sciendo
Control and Cybernetics
vol. 53 (2024) No. 1
pages: 77-107
DOI: 10.2478/candc-2024-0005

On the maximum principle when the endpoints of the optimal trajectory lie on the state boundary*

by

Andrei V. Dmitruk[1] **and Nikolai P. Osmolovskii**[2]

[1]Russian Academy of Sciences, Central Economics and Mathematics Institute,
Moscow, Russia
optcon@mail.ru.
[2]Systems Research Institute, Polish Academy of Science
Warszawa, Poland
nikolai.osmolovskii@ibspan.waw.pl.

Abstract: We consider an optimal control problem in the Mayer form with an autonomous control system, a bounded control set U, endpoint equality constraints, and one pointwise state inequality constraint. We analyze the case when the starting point of the optimal trajectory belongs to the state boundary. A nontrivial maximum principle was obtained for this case by A.Ya. Dubovitskii and V.A. Dubovitskii about 40 years ago, but the proof was written in an extremely condensed form. Here we offer a new proof of their result.

Keywords: optimal control, normed space, state constraint, Pontryagin function, Lebesgue–Stieltjes measure, singular measure, costate equation

1. Introduction

We consider an autonomous optimal control problem in the Mayer form on a non-fixed time interval $[t_0, t_1]$:

$$\mathcal{J}(x,u) := J(x(t_0), x(t_1)) \to \min, \tag{1}$$

$$K(x(t_0), x(t_1)) = 0, \tag{2}$$

$$\dot{x}(t) = f(x(t), u(t)) \quad \text{for a.e.} \quad t \in [t_0, t_1], \tag{3}$$

$$u(t) \in U \quad \text{for a.e.} \quad t \in [t_0, t_1], \tag{4}$$

*Submitted: May 2024; Accepted: June 2024

$$\Phi(x(t)) \leqslant 0 \quad \text{for all} \quad t \in [t_0, t_1], \tag{5}$$

where the mappings $J : \mathbb{R}^{2d(x)} \to \mathbb{R}$, $K : \mathbb{R}^{2d(x)} \to \mathbb{R}^{d(K)}$, and $f : \mathbb{R}^{d(x)+d(u)} \to \mathbb{R}^{d(x)}$ are continuously differentiable, the mapping $\Phi : \mathbb{R}^{d(x)} \to \mathbb{R}$ is twice continuously differentiable, and $U \subset \mathbb{R}^{d(u)}$ is a bounded set. By $d(a)$ we denote the dimension of the vector a, and we also denote $x_0 = x(t_0)$, $x_1 = x(t_1)$, $p = (x_0, x_1)$. Problem (1)–(5) will be called Problem A.

We emphasize that in this problem the state variable x and the control u must be optimally chosen together with their common domain $[t_0, t_1]$, since the latter is not fixed.

We say that a process $(x(t), u(t) \mid t \in [t_0, t_1])$ is *admissible* if the function $x : [t_0, t_1] \to \mathbb{R}^{d(x)}$ is Lipschitz continuous, the function $u : [t_0, t_1] \to \mathbb{R}^{d(u)}$ is measurable and essentially bounded, and all constraints (2)–(5) are satisfied on $[t_0, t_1]$. A process that gives the minimal value to the functional J among all admissible processes is called *(globally) optimal*. A process $(\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1])$ that gives the minimal value to the functional J among all admissible processes $(x(t), u(t) \mid t \in [t_0, t_1])$ such that

$$|t_0 - \hat{t}_0| < \varepsilon, \quad |t_1 - \hat{t}_1| < \varepsilon, \quad \max_{t \in [\hat{t}_0, \hat{t}_1] \cap [t_0, t_1]} |x(t) - \hat{x}(t)| < \varepsilon,$$

with some $\varepsilon > 0$, is called a *strong minimum*.

Let us recall the formulation of the *maximum principle* (MP) for a strong minimum process $(\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1])$. For a while, we additionally assume that

$$\Phi(\hat{x}(\hat{t}_0)) < 0, \qquad \Phi(\hat{x}(\hat{t}_1)) < 0, \tag{6}$$

i.e., roughly speaking, the endpoints of the optimal trajectory $\hat{x}(t)$ do not belong to the boundary of the state constraint.

Denote by $\mathbb{R}^{n*}$ the space of row vectors of dimension n.

By definition, the maximum principle for the process $(\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1])$ means that there is a tuple $\lambda = (\alpha_0, \beta, \psi, \mathrm{d}\mu)$ of Lagrange multipliers, where $\alpha_0 \in \mathbb{R}$, $\beta \in \mathbb{R}^{d(K)*}$, $\psi : [\hat{t}_0, \hat{t}_1] \to \mathbb{R}^{d(x)*}$ is a function of bounded variation*, $\mathrm{d}\mu$ is a Lebesgue–Stieltjes measure on $[\hat{t}_0, \hat{t}_1]$, such that the following conditions are satisfied

(i) non-negativity conditions $\alpha_0 \geqslant 0$, $\mathrm{d}\mu \geqslant 0$,

(ii) non-triviality condition $\alpha_0 + |\beta| + \int_{\hat{t}_0}^{\hat{t}_1} \mathrm{d}\mu > 0$,

(iii) complementary slackness condition $\Phi(\hat{x}(t))\,\mathrm{d}\mu(t) = 0, \quad t \in [\hat{t}_0, \hat{t}_1]$,

*Below we denote by $\mathrm{d}\psi$ the corresponding Lebesgue–Stieltjes measure on $[\hat{t}_0, \hat{t}_1]$.

(iv) adjoint equation, understood as equality between measures

$$-\,\mathrm{d}\psi(t) = \psi(t) f_x(\hat{x}(t), \hat{u}(t))\,\mathrm{d}t - \mathrm{d}\mu(t)\Phi'(\hat{x}(t)), \qquad t \in [\hat{t}_0, \hat{t}_1],$$

(v) transversality conditions

$$\psi(\hat{t}_0) = l_{x_0}(\hat{p}), \qquad -\psi(\hat{t}_1) = l_{x_1}(\hat{p}),$$

where $l(x_0, x_1) = \alpha_0 J(x_0, x_1) + \beta K(x_0, x_1)$ is the *endpoint Lagrange function*, and $\hat{p} = (\hat{x}(\hat{t}_0), \hat{x}(\hat{t}_1))$,

(vi) the constancy of the *Pontryagin function* $H = \psi f(x, u)$:

$$\psi(t) f(\hat{x}(t), \hat{u}(t)) = 0 \quad \text{a.e. on} \quad [\hat{t}_0, \hat{t}_1],$$

(vii) the maximality of the Pontryagin function:

$$\psi(t) f(\hat{x}(t), u) \leqslant 0 \quad \forall\, u \in U, \quad \forall\, t \in [\hat{t}_0, \hat{t}_1].$$

Under assumptions (6), the following theorem is valid, see Dubovitskii and Milyutin (1965).

Theorem 1 *If* $(\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1])$ *is a strong minimum in Problem A, then there exists a tuple* $\lambda = (\alpha_0, \beta, \psi, \mathrm{d}\mu)$ *of Lagrange multipliers that satisfies all conditions (i)–(vii) of the maximum principle.*

Assumptions (6) and the complementary slackness condition (iii) imply that the measure $\mathrm{d}\mu$ is equal to zero in some neighborhoods of the points $\hat{t}_0$ and $\hat{t}_1$, whence the adjoint equation (iv) yields that the function ψ is Lipschitz continuous in those neighborhoods.

Now, let us abandon assumptions (6). Then, one of the endpoints of the process $(\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1])$ may violate them. For definiteness, suppose that

$$\Phi(\hat{x}(\hat{t}_0)) \;=\; 0. \tag{7}$$

In this case, one can easily note that the tuple of multipliers

$$\alpha_0 = 0, \quad \beta = 0, \quad \psi(t) \equiv 0, \quad \mathrm{d}\mu = \delta(t - \hat{t}_0),$$

where $\delta(t)$ is the δ-function, satisfies all conditions (i)–(vii) of the maximum principle (MP). This tuple is nontrivial only because $\mathrm{d}\mu \neq 0$, and actually the maximum principle holds trivially in this case *for any admissible process satisfying* (7)*,* thus giving no information about the optimal process.

A.Ya. Dubovitskii and V.A. Dubovitskii (father and son) proposed in Dubovitskii and Dubovitskii (1985, 1987, 1995) quite simple conditions that guarantee the non-triviality of MP without assumptions (6). However, the proof they proposed for Problem A in Dubovitskii and Dubovitskii(1987) (which is practically the only available proof of their result), is far from being simple; moreover, it was given in an extremely concise form with a number of confusing typos. So, their proof is hardly accessible even for specialists. We have revised and partially changed that proof, and here we aim to present it in detail.

The theorem in question is as follows. Let $(\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1])$ be a process that is a strong minimum in Problem A. Suppose the following:

(a) there is an element $u^* \in U$ such that

$$\Phi'(\hat{x}(\hat{t}_0))f(\hat{x}(\hat{t}_0), u^*) < 0, \tag{8}$$

(b) for any pair $(x_0, x_1) \in \mathbb{R}^{2d(x)}$ the following implication is true

$$K(x_0, x_1) = 0 \quad \Longrightarrow \quad \Phi(x_0) = 0. \tag{9}$$

(c) the mapping, $\mathbb{R}^{2d(x)} \to \mathbb{R}^{d(K)}$, defined by $(\bar{x}_0, \bar{x}_1) \mapsto K_{x_0}(\hat{p})\bar{x}_0 + K_{x_1}(\hat{p})\bar{x}_1$, is onto; i.e., the rank of matrix $K'(\hat{p})$ is equal to $d(K)$.

Condition (a) is called the *controllability assumption*, (b) is the *consistency assumption*†, and (c) is the *regularity* of endpoint constraints.

In view of (b) we have $\Phi(\hat{x}(\hat{t}_0)) = 0$ (that is, the endpoint constraints force the left endpoint of the trajectory to lie on the "state boundary"). For the right endpoint of the trajectory, we assume for simplicity that $\Phi(\hat{x}(\hat{t}_1)) < 0$. Then, the adjoint variable ψ (which is a function of bounded variation) is continuous at $\hat{t}_1$.

For definiteness, it will be convenient to assume in what follows that the function $\psi(t)$ of bounded variation is right continuous at all $t \in [\hat{t}_0, \hat{t}_1)$, and also has some value $\psi(\hat{t}_0 - 0)$. Then, this function defines a measure

$$\mathrm{d}\psi(\{\hat{t}_0\}) = [\psi](\hat{t}_0) := \psi(\hat{t}_0) - \psi(\hat{t}_0 - 0)$$

of the one-point set $\{t_0\}$, where $[\psi](\hat{t}_0)$ denotes the jump of the function ψ at $\hat{t}_0$. This also applies to the non-decreasing function $\mu(t)$, which defines the measure $\mathrm{d}\mu$.

The main result is the following

Theorem 2 *If a process $(\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1])$ is a strong minimum in Problem A, and assumptions (a), (b), (c) hold true, then there is a tuple of Lagrange multipliers $(\alpha_0, \beta, \psi, \mathrm{d}\mu)$ such that all conditions (i) – (vii) of the maximum principle are fulfilled and* $\mathrm{d}\mu(\{\hat{t}_0\}) = 0$.

†Before the papers of Dubovitskii and Dubovitskii (1985, 1987), assumption (a) was used in Arutynuov and Tynyanskiy (1984), but assumption (b) was missing there, without which the result is incorrect.

The last relation means that the tuple of multipliers is non-trivial not for the reason that $\mathrm{d}\mu(\{\hat{t}_0\}) \neq 0$, i.e., the non-triviality condition has the form

$$\alpha_0 + |\beta| + \int_{(\hat{t}_0,\, \hat{t}_1]} \mathrm{d}\mu \,>\, 0.$$

2. Proof of the main result

The proof will be made in several steps.

Index θ

Let a process $(\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1])$ be a strong minimum in Problem A. With this process, we associate a family of optimization problems $\{B^\theta\}$ and their optimal solutions, labeled by a certain "index" θ.

Without loss of generality, we set $\hat{t}_0 = 0$.

By the "index" we mean a finite set of time instants t^s and control values u^s, i.e.,

$$\theta = \{(t^1, u^1),\, \ldots,\, (t^d, u^d)\},$$

where $0 < t^1 \leqslant \ldots \leqslant t^d < \hat{t}_1$, and $u^s \in U$, $s = 1, \ldots, d$ are arbitrary. The "length" d of the index depends on θ. For convenience, we set $t^0 = 0$ and $t^{d+1} = \hat{t}_1$.

Next, let us define a segment $[0, \tau_1]$ as follows: take the segment $[0, \hat{t}_1]$ and, at the points $0, t^1, \ldots, t^d$, insert segments of unit length, always preserving the position of the initial point 0. As a result, we obtain the segment $[0, \tau_1]$ with $\tau_1 = \hat{t}_1 + d + 1$, while the inserted segments have the form

$$\Delta^0 = [0, 1], \quad \Delta^1 = [t^1 + 1,\, t^1 + 2], \quad \Delta^2 = [t^2 + 2,\, t^2 + 3],\; \ldots,$$

$$\ldots,\; \Delta^d = [t^d + d,\, t^d + (d + 1)].$$

Set

$$E_0 = \bigcup_{k=0}^{d} \Delta^s, \qquad E_+ = [0, \tau_1] \setminus E_0\,,$$

and define the functions

$$v^\theta(\tau) = \begin{cases} 0, & \tau \in E_0, \\ 1, & \tau \in E_+, \end{cases} \qquad t^\theta(\tau) = \int_0^\tau v^\theta(r)dr, \qquad \tau \in [0, \tau_1]. \tag{10}$$

Then we have

$$\frac{\mathrm{d}t^\theta(\tau)}{\mathrm{d}\tau} = v^\theta(\tau) \quad \text{a.e. on} \quad [0,\tau_1], \qquad t^\theta(0)=0\,, \qquad t^\theta(\tau_1)=\hat{t}_1\,.$$

Thus, $t^\theta(\tau)$ is a piecewise linear, continuous, non-decreasing function that maps $[0,\tau_1]$ onto $[0,\hat{t}_1]$ being constant on each Δ^s. Moreover, $t^\theta(\Delta^s) = t^s$ for all $s = 0,1,\ldots,d$, where $t^0 = 0$.

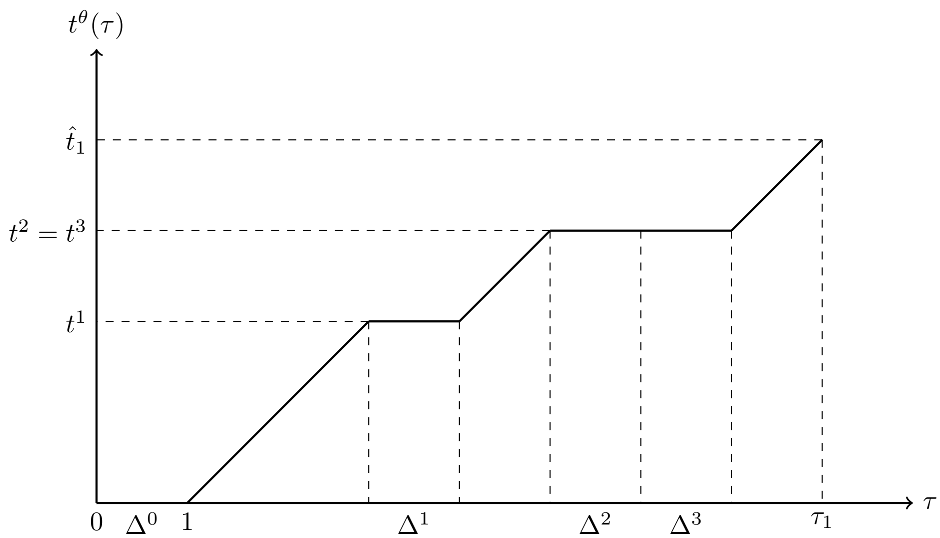

Figure 1: Function $t^\theta(\tau)$ for the case of $d=3$ and $t^1 < t^2 = t^3$

For convenience in notation, we also set $u^0 = u^*$, where the value of u^* is the same as in the controllability assumption (a). For a given index θ we define the functions

$$u^\theta(\tau) = \begin{cases} \hat{u}(t^\theta(\tau)), & \tau \in E_+, \\ u^s, \quad \tau\in\Delta^s, & s=0,1,\ldots,d, \end{cases} \qquad x^\theta(\tau) = \hat{x}(t^\theta(\tau)). \tag{11}$$

Obviously, the function $u^\theta(\tau)$ is measurable and essentially bounded with $u^\theta(\tau) \in U$ a.e. in $[0,\tau_1]$, and the function $x^\theta(\tau)$ is Lipschitz continuous. Moreover,

$$\frac{\mathrm{d}x^\theta(\tau)}{\mathrm{d}\tau} = v^\theta(\tau)\, f(x^\theta(\tau), u^\theta(\tau)) \quad \text{a.e. on} \quad [0,\tau_1], \quad x^\theta(0)=\hat{x}(0), \quad x^\theta(\tau_1) = \hat{x}(\hat{t}_1),$$

hence, the endpoints of the new trajectory $x^\theta(\tau)$ are the same as those of the original trajectory $\hat{x}(t)$. Note also that

$$x^\theta(\tau) \,=\, \hat{x}(0) \qquad \forall\, \tau \in [0,1]. \tag{12}$$

Note that some points $t^s > 0$ may coincide, for example, $t^{s'} = \ldots = t^{s''} = t^*$, therefore, at such a point t^* several segments of unit length are inserted sequentially, on each of which we have $v^\theta(\tau) = 0$ and $u^\theta(\tau) = u^s$, where the value u^s corresponds to the segment Δ^s. Note that the segment $\Delta^0 = [0,1]$ is separated from Δ^1 because $t^0 = 0 < t^1$.

The set E_0 is a finite union of closed intervals Δ^s, $s = 0,1,\ldots,d$, and the set E_+ is a finite union of intervals or semi-intervals. Consider the collection of all the intervals and semi-intervals of the sets E_0 and E_+, order it, and denote the elements of this collection by σ_k, $k = 0,1,\ldots,m$. Thus, $[0,\tau_1] = \sigma_0 \cup \sigma_1 \cup \cdots \cup \sigma_m$, where $\sigma_0 = \Delta^0 = [0,1]$ and different σ_k do not overlap.

Denote by $\chi_k(\tau)$ the characteristic function of the set $\sigma_k,\ k = 0,1,\ldots,m.$

Problem B^θ of the index θ

Given the index θ, we fix the constructed interval $\Delta = [0,\tau_1]$ and the function $u^\theta(\tau)$, which will not be varied.

We introduce the space

$$\mathcal{W} = \,\mathrm{Lip}\,(\Delta,\mathbb{R}^{d(x)}) \times L_\infty(\Delta^0,\mathbb{R}) \times \mathbb{R}^m$$

of elements $w = (x(\cdot),\, v_0(\cdot),\, z)$, where

$$x(\cdot) \in \mathrm{Lip}\,(\Delta,\mathbb{R}^{d(x)}), \quad v_0(\cdot) \in L_\infty(\Delta^0,\mathbb{R}), \quad z = (z_1,\,\ldots,\,z_m) \in \mathbb{R}^m.$$

The functions v_0, formally defined only on $\Delta^0 = [0,1]$, will be also considered on the whole interval $\Delta = [0,\tau_1]$, assuming that $v_0(\tau) = 0$ a.e. on $[1,\tau_1]$. We will apply this remark to all functions from $L_\infty(\Delta^0,\mathbb{R})$ that will be used.

For brevity, let us put $L_\infty(\Delta^0) := L_\infty(\Delta^0,\mathbb{R})$ and $L_\infty(\Delta) := L_\infty(\Delta,\mathbb{R})$, and let $L_\infty^-(\Delta)$ be the cone of nonpositive functions in the space $L_\infty(\Delta)$.

Given an element $w = (x(\cdot),\, v_0(\cdot),\, z)$, we define the function

$$v(\tau) = \, v_0(\tau) + \sum_{k=1}^{m} z_k\, \chi_k(\tau), \tag{13}$$

i.e., $v_0(\tau)$ defines the values of $v(\tau)$ on the interval Δ^0, and z_k is the value of $v(\tau)$ on the set σ_k , $k = 1,\ldots,m$.

In the space $\mathcal{W}$, consider the following *Problem* B^θ:

$$J(x(0), x(\tau_1)) \to \min, \tag{14}$$

$$v_0(\cdot) \geqslant 0, \qquad z \geqslant 0, \tag{15}$$

$$S(x)\Big|_\tau := \frac{\Phi(x(\tau)) - \Phi(x(0))}{\tau} \in L_\infty^-(\Delta), \tag{16}$$

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} - v(\tau) f(x(\tau), u^\theta(\tau)) = 0, \tag{17}$$

$$K(x(0), x(\tau_1)) = 0, \tag{18}$$

where $v(\tau)$ is defined by (13). We call it *the associated problem*, which corresponds to the index θ and the control $u^\theta(\cdot)$.

The quite unexpected representation of the state constraint in the form (16), proposed in Dubovitskii and Dubovitskii (1987), is the key point of the proof. This is what made it possible to perform a variational analysis of the situation with "endpoint at the state boundary" and obtain a non-trivial maximum principle for it.

Define the optimal point $w^\theta = (x^\theta,\, v_0^\theta,\, z^\theta) \in \mathcal{W}$ of the Problem B^θ. Set $v_0^\theta(\tau) \equiv 0$ and $z^\theta = (z_1^\theta, \ldots, z_m^\theta)$, where

$$z_k^\theta = \begin{cases} 0, & \sigma_k \in E_0 \\ 1, & \sigma_k \in E_+, \end{cases} \qquad k = 1, \ldots, m.$$

Then, according to (13),

$$v^\theta(\tau) = \sum_{k=1}^{m} z_k^\theta\, \chi_k(\tau).$$

It is easy to prove that, since $(\hat{x}, \hat{u} \mid t \in [0, \hat{t}_1])$ is a strong minimum in Problem A, the point $w^\theta = (x^\theta,\, v_0^\theta,\, z^\theta) \in \mathcal{W}$ is a local minimum in Problem B^θ. (We leave this to the reader.) The optimality conditions of the latter can be derived from the *general Lagrange multipliers rule* (see the below Theorem 3 in Appendix), and so, first we have to check its assumptions.

The operator $S :\, \mathrm{Lip}\,(\Delta, \mathbb{R}^{d(x)}) \to L_\infty(\Delta, \mathbb{R})$ is a composition of operators

$$\Phi :\, \mathrm{Lip}\,(\Delta, \mathbb{R}^{d(x)}) \to \mathrm{Lip}\,(\Delta, \mathbb{R}), \qquad x(\tau) \mapsto y(\tau) = \Phi(x(\tau)),$$

and

$$P :\, \mathrm{Lip}\,(\Delta, \mathbb{R}) \to L_\infty(\Delta, \mathbb{R}), \qquad y(\tau) \mapsto \frac{y(\tau) - y(0)}{\tau},$$

that is, $S(x) = P(\Phi(x))$. Obviously both Φ and P are Fréchet differentiable. Therefore, S is also Fréchet differentiable with the Fréchet derivative at the point x^θ of the form

$$S'(x^\theta)\,\bar{x}\Big|_\tau = \frac{\Phi'(x^\theta(\tau))\,\bar{x}(\tau) - \Phi'(x^\theta(0))\,\bar{x}(0)}{\tau} \qquad \forall\,\bar{x} \in \mathrm{Lip}\,(\Delta, \mathbb{R}^{d(x)}).$$

Here $x^\theta(0) = \hat{x}(0)$. For brevity, we put $\hat{x}_0 = \hat{x}(0)$.

The equality operator defined by equations (17) and (18) maps the space $\mathcal{W}$ to the space $L_\infty(\Delta, \mathbb{R}^{d(x)}) \times \mathbb{R}^{d(K)}$. As is known, the image of its derivative is closed, since the derivative of operator (17) is surjective, and operator (18) is finite-dimensional‡.

Recall that $\Phi(x^\theta(0)) = \Phi(\hat{x}(0)) = 0$. For every $\delta > 0$, we define a measurable set

$$M_\delta = \{\tau \in \Delta \;:\; S(x^\theta(\tau)) \geqslant -\delta\,\}.$$

Local maximum principle for the index θ

The point $(x^\theta(\cdot),\, v_0^\theta(\cdot),\, z^\theta)$ of local minimum in problem B^θ satisfies the *local maximum principle* (LMP), which we will formulate using Lagrange multipliers $\alpha_0, \beta, \rho, \xi, \eta, \nu$. Here, the multiplier α_0 corresponds to the cost (14), the row vector β – to the endpoint equality (18), the functional ν and the row vector η – to the inequality constraints (15), the functional ρ refers to the control system (17), and the functional ξ – to the "state constraint" (16).

Let $L_\infty^*(\Delta)$ be the dual to the space $L_\infty(\Delta)$, and $L_\infty^*(\Delta^0)$ be the space of functionals in $L_\infty^*(\Delta)$, concentrated on $\Delta^0 = [0,1]$ (i.e. vanishing on $[1, \tau_1]$.)

By Theorem 3, for the point $(x^\theta,\, v_0^\theta,\, z^\theta)$ there exists a tuple $(\alpha_0, \beta, \rho, \xi, \eta, \nu)$, where

$$\alpha_0 \in \mathbb{R},\; \beta \in \mathbb{R}^{d(K)*},\; \eta \in \mathbb{R}^{m*},\; \rho \in L_\infty^*(\Delta, \mathbb{R}^{d(x)}),\; \xi \in L_\infty^*(\Delta),\; \nu \in L_\infty^*(\Delta^0),$$

$$\alpha_0 \geqslant 0, \quad \eta \geqslant 0, \quad \eta\, z^\theta = 0, \quad \xi \geqslant 0, \quad \nu \geqslant 0, \tag{19}$$

the functional ξ is concentrated on the set M_δ for any $\delta > 0$, the functional ν is concentrated on the interval $[0,1]$, such that the normalization condition holds:

$$\alpha_0 + |\beta| + |\eta| + \|\rho\| + \|\xi\| = 1, \tag{20}$$

and the Lagrange function

$$L(w) = l(x_0, x_1) \,-\eta\, z \,-\langle \nu, v_0\rangle \,+\, \Big\langle \rho,\, \frac{\mathrm{d}x}{\mathrm{d}\tau} - v(\tau) f(x(\tau), u^\theta(\tau))\Big\rangle +$$

‡In this case, the equality operator of problem B^θ is said to satisfy *the weakened regularity condition*, which is a minimal assumption concerning the equality constraints.

$$+\Big\langle \xi,\ \frac{\Phi(x(\tau))-\Phi(x(0))}{\tau}\Big\rangle$$

is stationary at the point $(x^\theta,\ v_0^\theta,\ z^\theta)$, i.e. for any test element $\bar w=(\bar x,\bar v_0,\bar z)\in \mathcal{W}$, the following equality holds:

$$l'(\hat p)\,\bar p\ -\eta\,\bar z\ -\langle \nu,\bar v_0\rangle\ +\ \big\langle \rho,\ \dot{\bar x}-\bar v f^\theta-v^\theta f_x^\theta\bar x\big\rangle +$$

$$+\ \Big\langle \xi,\ \frac{1}{\tau}\Big(\Phi'(x^\theta(\tau))\,\bar x(\tau)-\Phi'(x^\theta(0))\,\bar x(0)\Big)\Big\rangle\ =\ 0, \tag{21}$$

where

$$l=\ \alpha_0 J+\beta K,\quad \hat p=(\hat x(0),\hat x(t_1)),\quad \bar p=(\bar x(0),\bar x(\tau_1)),$$

$$l'(\hat p)\bar p=\ l_{x_0}(\hat p)\,\bar x(0)+l_{x_1}(\hat p)\,\bar x(\tau_1),\quad \bar v(\tau)=\ \bar v_0(\tau)+\sum_{k=1}^m \bar z_k\,\chi_k(\tau),$$

$$f^\theta=f(x^\theta,u^\theta),\qquad f_x^\theta=f_x(x^\theta,u^\theta).$$

Here and below we omit the dependence of L and l on the Lagrange multipliers.

Note that condition (20) ensures that not all multipliers $\alpha_0,\beta,\rho,\xi,\eta,\nu$ are zero. Indeed, if $\alpha_0+|\beta|+|\eta|+\|\rho\|+\|\xi\|=0$, then $\nu=0$ by (21).

Preliminary analysis of LMP for the index θ

By putting $\bar x=0$ into equation (21), we get $\eta\bar z+\langle\nu,\bar v_0\rangle+\langle\rho,\ \bar v f^\theta\rangle=0$ for all $\bar v_0\,,\bar z$; in more detail,

$$\sum_{j=1}^m \eta_j\,\bar z_j+\langle\nu,\bar v_0\rangle+\langle\rho,\ (\bar v_0+\sum_{j=1}^m \bar z_j\,\chi_j)f^\theta\rangle=0\qquad \forall\,\bar v_0,\ \bar z,$$

whence

$$\eta_j+\langle\rho,\chi_j f^\theta\rangle=0,\qquad j=1,\dots,m, \tag{22}$$

and

$$\langle\nu,\bar v_0\rangle\ +\ \langle\rho,\bar v_0 f^\theta\rangle\ =\ 0\qquad \forall\,\bar v_0\in L_\infty(\Delta^0). \tag{23}$$

Let us note that conditions (22) allow us to exclude $|\eta|$ from the normalization condition (20). Indeed, if $\alpha_0+|\beta|+\|\rho\|+\|\xi\|=0$, then by (22) also $\eta=0$. Therefore, we will further use the following equivalent normalization:

$$\alpha_0+|\beta|+\|\rho\|+\|\xi\|=1. \tag{24}$$

Since $\eta \geqslant 0$ and $\eta\, z^\theta = 0$, we have: $\eta_k = 0$ if $z_k^\theta > 0$, and $\eta_k \geqslant 0$ if $z_k^\theta = 0$. In other words, $\eta_k = 0$ if $\sigma_k \subset E_+$ and $\eta_k \geqslant 0$ if $\sigma_k \subset E_0$. Note also that $\chi_k f^\theta = f(x^\theta, u^k)$ on $\sigma_k \subset E_0$. Therefore, condition (22) implies

$$\begin{aligned} &\langle \rho, \chi_k f(x^\theta, u^\theta)\rangle = 0 \quad \text{if} \quad \sigma_k \subset E_+\,, \\ &\langle \rho, \chi_k f(x^\theta, u^k)\rangle \leqslant 0 \quad \text{if} \quad \sigma_k \subset E_0\,, \end{aligned} \qquad k = 1, \dots, m. \tag{25}$$

Condition (23) will be analyzed later.

Now, upon putting $\bar v_0 = 0$ and $\bar z = 0$ (hence $\bar v = 0$ as well) in (21), we have

$$l'(\hat p)\,\bar p + \langle \rho, \dot{\bar x} - v^\theta f_x^\theta\, \bar x\rangle + \Big\langle \xi,\ \frac{1}{\tau}\Big(\Phi'(x^\theta(\tau))\,\bar x(\tau) - \Phi'(x^\theta(0))\,\bar x(0)\Big)\Big\rangle \ = \ 0 \tag{26}$$

for all $\bar x \in \mathrm{Lip}\,(\Delta, \mathbb{R}^{d(x)})$.

In the next few steps of the LMP analysis, up to the end of the proof of Lemma 2, we basically follow Dubovitskii and Dubovitskii (1987), with minor modifications. These steps, especially Lemmas 1 and 2, can be definitely called the "pearls" of optimal control theory.

Functional ρ

Consider first the functional $\rho \in L_\infty^*(\Delta, \mathbb{R}^{d(x)})$ related to the control system (17). For any small $\varepsilon \in (0,1)$ we define the interval $\Delta_\varepsilon := [\varepsilon,\, \tau_1]$, and (for $p = \infty$ or $p = 1$) denote by $L_p(\Delta_\varepsilon, \mathbb{R}^{d(x)})$ the subspace of functions $\omega \in L_p(\Delta, \mathbb{R}^{d(x)})$ that vanish on the interval $[0, \varepsilon]$ (as usual, $L_1(\Delta, \mathbb{R}^{d(x)})$ stands for the space of summable functions $\omega : \Delta \to \mathbb{R}^{d(x)}$ with the norm $\|\omega\|_1 = \int_\Delta |\omega(\tau)|\, \mathrm{d}\tau$).

For any small $\varepsilon > 0$ we denote by ρ_ε the restriction of the functional ρ to the interval Δ_ε, more precisely, the restriction to the space $L_\infty(\Delta_\varepsilon, \mathbb{R}^{d(x)})$. Take an arbitrary $\bar\omega \in L_\infty(\Delta_\varepsilon, \mathbb{R}^{d(x)})$ and let $\bar x \in \mathrm{Lip}\,(\Delta, \mathbb{R}^{d(x)})$ be such that $\dot{\bar x} = \bar\omega$ and $\bar x(0) = 0$. Note that $\bar x(\tau) = 0$ on $[0, \varepsilon]$ and $\|\bar x\|_\infty \leqslant \|\bar\omega\|_1$. As we substitute this $\bar x$ in the relation (26), we get

$$\langle \rho_\varepsilon, \bar\omega\rangle = \, -l'_{x_1}(\hat p)\, \bar x(\tau_1) + \langle \rho_\varepsilon, v^\theta f_x^\theta\, \bar x\rangle - \big\langle \xi, \frac{1}{\tau}\, \Phi'(x^\theta)\, \bar x\big\rangle,$$

whence

$$|\langle \rho_\varepsilon, \bar\omega\rangle| \ \leqslant\ K(\varepsilon)\, \|\bar x\|_\infty \ \leqslant\ K(\varepsilon) \int_{\Delta_\varepsilon} |\bar\omega|\, \mathrm{d}\tau \qquad \forall\, \bar\omega \in L_\infty(\Delta_\varepsilon, \mathbb{R}^{d(x)})$$

with some $K(\varepsilon)$. By the Hahn–Banach theorem, this estimate implies that ρ_ε can be extended to a functional over the space $L_1(\Delta_\varepsilon, \mathbb{R}^{d(x)})$, therefore there is a function $\psi_\varepsilon \in L_\infty(\Delta_\varepsilon, \mathbb{R}^{d(x)^*})$ such that

$$\langle \rho_\varepsilon, \omega\rangle = \int_{\Delta_\varepsilon} \psi_\varepsilon\, \omega\; \mathrm{d}\tau \qquad \forall\, \omega \in L_1(\Delta_\varepsilon, \mathbb{R}^{d(x)}). \tag{27}$$

Obviously, if $\varepsilon' < \varepsilon$, then $\Delta_\varepsilon \subset \Delta_{\varepsilon'}$ and $L_1(\Delta_\varepsilon, \mathbb{R}^{d(x)}) \subset L_1(\Delta_{\varepsilon'}, \mathbb{R}^{d(x)})$, hence $\rho_{\varepsilon'} = \rho_\varepsilon$ on $L_1(\Delta_\varepsilon, \mathbb{R}^{d(x)})$ and the family $\{\psi_\varepsilon\}$ has a similar property, that is, $\psi_{\varepsilon'}\chi_{\Delta_\varepsilon} = \psi_\varepsilon$. Therefore, there exists a measurable function $\psi(\tau)$, defined on the entire segment Δ, such that $\forall\, \omega \in L_\infty(\Delta)$, $\forall\, \varepsilon > 0$, for $\omega_\varepsilon = \omega\, \chi_{\Delta_e}$ we have

$$\langle \rho,\, \omega_\varepsilon\rangle = \int_\Delta \psi(\tau)\, \omega_\varepsilon(\tau)\; \mathrm{d}\tau. \tag{28}$$

From (27) it follows that, for any $\varepsilon > 0$

$$\int_{\Delta_\varepsilon} |\psi|\, dt \;=\; \int_{\Delta_\varepsilon} |\psi_\varepsilon|\, dt \;=\; \|\rho_\varepsilon\| \leqslant \|\rho\|,$$

whence $\psi \in L^1(\Delta, \mathbb{R}^{d(x)^*})$ and $\|\psi\|_1 \leqslant \|\rho\|$.

Let a functional $\rho_0 \in L^*_\infty(\Delta, \mathbb{R}^{d(x)})$ be such that

$$\langle \rho,\, \omega\rangle = \langle \rho_0,\, \omega\rangle + \int_\Delta \psi(\tau)\, \omega(\tau)\, d\tau \qquad \forall\, \omega \in L_\infty(\Delta, \mathbb{R}^{d(x)}), \tag{29}$$

that is, $\rho = \rho_0 + \psi$ (here ψ is the functional over $L_\infty(\Delta, \mathbb{R}^{d(x)})$, defined by the function ψ). Later we will show that $\rho_0 = 0$, but for now we just note that, due to (28), ρ_0 is concentrated on any interval $[0, \varepsilon]$.

Functional ξ

Now, let us deal with the functional $\xi \in L^*_\infty(\Delta)$ related to the state constraint (16). Consider its restriction to the subspace

$$C_0 := \{y \in C(\Delta) \;:\; y(0) = 0\},$$

where $C(\Delta)$ is the space of continuous functions $y : \Delta \to \mathbb{R}$. Using the Riesz theorem, we find that there is a measure $\mathrm{d}m \in C^*(\Delta)$ such that

$$\langle \xi, y\rangle \;= \int_\Delta y(\tau)\; \mathrm{d}m(\tau) \qquad \forall\, y \in C_0, \quad \text{and} \quad \mathrm{d}m(\{0\}) = 0. \tag{30}$$

On the entire space $L_\infty(\Delta)$, we have the equality

$$\langle \xi, y\rangle = \langle \xi_0, y\rangle + \int_\Delta y(\tau)\; \mathrm{d}m(\tau) \qquad \forall\, y \in L_\infty(\Delta) \tag{31}$$

with some $\xi_0 \in L^*_\infty(\Delta)$. We will briefly write this equality as $\xi = \xi_0 + \mathrm{d}m$.

Condition (30) implies that for any $\varepsilon > 0$ we have $\xi = \mathrm{d}m$ on Δ_ε; the latter means that

$$\langle \xi, \chi_{\Delta_\varepsilon} y\rangle = \int_{\Delta_\varepsilon} y(\tau)\, \mathrm{d}m(\tau) \qquad \forall\, y \in L_\infty(\Delta).$$

Consequently, $\xi_0 = 0$ on Δ_ε for all $\varepsilon > 0$, and therefore ξ_0 is concentrated on any segment $[0,\, \varepsilon]$. Since $\xi \geqslant 0$, then $\xi_0 \geqslant 0$ and $\mathrm{d}m \geqslant 0$ as well. Moreover, $\|\,\mathrm{d}m\| \leqslant \|\xi\| \leqslant 1$, and

$$\mathrm{d}m((0,\delta]) \to 0 \qquad \text{as} \quad \delta \to 0+\,. \tag{32}$$

The latter is true, because $\mathrm{d}m(\{0\}) = 0$.

Note that notations ρ_0 and ξ_0 do not refer to the interval Δ^0 (as might be wrongly understood); both these functionals are concentrated on any segment $[0,\, \varepsilon]$ with $\varepsilon > 0$, and hence they are singular functionals.

An important step in the proof is the following

Lemma 1 *Both $\rho_0 = 0$ and $\xi_0 = 0$.*

Proof. a) Consider the interval $\sigma_0 = \Delta^0 = [0,1] \subset E_0$. Recall that

$$x^\theta(\tau) = x^\theta(0) = \hat{x}_0 \quad \text{ and } \quad u^\theta(\tau) = u^* \quad \text{on } \Delta^0. \tag{33}$$

Hence $f^\theta(\tau) := f(x^\theta(\tau), u^\theta(\tau)) = f(\hat{x}_0, u^*)$ is a constant vector on Δ^0.

Take any $\varepsilon \in (0,1)$ and put $\bar{v}_0 = \chi_{[0,\varepsilon]}$ in equation (23). Recalling that $\nu \geqslant 0$, we get $\langle \rho,\, \chi_{[0,\varepsilon]} f(\hat{x}(0), u^*)\rangle \leqslant 0$, and since $\rho = \rho_0 + \psi$ by (29), we have

$$\langle \rho_0,\, \chi_{[0,\varepsilon]} f(\hat{x}_0, u^*)\rangle + \int_{[0,\varepsilon]} \psi(\tau) f(\hat{x}_0, u^*)\, \mathrm{d}\tau \leqslant 0.$$

Since ρ_0 is concentrated on $[0,\varepsilon]$ for any $\varepsilon > 0$, we obtain

$$\langle \rho_0, f(\hat{x}_0, u^*)\rangle \leqslant 0. \tag{34}$$

b) Let us get back to relation (26). It implies that, for any $\bar{x} \in \mathrm{Lip}\,(\Delta, \mathbb{R}^{d(x)})$ with $\bar{x}(0) = 0$, we have

$$l'_{x_1}(\hat{p})\,\bar{x}(\tau_1) + \langle \rho,\, \dot{\bar{x}} - v^\theta f^\theta_x\, \bar{x}\rangle + \left\langle \xi,\, \frac{1}{\tau}\, \Phi'(x^\theta)\, \bar{x} \right\rangle = 0. \tag{35}$$

Note that here the function $\frac{1}{\tau}\, \bar{x}(\tau)$ is bounded, since the function $\bar{x}$ is Lipschitz continuous with $\bar{x}(0) = 0$.

We claim that

$$\langle \rho_0, \dot{\bar{x}} \rangle \;+\; \Big\langle \xi_0,\; \frac{1}{\tau}\, \Phi'(x^\theta)\, \bar{x} \Big\rangle \;=0. \tag{36}$$

Indeed, fix any $\bar{x} \in \mathrm{Lip}\,(\Delta, \mathbb{R}^{d(x)})$ with $\bar{x}(0) = 0$, and define the functions

$$\bar{x}_\varepsilon(\tau) = \begin{cases} \bar{x}(\tau), & \tau \in [0, \varepsilon], \\ \bar{x}(\varepsilon), & \tau \in [\varepsilon, \tau_1], \end{cases} \qquad \varepsilon > 0.$$

Since $\|\bar{x}_\varepsilon\|_\infty \to 0$ as $\varepsilon \to 0$, equality (35) yields

$$\langle \rho,\, \dot{\bar{x}}_\varepsilon \rangle \;+\; \Big\langle \xi,\; \frac{1}{\tau}\, \Phi'(x^\theta)\, \bar{x}_\varepsilon \Big\rangle \;\to\; 0 \quad \text{as} \quad \varepsilon \to 0. \tag{37}$$

Since $\dot{\bar{x}}_\varepsilon = 0$ on $[\varepsilon, \tau_1]$, the first summand here can be replaced by $\langle \rho_0,\, \dot{\bar{x}}_\varepsilon \rangle$.

Now recall that $\xi = \xi_0 + \mathrm{d}m$. We claim that

$$\int_{0+}^{\tau_1} \frac{1}{\tau}\, \Phi'(x^\theta(\tau))\, \bar{x}_\varepsilon(\tau)\; \mathrm{d}m \;\to\; 0 \qquad \text{as} \quad \varepsilon \to 0. \tag{38}$$

Since the function $\Phi'(x^\theta(\tau))$ is bounded, it suffices to prove that

$$\int_{0+}^{\tau_1} \frac{1}{\tau}\, |\bar{x}_\varepsilon(\tau)|\; \mathrm{d}m \;\to\; 0 \qquad \text{as} \quad \varepsilon \to 0. \tag{39}$$

Fix any $\delta > 0$. Since $|x_\varepsilon(\tau)| \leqslant B\tau$ for some B (depending only on the Lipschitz constant of the function $\bar{x}$), we have

$$\int_{0+}^{\delta} \frac{1}{\tau}\, |\bar{x}_\varepsilon|\, \mathrm{d}m \;\leqslant\; B \cdot \mathrm{d}m((0, \delta]) \qquad \text{for all} \quad \varepsilon > 0,$$

and since the function $1/\tau$ is bounded on $[\delta, \tau_1]$, we have $\displaystyle\int_{\delta}^{\tau_1} \frac{1}{\tau}\, |\bar{x}_\varepsilon|\, \mathrm{d}m \to 0$ as $\varepsilon \to 0$, whence

$$\int_{0+}^{\tau_1} \frac{1}{\tau}\, |\bar{x}_\varepsilon(\tau)|\, \mathrm{d}m \;\leqslant\; B \cdot \mathrm{d}m((0, \delta]) \;+\; o(1) \qquad \text{as} \;\; \varepsilon \to 0.$$

In view of (32), this obviously implies (39), and so, relation (38) is proven.

Therefore, in the second summand of (37) we can replace ξ by ξ_0, thus obtaining

$$\langle \rho_0,\, \dot{\bar{x}}_\varepsilon \rangle \;+\; \Big\langle \xi_0,\; \frac{1}{\tau}\, \Phi'(x^\theta)\, \bar{x}_\varepsilon \Big\rangle \;\to\; 0 \qquad \text{as} \;\; \varepsilon \to 0.$$

But both ρ_0 and ξ_0 allows one to replace $\bar{x}_\varepsilon$ by $\bar{x}$, so that we get (36). Recall that equality (36) is obtained for any $\bar{x} \in \mathrm{Lip}\,(\Delta, \mathbb{R}^{d(x)})$ with $\bar{x}(0) = 0$.

c) Now, by setting $\bar{x}(\tau) = f(\hat{x}_0, u^*)\,\tau$ in (36), we get

$$\langle \rho_0, f(\hat{x}_0, u^*)\rangle \,+\, \langle \xi_0,\, \Phi'(\hat{x}_0)\, f(\hat{x}_0, u^*)\rangle \,=0. \tag{40}$$

By (34), the first summand here is non-positive. Since $\xi_0 \geqslant 0$ and $\Phi'(\hat{x}_0)\, f(\hat{x}_0, u^*) < 0$ by the controllability assumption, the second summand is also non-positive. Therefore, both these terms are equal to zero:

$$\langle \rho_0, f(\hat{x}_0, u^*)\rangle = 0, \qquad \big\langle \xi_0,\, \Phi'(\hat{x}_0)\, f(\hat{x}_0, u^*)\big\rangle = 0.$$

Moreover, the last equality, in view of the controllability assumption, gives $\xi_0 = 0$, and then (36) obviously gives $\rho_0 = 0$. Lemma 1 is proven. □

Thus, the functionals ρ and ξ have the form

$$\langle \rho, \omega\rangle = \int_\Delta \psi(\tau)\,\omega(\tau)\; \mathrm{d}\tau \qquad \forall\, \omega \in L_\infty(\Delta, \mathbb{R}^{d(x)}),$$

$$\langle \xi, y\rangle = \int_\Delta y(\tau)\; \mathrm{d}m(\tau) \qquad \forall\, y \in L_\infty(\Delta, \mathbb{R}),$$

where $\psi \in L^1(\Delta, \mathbb{R}^{d(x)^*})$ and $\mathrm{d}m \in C^*(\Delta)$ with $\mathrm{d}m(\{0\}) = 0$.

In view of this, condition (25) means that, for $k = 1, \ldots, m$,

$$\int_{\sigma_k} \psi(\tau) f(x^\theta(\tau), u^\theta(\tau))\; \mathrm{d}\tau \; \begin{cases} = 0, & \text{if} \quad \sigma_k \subset E_+\,, \\ \leqslant 0, & \text{if} \quad \sigma_k \subset E_0\,. \end{cases} \tag{41}$$

As was already noted, since $\xi \geqslant 0$, then $\mathrm{d}m \geqslant 0$. Moreover, since ξ is concentrated on the set M_δ for any $\delta > 0$, we obtain the complementary slackness condition

$$\Phi(x^\theta(\tau))\; \mathrm{d}m(\tau) = 0 \qquad \text{a.e. on} \quad \Delta.$$

Finally, relation (23) now becomes

$$\int_{\Delta^0} \psi(\tau) f(\hat{x}_0, u^*)\, \bar{v}_0(\tau)\, \mathrm{d}\tau + \langle \nu, \bar{v}_0\rangle = 0 \qquad \forall\, \bar{v}_0 \in L_\infty(\Delta^0).$$

Since $\nu \geqslant 0$ by (19), then ν is in fact an integral functional defined by an integrable function $\tilde{\nu}(\tau) \geqslant 0$ such that

$$\psi(\tau) f(\hat{x}_0, u^*) \,+\, \tilde{\nu}(\tau) \,= 0 \quad \text{for a.a. } \tau \in [0,1].$$

This relation is equivalent to the inequality

$$\psi(\tau) f(\hat{x}_0, u^*) \leqslant 0 \quad \text{for a.a. } \tau \in [0,1]. \tag{42}$$

So, the collection of multipliers in LMP is now as follows:

$$\alpha \geqslant 0, \quad \beta \in \mathbb{R}^{d(K)^*}, \quad \psi \in L_1(\Delta, \mathbb{R}^{d(x)^*}), \quad \mathrm{d}m \in C^*(\Delta),$$

and the following conditions are met:

$$\mathrm{d}m \geqslant 0, \quad \mathrm{d}m(\{0\}) = 0, \quad \Phi(x^\theta(\tau))\, \mathrm{d}m(\tau) = 0, \quad \text{conditions (41), (42),}$$

normalization condition (see (24)):

$$\alpha_0 + |\beta| + \int_\Delta |\psi(\tau)|\, \mathrm{d}\tau + \int_\Delta \mathrm{d}m(\tau) = 1, \tag{43}$$

and, according to (26), for any $\bar{x} \in \operatorname{Lip}(\Delta, \mathbb{R}^{d(x)})$

$$l'(\hat{p})\,\bar{p} + \int_\Delta \psi(\dot{\bar{x}} - v^\theta f_x^\theta\, \bar{x})\, \mathrm{d}\tau +$$

$$+ \int_\Delta \frac{1}{\tau} \Big(\Phi'(x^\theta(\tau))\, \bar{x}(\tau) - \Phi'(\hat{x}_0)\, \bar{x}(0) \Big)\, \mathrm{d}m = 0. \tag{44}$$

(We exclude the multiplier ν, since it brings no more information.)

Adjoint equation and transversality conditions in the LMP

Let us analyze condition (44). To do this, we introduce a right continuous function $\gamma(\tau)$ as a solution to equation

$$\mathrm{d}\gamma = -v^\theta \psi f_x^\theta\, \mathrm{d}\tau + \frac{1}{\tau}\, \Phi'(x^\theta)\, \mathrm{d}m, \qquad \gamma(\tau_1) = -l_{x_1}, \tag{45}$$

where $l_{x_1} = l_{x_1}(\hat{p})$. Recall that by assumption, $\Phi(\hat{x}(\hat{t}_1)) = \Phi(x^\theta(\tau_1)) < 0$, therefore, $\mathrm{d}m(\tau) = 0$ in some neighborhood of the point τ_1, and so, $\gamma(\tau)$ is continuous at τ_1.

Let $BV(\Delta)$ denote the space of functions of bounded variation $\varphi : [0, \tau_1] \to \mathbb{R}$ for which the values $\varphi(0-)$ and $\varphi(\tau_1+)$ are defined. Hence, the jumps

$$[\varphi(0)] := \varphi(0+) - \varphi(0-), \qquad [\varphi(\tau_1)] := \varphi(\tau_1+) - \varphi(\tau_1-)$$

are also defined. Each such function determines a Lebesgue–Stieltjes measure $\mathrm{d}\varphi$ on $[0,\tau_1]$. We will assume that φ is right continuous, so $\varphi(0+)=\varphi(0)$ and $\varphi(\tau_1+)=\varphi(\tau_1)$.

Since $\psi\in L_1(\Delta,\mathbb{R}^{d(x)^*})$, and $m\in BV(\Delta)$, then for any $\varepsilon>0$ the function γ has bounded variation on the segment $\Delta_\varepsilon=[\varepsilon,\tau_1]$. (But, perhaps, not on the entire segment $\Delta=[0,\tau_1]$, since, a priori, the measure $\frac{1}{\tau}\,\mathrm{d}m$ can have an infinite integral there.)

Fix any $\varepsilon\in(0,1]$ and consider equation (44) for all $\bar x\in\mathrm{Lip}\,[0,\tau_1]$ such that $\bar x(t)=\mathrm{const}$ on $[0,\varepsilon]$. Notice that since $\dot{\bar x}=0$ on $[0,\varepsilon]$ and $v^\theta=0$ on $[0,1]$, we have

$$\int_{[0,\varepsilon]}\psi\big(\dot{\bar x}-v^\theta f_x^\theta\,\bar x\big)\;\mathrm{d}\tau\;=\;0.$$

Moreover, since $x^\theta(\tau)=\hat x_0$ and $\bar x(\tau)=\bar x(0)$ on $[0,\varepsilon]$, we also have

$$\int_{[0,\varepsilon]}\frac{1}{\tau}\,\Big(\Phi'(x^\theta(\tau))\,\bar x(\tau)-\Phi'(\hat x_0)\,\bar x(0)\Big)\;\mathrm{d}m\;=\;0.$$

Therefore, both integrals in (44) should be taken over the interval $[\varepsilon,\tau_1]$ only, and so, we have the equality

$$l'_{x_0}\,\bar x(0)+l'_{x_1}\,\bar x(\tau_1)\;+\;\int_\varepsilon^{\tau_1}\psi\,\dot{\bar x}\,d\tau\,+$$
$$+\int_\varepsilon^{\tau_1}\bar x\;\mathrm{d}\gamma\;-\int_\varepsilon^{\tau_1}\frac{1}{\tau}\,\Phi'(\hat x_0)\,\bar x(0)\;\mathrm{d}m\;=\;0.$$

The last integral here is finite, since $\tau\geqslant\varepsilon$. Upon integrating the second integral by parts and taking into account the terminal condition in (45), we obtain

$$l_{x_0}\bar x(0)-\gamma(\varepsilon)\,\bar x(\varepsilon)\;+\;\int_\varepsilon^{\tau_1}(\psi-\gamma)\,\dot{\bar x}\,d\tau\;\;-\int_\varepsilon^{\tau_1}\frac{1}{\tau}\,\Phi'(\hat x_0)\,\bar x(0)\;\mathrm{d}m\;=\;0.$$

Since $\bar x(0)=\bar x(\varepsilon)$, we have

$$\Big(l_{x_0}-\gamma(\varepsilon)\;-\int_\varepsilon^{\tau_1}\frac{\mathrm{d}m}{\tau}\;\Phi'(\hat x_0)\Big)\,\bar x(\varepsilon)\;+\;\int_\varepsilon^{\tau_1}(\psi-\gamma)\,\dot{\bar x}\,d\tau\;\;=\;0.$$

This equation holds for any $\bar x\in\mathrm{Lip}\,[\varepsilon,\tau_1]$, that is, for any $\bar x(\varepsilon)\in\mathbb{R}^{d(x)}$ and $\dot{\bar x}\in L_\infty[\varepsilon,\tau_1]$, independent each of the other. Therefore, the coefficients at both of these variables should vanish, whence $\psi=\gamma$ almost everywhere on $[\varepsilon,\tau_1]$, and so, changing ψ on a set of zero measure, we may assume that it is right continuous and, in view of (45), satisfies the relations

$$\mathrm{d}\psi=\;-\,v^\theta\psi f_x^\theta\,\mathrm{d}\tau+\frac{1}{\tau}\,\Phi'(x^\theta)\;\mathrm{d}m, \tag{46}$$

$$\psi(\tau_1) \;=\; -l_{x_1}\,, \tag{47}$$

$$\psi(\varepsilon) \;=\; l_{x_0} \;-\int_{\varepsilon}^{\tau_1} \frac{1}{\tau}\,\Phi'(\hat{x}_0)\;\mathrm{d}m, \tag{48}$$

where equation (46) holds on the interval $[\varepsilon, \tau_1]$. But since $\varepsilon \in (0,1]$ is arbitrary, this equation holds for all $\tau \in (0, t_1]$, while relation (48) gives

$$\psi(\tau) \;=\; l_{x_0} \;-\Big(\int_{\tau}^{\tau_1} \frac{\mathrm{d}m}{s}\Big)\,\Phi'(\hat{x}_0) \qquad \forall\,\tau \in (0,\tau_1]. \tag{49}$$

Yet, we cannot assume that $\psi(0+)$ is finite.

The measure $\frac{1}{\tau}\;\mathrm{d}m$

Now consider the measure $\frac{1}{\tau}\;\mathrm{d}m$. If we show that it is finite, then $\psi(0+)$ would be finite, and, by (46), ψ would also have a bounded variation. Recall that the control set U in our problem is bounded, and that u^θ depends on the choice of values $u^k \in U$ on the intervals Δ^k, $k = 1,\dots,d$.

Lemma 2 *There exists a constant B, common to all values $u \in U$, such that*

$$\int_{\Delta} \frac{\mathrm{d}m}{\tau} \;\leqslant\; B \quad \textit{and} \quad \mathit{Var}\,\psi\Big|_{\Delta} \;\leqslant\; B.$$

Proof. According to (42), for almost all $\tau \in \Delta^0 = [0,1]$ we have $\psi(\tau)f(\hat{x}_0,u^*) \leqslant 0$. Then, by multiplying (49) by $f(\hat{x}_0,u^*)$, we obtain for a.a. $\tau \in \Delta^0$:

$$\psi(\tau)f(\hat{x}_0,u^*) \;=\; l_{x_0}f(\hat{x}_0,u^*) \;-\Big(\int_{\tau}^{\tau_1} \frac{1}{s}\;\mathrm{d}m\Big)\,\Phi'(\hat{x}_0)\,f(\hat{x}_0,u^*) \;\leqslant 0.$$

By the controllability assumption, $A := \,-\Phi'(\hat{x}_0)f(\hat{x}_0,u^*) > 0$, and therefore,

$$A\int_{\tau}^{\tau_1} \frac{\mathrm{d}m}{s} \;\leqslant\; -\,l_{x_0}f(\hat{x}_0,u^*) \qquad \text{for a.a.} \quad \tau \in [0,1].$$

Since $\alpha_0 + |\beta| \leqslant 1$, the right hand side of this inequality is uniformly bounded by some constant B, and hence

$$A\int_{\tau}^{\tau_1} \frac{\mathrm{d}m}{s} \;\leqslant\; B \qquad \forall\,\tau \in [0,1],$$

which yields

$$\int_{0}^{\tau_1} \frac{\mathrm{d}m}{s} \;\leqslant\; B/A\,.$$

This, and (46), imply (by the Gronwall lemma) the estimate $\|\psi\|_\infty \leqslant B'$, whence also $\mathrm{Var}\psi|_\Delta \leqslant B''$ for some constants B', B'' common for all choices of $u \in U$.

Lemma 2 is proven. □

From this lemma and (49) it follows that $\psi(0+)$ is finite and, since ψ is right continuous, we obtain

$$\psi(0) \;=\; \psi(0+) \;=\; l_{x_0} \;-\Big(\int_0^{\tau_1} \frac{\mathrm{d}m}{\tau}\Big)\,\Phi'(\hat{x}_0). \tag{50}$$

Thus, for the collection $\alpha_0 \geqslant 0,\ \beta,\ \psi \in BV(\Delta, \mathbb{R}^{d(x)^*}),\ \mathrm{d}m \in C^*(\Delta)$, we can impose, instead of (43), the normalization

$$\alpha_0 \,+\, |\beta| \,+ \int_0^{\tau_1} \frac{\mathrm{d}m(\tau)}{\tau} \;=\; 1. \tag{51}$$

Indeed, if $\alpha_0 = 0,\ \ \beta = 0$, and $\mathrm{d}m = 0$, then $l = 0$ and equation (49) obviously implies $\psi(\tau) \equiv 0$, which contradicts (43).

Let us now define a measure $\mathrm{d}\mu(\tau)$ on $[0, \tau_1]$ such that

$$\mathrm{d}\mu(\{0\}) = 0, \quad \text{and} \quad \mathrm{d}\mu(\tau) = \frac{1}{\tau}\,\mathrm{d}m(\tau) \quad \text{on} \;\; (0, \tau_1].$$

Then, relations (46), (47), and (50) transform, respectively, into

$$\mathrm{d}\psi = \,-\,v^\theta \psi f_x^\theta\,\mathrm{d}\tau + \,\Phi'(x^\theta)\,\mathrm{d}\mu, \tag{52}$$

$$\psi(\tau_1) \;=\; -l_{x_1}\,, \tag{53}$$

$$\psi(0) \;=\; l_{x_0} \;-\left(\int_0^{\tau_1} \mathrm{d}\mu\right)\,\Phi'(\hat{x}_0)\,, \tag{54}$$

while the normalization becomes

$$\alpha_0 \,+\, |\beta| \,+ \int_0^{\tau_1} \mathrm{d}\mu(\tau) \;=\; 1. \tag{55}$$

From now on, the Lagrange multipliers ψ and $\mathrm{d}\mu$ will be also marked by the superscript θ.

As this was shown above, at the point $(x^\theta,\, v_0^\theta\,,\, z^\theta) \in \mathcal{W}$ in Problem B^θ, the following necessary optimality conditions hold:

there exists a collection of Lagrange multipliers $(\alpha_0, \beta, \psi^\theta, \mathrm{d}\mu^\theta)$ such that

$$\alpha_0 \in \mathbb{R}, \quad \beta \in \mathbb{R}^{d(K)*}, \quad \psi^\theta \in BV(\Delta, \mathbb{R}^{d(x)*}), \quad \mathrm{d}\mu^\theta \in C^*(\Delta),$$

$$\alpha_0 \geqslant 0, \quad \mathrm{d}\mu^\theta \geqslant 0, \quad \Phi(x^\theta(\tau))\,\mathrm{d}\mu^\theta(\tau) = 0, \quad \mathrm{d}\mu^\theta(\{0\}) = 0, \tag{56}$$

$$\alpha_0 + |\beta| + \int_0^{\tau_1} \mathrm{d}\mu^\theta(\tau) \;=\; 1, \tag{57}$$

$$\mathrm{d}\psi^\theta = \; -v^\theta \psi^\theta f_x^\theta \,\mathrm{d}\tau + \Phi'(x^\theta)\,\mathrm{d}\mu^\theta, \tag{58}$$

$$\psi^\theta(0) \;=\; l_{x_0} - \left(\int_0^{\tau_1} \mathrm{d}\mu^\theta\right) \Phi'(\hat{x}_0)\,, \qquad \psi^\theta(\tau_1) = -l_{x_1}, \tag{59}$$

$$\int_{\sigma_k} \psi^\theta(\tau) f(x^\theta(\tau), u^\theta(\tau))\,\mathrm{d}\tau \begin{cases} = 0 & \text{if } \sigma_k \subset E_+\,, \\ \leqslant 0 & \text{if } \sigma_k \subset E_0\,, \end{cases} \qquad k = 1, \ldots, m, \tag{60}$$

$$\psi^\theta(\tau) f(\hat{x}_0, u^*) \leqslant 0, \qquad \tau \in \sigma_0 = [0,1]. \tag{61}$$

Recall that, by assumption, the functions ψ^θ and μ^θ are right continuous. Since $\mathrm{d}\mu^\theta(\{0\}) = 0$, by the adjoint equation (58) we have $\mathrm{d}\psi^\theta(\{0\}) = 0$. Consequently,

$$\mu^\theta(0-) = \mu^\theta(0) = \mu^\theta(0+) \quad \text{and} \quad \psi^\theta(0-) = \psi^\theta(0) = \psi^\theta(0+).$$

Moreover, both functions ψ^θ and μ^θ are continuous at τ_1.

Pre-maximality condition

Consider in more detail the second condition in (60), which refers to the set E_0. Take an arbitrary interval $\sigma_k \subset E_0$ with $k > 0$, that is $\sigma_k \neq [0,1]$. Let us denote it as $\sigma_k = [\tau', \tau'']$. On this interval, $u^\theta(\tau) = u^k$ is a constant vector, and $v^\theta(\tau) = 0$, whence the value $x^\theta(\tau)$ is also constant, which we denote by x^k. Then, by virtue of (58),

$$\mathrm{d}\psi^\theta(\tau) \;=\; \Phi'(x^k)\,\mathrm{d}\mu^\theta(\tau) \quad \text{on } [\tau', \tau''].$$

Therefore, $\psi^\theta(\tau)$ moves in the space $\mathbb{R}^{d(x)*}$ in a constant direction $\Phi'(x^k)$ on σ_k. According to the second condition in (60),

$$\int_{\tau'}^{\tau''} \psi^\theta(\tau) f(x^k, u^k)\,\mathrm{d}\tau \leqslant 0. \tag{62}$$

There can be other intervals in E_0, adjoining the interval $[\tau', \tau'']$ on the left or the right. (Under the mapping $\tau \mapsto t$, they all are taken to the same point t^k.) Let $[\tau'_*, \tau''_*]$ be the union of the interval $[\tau', \tau'']$ with all adjacent intervals from E_0, if any. Then, $v^\theta(\tau) = 0$ on this whole united interval; so, as before, $x^\theta(\tau) = x^k$ is constant, $\mathrm{d}\psi^\theta(\tau) = \Phi'(x^k)\,\mathrm{d}\mu^\theta(\tau)$, and therefore, $\psi^\theta(\tau)$ still changes in the same constant direction $\Phi'(x^k)$ on $[\tau'_*, \tau''_*]$.

Let us consider the integrand $h^k(\tau) := \psi^\theta(\tau) f(x^k, u^k)$ from (62) on the interval $[\tau'_*, \tau''_*]$. Note that, on this interval, the control $u^\theta(\tau)$, in general, is not constant, but only piecewise constant. However, we fix the value u^k and consider the function $h^k(\tau)$ with this value even on the "alien" intervals from E_0, adjacent to $[\tau', \tau'']$. Like $\psi^\theta(\tau)$, it is a function of bounded variation. On the whole interval $[\tau'_*, \tau''_*]$, we have

$$\mathrm{d}h^k(\tau) = \mathrm{d}\psi^\theta(\tau) f(x^k, u^k) = \Phi'(x^k)\, f(x^k, u^k)\, \mathrm{d}\mu^\theta(\tau).$$

For definiteness, let the constant $\Phi'(x^k)\, f(x^k, u^k)$ be nonnegative. Then $h^k(\tau)$ does not decrease, whence (62) implies $h^k(\tau'+0) \leqslant 0$. Moreover, $h^k(\tau'_*+0) \leqslant 0$, since $\tau'_* \leqslant \tau'$ and $\mathrm{d}\mu \geqslant 0$.

We claim that in this case also $h^k(\tau'_*-0) \leqslant 0$. Indeed, by the adjoint equation (58), the jump $[h^k](\tau'_*)$ of the function $h^k(\tau)$ at the point τ'_* occurs in the same direction:

$$[h^k](\tau'_*) := h^k(\tau'_* + 0) - h^k(\tau'_* - 0) = \Phi'(x^k) f(x^k, u^k)\, [\mu^\theta](\tau'_*) \geqslant 0.$$

Hence $h^k(\tau'_* - 0) = h^k(\tau'_* + 0) - [h^k](\tau'_*) \leqslant 0$, q.e.d. □

The case, in which $\Phi'(x^k)\, f(x^k, u^k) \leqslant 0$, similarly gives $h^k(\tau''_* + 0) \leqslant 0$.

So, we have shown that the second condition from (60) guarantees the fulfillment of at least one of the inequalities

$$h^k(\tau'_* - 0) \leqslant 0 \quad \text{or} \quad h^k(\tau''_* + 0) \leqslant 0.$$

In other words, at least one of the inequalities[§]

$$\psi^\theta(\tau'_* - 0) f(x^k, u^k) \leqslant 0 \quad \text{or} \quad \psi^\theta(\tau'_* + 0) f(x^k, u^k) \leqslant 0 \tag{63}$$

is satisfied, where $[\tau'_*, \tau''_*]$ is the maximal interval containing the given interval $\sigma_k \subset E_0$, on which the value $v^\theta(\tau) = 0$ is preserved.

Let us now rewrite the obtained conditions (56)–(63) in terms of the original time t. This will make it possible to consider the conditions obtained for different indices θ on one and the same original interval $[0, \hat{t}_1]$.

[§]Here we write $\psi^\theta(\tau'_* + 0)$ for clarity, despite the fact that, by agreement, ψ^θ is right continuous, so we could simply write $\psi^\theta(\tau'_*)$.

A finite-valued maximum principle of index θ

Thus, on the interval $\Delta = [0, \tau_1]$ we have a piecewise linear, continuous, non-decreasing function $t^\theta(\tau)$, which maps it onto the interval $[0, \hat{t}_1]$, being constant on each segment $\sigma_k \subset E_0$. In particular, $t^\theta(\tau) = 0$ for $\tau \in \sigma_0 = [0, 1]$.

Moreover, on $[0, \tau_1]$ there exist functions $u^\theta(\tau)$ and $x^\theta(\tau)$, generated by the original $\hat{u}(t)$ and $\hat{x}(t)$ according to formulas (11), there is also a right continuous function of bounded variation $\psi^\theta(\tau)$, there is a non-negative measure $\mathrm{d}\mu^\theta(\tau)$, corresponding to a non-decreasing right continuous function $\mu^\theta(\tau)$.

For each $t \in [0, \hat{t}_1]$, let $\tau^\theta(t)$ be the largest root of the equation $t^\theta(\tau) = t$, and let $\tau^\theta(0-) = 0$. Since $\sigma_0 = [0, 1] \subset E_0$, we have $\tau^\theta(0) = \tau^\theta(0+) = 1$, so

$$[\tau^\theta](0) := \tau^\theta(0+) - \tau^\theta(0-) = 1.$$

Clearly, the function $\tau^\theta : (0, \hat{t}_1] \to (1, \tau_1]$ is strictly increasing and right continuous. It has discontinuities at the given points t^s, $s = 1, \dots, d$ (and only at them):

$$[\tau^\theta](t^s) = \tau''_{s*} - \tau'_{s*}, \qquad s = 1, \dots, d,$$

where $[\tau'_{s*}, \tau''_{s*}]$ is the above defined maximal segment corresponding to the point t^s. (Recall also that some points t^s may coincide.)

It is easily seen that $x^\theta(\tau^\theta(t)) = \hat{x}(t)$ for all $t \in [0, \hat{t}_1]$, and $u^\theta(\tau^\theta(t)) = \hat{u}(t)$ for almost all $t \in [0, \hat{t}_1]$, i.e., by the time transformation $t \mapsto \tau^\theta(t)$ we get back to the original $\hat{x}(t)$ and $\hat{u}(t)$.

Further, we set

$$\mu(t) = \mu^\theta(\tau^\theta(t)), \qquad \psi(t) = \psi^\theta(\tau^\theta(t)), \qquad t \in [0, \hat{t}_1].$$

Then, $\mu(0-) = \mu^\theta(\tau^\theta(0-)) = \mu^\theta(0)$ and

$$\psi(0-) = \psi^\theta(\tau^\theta(0-)) = \psi^\theta(0) = l_{x_0} - \left(\int_0^{\tau_1} \mathrm{d}\mu^\theta \right).$$

It is easy to check that $\mu(t)$ is still a non-decreasing function that has jumps at the points t^s:

$$[\mu](t^s) = \mu^\theta(\tau''_{s*}+) - \mu^\theta(\tau'_{s*}-), \qquad s = 1, \dots, d,$$

$$[\mu](0) = \mu^\theta(\tau^\theta(0+)) - \mu^\theta(\tau^\theta(0-)) = \mu^\theta(1) - \mu^\theta(0) = \int_{[0,1]} \mathrm{d}\mu^\theta,$$

and $\psi(t)$ is a function of bounded variation (with the value $\psi(0-)$) satisfying the equation

$$\mathrm{d}\psi(t) = -\psi(t) f_x(\hat{x}(t), \hat{u}(t))\,\mathrm{d}t + \Phi'(\hat{x}(t))\,\mathrm{d}\mu(t), \qquad t \in [0, \hat{t}_1],$$

and having the same terminal values as $\psi^\theta(\tau)$. (Here we also took into account the fact that for $\tau \in E_+$ and the corresponding $t = t^\theta(\tau)$, we have $\mathrm{d}\mu(t) = \mathrm{d}\mu^\theta(\tau)$).

Clearly, the complementary slackness condition $\Phi(x^\theta(\tau))\,\mathrm{d}\mu^\theta(\tau) = 0$, $\tau \in [0, \tau_1]$ for the measure $\mathrm{d}\mu^\theta$ implies the same condition for the measure $\mathrm{d}\mu$:

$$\Phi(\hat{x}(t))\,\mathrm{d}\mu(t) = 0, \qquad t \in [0, \hat{t}_1].$$

Due to this condition and our assumption $\Phi(\hat{x}(\hat{t}_1)) < 0$, in the vicinity of $\hat{t}_1$ the measure "does not work": $\mathrm{d}\mu(t) = 0$, and so, the function $\psi(t)$ is continuous at $\hat{t}_1$.

Thus, conditions (56)–(61) can be rewritten in the original time $t \in [0, \hat{t}_1]$ as follows.

For any index θ, there is a tuple $\lambda = (\alpha_0, \beta, \mathrm{d}\mu(t))$ and a corresponding function of bounded variation $\psi(t)$, such that the following conditions are met:

(i) $\alpha_0 \geqslant 0,\ \mathrm{d}\mu \geqslant 0,$

(ii) $\alpha_0 + |\beta| + \int\limits_{[0,\hat{t}_1]} \mathrm{d}\mu = 1,$

(iii) $\Phi(\hat{x}(t))\,\mathrm{d}\mu(t) = 0, \quad t \in [0, \hat{t}_1],$

(iv) $\mathrm{d}\psi(t) = -\psi(t) f_x(\hat{x}(t), \hat{u}(t))\,\mathrm{d}t + \Phi'(\hat{x}(t))\,\mathrm{d}\mu(t), \quad \mathrm{d}\psi(\{0\}) = 0,$

(v) $\psi(0-) = l_{x_0} - \Phi'(\hat{x}(\hat{t}_0)) \int\limits_{[0,\hat{t}_1]} \mathrm{d}\mu, \qquad \psi(\hat{t}_1) = -l_{x_1},$

(vi) $\int\limits_{t^s}^{t^{s+1}} \psi(t)\, f(\hat{x}(t), \hat{u}(t))\,\mathrm{d}t = 0, \qquad s = 0, 1, \ldots, d,$

(vii) for any pair (t^s, u^s) from the index θ, at least one of the inequalities is satisfied

$$\psi(t^s - 0) f(\hat{x}(t^s), u^s) \leqslant 0 \quad \text{or} \quad \psi(t^s + 0) f(\hat{x}(t^s), u^s) \leqslant 0.$$

Condition (vi) is obtained here from the first condition (60) in view of the fact that, on each $\sigma = [t^s, t^{s+1}] \subset E_+$ with $t^s < t^{s+1}$, $s = 0, 1, \ldots, d$, the mapping $\tau \to t$ is one-to-one, $v^\theta(\tau) = 1$, and therefore $\mathrm{d}\tau = \mathrm{d}t$.

Condition (vii) follows from (63). The set of all conditions (i)–(vii) forms a "partial maximum principle" corresponding to the given index θ.

Thus, for any θ, we have obtained a tuple of Lagrange multipliers which generate the function $\psi(t)$, so that conditions (i)–(vii) are satisfied. These Lagrange multipliers, in general, depend on the index θ. Conditions (i)–(v) are the same for all indices, and conditions (vi)–(vii) are connected with each individual index. Our aim now is to pass to conditions (vi)–(vii) with a tuple of multipliers independent of θ.

Passage to a global maximum principle

Now, we will "organize" the obtained family of partial maximum principles.

For a given index θ, denote by Λ^θ the set of all tuples

$$\lambda = \ (\alpha_0, \beta,\, \mathrm{d}\mu) \,\in \mathbb{R} \times \mathbb{R}^{d(K)*} \times C^*([0,\hat{t}_1], \mathbb{R}),$$

for which there exists a function of bounded variation ψ such that conditions (i)–(vii) of the "finite-valued maximum principle of index θ" hold. This is a set in the space

$$Y^* = \ \mathbb{R} \times \mathbb{R}^{d(K)*} \times C^*([0,\hat{t}_1], \mathbb{R}),$$

dual to the space $Y = \mathbb{R} \times \mathbb{R}^{d(K)} \times C([0,\hat{t}_1], \mathbb{R})$.

The key fact in our proof is that the set Λ^θ is compact in the weak* topology of Y (cf. the proof of the compactness of such a set in Dmitruk and Osmolovskii, 2018).

Lemma 3 Λ^θ *is a weak* compact set, that is, a compact set with respect to the usual convergence of finite-dimensional vectors* (α_0, β) *and the weak* convergence of measures* $\mathrm{d}\mu(t)$ *in the space* $C^*([0,\hat{t}_1], \mathbb{R})$.

Proof. Let a tuple $\lambda = (\alpha_0, \beta,\, \mathrm{d}\mu(t))$ and the corresponding function of bounded variation $\psi(t)$ be given. First of all, we show that $\psi(t)$ is uniquely determined (up to values at its discontinuity points) by the vector (α_0, β), the adjoint equation (iv), and any of the end conditions (v), for example, the left one.

Let $W(t)$ be the fundamental matrix of the homogeneous equation

$$\dot{W} \ = \ -W f_x(\hat{x}, \hat{u})\,, \qquad W(0) = I,$$

where I is the identity matrix. Then, the Cauchy formula says that, for all t,

$$\psi(t-0) = \Big(\psi(0-) + \int_0^{t-0} \Phi'(\hat{x}(s)) W^{-1}(s)\, \mathrm{d}\mu(s)\Big) W(t) \tag{64}$$

and a similar formula is valid for $\psi(t+0)$.

Since the weak-* topology of the space Y^* is metrizable (because Y is separable), then to establish compactness it is sufficient to consider an arbitrary sequence of its elements λ_n. We can assume that the vectors (α_{0n}, β_n) converge to some(α_0, β).

Since the norms of all measures $\mathrm{d}\mu_n$ are uniformly bounded due to (ii), then it follows from (64) and the boundedness of all initial values $\psi_n(0-)$, that the functions $\psi_n(t)$ are uniformly bounded by a common constant. Hence, passing to a subsequence, we can assume that the measures $\mathrm{d}\mu_n$ and $\mathrm{d}\psi_n$ weakly* converge to some measures $\mathrm{d}\mu$ and $\mathrm{d}\psi$, respectively, the limit measures are related by equations (iv), and, according to the Helly theorems (see, for example, Kolmogorov and Fomin, 1999, Ch. 6), the corresponding functions $\psi_n(t)$ converges for each t to some function of bounded variation $\psi(t)$. If needed, we can change its values at a countable set of its discontinuity points t, so that ψ would become right continuous on $[0, \hat{t}_1]$. Under this operation, the measure $\mathrm{d}\psi$ would not change, and now the limit pair ψ and μ would satisfy relation (64) for all t.

Obviously, conditions (i)–(vi) are preserved when passing to the limit. In particular, (iii) is equivalent to the fact that, for any continuous test function $\zeta(t)$, we have

$$\int\limits_{[0,\hat{t}_1]} \zeta(t)\, \Phi(\hat{x}(t))\, \mathrm{d}\mu(t) \;=\; 0.$$

Obviously, this property is preserved when passing to the weak* limit.

Consider now condition (vii). Let us fix any $s \in \{1, \dots d\}$, where d corresponds to θ. We hope that there will be no confusion between this s and numbering n within the sequence. Recall that u^s is any predetermined point in U.

Consider the sequence of functions of bounded variation

$$h_n(t) = \;\psi_n(t)\, f(x^s, u^s).$$

For it, we also have convergence

$$h_n(t) \to h(t) := \;\psi(t)\, f(x^s, u^s) \qquad \forall\, t$$

and, in addition, taking into account (iv) (which is written for $\psi(t) = \psi_n(t)$ and then multiplied by $f(x^s, u^s)$), we get

$$\mathrm{d}h_n(t) = \; -\psi_n(t)\, f_x(\hat{x}(t), \hat{u}(t))\, f(x^s, u^s)\, \mathrm{d}t + \Phi'(\hat{x}(t))\, f(x^s, u^s)\, \mathrm{d}\mu_n(t). \quad (65)$$

Moreover, the jump at the point t^s is

$$[h_n](t^s) = \;\Phi'(x^s)\, f(x^s, u^s)\, [\mu_n](t^s).$$

According to (vii), at least one of the inequalities holds:

$$h_n(t^s - 0) \leqslant 0 \quad \text{or} \quad h_n(t^s + 0) \leqslant 0.$$

Without loss of generality, we assume that $\Phi'(x^s)\,f(x^s, u^s) \geqslant 0$. Then $[h_n](t^s) \geqslant 0$, and therefore $h_n(t^s - 0) \leqslant 0$ for all n. Let us show that the limit function h satisfies the same inequality: $h(t^s - 0) \leqslant 0$.

Fix any $\varepsilon > 0$. From the uniform boundedness of all $\psi_n(t)$ and the continuity of the function $\Phi'(\hat{x}(t))$ it follows, by virtue of (65), that there exist a $\delta > 0$ and a constant c such that

$$\mathrm{d}h_n(t) \geqslant -c\,\mathrm{d}t - \varepsilon\,\mathrm{d}\mu_n(t) \qquad \forall\, n$$

on the interval $(t^s - \delta,\, t^s)$. Then, by integrating over interval (t, t^s) and taking into account normalization (ii), we have on this interval

$$h_n(t^s - 0) - h_n(t) \;\geqslant\; -c\,\delta \,-\, \varepsilon \int\limits_{(t^s-\delta,\; t^s)} \mathrm{d}\mu_n(t) \;\geqslant\; -c\,\delta - \varepsilon.$$

Reducing δ, if necessary, we have

$$h_n(t^s - 0) \,-\, h_n(t) \geqslant -2\,\varepsilon,$$

i.e.,

$$h_n(t) \leqslant h_n(t^s - 0) \,+\, 2\,\varepsilon.$$

Since $h_n(t^s - 0) \leqslant 0$, we obtain $h_n(t) \leqslant 2\,\varepsilon$ on the interval $(t^s - \delta,\, t^s)$.

Since ε and δ do not depend on the number n, the same inequality holds for the pointwise limit: $h(t) \leqslant 2\,\varepsilon$ on the same interval, and therefore $h(t^s - 0) \leqslant 2\,\varepsilon$. Since $\varepsilon > 0$ is arbitrary, we get $h(t^s - 0) \leqslant 0$. Lemma 3 is proven. □

Thus, taking all possible indices θ, we have a nonempty compact set Λ^θ for each of them. Let us show that the family of all these compacta form a centered system (i.e., have the finite intersection property). To this end, we introduce a partial order in the set of all indices. We say that $\theta_1 \subset \theta_2$ if each pair (t^s, u^s) of θ_1 belongs to θ_2. For any two indices θ_1 and θ_2, there is a third index containing each of them, for example, their union. Obviously, when an index θ expands, the set Λ^θ narrows, i.e., the inclusion $\theta_1 \subset \theta_2$ implies the inverse inclusion $\Lambda^{\theta_1} \supset \Lambda^{\theta_2}$.

Now, let be given any finite collection of compacta $\Lambda^{\theta_1}, \ldots, \Lambda^{\theta_r}$. Take any index θ containing all indices $\theta_1, \ldots, \theta_r$. Then the nonempty compact set Λ^θ is contained in each of compacta $\Lambda^{\theta_1}, \ldots, \Lambda^{\theta_r}$, and consequently, is contained in their intersection. Thus, the family $\{\Lambda^\theta\}$ is centered, and, therefore, its total intersection is nonempty:

$$\Lambda^* := \bigcap_\theta \Lambda^\theta \neq .$$

Now, take an arbitrary tuple of multipliers $\lambda = (\alpha_0, \beta, \mathrm{d}\mu) \in \Lambda^*$ and let ψ be the corresponding adjoint function. By definition, conditions (i)—(vii) hold for this tuple.

The fulfillment of condition (vi) in any index means that for any interval (t', t'')

$$\int_{t'}^{t''} \psi(t)\, f(\hat{x}(t), \hat{u}(t))\, \mathrm{d}t = 0,$$

(since there is an index containing the points t', t'') and this is equivalent to the fulfillment of the equality

$$\psi(t)\, f(\hat{x}(t), \hat{u}(t)) = 0 \qquad \text{a.e. on } [0, \hat{t}_1]. \tag{66}$$

Finally, the fulfillment of (vii) for the chosen “universal” tuple means that for any point $t \in (t', t'')$ and any $u \in U$, at least one of the inequalities holds:

$$\psi(t-0)\, f(\hat{x}(t), u) \leqslant 0 \quad \text{or} \quad \psi(t+0)\, f(\hat{x}(t), u) \leqslant 0.$$

This is obviously equivalent to the statement that $\psi(t) f(\hat{x}(t), u) \leqslant 0$ for all points of continuity of the function ψ on the interval $(0, \hat{t}_1)$, and then for its boundary points (since the function ψ is continuous at these points), while for all discontinuity points both of the above inequalities are satisfied. Thus, for any $u \in U$ we have

$$\psi(t-0)\, f(\hat{x}(t), u) \leqslant 0 \quad \text{and} \quad \psi(t+0)\, f(\hat{x}(t), u) \leqslant 0 \quad \forall\, t \in [0, \hat{t}_1]. \tag{67}$$

Hence, the obtained tuple (α_0, β, ψ) ensures the fulfillment of conditions (i)–(v), and (66)–(67). The only point where all these conditions differ from those claimed in Theorem 2 is the “incorrect” left transversality condition in (v). We have to correct it.

To this aim, consider separately the costate equation

$$\mathrm{d}\psi = -\psi f_x(\hat{x}, \hat{u})\, \mathrm{d}t + \Phi'(\hat{x})\, \mathrm{d}\mu, \tag{68}$$

and the transversality conditions

$$\psi(0-) = l_{x_0} - c\,\Phi'(\hat{x}_0), \qquad \psi(\hat{t}_1) = -l_{x_1}, \quad \text{where} \quad c = \int_{[0,\hat{t}_1]} \mathrm{d}\mu\,. \tag{69}$$

Refined optimality conditions

Now, let us take into account the consistency assumption (9). It means that

$$K(x_0,\, x_1) = 0 \quad \Longrightarrow \quad \Phi(x_0) = 0$$

for all (x_0, x_1) sufficiently close to the reference point $(\hat{x}_0, \hat{x}_1) := (\hat{x}(0), \hat{x}(\hat{t}_1))$. Linearizing at this point (with account of regularity assumption (c) of the endpoints constraints), we have the implication

$$K'_{x_0}(\hat{x}_0, \hat{x}_1)\, \bar{x}_0 + K'_{x_1}(\hat{x}_0, \hat{x}_1)\, \bar{x}_1 = 0 \quad \Longrightarrow \quad \Phi'(\hat{x}_0)\, \bar{x}_0 \,=\, 0$$

for all $(\bar{x}_0, \bar{x}_1) \in \mathbb{R}^{2d(x)}$. This, in turn, implies that there exists a row vector $\beta_* \in \mathbb{R}^{d(K)*}$ such that

$$\Phi'(\hat{x}_0)\, \bar{x}_0 = \beta_* \left(K'_{x_0}(\hat{x}_0, \hat{x}_1)\, \bar{x}_0 + K'_{x_1}(\hat{x}_0, \hat{x}_1)\, \bar{x}_1\right) \qquad \forall\ (\bar{x}_0, \bar{x}_1) \in \mathbb{R}^{2d(x)},$$

whence

$$\Phi'(\hat{x}_0) = \beta_*\, K'_{x_0}(\hat{x}_0, \hat{x}_1) \quad \text{and} \quad \beta_*\, K'_{x_1}(\hat{x}_0, \hat{x}_1) = 0. \tag{70}$$

Let us change the multiplier β to $\widetilde{\beta} = \beta - c\,\beta_*$. Then, the endpoint function $l = \alpha_0 J + \beta K$ will be correspondingly changed to $\widetilde{l} = \alpha_0 J + (\beta - c\,\beta_*)K = l - c\,\beta_* K$, while the endpoint conditions of the costate function ψ (the same one, we do not change it!) will have the following "proper" form:

$$\psi(0-) = \, l_{x_0} - c\,\Phi'(\hat{x}_0) \,=\, l_{x_0} - c\,\beta_* K'_{x_0} \,=\, \widetilde{l}_{x_0}\,, \qquad \psi(\hat{t}_1) = -\widetilde{l}_{x_1}\,. \tag{71}$$

Note that the new triple $(\alpha_0, \widetilde{\beta}, \mathrm{d}\mu)$ is nontrivial. Otherwise $c = 0$, then $\widetilde{\beta} = \beta$, and so, the initial triple is trivial, which contradicts its normalization.

Finally, it remains to obtain the condition $\mathrm{d}\mu(\{0\}) = 0$. First, we claim that

$$\alpha_0 \,+\, |\widetilde{\beta}| \,+ \int_{(0,\hat{t}_1]} \mathrm{d}\mu \,>0.$$

Indeed, suppose it is equal to zero. Then $\widetilde{l} = 0$, so $\psi(\hat{t}_1) = 0$, and $\mathrm{d}\psi = 0$ on $(0, \hat{t}_1]$, whence $\psi(0+) = 0$. By the left transversality, also $\psi(0-) = 0$, and so $[\psi](0) = 0$. But (68) implies $[\psi](0) = [\mu](0)\, \Phi'(\hat{x}_0)$. Since $\Phi'(\hat{x}_0) \neq 0$ (by the controlability assumption (a)), we have $[\mu](0) = 0$, hence $\mathrm{d}\mu \equiv 0$ totally, and all the new triple is trivial, a contradiction.

Now, we restrict the obtained ψ and $\mathrm{d}\mu$ on the interval $(0, \hat{t}_1]$, and consider the right hand value of ψ at 0. Since $\psi(0-) = \widetilde{l}_{x_0}$, we have

$$\psi(0+) \,=\, \widetilde{l}_{x_0} + [\mu](0)\, \Phi'(\hat{x}_0).$$

Repeating once again(!) the above trick with changing $\widetilde{\beta}$ by some $\widetilde{\widetilde{\beta}} = \widetilde{\beta} + c'\beta'_*$, we arrive at the triple $(\alpha, \widetilde{\widetilde{\beta}}, d\mu)$, which is nontrivial and provides the endpoint conditions

$$\psi(0+) = \widetilde{\widetilde{l}}_{x_0}, \qquad \psi(\hat{t}_1) = -\widetilde{\widetilde{l}}_{x_1}.$$

Finally, upon setting $\psi(0-) := \psi(0) = \psi(0+)$ and $\mu(0-) := \mu(0) = \mu(0+)$, we have the fulfillment of all the conditions of MP on the interval $[0, \hat{t}_1]$ with $d\mu(\{0\}) = 0$, q.e.d.

Theorem 2 is completely proven. □

3. Appendix: Lagrange multipliers rule

Let X, Y, and Z_i, $i = 1, \ldots, \nu$ be Banach spaces, $\mathcal{D} \subset X$ an open set, and $Q_i \subset Z_i$, $i = 1, \ldots, \nu$ closed convex sets with nonempty interiors. (Particularly, the latter sets can be cones.) Let $F_0 : \mathcal{D} \to \mathbb{R}$, $g : \mathcal{D} \to Y$, and $f_i : \mathcal{D} \to Z_i$, $i = 1, \ldots, \nu$, be given mappings. Consider the following optimization problem:

$$\left.\begin{array}{l} F_0(x) \to \min, \\ f_i(x) \in Q_i, \quad i = 1, \ldots, \nu, \\ g(x) = 0. \end{array}\right\} \qquad (72)$$

Let us impose the following

Assumptions. 1) The objective function F_0 and the mappings f_i are Fréchet differentiable at x_0; the operator g is strictly differentiable at x_0 (smoothness of the data functions), 2) the image of the derivative $g'(x_0)$ is closed in Y (weak regularity of equality constraint).

For any i denote by $T_i(z_i)$ and by $N_i(z_i)$ the cones of tangent and outer normal directions, respectively, to the convex set Q_i at a point $z_i \in Z_i$. Obviously, $N_i(z_i) = -T_i^*(z_i)$ and $T_i(z_i) = -N_i^*(z_i)$, where asterisk stands for the conjugate cone. Moreover, $\operatorname{int} T_i(z_i) = \operatorname{con}(\operatorname{int} Q_i - z_i)$. Clearly, if $z_i \in \operatorname{int} Q_i$, then $T_i(z_i) = Z_i$ and $N_i(z_i) = \{0\}$.

By $\langle z_i^*, z \rangle$ we denote the duality pairing between Z_i and its dual space Z_i^*.

The following theorem gives a generalized Lagrange multipliers rule for problem (72), see Dmitruk and Osmolovskii (2017, 2018, 2020).

Theorem 3 *a) Let x_0 be a local minimum in problem (72). Then there exist Lagrange multipliers $\alpha_0 \geqslant 0$, $z_i^* \in N_i(f_i(x_0))$, $i = 1, \ldots, \nu$, and $y^* \in Y^*$, not* *all equal zero, such that the Lagrange function*

$$L(x) = \alpha_0 F_0(x) + \sum_{i=1}^{\nu} \langle z_i^*,\, f_i(x)\rangle + \langle y^*,\, g(x)\rangle$$

is stationary at x_0*:* $L'(x_0) = 0$, *i.e.,*

$$\alpha_0 F_0'(x_0) \;+\; \sum_{i=1}^{\nu} z_i^* f_i'(x_0) \;+\; y^* g'(x_0) \;=\; 0, \tag{73}$$

where $y^* g'(x_0)$ *is a linear functional on* X *acting by the rule* $\bar{x} \mapsto \langle y^*,\, g'(x_0)\bar{x}\rangle$ *and the expressions* $z_i^* f_i'(x_0)$ *have the similar sense.*

b) If the equality constraint is regular at the point x_0, *i.e.* $g'(x_0)X = Y$ *(the Lyusternik condition), and* $\exists\, \bar{x}$ *such that* $g'(x_0)\,\bar{x} = 0$ *and* $f_i'(x_0)\,\bar{x} \in \operatorname{int} T_i(f_i(x_0))$ *for all* $i = 1, \ldots, \nu$ *(a generalized Slater condition), then necessarily* $\alpha_0 > 0$.

Note that the usual complementary slackness conditions are absent here. If we call the active indices those i for which $f_i(x_0) \in \partial Q_i$, then for any inactive i we have $f_i(x_0) \in \operatorname{int} Q_i$, whence automatically $z_i^* = 0$, since $N_i(f_i(x_0)) = \{0\}$, so the inactive indices do not enter into the Lagrange function.

References

Arutyunov, A. and Tynyanskiy, N. (1984) The maximum principle in a problem with phase constraints. *Izv. Akad. Nauk SSSR Tekhn. Kibernet.*, 0-060-68, 235.

Dubovitskii, A.Ya. and Dubovitskii, V.A. (1985) Necessary conditions for a strong minimum in optimal control problems with degeneracy of the end-point and phase constraints. *Russian Math. Surveys*, **40**, 2, 209–210.

Dubovitskii, A.Ya. and Dubovitskii, V.A. (1987) The maximum principle in regular optimal control problems where the ends of the phase path are on the boundary of the phase constraint. *Avtomatika i Telemekhanika*, 12, 25–33 (in Russian).

Dubovitskii, A.Ya. and Dubovitskii, V.A. (1995) A criterion for the existence of a meaningful maximum principle in a problem with phase constraints. *Diff. Equat.*, **31**, 10, 1595–1602.

Dubovitskii, A. Ya. and Milyutin, A.A. (1965) Extremum problems in the presence of restrictions. *USSR Comput. Math. and Math. Phys.*, **5**, 3, 1–80.

Dmitruk, A. V. and Osmolovskii, N. P. (2018) Variations of the type of v–change of time in problems with state constraints. *Proc. of the Institute of Mathematics and Mechanics, the Ural Branch of Russian Academy of Sciences*, 24, 76–92 (in Russian).

Dmitruk, A. V. and Osmolovskii, N. P. (2019) Proof of the maximum principle for a problem with state constraints by the v-change of time variable. *Discrete and Continuous Dynamical Systems, Ser. B*, **24**, 5, 2189–2204. doi:10.3934/dcdsb.2019090.

Dmitruk, A. V. (2009) On the development of Pontryagin's Maximum principle in the works of A.Ya. Dubovitskii and A.A. Milyutin. *Control and Cybernetics* **38**, 4a, 923–958.

Dmitruk, A. V. and Osmolovskii, N. P. (2017) A General Lagrange Multipliers Theorem. Constructive Nonsmooth Analysis and Related Topics (CNSA-2017). *IEEE Xplore Digital Library.* doi:10.1109/CNSA.2017.7973951.

Dmitruk, A. V. and Osmolovskii, N. P. (2018) A General Lagrange Multipliers Theorem and Related Questions. In: G. Feichtinger et al., eds., *Control Systems and Math. Methods in Economics. Lecture Notes in Economics and Mathematical Systems*, **687**, 165–194, Springer.

Dmitruk, A. V. and Osmolovskii, N. P. (2020) Lagrange Multipliers Rule for a General Optimization Problem with an Infinite Number of Constraints. In: A. Vasin and F. Aleskerov, eds., *Recent Advances of the Russian Operations Research Soc.* Cambridge Scholars Publishing, 212–232.

Kolmogorov, A. N. and Fomin, S. V. (1999) *Elements of the Theory of Functions and Functional Analysis. Dover Books on Mathematics*, Russian 4th Edition: Nauka, Moscow.

Milyutin, A. A., Dmitruk, A. V. and Osmolovskii, N. P. (2004) *Maximum Principle in Optimal Control.* Moscow State University, Faculty of Mechanics and Mathematics, Moscow (in Russian), **168**.