it gives a strong minimum, and if w^0 gives a strong minimum, then it gives a Pontryagin minimum.

To obtain optimality conditions in Problem A, we will reduce it to the following *canonical autonomous optimal control problem of Pontryagin type* on a fixed time interval $[0, T]$.

$$\textbf{Problem } K: \qquad \begin{cases} \dot{x} = f(x, u), \\ u \in U, \quad (x, u) \in \mathcal{Q}, \\ \eta_j(p) = 0, \quad j = 1, \ldots, q, \\ \varphi_i(p) \le 0, \quad i = 1, \ldots, m, \\ J = \varphi_0(p) \to \min . \end{cases}$$

Here $p = (x(0), x(T)) \in \mathbf{R}^{2n}$ is a vector of terminal values of the trajectory $x(t)$, and $\mathcal{Q}$ is an open set in $\mathbf{R}^{n+r}$. The condition $(x, u) \in \mathcal{Q}$ should be regarded not as a constraint, but as definition of an open domain in the space (x, u) where the problem is considered (see, e.g., [7]). Note that Problem K is a special case of Problem A when $\nu = 1$, i.e., when the intermediate points are absent.

Suppose that Problem K satisfies assumptions similar to $A1 - A3$. Then the following theorem holds (see, e.g., [1], [2], [7]).

Theorem 1. (Pontryagin Maximum Principle for Problem K). *Let a process $w^0 = (x^0(t), u^0(t))$ give a Pontryagin minimum in Problem K. Then there exists a collection $\lambda = (\alpha, \beta, c, \psi(\cdot))$, where $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_m) \geq 0$, $\beta = (\beta_1, \beta_2, \ldots, \beta_q) \in \mathbf{R}^q$, $c \in \mathbf{R}^1$, $\psi(\cdot)$ is an $n-$ dimensional Lipschitz function on $[0,T]$, which generates the Pontryagin function $H(\psi, x, u) = \langle \psi, f(x,u) \rangle$, the terminal Lagrange function $l(p) = \sum_{i=0}^{m} \alpha_i \, \varphi_i(p) + \sum_{j=1}^{q} \beta_j \, \eta_j(p)$, and satisfies the following conditions:*

a) nontriviality condition: $(\alpha, \beta) \neq (0,0)$;

b) conditions of complementary slackness: $\alpha_i \varphi_i(p^0) = 0, \quad i = 1, \ldots, m$;

c) adjoint equation:

$$\dot{\psi}(t) = -H_x^0 = -\psi(t) f_x(x^0(t), u^0(t)) \quad \text{a.e. on } [0,T];$$

d) transversality conditions: $\psi(0) = l_{x(0)}(p^0), \quad \psi(T) = -l_{x(T)}(p^0)$;

e) constancy of function H condition: for a.e. $t \in [0,T]$

$$H(\psi(t), x^0(t), u^0(t)) = c\,;$$

f) maximality condition: for all $t \in [0,T]$ $\quad \max_{u \in U^0(t)} H(\psi(t), x^0(t), u) = c,$

where $U^0(t) = \Big\{ u \in U \;\Big|\; (x^0(t), u) \in \mathcal{Q} \Big\}$.

3 Obtaining the Hybrid Maximum Principle

We will pass from Problem A to some problem of canonical type K and establish a correspondence between the admissible and optimal processes in these problems. The idea of such passage is quite natural: one should reduce all the state and control variables to a common fixed time interval, for example, to $[0,1]$.

Let $(\theta, x(t), u(t))$ be an arbitrary admissible process in Problem A.
Introduce a new time $\tau \in [0,1]$ and define functions $\rho_k : [0,1] \to \Delta_k$, $k = 1, \ldots, \nu$, from the equations:

$$\frac{d\rho_k}{d\tau} = z_k(\tau), \qquad \rho_k(0) = t_{k-1},$$

where $z_k(\tau) > 0$ are arbitrary measurable essentially bounded functions on $[0,1]$ such that $\rho_k(1) = t_k$, i.e. $\int_0^1 z_k(\tau)\, d\tau = |\Delta_k|$. The functions ρ_k play the role of time t on the intervals Δ_k. Define also functions $y_k(\tau) = x_k(\rho_k(\tau))$ and $v_k(\tau) = u_k(\rho_k(\tau))$, $k = 1, \dots, \nu$, $\tau \in [0,1]$. They obviously satisfy the relations:

$$\begin{cases} \dfrac{dy_k}{d\tau} = z_k f(\rho_k, y_k, v_k), \quad v_k \in U_k, \\ \dfrac{d\rho_k}{d\tau} = z_k, \quad k = 1, \dots, \nu, \end{cases} \tag{1}$$

$$(\rho_k, y_k, v_k) \in \mathcal{Q}_k, \quad z_k > 0, \tag{2}$$

$$\begin{cases} \rho_2(0) - \rho_1(1) = 0, \\ \rho_3(0) - \rho_2(1) = 0, \\ . \quad . \quad . \quad . \\ \rho_\nu(0) - \rho_{\nu-1}(1) = 0, \end{cases} \tag{3}$$

$$\begin{cases} \eta_j(\hat{p}) = 0, \quad j = 1, \dots, q, \\ \varphi_i(\hat{p}) \le 0, \quad i = 1, \dots, m, \end{cases} \tag{4}$$

where, for simplicity, we use the notation

$$\hat{p} = \hat{p}(\rho, y) = \Big(\rho_1(0), (\rho_1(1), y_1(0), y_1(1)), (\rho_2(1), y_2(0), y_2(1)), \dots, (\rho_\nu(1), y_\nu(0), y_\nu(1))\Big).$$

Obviously, this vector coincides with the initial vector $p(t,x)$ and is a part of the full vector $\tilde{p}(\rho, y)$ of terminal values.

For brevity, define the vectors $\rho = (\rho_1, \rho_2, \dots, \rho_\nu)$, $y = (y_1, y_2, \dots, y_\nu)$, $v = (v_1, v_2, \dots, v_\nu)$, and $z = (z_1, z_2, \dots, z_\nu)$.

On the set of admissible processes $\tilde{w} = (\rho(\tau), y(\tau), v(\tau), z(\tau))$ satisfying constraints (1)–(4), we will minimize the functional

$$\tilde{J}(\tilde{w}) = \varphi_0(\hat{p}) \to \min. \tag{5}$$

The obtained optimal control problem will be called **Problem** $\tilde{A}$. Here, the state variables are ρ_k and y_k, the controls are v_k and z_k, $k = 1, \dots, \nu$, and the time interval $[0,1]$ is fixed. The open set $\widetilde{\mathcal{Q}}$ consists of all vectors (ρ_k, y_k, v_k, z_k) satisfying (2). The open set $\widetilde{\mathcal{P}}$ consists of all vectors $\tilde{p}$ for which the "truncated" vector $\hat{p} \in \mathcal{P}$. It is easy to see that Problem $\tilde{A}$ is a problem of type K.

The following two correspondences can be established between the admissible processes of Problems A and $\tilde{A}$. As was shown above, any admissible process $w = (\theta, x(t), u(t))$ of Problem A can be transformed to an admissible process $\tilde{w} = (\rho(\tau), y(\tau), v(\tau), z(\tau))$ of Problem $\tilde{A}$. This transformation is not defined uniquely, since it depends on the choice of functions $z_k(\tau)$. In order to make it unique, let us set, for example, $z_k(\tau) = |\Delta_k|$. Denote the obtained mapping by F.

Construct also the mapping G that transforms a process $\tilde{w}$ into w. To do this we first define time moments $t_0 = \rho_1(0)$, $t_1 = \rho_2(0), \dots,$ $t_{\nu-1} = \rho_\nu(0)$, $t_\nu = \rho_\nu(1)$ (and so, define a vector θ), and also time intervals $\Delta_k = [t_{k-1}, t_k]$, $k = 1, \dots, \nu$.

Now, introduce the functions $x_k(t) = y_k(\rho_k^{-1}(t))$ and $u_k(t) = v_k(\rho_k^{-1}(t))$ defined on the corresponding intervals Δ_k, and the mappings $x(t) = x_k(t)$ and $u(t) = u_k(t)$ for $t \in \Delta_k$ defined on the whole Δ. One can easily show that the process $w = G(\tilde{w}) = (\theta, x(t), u(t))$ is admissible in Problem A.

An important property of both these mappings is that they preserve the value of cost functional. Note that the constructed mappings F and G are not inverse to each other (GF is the identity, but FG is not). Nevertheless, the very fact that there exist two mappings that put in correspondence to any admissible process of one problem an admissible process of another problem with the same value of the cost functional, readily implies the next statement.

Theorem 2. *If a process w^0 is optimal (i.e. globally minimal) in Problem A, then the process $\tilde{w}^0 = F(w^0)$ is optimal in Problem $\tilde{A}$; and vice versa, if a process $\tilde{w}^0$ is optimal in Problem $\tilde{A}$ then the process $w^0 = G(\tilde{w}^0)$ is optimal in Problem A.*

Indeed, let us prove the first implication. Suppose a process w^0 is optimal in Problem A. If the process $\tilde{w}^0 = F(w^0)$ is not optimal in Problem $\tilde{A}$, there exists another admissible process $\tilde{w}'$ in this problem such that $\tilde{J}(\tilde{w}') < \tilde{J}(\tilde{w}^0)$. Then the corresponding process $w' = G(\tilde{w}')$ is admissible in Problem A and satisfies the relations $J(w') = \tilde{J}(\tilde{w}') < \tilde{J}(\tilde{w}^0) = J(w^0)$, which lead to a contradiction with the optimality of the process w^0. The inverse implication is proved in the same way. $\square$

This theorem asserts the invariance of a rather rough property (global minimality); it does not take into account the specificity of the problems and the mappings F, G. For our Problems A, $\tilde{A}$ and the above mappings F, G, the following refined statement holds [8].

Theorem 3. *If a process w^0 gives a strong (Pontryagin) minimum in Problem A, then the process $\tilde{w}^0 = F(w^0)$ gives a strong (respectively, Pontryagin) minimum in Problem $\tilde{A}$; and vice versa, if a process $\tilde{w}^0$ gives a strong (Pontryagin) minimum in Problem $\tilde{A}$, then the process $w^0 = G(\tilde{w}^0)$ gives a strong (respectively, Pontryagin) minimum in Problem A.*

Thus, the study of optimality (in any of the three above senses) of a process w^0 in Problem A reduces to the study of optimality of the corresponding process $\tilde{w}^0$ in Problem $\tilde{A}$.

Now, let a process $w^0 = (\theta^0, x^0(t), u^0(t))$ give a Pontryagin minimum in Problem A. Then by Theorem 3 the corresponding process $\tilde{w}^0 = (\rho^0(\tau),\, y^0(\tau),\, v^0(\tau), z^0(\tau))$ gives a Pontryagin minimum in Problem $\tilde{A}$, and therefore, according to Theorem 1, it satisfies the Pontryagin MP, i.e. there exists a collection $\lambda = (\alpha, \beta, \delta, c, \psi_y(\cdot), \psi_\rho(\cdot))$, where $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_m) \ge 0$, $\beta = (\beta_1, \beta_2, \ldots, \beta_q) \in \mathbf{R}^q$, $\delta = (\delta_1, \delta_2, \ldots, \delta_{\nu-1}) \in \mathbf{R}^{\nu-1}$, $c \in \mathbf{R}^1$, $\psi_y = (\psi_{y_1}, \psi_{y_2}, \ldots, \psi_{y_\nu})$, $\psi_\rho = (\psi_{\rho_1}, \psi_{\rho_2}, \ldots, \psi_{\rho_\nu})$, ψ_{y_k} and ψ_{ρ_k} are Lipschitz functions on $[0,1]$, which generates the Pontryagin function

$$\tilde{H}(\psi_\rho, \psi_y, \rho, y, v, z) = \sum_{k=1}^{\nu} z_k \Big(\psi_{y_k} f(\rho_k, y_k, v_k) + \psi_{\rho_k}\Big) = \sum_{k=1}^{\nu} z_k\, \Pi_k(\psi_{\rho_k}, \psi_{y_k}, \rho_k, y_k, v_k),$$

where $\Pi_k(\psi_{\rho_k}, \psi_{y_k}, \rho_k, y_k, v_k) = \psi_{y_k} f(\rho_k, y_k, v_k) + \psi_{\rho_k}$, the terminal Lagrange function

$$\tilde{l}(\tilde{p}) = l(\hat{p}) + \sum_{k=1}^{\nu-1} \delta_k(\rho_{k+1}(0) - \rho_k(1)), \quad \text{where} \quad l(\hat{p}) = \sum_{i=0}^{m} \alpha_i \varphi_i(\hat{p}) + \sum_{j=1}^{q} \beta_j \eta_j(\hat{p}),$$

and satisfies the nontriviality condition $(\alpha, \beta, \delta) \neq (0,0,0)$, the complementary slackness conditions $\alpha_i \varphi_i(\hat{p}^0) = 0, \quad i = 1, \ldots, m,$ the adjoint equations:

$$\begin{aligned} \dot{\psi}_{y_k}(\tau) &= -\tilde{H}^0_{y_k} = -z^0_k(\tau)\psi_{y_k}(\tau) f_{kx}(\rho^0_k(\tau), y^0_k(\tau), v^0_k(\tau)), \\ \dot{\psi}_{\rho_k}(\tau) &= -\tilde{H}^0_{\rho_k} = -z^0_k(\tau)\psi_{y_k}(\tau) f_{kt}(\rho^0_k(\tau), y^0_k(\tau), v^0_k(\tau)), \quad k = 1, \ldots, \nu; \end{aligned} \tag{6}$$

the transversality conditions:

$$\psi_{y_k}(0) = l_{y_k(0)}, \qquad \psi_{y_k}(1) = -l_{y_k(1)}, \qquad k = 1, \ldots, \nu;$$

$$\begin{cases} \psi_{\rho_1}(0) = l_{\rho_1(0)}, & \psi_{\rho_1}(1) = -l_{\rho_1(1)} + \delta_1, \\ \psi_{\rho_2}(0) = \delta_1, & \psi_{\rho_2}(1) = -l_{\rho_2(1)} + \delta_2, \\ \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \\ \psi_{\rho_{\nu-1}}(0) = \delta_{\nu-2}, & \psi_{\rho_{\nu-1}}(1) = -l_{\rho_{\nu-1}(1)} + \delta_{\nu-1}, \\ \psi_{\rho_\nu}(0) = \delta_{\nu-1}; & \psi_{\rho_\nu}(1) = -l_{\rho_\nu(1)}, \end{cases}$$

(all the derivatives of function $l(\hat{p})$ are taken at the point $\hat{p}^0$);

the constancy of function $\tilde{H}$ condition: for a.e. $\tau \in [0,1]$

$$\tilde{H}(\psi_\rho(\tau), \psi_y(\tau), \rho^0(\tau), y^0(\tau), v^0(\tau), z^0(\tau)) = c,$$

and the maximality condition: for all $\tau \in [0,1]$

$$\max_{(v,z)\in \tilde{U}^0(\tau)} \tilde{H}(\psi_\rho(\tau), \psi_y(\tau), \rho^0(\tau), y^0(\tau), v, z) = c,$$

$$\text{where} \quad \tilde{U}^0(\tau) = \tilde{U}^0_1(\tau) \times \cdots \times \tilde{U}^0_\nu(\tau), \qquad \tilde{U}^0_k(\tau) = V^0_k(\tau) \times Z_k,$$

$$V^0_k(\tau) = \Big\{ v_k \in U_k \;\Big|\; (\rho^0_k(\tau), y^0_k(\tau), v_k) \in \mathcal{Q}_k \Big\}, \quad Z_k = \{ z \in \mathbf{R} \;\big|\; z > 0 \}.$$

Let us analyze these conditions.

1) From the constancy of $\tilde{H}$ and the maximality condition we conclude, fixing the control $v^0_k(\tau)$, that $\forall\, k$ for a.e. τ, the function $\tilde{H}$ reaches its maximum over $z_k \in Z_k$ at the point $z^0_k(\tau) > 0$. Therefore, since $\tilde{H}$ is linear in all z_k, we have

$$\frac{\partial \tilde{H}}{\partial z_k}(\psi_\rho(\tau), \psi_y(\tau), \rho^0(\tau), y^0(\tau), v^0(\tau), z^0(\tau)) \;=\; \Pi^0_k(\tau) = 0 \tag{7}$$

for a.e. $\tau \in [0,1]$, from which we get $c = 0$, and then, taking into account that all the control pairs $(z_k,\, v_k)$ come separately into $\tilde{H}$, we obtain from the maximality condition that for all $\tau \in [0,1]$

$$\max_{v_k \in V^0_k(\tau)} \Pi_k\,(\psi_{\rho_k}, \psi_{y_k}, \rho^0_k(\tau), y^0_k(\tau), v_k) = 0. \tag{8}$$

2) Problem $\tilde{A}$, as a result of transformation of Problem A, involves additional equality constraints (3), and therefore the MP for Problem $\tilde{A}$ involves the corresponding Lagrange multipliers δ_k. Let us show that these additional multipliers can be excluded from the formulation of MP without loss of information. To this end we note that they come only in the nontriviality and transversality conditions for ψ_ρ. But the nontriviality condition is not influenced by these multipliers as the following simple lemma shows.

Lemma 1. $(\alpha, \beta, \delta) \neq (0,0,0)$ if and only if $(\alpha, \beta) \neq (0,0)$.

Indeed, if $\alpha = \beta = 0$, then $l(\hat{p}) \equiv 0$, hence the adjoint equation and transversality conditions imply that all $\psi_{y_k} \equiv 0$ and all ψ_{ρ_k} are constants. Further, since $\psi_{\rho_1}(0) = 0$, we get $\psi_{\rho_1}(1) = 0$, whence $\delta_1 = 0$, therefore $\psi_{\rho_2}(0) = 0$, and so on, hence all $\delta_k = 0$. □

Moreover, all δ_k can be also excluded from transversality conditions for ψ_ρ by rewriting these conditions in the form:

$$\begin{cases} \psi_{\rho_1}(0) = l_{\rho_1(0)}(\hat{p}^0), \\ \psi_{\rho_{k+1}}(0) - \psi_{\rho_k}(1) = l_{\rho_{k+1}(0)}(\hat{p}^0), \quad k = 1, \ldots, \nu-1, \\ \psi_{\rho_\nu}(1) = -l_{\rho_\nu(1)}(\hat{p}^0). \end{cases} \tag{9}$$

It can be easily shown that the obtained conditions (9) are equivalent to the initial ones. Thus, all multipliers δ_k can be totally excluded from the MP for Problem $\tilde{A}$.

3) Define the functions $\pi_k^0(t) = (\rho_k^0)^{-1}(t)$, $x_k^0(t) = y_k^0(\pi_k^0(t))$, $u_k^0(t) = v_k^0(\pi_k^0(t))$, $\psi_t(t) = \psi_{\rho_k}(\pi_k^0(t))$, $\psi_{x_k}(t) = \psi_{y_k}(\pi_k^0(t))$ for $t \in \Delta_k^0$, $k = 1, \ldots, \nu$. Since all $\psi_{y_k}(\tau)$ are Lipschitz continuous on $[0,1]$, all $\psi_{x_k}(t)$ are Lipschitz continuous on their Δ_k^0. Since all $\psi_{\rho_k}(\tau)$ are Lipschitz continuous on $[0,1]$, the function $\psi_t(t)$ is Lipschitz continuous on every Δ_k^0 (such functions will be called piecewise Lipschitz), and at the points t_k^0 it may have jumps defined by conditions (9). In view of relations $dt = z_k^0\, d\tau$ on Δ_k^0, the adjoint equations (6) can be rewritten in the form

$$\frac{d\psi_t}{dt} = -\psi_{x_k} f_{kt}(t, x_k, u_k), \qquad \frac{d\psi_{x_k}}{dt} = -\psi_{x_k} f_{kx}(t, x_k, u_k), \quad t \in \Delta_k^0.$$

On every Δ_k^0 we construct the function $H_k(\psi_t, \psi_{x_k}, t, x_k, u_k) = \psi_{x_k} f_k(t, x_k, u_k) + \psi_t$. Along the optimal trajectory $x_k^0(t)$, $u_k^0(t)$ with corresponding $\psi_{x_k}(t)$, $\psi_t(t)$ it obviously coincides with

$$\Pi_k\left(\psi_{\rho_k}(\tau), \psi_{y_k}(\tau), \rho_k^0(\tau), y_k^0(\tau), v_k^0(\tau)\right) \qquad \text{for} \quad \tau = \pi_k^0(t),$$

therefore we conclude from (7) that $H_k(\psi_t(t), \psi_{x_k}(t), t, x_k^0(t), u_k^0(t)) = 0$ a.e. on Δ_k^0. In view of (8), this implies the maximality of H_k over all $u_k \in U_k$ such that $(t, x_k^0(t), u_k) \in \mathcal{Q}_k$.

Since $\hat{p}(\rho, y) = p(t, x)$, the conditions of complementary slackness are not changed.

Thus, the analysis of Pontryagin MP for Problem $\tilde{A}$ yields the following result.

Theorem 4. **(The Maximum Principle for hybrid problems).** *Let a process* $w^0 = (\theta^0, x^0(t), u^0(t))$ *give a Pontryagin minimum in Problem A. Then there exists a collection* $\lambda = (\alpha, \beta, \psi_x(\cdot), \psi_t(\cdot))$, *where* $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_m) \geq 0$, $\beta = (\beta_1, \beta_2, \ldots, \beta_q) \in \mathbf{R}^q$, $\psi_x = (\psi_{x_1}, \ldots, \psi_{x_\nu})$, $\psi_{x_k} : \Delta_k^0 \to \mathbf{R}^{n_k}$ *are Lipschitz functions,* ψ_t *is a piecewise Lipschitz function on* Δ^0, *which generates the Pontryagin functions*

$$H_k(\psi_t, \psi_{x_k}, t, x_k, u_k) = \langle \psi_{x_k}, f_k(t, x_k, u_k) \rangle + \psi_t, \qquad t \in \Delta_k^0,$$

the terminal Lagrange function $l(p) = \sum\limits_{i=0}^{m} \alpha_i \varphi_i(p) + \sum\limits_{j=1}^{q} \beta_j \eta_j(p)$,

and satisfies the following conditions:

a) nontriviality condition: $(\alpha, \beta) \neq (0, 0)$;

b) conditions of complementary slackness: $\alpha_i \varphi_i(p) = 0, \quad i = 1, \ldots, m$;

c) adjoint equations:

$$\left.\begin{aligned} \dot{\psi}_{x_k}(t) &= -\frac{\partial H_k^0}{\partial x_k} = -\psi_{x_k}(t) f_{kx}(t, x_k^0(t), u_k^0(t)) \\ \dot{\psi}_t(t) &= -\frac{\partial H_k^0}{\partial t} = -\psi_{x_k}(t) f_{kt}(t, x_k^0(t), u_k^0(t)) \end{aligned}\right\} \quad a.e. \ on \quad \Delta_k^0;$$

d) transversality conditions for ψ_x *and* ψ_t :

$$d1: \quad \begin{cases} \psi_{x_k}(t_{k-1}) = l_{x_k(t_{k-1})}(p^0), \\ \psi_{x_k}(t_k) = -l_{x_k(t_k)}(p^0), \qquad k = 1, \ldots, \nu; \end{cases}$$

$$d2: \quad \begin{cases} \psi_t(t_0) = l_{t_0}(p^0), \quad \psi_t(t_\nu) = -l_{t_\nu}(p^0), \\ \triangle \psi_t(t_k) = l_{t_k}(p^0), \quad k = 1 \ldots, \nu - 1; \end{cases}$$

e) for a.e. $t \in \Delta_k^0$, $\quad k = 1, \ldots, \nu$, $\quad H_k(\psi_t(t), \psi_{x_k}(t), t, x_k^0(t), u_k^0(t)) = 0$;

f) maximality condition: for all $t \in \Delta_k^0$, $\quad k = 1, \ldots, \nu$,

$$\max_{u_k \in U_k^0(t)} H_k(\psi_t(t), \psi_{x_k}(t), t, x_k^0(t), u_k) = 0,$$

$$\text{where} \quad U_k^0(t) = \left\{ u_k \in U_k \,\middle|\, (t, x_k^0(t), u_k) \in \mathcal{Q}_k \right\}.$$

Remark 1. For smooth control systems, theorem 4 coincides with the Hybrid MP obtained in [4], [5]. However, in these papers it was obtained as a new independent result, whereas it easily follows from the Pontryagin MP after a transformation of the hybrid problem to a standard optimal control problem. Note that a similar transformation was used in an old paper [3] for some hybrid problem like our Problem A. But this trick was not there clearly identified as such in rather cumbersome technical constructions overloaded by specific features of the studied problem, and therefore, it actually remained unnoticed. A detailed justification of this trick and its modifications for obtaining MP in different optimal control problems are given in [8].

4 Hybrid systems with a quasivariable control set

Along with hybrid control problems of the above "standard" type A, in literature there are also considered hybrid systems in which the control set U_k on each time interval Δ_k depends, in a special way, on the values of state variable at switching times t_k, namely, the control set on Δ_k is $\sigma_k(p)\,U_k$, where the functions $\sigma_k(p)$ have the derivative on $\mathcal{P}$. We will call these problems by problems with quasivariable control set or, for brevity, problems of type B. Such a problem was considered, for example, in [5]. In that paper, to obtain necessary optimality conditions, the authors introduce a notion of generalized needle-like variations, by means of which they prove, under rather restrictive assumptions about functions f_k (their twice smoothness), the so-called "Hybrid Necessary Principle", being an equation w.r.t. some variations, to which the optimal process must satisfy. The MP for this problem was not obtained in that paper.

However, the MP can be easily obtained if one notes that Problem B can also be reduced to a problem of type A. Indeed, the admissible control $u_k(t)$ on the interval Δ_k has the form $\sigma_k(p)v_k(t)$, where $v_k(t) \in U_k$. Considering $v_k \in U_k$ as a new control on Δ_k and introducing on every Δ_k a new phase variable s_k satisfying differential equation $\dot{s}_k = 0$ and initial condition $s_k(t_{k-1}) = \sigma_k(p)$, we can rewrite the problem of type B as a Problem $\tilde{B}$ which is a problem of type A.

In order to apply the hybrid MP (theorem 4) to the obtained Problem $\tilde{B}$, we impose an auxiliary assumption:

B1) *every function $f_k(t,x,u)$ has the partial derivative in u which is continuous on the coresponding $\mathcal{Q}_k$.*

In this case, the application of the hybrid MP to Problem $\tilde{B}$ yields the following theorem.

Theorem 5. (Maximum Principle for hybrid problems with quasivariable control set). *Let a process $w^0 = (\theta^0, x^0(t), s^0(t), v^0(t))$ give a Pontryagin minimum in Problem B. Then there exists a collection $\lambda = (\alpha, \beta, \gamma, \psi_x(\cdot), \psi_s(\cdot), \psi_t(\cdot))$, where $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_m) \geq 0$, $\beta = (\beta_1, \beta_2, \ldots, \beta_q) \in \mathbf{R}^q$, $\gamma = (\gamma_1, \ldots, \gamma_\nu) \in \mathbf{R}^\nu$, $\psi_x = (\psi_{x_1}, \ldots, \psi_{x_\nu})$, $\psi_s = (\psi_{s_1}, \ldots, \psi_{s_\nu})$, $\psi_{x_k} : \Delta_k^0 \to \mathbf{R}^{n_k}$ and $\psi_{s_k} : \Delta_k^0 \to \mathbf{R}^1$ are Lipschitz functions, ψ_t is a piecewise Lipschitz functions on Δ^0, which generates the Pontryagin functions*

$$H_k(\psi_t, \psi_{x_k}, \psi_{s_k}, t, x_k, s_k, v_k) = \langle \psi_{x_k}, f_k(t, x_k, s_k v_k) \rangle + \psi_t, \qquad t \in \Delta_k^0,$$

the terminal Lagrange function

$$l^B(p, p_s) = l^A(p) + \sum_{r=1}^{\nu} \gamma_r (s_r(t_{r-1}) - \sigma_r(p)), \text{ where } l^A(p) = \sum_{i=0}^{m} \alpha_i \varphi_i(p) + \sum_{j=1}^{q} \beta_j \eta_j(p),$$

(here p is the full vector of terminal values in Problem A, and

$$p_s = ((s_1(t_0), s_1(t_1)), (s_2(t_1), s_2(t_2)), \ldots, (s_\nu(t_{\nu-1}), s_\nu(t_\nu))\,)\ ),$$

and satisfies the following conditions:

a) nontriviality condition: $\quad (\alpha, \beta, \gamma) \neq (0,0,0)$;

b) conditions of complementary slackness: $\quad \alpha_i \varphi_i(p) = 0, \quad i = 1, \ldots, m$;

c) adjoint equations: $\quad \forall\, k = 1, \ldots, \nu,$

$$\left.\begin{aligned} \dot\psi_{x_k}(t) &= -\frac{\partial H_k^0}{\partial x_k} = \; -\psi_{x_k}(t) f_{kx}(t, x_k^0(t), s_k^0(t) v_k^0(t)) \\ \dot\psi_t(t) &= -\frac{\partial H_k^0}{\partial t} = \; -\psi_{x_k}(t) f_{kt}(t, x_k^0(t), s_k^0(t) v_k^0(t)), \\ \dot\psi_{s_k}(t) &= -\frac{\partial H_k^0}{\partial s_k} = \; -\psi_{x_k}(t) f_{ku}(t, x_k^0(t), s_k^0(t)\, v_k^0(t)) \cdot v_k^0(t), \end{aligned}\right. \qquad a.e. \;\; on \;\; \Delta_k^0\,;$$

d) transversality conditions:

$$d1: \; \begin{cases} \psi_{x_k}(t_{k-1}) = l^A_{x_k(t_{k-1})}(p^0) - \sum\limits_{r=1}^{\nu} \gamma_r \sigma'_{r x_k(t_{k-1})}(p^0), \\ \psi_{x_k}(t_k) = -l^A_{x_k(t_k)}(p^0) + \sum\limits_{r=1}^{\nu} \gamma_r \sigma'_{r x_k(t_k)}(p^0), \qquad k = 1, \ldots, \nu; \end{cases}$$

$$d2: \; \begin{cases} \psi_t(t_0) = l^A_{t_0}(p^0) - \sum\limits_{r=1}^{\nu} \gamma_r \sigma'_{r t_0}(p^0), \quad \psi_t(t_\nu) = -l^A_{t_\nu}(p^0) + \sum\limits_{r=1}^{\nu} \gamma_r \sigma'_{r t_\nu}(p^0), \\ \triangle \psi_t(t_k) = l^A_{t_k}(p^0) - \sum\limits_{r=1}^{\nu} \gamma_r \sigma'_{r t_k}(p^0), \quad k = 1 \ldots, \nu - 1; \end{cases}$$

$$d3: \;\; \psi_{s_k}(t_{k-1}) = \gamma_k, \quad \psi_{s_k}(t_k) = 0, \quad k = 1, \ldots, \nu;$$

e) for all $k = 1, \ldots, \nu$, *and almost all* $t \in \Delta_k^0$,

$$H_k(\psi_t(t), \psi_{x_k}(t), \psi_{s_k}(t), t, x_k^0(t), s_k^0(t), v_k^0(t)) = 0;$$

f) maximality condition: *for all* $k = 1, \ldots, \nu$, *and all* $t \in \Delta_k^0$,

$$\max_{v_k \in U_k^0(t)} H_k(\psi_t(t), \psi_{x_k}(t), \psi_{s_k}(t), t, x_k^0(t), s_k^0(t), v_k) = 0,$$

where $\quad U_k^0(t) = \Big\{ v_k \in U_k \;\Big|\; (t, x_k^0(t), s_k^0(t) v_k) \in \mathcal{Q}_k \Big\}.$

Remark 2. Theorem 5 involves the adjoint variables ψ_{s_k} related to the specific method of reduction of Problem B to a problem of type A. Let us show that these variables can be harmlessly removed from the MP for Problem B.

Indeed, since the functions H_k do not depend on ψ_{s_k}, conditions e) and f) do not actually involve ψ_{s_k}. The adjoint equations and transversality conditions *d3* for ψ_{s_k} can be replaced by equivalent conditions

$$\int_{t_{k-1}^0}^{t_k^0} \psi_{x_k}(t) f_{ku}(t, x_k^0(t), s_k^0(t) v_k^0(t))\; v_k^0(t)\, dt \;=\; \gamma_k, \qquad k = 1, \ldots, \nu. \tag{10}$$

The remaining conditions of MP do not involve ψ_{s_k} by definition. Thus, all ψ_{s_k} can be excluded from the formulation of MP by supplementing it with conditions (10).

Remark 3. The statement of Problem B obviously generalizes the statement of Problem A (one should put all $\sigma_k(p) \equiv 1$). It is easy to show that MP for Problem B, applied to Problem A, exactly gives the MP for Problem A.

5 Example: control of a car with two gears [6]

A car moves under the law $\dot{x} = y, \quad \dot{y} = u\, g_1(y), \quad u \in U$ on the time interval $\Delta_1 = [0, t_1]$, and under the law $\dot{x} = y, \quad \dot{y} = u g_2(y), \quad u \in \sigma(y(t_1))\, U$ on the time interval $\Delta_2 = [t_1, T]$.

The initial and final time moments $t_0 = 0$ and $t_2 = T$ are fixed, the moment t_1 is not fixed, the set $U = [0, 1]$, the functions g_1, g_2, σ are positive and differentiable in $\mathbf{R}^1$. The car starts from the point $(x^0, y^0) = (0, 0)$ and the state variables x and y are assumed to be continuous on the whole interval $\Delta = [0, T]$. It is required to maximize $x(T)$.

Rewrite this problem as a problem of type A. To this aim, on the interval Δ_2 introduce a new control $v \in U$ and a new state variable s satisfying equation $\dot{s} = 0$ and initial condition $s(t_1) = y(t_1)$. The control $u(t)$ takes then the form: $u(t) = \sigma(s(t))v(t)$. Thus, the control system on Δ_2 is

$$\begin{cases} \dot{x} = y, \\ \dot{y} = v\, \sigma(s)\, g_2(y), \\ \dot{s} = 0, \quad v \in U, \end{cases}$$

and the initial problem is reduced to a problem of type A. On the interval Δ_1 we have $\sigma(s) \equiv 1$, $u(t) = v(t)$, and so, on the whole interval Δ the control now is $v \in U$. The open sets $\mathcal{Q}_1$ and $\mathcal{Q}_2$ coincide here with the whole space.

The obtained problem possesses the convexity and compactness w.r.t. v, therefore, passing to a problem of type K and using standard existence theorems, one can show that this problem always has a solution.

Let us write out MP for this problem. The Pontryagin function is $H_1 = \psi_x y + \psi_y v g_1(y) + \psi_t$ on Δ_1, $H_2 = \psi_x y + \psi_y v \sigma(s) g_2(y) + \psi_t$ on Δ_2, the terminal Lagrange function is

$$l(p) = -\alpha_0 x(T) + \beta_{t_0} t_0 + \beta_T (t_2 - T) + \beta_x x(0) + \beta_y y(0) + \gamma(s(t_1) - y(t_1)).$$

For optimal process, we have $\alpha_0 \geq 0, \qquad (\alpha_0,\, \beta_{t_0},\, \beta_T,\, \beta_x,\, \beta_y,\, \gamma) \neq 0,$

- the adjoint system takes the form:

$$-\dot{\psi}_x = 0, \qquad -\dot{\psi}_y = \psi_x + v\, \psi_y\, g_1'(y), \qquad \dot{\psi}_t = 0 \qquad \text{on} \quad \Delta_1;$$

$$\left.\begin{array}{c} -\dot{\psi}_x = 0, \qquad -\dot{\psi}_y = \psi_x + v\, \psi_y\, \sigma(s)\, g_2'(y) \\ -\dot{\psi}_s = v\, \psi_y \sigma'(s)\, g_2(y), \qquad \dot{\psi}_t = 0 \end{array}\right\} \qquad \text{on} \quad \Delta_2;$$

- the transversality conditions are:

  at the left endpoint: $\quad \psi_x(0) = \beta_x, \qquad \psi_y(0) = \beta_y, \qquad \psi_t(0) = \beta_{t_0};$

  at the right endpoint: $\quad \psi_x(T) = \alpha_0, \qquad \psi_y(T) = 0, \qquad \psi_t(T) = -\beta_T;$

- the transversality conditions for ψ_s and jump conditions for ψ_x, ψ_y and ψ_t :

$$\psi_s(t_1) = \gamma, \qquad \psi_s(T) = 0,$$

$$\triangle\psi_x(t_1) = 0, \qquad \triangle\psi_y(t_1) = -\gamma, \qquad \triangle\psi_t(t_1) = 0;$$

- for a.e. $t \in \Delta_1 \quad H_1 = \psi_x y + \psi_y v\, g_1(y) + \psi_t = 0;$
  for a.e. $t \in \Delta_2 \quad H_2 = \psi_x y + \psi_y v\, \sigma(s)\, g_2(y) + \psi_t = 0;$

- the maximality condition: for all $t \in \Delta_1 \quad \max_{v\in U} (\, \psi_x y + \psi_y v\, g_1(y) + \psi_t) = 0,$
  for all $t \in \Delta_2 \quad \max_{v\in U} (\, \psi_x y + \psi_y v \sigma(s)\, g_2(y) + \psi_t) = 0.$

The transversality conditions for ψ_x and ψ_y at the point t_1 are replaced here by the jump conditions $\triangle\psi_x(t_1) = l_{x(t_1)}(p^0)$ and $\triangle\psi_y(t_1) = l_{y(t_1)}(p^0)$, because the state variables x and y are continuous at the intermediate points t_k (a detailed justification see in [8]).

Let us find all the extremals in this problem. The adjoint system and transversality conditions imply that $\psi_x(t) \equiv \alpha_0$ and $\psi_t \equiv \beta_{t_0}$.

Prove that $\alpha_0 \neq 0$. Indeed, if $\alpha_0 = 0$, then $\psi_x \equiv 0$ on the whole Δ. In that case, since the adjoint equation is homogeneous, we obtain $\psi_y \equiv 0$ on Δ_2. Therefore $\psi_t \equiv 0$ (since $H_2 = 0$), and in view of transversality conditions we get $\psi_s \equiv 0$. Therefore $\gamma = 0$ and ψ_y is continuous on the whole Δ. But then $\psi_y(t_1) = 0$ and, in view of homogeneity of the adjoint equation, $\psi_y \equiv 0$ on Δ_1, whence $\psi_y \equiv 0$ on the whole Δ, which easily implies that all Lagrange multipliers vanish, a contradiction. Thus, without loss of generality we can take $\alpha_0 = \psi_x = 1$.

Let us prove that, if $\psi_y(t') < 0$ for some $t' \neq t_1$, then $v \equiv 0$ on the "rest" of the corresponding interval Δ_k. Indeed, in this case by the continuity of ψ_y there exists a neighborhood $\mathcal{O}(t') \subset \Delta_k$ in which $\psi_y(t) < 0$. The maximality condition on $\mathcal{O}(t')$ gives $v \equiv 0$, and the adjoint equation $\dot\psi_y(t) = -\psi_x = -1$ implies that $\psi_y(t)$ stays decreasing further on Δ_k. Thus, $\psi_y(t) < 0$ on the "rest" of Δ_k and hence, $v \equiv 0$ there.

If $\psi_y(t') = 0$ for some $t' \neq t_1$, then $\dot\psi_2(t) < 0$ in a neighborhood $\mathcal{O}(t')$ and the above consideration gives that again $\psi_y < 0$ on the "rest" of Δ_k .

Therefore, ψ_y can change its sign only from plus to minus, and the optimal control is a piecewise constant function taking the values 1 or 0 with at most one switching from 1 to 0 on every Δ_k (which, in particular, implies that H_1 and H_2 are continuous everywhere except the discontinuity points of the control).

The extremals with $\psi_y \leq 0$ on $(t', T) \subset \Delta_2$ do not satisfy the MP, since they obviously contradict the transversality condition $\psi_y(T) = 0$. So, $\psi_y > 0$ on (t_1, T) and $v \equiv 1$ on the whole Δ_2 .

The extremals with $\psi_y \leq 0$ on $(t', t_1) \subset \Delta_1$ do not satisfy the MP too. Indeed, according to equalities $H_1(t_1 - 0) = 0$, $H_2(T) = 0$, and $\psi_y(T) = 0$, we get $y(t_1) + \psi_t = 0$, $y(T) + \psi_t = 0$, and then $y(t_1) = y(T)$. But since $v \equiv 1$ on Δ_2 , we have $\dot y > 0$ on Δ_2 , whence $y(t_1) < y(T)$, a contradiction.

Thus, $\psi_y > 0$ on $[0, t_1)$, and the problem possesses a single extremal with $v \equiv 1$ on the whole Δ, which is then optimal. Let us find it completely.

By a direct calculation it is easy to show [6] that ψ_y takes the form

$$\psi_y(t) = \begin{cases} \dfrac{C - y(t)}{g_1(y(t))}, \quad t \in \Delta_1\,, \\[2ex] \dfrac{y(T) - y(t)}{\sigma(s)\; g_2(y(t))}, \quad t \in \Delta_2\,. \end{cases} \tag{11}$$

Consider equalities $H_1(0) = 0$ and $H_2(T) = 0$. The first one implies that $\psi_y(0)\, g_1(0) + \psi_t = 0$, and the second one that $y(T) + \psi_t = 0$. Therefore, in view of (11) and the initial conditions, we obtain $y(T) = C$. Thus, the values $\psi_y(t_1 \pm 0)$ are explicitly expressed in terms of $y(t_1) = s$ and $y(T)$. Then, the jump condition for ψ_y at the point t_1 takes the form:

$$\frac{y(T) - s}{\sigma(s) g_2(s)} - \frac{y(T) - s}{g_1(s)} = -\gamma. \tag{12}$$

The adjoint equation and transversality conditions for ψ_s imply that $\sigma'(s) \int\limits_{t_1}^{T} \psi_y(t) g_2(y(t))\, dt = \gamma$. In view of (11) this equation reduces to the form:

$$\frac{\sigma'(s)}{\sigma(s)} \int_{t_1}^{T} (y(T) - y(t))\, dt \;=\; \gamma. \tag{13}$$

For a fixed t_1 the function $y(t)$ is determined as the solution to Cauchy problems $\dot y = g_1(y)$, $y(0) = 0$ on Δ_1 (which yields $s = y(t_1)$) and $\dot y = \sigma(s)\, g_2(y)$, $y(t_1) = s$ on Δ_2. Here

$$s = \int_0^{t_1} g_1(y(t))\, dt, \tag{14}$$

$$y(T) = s + \sigma(s) \int_{t_1}^{T} g_2(y(t))\, dt. \tag{15}$$

Substituting s and $y(T)$ into (12), we can find $\gamma = \gamma(t_1)$, and then (13) is an equation with a single unknown t_1.

Thus, relations (12)–(15) is a system of four equations with four unknown $t_1, s, y(T), \gamma$, which can be resolved in the non-singular case, and so, the complete extremal can be found.

Remark 4. This example was taken from paper [6], in which the authors investigate the extremal with $u \equiv 1$ by means of Hybrid Neccesary Principle, claiming that in this case it gives stronger results than MP does. Imposing assumptions on the functions g_1, g_2, σ, which guarantee the equality $\gamma = 0$, and using Hybrid Neccesary Principle, the authors obtain the optimality condition $\sigma'(s) \geq 0$. However, in the case $\gamma = 0$ the MP gives a stronger result $\sigma'(s) = 0$ (it follows from (13) and inequality $y(t) < y(T)$.) Moreover, when $\gamma = 0$ we obtain from (12) that the value of s also satisfies the equation $g_1(s) - \sigma(s)\, g_2(s) = 0$, which is not typical for functions

g_1, g_2, σ of generic type. For example, if σ strictly increases, then by virtue of (13) we necessarily have $\gamma > 0$, and therefore, the case $\gamma = 0$ is not typical.

Acknowledgments. This work was supported by the Russian Foundation for Basic Research, project no. 04-01-00482. The authors thank B.M. Miller and V.A. Dykhta for valuable discussions.

References

[1] *L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, E.F. Mischenko*, *Mathematical Theory of Optimal Processes (in Russian), M., Nauka, 1961.*

[2] *V.M. Alekseev, V.M. Tikhomirov, S.V. Fomin*, *Optimal Control (in Russian). M., Nauka, 1979.*

[3] *U.M. Volin, G.M. Ostrovskii*, *Maximum principle for discontinuous systems and its application to problems with state constraints. (in Russian) // Izvestia vuzov. Radiophysics, 1969, v. 12, no. 11, p. 1609–1621.*

[4] *H.J. Sussmann*, *A maximum principle for hybrid optimal control problems. // Proc. of 38th IEEE Conference on Decision and Control, Phoenix, 1999.*

[5] *M. Garavello, B. Piccoli*, *Hybrid necessary principle. // SIAM J. on Control and Optimization, 2005, v. 43, no. 5, p. 1867-1887.*

[6] *C. D'Apice, M. Garavello, R. Manzo, B. Piccoli*, *Hybrid optimal control: case study of a car with gears. // International Journal of Control, v. 76 (2003), 1272-1284.*

[7] *A.A. Milutin, A.V. Dmitruk, N.P. Osmolovskii*, *The maximum principle in optimal control (in Russian), M., Mech.-Math. Dept. of Moscow State University, 2004.*

[8] *A.V. Dmitruk, A.M. Kaganovich*, *Maximum principle for optimal control problems with intermediate constraints (in Russian). // In "Nonlinear Dynamics and Control", vol. 6, M., Nauka, 2006 (in Russian), to appear.*