Generalized Implicit Function Theorem and its Application to Parametric Optimal Control Problems *

Olga Kostyukova† and Ekaterina Kostina‡

Abstract

Many problems in physics and mathematics may be reduced to solving equations depending on a parameter. The justification of the existence of solutions to the equations and the sensitivity analysis maybe conducted based on Implicit Function Theorem (IFT) under certain regularity assumptions. Provided the regularity assumptions do not hold, generalizations of IFT are needed in order to study solutions to the equations. The paper focuses on a particular generalization of IFT which is then applied to a parametric linear time-optimal control problem.

1 Introduction

In many problems in physics and mathematics equations depending on a parameter arise. The equations may be formally written in the form

$$F(y, \tau) = 0 \tag{1}$$

where $F : R^s \times R \to R^s$. Basically, one would like to know

- for which values of τ does the equation (1) have a solution;
- how many solutions exist for a given value of τ;
- how do the solutions vary as the parameter τ varies.

In some situations, the answers to these questions can be derived based on the classical implicit function theorem (IFT). The classical IFT states, that when a continuously differentiable function $F(y, \tau)$ vanishes at a point (y_0, τ_0) with the nonsingular Jacobian $\dfrac{\partial F(y_0, \tau_0)}{\partial y}$, there exist a number $\delta_0 > 0$ and a unique function $y(\tau)$ satisfying equation (1) for $\tau \in [\tau_0 - \delta_0, \tau_0 + \delta_0]$ and the initial condition

$$y(\tau_0) = y_0. \tag{2}$$

*The research was supported by the Institute of Mathematics, National Academy of Sciences of Belarus, through the State Program of Fundamental Research "Mathematical Structures 15".

†Institute of Mathematics, National Academy of Sciences of Belarus, Surganov str. 11, 220072, Minsk, Belarus

‡Interdisciplinary Center for Scientific Computing (IWR), University of Heidelberg, Im Neuenheimer Feld 368, D-69120 Heidelberg, Germany

We mention few examples of applications of classical IFT. Robinson [17] and Fiacco [3] proposed independently to use the classical IFT for showing Frechet differentiability of solutions to finite-dimensional parametric programming. The classical IFT is used in [11, 12] to investigate differentiability of solutions to parametric optimal control problems in a neighborhood of a regular parameter value.

However, in many important physical problems the classical IFT does not apply because conditions of this theorem do not hold true due to several reasons: mapping $F(y,\tau)$ can be mapping in Banach spaces or can be nonsmooth or the linear transformation given by $L := \dfrac{\partial F(y,\tau)}{\partial y}$ has a nontrivial kernel. That is why, the classical IFT has been extended in various directions, e.g., to Banach spaces [13], to multivalued mappings [2, 16], to nonsmooth functions [1, 15, 19]. In the case dim Ker $L \neq 0$ the existence of solutions to (1) and investigation of properties of the solutions can be reduced to an application of the IFT for a new bifurcation function which is obtained, e.g., from Lyapunov-Schmidt reduction. These methods are explained in detail in [5].

Since our setting here is finite dimensional and functions, that we are interested in, are single-valued, but dim Ker $L \neq 0$, we focus in this paper on a particular generalization of the IFT which is sufficient in many applications including presented below. This generalization allows us to get special representations of solution function $y(\tau), \tau \in [\tau_0, \tau_0 + \delta]$, $\delta_0 > 0$, to the equation (1) with the initial condition (2). This representation will be used in the later applications. Our results are related methodologically to the results from [10], [6]. However, our results differ by more detailed form of presentation which is suitable for further applications to optimal control problems.

The paper is organized as follows. In Section 2, we formulate and prove a generalized implicit function theorem which states that there exist $2\bar{m}$ solution functions $y^p(\tau), \tau \in [\tau_0 - \delta_0, \tau_0 + \delta_0], = 1, \dots, 2\bar{m}$, satisfying the equation (1) and the initial condition (2) if a corresponding algebraic system of $k :=$ dim Ker L equations has $2\bar{m}$ solutions. We show that the functions $y^p(\tau), \tau \in [\tau_0 - \delta_0, \tau_0 + \delta_0], = 1, \dots, 2\bar{m}$, can be presented in a special form. On the base of these results we investigate changes of the structure of solutions to parametric time-optimal control problems in a neighborhood of an irregular point τ_0 in Section 3.

2 Implicit function theorem

Let $F(y,\tau), y \in R^s, \tau \in R$, be a sufficiently smooth function. Let $y_0 \in R^s, \tau_0 \in R$ be such that

$$F(y_0, \tau_0) = 0. \tag{3}$$

Denote

$$L := \frac{\partial F(y_0, \tau_0)}{\partial y},\ b := \frac{\partial F(y_0, \tau_0)}{\partial \tau}. \tag{4}$$

Suppose, that the matrix L is singular. Denote by $\varphi(i) \in R^s$, $i = 1, \dots, k$; $\psi(i) \in R^s$, $i = 1, \dots, k$; bases of the spaces Ker L and Ker L^T respectively,

$$\Phi := (\varphi(i), i = 1, \dots, k) \in R^{s \times k},\ \Phi_* := (\psi(i), i = 1, \dots, k) \in R^{s \times k}, \tag{5}$$

$$A_j := \frac{\partial}{\partial y}\Big(\frac{\partial F(y,\tau)}{\partial y}\varphi(j)\Big)\Big|_{y=y_0,\tau=\tau_0} \in R^{s\times s},\ j=1,\dots,k.$$

Theorem 1 *Let $\beta^0 = (\beta_i^0, i=1,\dots,k)$ be a solution to the system of k equations*

$$\Phi_*^T[b + \frac{1}{2}S(\beta)\beta] = 0, \tag{6}$$

where

$$S(\beta) := (A_j\Phi\beta, j=1,\dots,k) \in R^{s\times k}, \tag{7}$$

with respect to k unknowns $\beta = (\beta_i, i=1,\dots,k)$. Suppose that

$$det\ \Phi_*^T S(\beta^0) \neq 0. \tag{8}$$

Then there exist a number $\delta > 0$ and continuous function $\bar{y}(\varepsilon)$, $\varepsilon \in [-\delta,\delta]$, such that

$$F(\bar{y}(\varepsilon), \tau_0 + \varepsilon^2) \equiv 0,\ \varepsilon \in [-\delta,\delta], \tag{9}$$
$$y(0) = y_0. \tag{10}$$

The function $\bar{y}(\varepsilon)$, $\varepsilon \in [-\delta,\delta]$, can be presented as follows

$$\bar{y}(\varepsilon) = y_0 + \varepsilon\sum_{i=1}^{k}\beta_i(\varepsilon)\varphi(i) + \varepsilon^2 v(\varepsilon),\ \varepsilon \in [-\delta,\delta], \tag{11}$$

where $\varphi(i) \in R^s$, $i=1,\dots,k$, is a basis of the space Ker L, $\beta(\varepsilon) = (\beta_i(\varepsilon),\ i=1,\dots,k)$; $v(\varepsilon)$, $\varepsilon \in [-\delta,\delta]$, are some continuous functions with the initial conditions $\beta(0)=\beta^0$.

Proof. Let us consider an s-vector-function of the form (11) and show that there exist continuous functions

$$\beta(\varepsilon) = (\beta_i(\varepsilon),\ i=1,\dots,k),\ v(\varepsilon) \in R^s,\ \varepsilon \in [0,\delta], \tag{12}$$

and parameters

$$\beta^0 = (\beta_i^0,\ i=1,\dots,k),\ v_0 \in R^s, \tag{13}$$

such that the relations hold

$$F(\bar{y}(\varepsilon), \tau_0 + \varepsilon^2) \equiv 0, \tag{14}$$
$$\varphi^T(i)v(\varepsilon) \equiv 0,\ i=1,\dots,k,\ \varepsilon \in [-\delta,\delta], \tag{15}$$
$$\beta(0) = \beta^0,\ v(0) = v_0. \tag{16}$$

We consider two new functions

$$\hat{F}(\beta, v, \varepsilon) = F(y_0 + \varepsilon\Phi\beta + \varepsilon^2 v, \tau_0 + \varepsilon^2),$$
$$\tilde{F}(\beta, v, \varepsilon) = \frac{1}{\varepsilon^2}\hat{F}(\beta, v, \varepsilon),\ \beta \in R^k,\ v \in R^s. \tag{17}$$

At $\varepsilon = 0$ we define the function $\tilde{F}(\beta, v, \varepsilon)$ and its derivatives by continuity, namely

$$\tilde{F}(\beta, v, 0) = \frac{1}{2}\frac{\partial^2 \hat{F}(\beta, v, 0)}{\partial \varepsilon^2} = \frac{\partial F}{\partial y}v + \frac{\partial F}{\partial \tau} + \frac{1}{2}\sum_{j=1}^{k}\beta_j\Big[\frac{\partial}{\partial y}\Big(\frac{\partial F}{\partial y}\varphi(i)\Big)\Big]\Phi\beta, \tag{18}$$

$$\frac{\partial \tilde{F}(\beta, v, 0)}{\partial \beta} = S(\beta), \quad \frac{\partial \tilde{F}(\beta, v, 0)}{\partial v} = \frac{\partial F}{\partial y}. \tag{19}$$

Here $F = F(y, \tau)$ and the function and all its derivatives are calculated at the point (y_0, τ_0).

Taking into account the notations (4), (5) and (7), we can rewrite (18) as follows

$$\tilde{F}(\beta, v, 0) \quad = \quad Lv + b + \frac{1}{2}\sum_{j=1}^{k}\beta_j A_j \Phi\beta = Lv + b + \frac{1}{2}S(\beta)\beta. \tag{20}$$

Obviously, the equations (14), (15) are equivalent to the following equations

$$\tilde{F}(\beta(\varepsilon), v(\varepsilon), \varepsilon) \equiv 0, \quad \Phi^T v(\varepsilon) \equiv 0, \ \varepsilon \in [-\delta, \delta]. \tag{21}$$

We define the parameters (13) as a solution of the system

$$\tilde{F}(\beta^0, v_0, 0) = 0, \tag{22}$$

$$\Phi^T v_0 = 0. \tag{23}$$

Taking into account (20), (22) may be rewritten in the form

$$Lv_0 + b + \frac{1}{2}S(\beta^0)\beta^0 = 0. \tag{24}$$

Let us show that the system (23), (24) has a solution. Multiplying the both sides of the equation (24) by the matrix Φ_*^T yields the system (6) of k equations with respect to the unknowns $\beta = (\beta_i, \ i = 1, \dots, k)$. By assumption, this system has a solution $\beta^0 = (\beta_i^0, i = 1, \dots, k)$.

In the next we make use of the following Proposition (see results in [20] or Lemma 3 in [8])

Proposition 1 *Let $\bar{b} \in R^s$ be such a vector that $\Phi_*^T \bar{b} = 0$. Then there exists a unique solution $v \in R^s$ to the system $Lv = \bar{b}$, $\Phi^T v = 0$.*

From Proposition 1 we may conclude that, for any vector $\beta^0 = (\beta_i^0, i = 1, \dots, k)$ satisfying (6), there is a unique vector $v_0 = v_0(\beta^0)$ satisfying (23), (24). Hence, the system (23), (24) has the solution (β^0, v_0). It is easy to check that the function $\tilde{F}(\beta, v, \varepsilon)$ and its derivatives $\partial \tilde{F}(\beta, v, \varepsilon)/\partial \beta$, $\partial \tilde{F}(\beta, v, \varepsilon)/\partial v$ are continuous in a neighborhood of the point $(\beta^0, v_0, \varepsilon = 0) \in R^k \times R^s \times R$ under assumption that the function $F(y, \tau)$, $y \in R^s, \tau \in R$, is sufficiently smooth.

Now let us compute the Jacobian Ω of the equations (21) with respect to the variables β, v at the point $(\beta^0, \, v_0, \, \varepsilon = 0)$:

$$\Omega = \begin{pmatrix} S(\beta^0) & L \\ 0 & \Phi^T \end{pmatrix}, \tag{25}$$

where Φ and $S(\beta)$ are defined in (5), (7). Let us show that under the condition (8) the matrix Ω is not singular:

$$\det\Omega \neq 0. \tag{26}$$

Suppose the contrary. Then there is a vector $(x, z) \neq 0$, $x \in R^s$, $z \in R^k$, such that

$$x^T L + z^T \Phi^T = 0, \quad x^T S(\beta^0) = 0. \tag{27}$$

Multiplying the first equality from (27) by Φ, we get $z^T \Phi^T \Phi = 0$. Consequently, $z = 0$ and the system (27) takes the form

$$x^T L = 0, \quad x^T S(\beta^0) = 0. \tag{28}$$

Since $x^T L = 0$, then $x \in \text{Ker } L^T$, and consequently, $x = \Phi_* \gamma$, where $\gamma \in R^k$. Taking this fact into account, we may conclude that (28) is equivalent to the system

$$\gamma^T {\Phi_*}^T S(\beta^0) = 0. \tag{29}$$

The assumption (8) and the equality (29) result in $\gamma = 0$, and, consequently, $x = 0$. Thus, we have shown that $x = 0$, $z = 0$. However this contradicts to the assumption that $(x, z) \neq 0$. Hence, the relation (26) follows from (8).

To finish the proof we apply the classical IFT in the formulation of [18].

Theorem 2 *Consider a continuous function $\mathcal{F}(z, \varepsilon) : R^l \times R \to R^l$. Assume that in a some neighborhood $W \subset R^l \times R$ of a point $(z_0, \varepsilon_0) \in R^l \times R$ there exists the continuous derivative $\dfrac{\partial \mathcal{F}(z, \varepsilon)}{\partial z}$. Assume further that*

$$\mathcal{F}(z_0, \varepsilon_0) = 0, \; det\, \frac{\partial \mathcal{F}(z_0, \varepsilon_0)}{\partial z} \neq 0.$$

Then there exist a number $\delta > 0$ and a unique continuous l-vector-function $z(\varepsilon)$, $\varepsilon \in [\varepsilon_0 - \delta, \varepsilon_0 + \delta]$, such that

$$\mathcal{F}(z(\varepsilon), \varepsilon) \equiv 0, \; \varepsilon \in [\varepsilon_0 - \delta, \varepsilon_0 + \delta], \; z(\varepsilon_0) = z_0.$$

The relations (22), (23) and (26) allow us to apply Theorem 2 to the system (21). According to the theorem there are unique functions (12) satisfying (21) and (16). Consequently, there is a continuous function $\bar{y}(\varepsilon)$, $\varepsilon \in [-\delta, \delta]$, of the form (11) satisfying (14)-(16). This finishes the proof of the theorem. $\diamond$

Corollary 1 *Let $\bar{y}(\varepsilon), \varepsilon \in [-\delta, \delta]$, be a function from Theorem 1. Then there exists another continuous function*

$$z(\varepsilon) := \bar{y}(-\varepsilon) \not\equiv \bar{y}(\varepsilon), \varepsilon \in [-\delta, \delta],$$

such that

$$F(z(\varepsilon), \tau_0 + \varepsilon^2) \equiv 0, \; \varepsilon \in [-\delta, \delta], \; z(0) = y_0.$$

Corollary 2 *Let the system (6) with respect to k unknowns $\beta = (\beta_i, i = 1, \ldots, k)$ have $2\bar{m}$ different solutions $\beta^p \in R^s$, $p = 1, \ldots, 2\bar{m}$, satisfying the conditions*

$$det\ \Phi_*^T S(\beta^p) \neq 0,\ p = 1, \ldots, 2\bar{m}. \tag{30}$$

Then there exist a number $\delta_0 > 0$ and $2\bar{m}$ different functions $y^p(\tau), \tau \in [\tau_0, \tau_0 + \delta_0]$, $p = 1, \ldots, 2\bar{m}$, such that

$$F(y^p(\tau), \tau) \equiv 0,\ \tau \in [\tau_0, \tau_0 + \delta_0],\ y^p(\tau_0) = y_0. \tag{31}$$

Each function $y^p(\tau)$, $\tau \in [\tau_0, \tau_0 + \delta_0]$, $p = 1, \ldots, 2\bar{m}$, can be presented as follows

$$y^p(\tau) = y^p(\tau_0 + \Delta\tau) = y_0 + \sqrt{\Delta\tau} \sum_{i=1}^{k} \beta_i^p(\sqrt{\Delta\tau})\varphi(i) + \Delta\tau v^p(\sqrt{\Delta\tau}),\ \Delta\tau \in [0, \delta_0].$$

Here $\beta^p(\varepsilon) = (\beta_i^p(\varepsilon),\ i = 1, \ldots, k)$; $v^p(\varepsilon)$, $\varepsilon \in [0, \sqrt{\delta_0}]$, $p = 1, \ldots, \bar{m}$, are some continuous functions connected by the relations

$$\beta^p(\varepsilon) = -\beta^{\bar{m}+p}(-\varepsilon),\ v^p(\varepsilon) = v^{\bar{m}+p}(-\varepsilon),\quad \varepsilon \in [0, \sqrt{\delta_0}], p = 1, \ldots, \bar{m},$$

with the initial conditions $\beta^p(0) = -\beta^{\bar{m}+p}(0) = \beta^p$, $v^p(0) = v^{\bar{m}+p}(0) = v^p$, where $v^p \in R^s$ is a unique solution to the system

$$Lv^p + b + \frac{1}{2}S(\beta^p)\beta^p = 0,\ \Phi^T v^p = 0,$$

for $p = 1, \ldots, \bar{m}$.

Let us give some examples illustrating the assumptions of Theorem 1 and Corollary 2.

First, let us show that the system (6) may have "many" solutions. We consider the following system

$$s = 2,\ F(y, \tau) = \begin{pmatrix} y_1^2 - \alpha_1\tau \\ y_2^2 - \alpha_2\tau - \alpha_3\tau^2 \end{pmatrix}, \tau \geq 0,\ y = (y_1, y_2),\ y_0 = (0, 0),\ \tau_0 = 0.$$

Consequently, $F(y_0, \tau_0) = 0$,

$$L := \frac{\partial F(y, \tau)}{\partial y}\Big|_{y=y_0, \tau=\tau_0} = \begin{pmatrix} 2y_1 & 0 \\ 0 & 2y_2 \end{pmatrix}\Big|_{y=y_0, \tau=\tau_0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, b = \begin{pmatrix} -\alpha_1 \\ -\alpha_2 \end{pmatrix},$$

$$\psi(1) = \varphi(1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ \psi(2) = \varphi(2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ \Phi = \Phi_* = I,$$

$$\frac{\partial F(y, \tau)}{\partial y}\varphi(1) = \begin{pmatrix} 2y_1 \\ 0 \end{pmatrix},\ \frac{\partial F(y, \tau)}{\partial y}\varphi(2) = \begin{pmatrix} 0 \\ 2y_2 \end{pmatrix},$$

$$A_1 := \frac{\partial}{\partial y}\Big(\frac{\partial F(y, \tau)}{\partial y}\varphi(1)\Big) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix},\ A_2 := \frac{\partial}{\partial y}\Big(\frac{\partial F(y, \tau)}{\partial y}\varphi(2)\Big) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

For this example, we have

$$S(\beta) = (A_1\beta, A_2\beta) = \begin{pmatrix} 2\beta_1 & 0 \\ 0 & 2\beta_2 \end{pmatrix}.$$

The equations (6) take the form

$$-\alpha_1 + \beta_1^2 = 0,\quad -\alpha_2 + \beta_2^2 = 0. \tag{32}$$

Consider the following situations.

A) $\alpha_1 > 0, \alpha_2 > 0$. In this case the system (32) has 4 solutions :

$$\beta^1 = (\sqrt{\alpha_1}, \sqrt{\alpha_2}),\ \beta^2 = -\beta^1,\ \beta^3 = (\sqrt{\alpha_1}, -\sqrt{\alpha_2}),\ \beta^4 = -\beta^3.$$

For any $\beta^p, p = 1, \dots, 4$, the matrix $\Phi_*^T S(\beta^p) = S(\beta^p)$ is nonsingular. There are four functions $y^p(\tau), \tau \in [0, \delta_0]$, $p = 1, \dots, 4$, satisfying (31).

B) $\alpha_1 > 0, \alpha_2 = 0$. In this case the system (32) has 2 solutions :

$$\beta^1 = (\sqrt{\alpha_1}, 0),\ \beta^2 = -\beta^1$$

but for any $\beta^p, p = 1, 2$, the matrix $\Phi_*^T S(\beta^p)$ is singular. If $\alpha_3 < 0$ then there is no function $y^p(\tau), \tau \in [0, \delta_0]$, satisfying (31).

C) $\alpha_1 < 0$ or $\alpha_2 < 0$. In this case the system (32) has no solution.

Above we gave an example when the system (6) has more than 2 solutions and the matrix $\Phi_*^T S(\beta^p)$ is non-singular for all p as well as an example when the system (6) has 2 solutions and the matrix $\Phi_*^T S(\beta^p)$ is singular for $p = 1, 2$.

Now, let us give an example when the system (6) has only 2 solutions $\beta^p, p = 1, 2$, and the matrix det $\Phi_*^T S(\beta^p)$ is nonsingular for $p = 1, 2$,

$$s = 2,\ F(y, \tau) = \begin{pmatrix} y_1^2 - \alpha_1 \tau \\ y_1 y_2 - \alpha_2 \tau \end{pmatrix}, \tau \geq 0,\ y = (y_1, y_2),\ y_0 = (0, 0),\ \tau_0 = 0.$$

In this example, $F(y_0, \tau_0) = 0$,

$$L := \frac{\partial F(y, \tau)}{\partial y}\Big|_{y=y_0, \tau=\tau_0} = \begin{pmatrix} 2y_1 & 0 \\ y_1 & y_2 \end{pmatrix}\Big|_{y=y_0, \tau=\tau_0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, b = \begin{pmatrix} -\alpha_1 \\ -\alpha_2 \end{pmatrix},$$

$$\psi(1) = \varphi(1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ \psi(2) = \varphi(2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ \Phi = \Phi_* = I,$$

$$\frac{\partial F(y, \tau)}{\partial y}\varphi(1) = \begin{pmatrix} 2y_1 \\ y_2 \end{pmatrix},\ \frac{\partial F(y, \tau)}{\partial y}\varphi(2) = \begin{pmatrix} 0 \\ y_1 \end{pmatrix},$$

$$A_1 := \frac{\partial}{\partial y}\Big(\frac{\partial F(y, \tau)}{\partial y}\varphi(1)\Big) = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix},\ A_2 := \frac{\partial}{\partial y}\Big(\frac{\partial F(y, \tau)}{\partial y}\varphi(2)\Big) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

$$S(\beta) = \begin{pmatrix} 2\beta_1 & 0 \\ \beta_2 & \beta_1 \end{pmatrix},\ \Phi_*^T S(\beta) = S(\beta).$$

In this example the system (6) takes the form

$$-\alpha_1 + \beta_1^2 = 0, \quad -\alpha_2 + \beta_1 \beta_2 = 0. \tag{33}$$

Consider several situations.

A1) $\alpha_1 > 0, \alpha_2 \in R$. In this case the system (33) has 2 solutions :

$$\beta^1 = (\sqrt{\alpha_1}, \alpha_2/\sqrt{\alpha_1}),\ \beta^2 = -\beta^1.$$

For any $\beta^p, p = 1, 2$, the matrix $\Phi_*^T S(\beta^p)$ is nonsingular.

B1) $\alpha_1 = 0, \alpha_2 \neq 0$. In this case the system (33) has no solution.

C1) $\alpha_1 = 0, \alpha_2 = 0$. In this case the system (33) has a unique solution : $\beta^1 = (0, 0)$. It is evident that the matrix $\Phi_*^T S(\beta^p)$ is singular.

These examples show that all assumptions of Theorem 1 and Corollary 2 are essential.

3 A parametric time-optimal control problem

The necessity of constructing and investigating a function $y(\tau)$, $\tau \geq \tau_0$, which is implicitly defined by the relations

$$F(y(\tau),\tau) \equiv 0,\ \tau \in [\tau_0, \tau_0+\delta_0],\ y(\tau_0)=y_0, \tag{34}$$

arises when we study parametric optimal control problems. In this case, provided the value of the parameter $\tau=\tau_0$ is an irregular (or a bifurcation) point (see Definition in [8] or below) then the matrix $L := \dfrac{\partial F(y_0,\tau_0)}{\partial y}$ is singular and we can not apply the classical IFT for justification of the existence of the functions $y(\tau)$, $\tau \geq \tau_0$, satisfying (34).

In what follows, we give an example of a parametric optimal control problem where Theorem 1 and Corollary 2 allow us to justify the existence of the problem solution.

3.1 Problem statement and optimality conditions

In the class of piecewise-continuous controls, we consider the classical time-optimal control problem in which the initial state depends on a parameter $\tau \in [0,\tau^*]$:

$$TO(\tau): \quad \begin{cases} \min t^*, \\ \dot{x} = Ax+bu,\ x(0)=z(\tau),\ x(t^*)=0, \\ |u(t)| \leq 1,\ t \in [0,t^*]. \end{cases} \tag{35}$$

Here $x=x(t) \in R^n$ is a state vector, $u=u(t) \in R$ is a control, $z(\tau)$, $\tau \in [0,\tau^*]$, is known continuous piecewise-smooth $n-$vector function. In the sequel we assume that

$$\text{rank}\{b, Ab, \ldots, A^{n-1}b\} = n \tag{36}$$

and the function $z(\tau)$, $\tau \in [0,\tau^*]$, is such that $\|z(\tau)\|_2 \neq 0$ and the problem (35) has a solution for any $\tau \in [0,\tau^*]$.

For a fixed τ, we denote the optimal time by $t^*(\tau)$, the control interval by $T_\tau = [0, t^*(\tau)]$, and the optimal control and the optimal trajectory by $u^0_\tau(\cdot) = (u^0_\tau(t),\ t \in T_\tau)$ and $x^0_\tau(\cdot) = (x^0_\tau(t),\ t \in T_\tau)$.

The following necessary optimality conditions known as the maximum principle are proved in [14, 9].

Theorem 3 Maximum Principle *A feasible control $u^0_\tau(\cdot) = (u^0_\tau(t),\ t \in T_\tau)$ is optimal in the problem $TO(\tau)$ if and only if there is a vector $q(\tau) \in R^n$ such that along the solution $\psi(q(\tau),t)$, $t \in T_\tau$, of the adjoint system*

$$\dot{\psi} = -A^T\psi,\ \psi(0)=q(\tau) \tag{37}$$

the following conditions are satisfied

$$q^T(\tau)z(\tau) = -1, \quad \psi^T(q(\tau),t)bu^0_\tau(t) = \max_{|u|\leq 1} \psi^T(q(\tau),t)bu,\ t \in T_\tau. \tag{38}$$

It is easy to show that under the assumption (36) the problem $TO(\tau)$ has a unique optimal control $u^0_\tau(\cdot)$ and the control has a bang-bang form

$$u^0_\tau(t) = sign\ \sigma(q(\tau), t),\ t \in T_\tau. \tag{39}$$

Here $\sigma(q(\tau), t),\ t \in T_\tau$, denotes a switching function

$$\sigma(q(\tau), t) = \psi^T(q(\tau), t)b = q(\tau)^T f(t),\ t \in T_\tau, \tag{40}$$

$$f(t) = F^{-1}(t)b,\ \dot{F} = AF,\ F(0) = I. \tag{41}$$

Let $t_j(\tau), j = 1, \ldots, p(\tau)$, be zeroes of the switching function. Due to the maximum principle conditions (38), the problems $TO(\tau)$, $\tau \in [0, \tau^*]$, are considered to be solved if a vector of data (the zeroes of the switching function, and optimal time, and the Lagrange vector)

$$P(\tau) = (\ t_j(\tau),\ j = 1, \ldots, p(\tau),\ t^*(\tau),\ q(\tau)\ ) \tag{42}$$

is known for $\tau \in [0, \tau^*]$.

We are interested in the dependence of a solution to the problem $TO(\tau)$ on the parameter τ and the expansion of the functions (42) in a neighborhood of a point τ_0 if a solution to $TO(\tau_0)$, or in other words a data set $P(\tau_0)$, is known.

Let $Q(\tau)$ be the set of all vectors $q(\tau)$ satisfying (37), (38). The properties of the *point-set* mapping $\tau \rightarrow Q(\tau),\ \tau \in [0, \tau^*]$, and of the *time-optimal function* $t^*(\tau),\ \tau \in [0, \tau^*]$, are studied in [8].

Let us take an arbitrary vector $q(\tau) \in Q(\tau)$. A vector $q(\tau)$ is called *basic*, if

$$\text{rank}\Big(f^{(i)}(t_j(\tau)),\ i = 0, ..., s_j - 1;\ j = 1, ..., p(\tau),\ z(\tau)\Big) = n, \tag{43}$$

where s_j is the order of zero $t_j(\tau)$. It is shown in [8] that under the assumption (36) there is always a basic vector $q(\tau) \in Q(\tau)$ among the vectors $q \in Q(\tau)$. Hence, without loss of generality we will consider only basic vectors $q(\tau) \in Q(\tau)$. Consider the corresponding switching function (40) and construct the set of the zeroes of the switching function

$$\begin{aligned}&\{t_j(\tau),\ j = 1, ..., p(\tau)\} = \{t \in T_\tau :\ \sigma(q(\tau), t) = 0\},\\ &t_j(\tau) < t_{j+1}(\tau),\ j = 1, ..., p(\tau) - 1.\end{aligned}$$

We denote $l_*(\tau) = 1$ if the time point 0 is a zero of the switching function, and $l_*(\tau) = 0$ otherwise. Similarly, $l^*(\tau) = 1$ if the optimal time $t^*(\tau)$ is a zero of the switching function and $l^*(\tau) = 0$ otherwise. Further we denote by $\mathcal{L}(\tau)$ the indices of the multiple zeroes, that is

$$\mathcal{L}(\tau) = \{j \in \{1, ..., p(\tau)\} : \frac{\partial \sigma(q(\tau), t_j(\tau))}{\partial t} = 0\}.$$

Finally, we denote $k(\tau) := \text{sign}\ q(\tau)^T f(+0) = u^0_\tau(+0)$.

Definition 1 *The sets of the data*

$$S(\tau) = \{p(\tau), k(\tau), l_*(\tau), l^*(\tau), \mathcal{L}(\tau)\}, \quad P(\tau) = (t_j(\tau),\ j = 1, ..., p(\tau);\ t^*(\tau), q(\tau)) \quad (44)$$

are called the structure and the defining elements of the solution to $TO(\tau)$, corresponding to the vector $q(\tau)$.

Having the data (44) been computed, we are able to construct the control $u^0_\tau(\cdot)$ by the rules

$$u^0_\tau(t) = (-1)^j k,\ t \in [t_j(\tau), t_{j+1}(\tau)[,\ j = 0, ..., p(\tau);\ t_0(\tau) \equiv 0,\ t_{p(\tau)+1}(\tau) \equiv t^*(\tau), \quad (45)$$

and to validate the maximum principle. In other words, the problem $TO(\tau)$ is considered to be solved if the data (44) is known.

Definition 2 *A control $u^0_\tau(\cdot)$ is called regular, and the parameter value $\tau \in [0, \tau_*]$ is a regular point, if the "degree of irregularity" $\beta_*(\tau) = l_*(\tau) + l^*(\tau) + |\mathcal{L}(\tau)|$ equals to zero.*

Remark 1 *It is shown in [8] that under the condition (36) the property of regularity (or irregularity) of a control $u^0_\tau(\cdot)$ is independent of the choice of the basic vector $q(\tau) \in Q(\tau)$.*

Figure 1 shows the switching functions for a regular (left) and an irregular (right) values of the parameter τ.

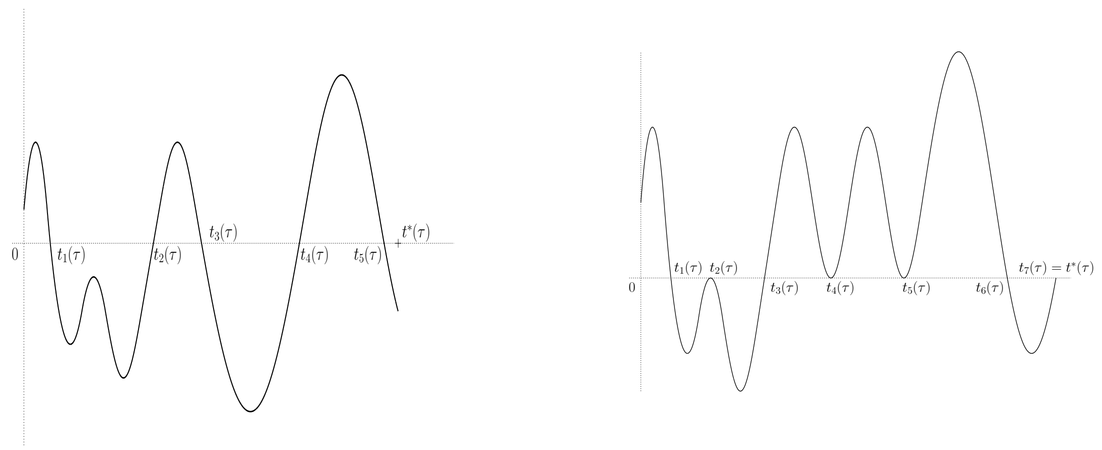

Figure 1: Switching functions for a regular (left) and an irregular (right) parameter values

As pointed out in [8] one of the main properties of a regular point is the stability of a solution structure and a "regular" behavior of the defining elements for small perturbations in the parameter τ. Irregular points do not possess such a property since a solution structure may change for any small perturbation of the parameter τ

$$S(\tau - 0) \neq S(\tau) \neq S(\tau + 0)$$

and the defining elements behave "irregularly", in particular the following situations may occur

$$q(\tau - 0) \neq q(\tau) \neq q(\tau + 0) \text{ or } |dt_j(\tau)/d\tau| = \infty \quad \text{for some } 1 \leq j \leq p(\tau).$$

Suppose that for some parameter value $\tau_0 \in [0, \tau_*]$ the optimal time $t^*(\tau_0)$, the optimal control $u^0_{\tau_0}(\cdot)$ and some vector $q(\tau_0) \in Q(\tau_0)$ are known. We denote by $S(\tau_0)$, $P(\tau_0)$ the data (44), corresponding to $q(\tau_0)$, and by $\mathcal{E}^+(\tau_0)$ a sufficiently small right-sided neighborhood of the point τ_0. Our aim is to determine a new solution structure $S(\tau_0+0)$ and new initial values $P(\tau_0+0)$ for the defining elements. As it was mentioned before, having this information we are able (see [8]) to describe the behavior of the solutions to the problems $TO(\tau)$ and to construct the optimal controls $u^0_\tau(t), t \in [0, t^*(\tau)]$, for $\tau \in \mathcal{E}^+(\tau_0)$.

If τ_0 is a regular point then in a neighborhood of τ_0 the solutions to $TO(\tau)$ are determined uniquely by the rules described in [4, 7].

3.2 Properties of the solutions in a neighborhood of an irregular parameter value

We suppose that τ_0 is an irregular parameter value and study the properties of the solutions to $TO(\tau)$ for $\tau \in \mathcal{E}^+(\tau_0)$. Due to the irregularity of τ_0 the set $Q(\tau_0)$ may contain more than one element and in general it may happen that the initial state for the adjoint system changes $q(\tau_0+0) \neq q(\tau_0)$. Moreover, $S(\tau_0+0) \neq S(\tau_0)$. Consequently, in order to know the properties of the solutions to $TO(\tau)$ in the neighborhood $\mathcal{E}^+(\tau_0)$ it is necessary to define the new vector $q(\tau_0+0)$, the new structure $S(\tau_0+0)$ and the new initial values $P(\tau_0+0)$ for the data $P(\tau)$.

Assume, that we determined the vector $q^* := q(\tau_0+0)$, e.g. by the rules, which are described in [8]. Denote by $\sigma^*(t) := q^{*T} f(t),\ t \in T_{\tau_0}$, the switching function corresponding to q^*. Let us make

Assumption 1 *The following conditions are satisfied*

$$\ddot{\sigma}^*(t^*_j) \neq 0 \quad \text{if} \ \ t^*_j \in (0, t^*_*),\ j \in J^* \setminus J^*_R;\ \dot{\sigma}^*(t^*_j) \neq 0 \quad \text{if} \ \ t^*_j = 0 \vee t^*_*, \quad \text{or}\ j \in J^*_R.$$

Here, the moments t^*_j are the zeroes of the switching function $\sigma^*(t)$, $t \in T_{\tau_0}$, the set J^* is a set of the indices of the zeroes, J^*_R is a set of the indices of the zeroes where the switching function changes its sign and consequently the corresponding control is discontinuous, the numbers d_j are absolute values of the derivatives of the switching function at its zeroes:

$$\begin{aligned}
&t^*_* = t^*(\tau_0),\ \{t^*_j,\ j = 1, ..., p^*\} = \{t \in [0, t^*_*] : \sigma^*(t) = 0\},\\
&t^*_j < t^*_{j+1},\ j = 1, ..., p^* - 1;\\
&J^* = \{1, 2, ..., p^*\},\ J^*_R = \{j \in J^* : t^*_j \in (0, t^*_*), u^0_{\tau_0}(t^*_j - 0) \neq u^0_{\tau_0}(t^*_j + 0)\}; \qquad (46)\\
&\alpha_j = -u^0_{\tau_0}(t^*_j + 0) \text{ if } t^*_j \neq t^*_*;\ \alpha_j = u^0_{\tau_0}(t^*_j - 0) \text{ if } t^*_j = t^*_*;\ d_j = |q^{*T} \dot{f}(t^*_j)|, j \in J^* \qquad (47)
\end{aligned}$$

Let us illustrate the notations (46), (47). Consider the switching function and the corresponding control $\sigma^*(t)$, $u^0_{\tau_0}(t)$, $t \in [0, t^*_*]$, as in Figure 2. Then according to the notations (46), (47)

$$p^* = 7,\ J^* = \{1, 2, \ldots, 7\};\ J^*_R = \{1, 3, 6\}, \tag{48}$$

$$\alpha_j = 1,\ j = 1, 2, 6;\ \alpha_j = -1,\ j = 3, 4, 5, 7;$$
$$d_j > 0 \text{ for } j = 1, 3, 6, 7;\ d_j = 0 \text{ for } j = 2, 4, 5.$$

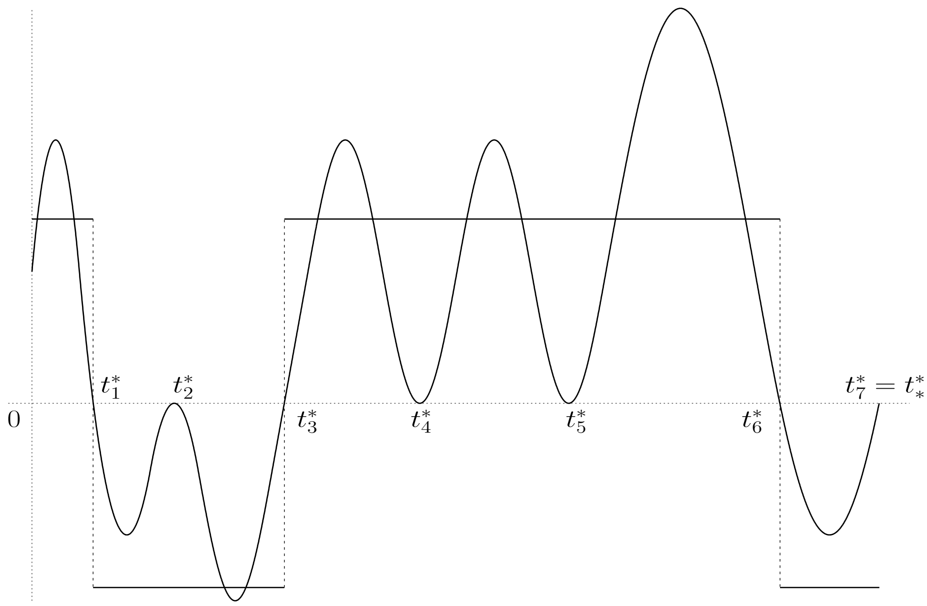

Figure 2: Switching function and the corresponding control at an irregular point.

Now let us discuss problems arising while constructing the new structure $S(\tau_0 + 0)$ and the new initial conditions $P(\tau_0 + 0)$.

According to the definition, the structure $S(\tau_0 + 0)$ and the initial conditions $P(\tau_0 + 0)$ are defined by the zeroes

$$t_j(\tau),\ j = 1, \ldots, p(\tau), \tag{49}$$

of the switching functions $\sigma(q(\tau), t)$, $t \in [0, t^*(\tau)]$, for $\tau \in \mathcal{E}^+(\tau_0)\backslash\tau_0$. Since the initial state for the adjoint system is equal to $q^* = q(\tau_0{+}0)$ then the new zeroes (49) may be generated only by the zeroes t_j^*, $j = 1, \ldots, p^*$, of the function $\sigma^*(t) = q^{*T} f(t)$, $t \in [0, t^*(\tau_0)]$. Let us define the indices $m(j)$, $j = 1, \ldots, p^*$, by the following rule:

$$m(j) \in \{1, \ldots, p(\tau)\} \text{ is the minimal index for which } t_{m(j)}(\tau_0 + 0) = t_j^*. \tag{50}$$

In other words, zero $t_{m(j)}(\tau)$ is the minimal zero from (49) generated by the zero t_j^*.

Obviously,

1. if $j \in J_R^*$ then the single zero t_j^* always generates one single zero $t_{m(j)}(\tau)$;
2. if $j \in J^*\backslash J_R^*$, $t_j^* \neq 0 \vee t_*^*$ then the double zero t_j^* may (see Fig. 3)
    a) either generate two single zeroes $t_{m(j)}(\tau)$ and $t_{m(j)+1}(\tau)$;
    b) either disappear, that is not generate any zeroes $t_j(\tau)$, $j = 1, \ldots, p(\tau)$;
    c) or generate one double zero $t_{m(j)}(\tau)$;

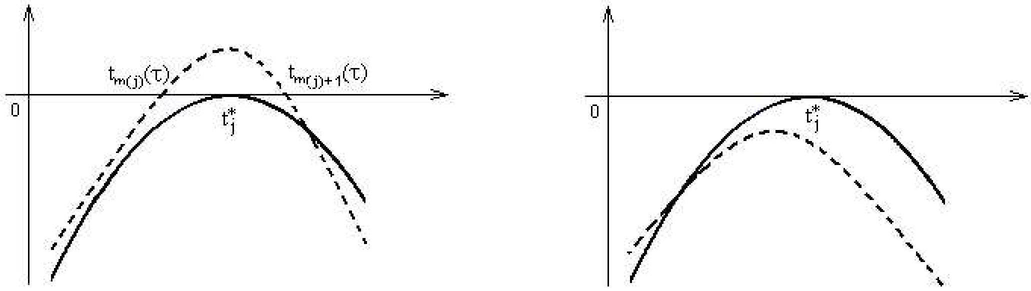

Figure 3: Situations 2a) (left), 2b) (right)

3. if $j \in J^* \backslash J_R^*$, $t_j^* = 0$ (or $t_j^* = t_*^*$) what means that t_j^* is a "boundary" zero, then it may (see Figures 4 and 5)

   a) either generate one single non-boundary zero $t_1(\tau) > 0$, $\tau \in \mathcal{E}^+(\tau_0)\backslash\tau_0$, $t_1(\tau_0 + 0) = 0$ (or $t_{p(\tau)}(\tau) < t^*(\tau)$, $\tau \in \mathcal{E}^+(\tau_0)\backslash\tau_0$, $t_{p(\tau)}(\tau_0 + 0) = t_*^*$);

   b) either disappear, that is not generate any zeroes $t_j(\tau)$, $j = 1, \ldots, p(\tau)$;

   c) or generate one boundary zero $t_1(\tau) \equiv 0$, (or $t_{p(\tau)}(\tau) \equiv t^*(\tau)$), $\tau \in \mathcal{E}^+(\tau_0)\backslash\tau_0$.

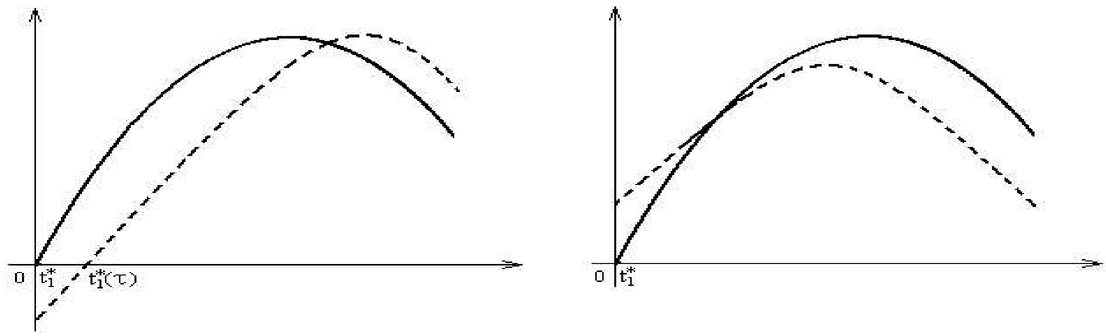

Figure 4: Situations 3a) (left), 3b) (right) for $t_j^* = 0$

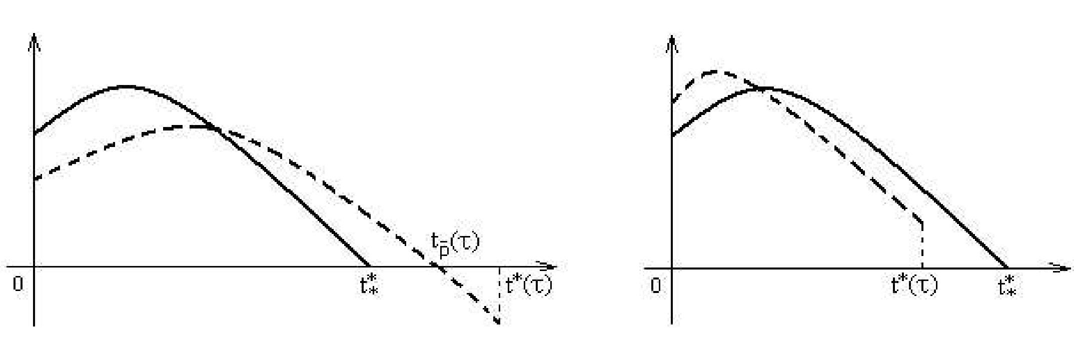

Figure 5: Situations 3a) (left), 3b) (right) for $t_j^* = t_*^*$

Thus, in order to define the new structure $S(\tau_0+0)$ and the new initial conditions $P(\tau_0+0)$ it is necessary to determine which of the situation 2a) – 3c) occurs for each zero t^*_j, $j \in J^*\backslash J^*_R$. In other words, we have to determine the index sets

$$\begin{aligned} J_{(0)} &= \{j \in J^*\backslash J^*_R :\ t^*_j \neq 0 \vee t^*_* \text{ and situation 2a) is true for } t^*_j\}, \\ J_{3a} &= \{j \in J^*\backslash J^*_R :\ t^*_j = 0 \text{ and situation 3a) is true for } t^*_j\}, \\ \bar{J}_{3a} &= \{j \in J^*\backslash J^*_R :\ t^*_j = t^*_* \text{ and situation 3a) is true for } t^*_j\}. \end{aligned} \tag{51}$$

Note, that according to [8], the set $\bar{J}_{3a}$ can be determined exactly, if $p^* \notin J^*$ or $p^* \in J^*$ and $q^{*T}\dot{z}(\tau_0+0) \neq 0$ by the following rule

$$\bar{J}_{3a} = \emptyset \text{ if } p^* \notin J^* \text{ or } p^* \in J^* \text{ and } q^{*T}\dot{z}(\tau_0+0) > 0,$$
$$\bar{J}_{3a} = \{p^*\} \text{ if } p^* \in J^* \text{ and } q^{*T}\dot{z}(\tau_0+0) < 0.$$

Without additional investigation we are not able to determine sets (51) in advance. But we may try to find the sets (51) by enumeration. For example, we may choose

$$J_{(0)} \subset \{j \in J^*\backslash J^*_R :\ t^*_j \neq 0 \vee t^*_*\},\ J_{3a} = \emptyset \text{ or } J_{3a} = \{1\};\ \bar{J}_{3a} = \emptyset \text{ or } \bar{J}_{3a} = \{p^*\}, \tag{52}$$
$$J_{(*)} := J^*_R \cup J_{(0)} \cup J_{3a} \cup \bar{J}_{3a}.$$

This means, that we assume that for $\tau \in \mathcal{E}^+(\tau_0)\backslash\tau_0$,

if $j \in J_{(0)}$, then double zero t^*_j will generate two single zeroes $t_{m(j)}(\tau)$ and $t_{m(j)+1}(\tau)$;

if $j \in J_{3a}$ then "boundary" zero $t^*_j = 0$ will generate one single non-boundary zero $t_1(\tau) > 0$,

if $j \in \bar{J}_{3a}$ then "boundary" zero $t^*_j = t^*_*$ will generate one single non-boundary zero $t_{p(\tau)}(\tau) > t^*_*$.

Hence, we assume that for $\tau \in \mathcal{E}^+(\tau_0)\backslash\tau_0$ the solution to the problem $TO(\tau)$ will have the following structure: optimal control $u^0_\tau(t), t \in [0, t^*(\tau)]$, will have $\bar{p}$ switching points where $\bar{p} := |J_{(*)}| + |J_{(0)}|$. Here, we take into account, that each double zero $t^*_j, j \in J_{(0)}$, of the old switching function $\sigma^*(t), t \in [0, t^*_*]$, will generate two single zeroes of the new switching function $\sigma(q(\tau), t), t \in T_\tau$.

Let us give some explanations using the switching function shown in Figure 2. The old function $\sigma^*(t), t \in [0, t^*_*]$, has 7 zeroes and the corresponding sets J^* and J_R defined by (48). Suppose we choose

$$J_{(0)} := \{2, 5\},\ J_{3a} := \emptyset,\ \bar{J}_{3a} := \{7\}. \tag{53}$$

Then $J_{(*)} = \{1, 2, 3, 5, 6, 7\}$ and $\bar{p} = 8$, and we suppose that the new switching function $\sigma(q(\tau), t), t \in T_\tau$, will have a form presented in Figure 6 with the dashed line (here $t_j := t_j(\tau), \tau \in \mathcal{E}^+(\tau_0)$).

Let us denote the zeroes of the new switching function $\sigma(q(\tau), t), t \in T_\tau$, by

$$t_j(\tau),\ j = 1, \ldots, \bar{p};\ \ t_j(\tau) < t_{j+1}(\tau),\ j = 1, \ldots \bar{p} - 1. \tag{54}$$

Since the new zeroes are generated by the old zeroes

$$t^*_j,\ j = 1, \ldots, p^*, t^*_j < t^*_{j+1},\ j = 1, \ldots, p^* - 1, \tag{55}$$

it is important to know the correspondence between the new and the old zeroes, that is, which old zero generates which new zero. In order to define this correspondence we use the index function (50) $m(j) : J_{(*)} \to \{1, \ldots, \bar{p}\}$, where $m(j)$ denotes the minimal index from $\{1, \ldots, \bar{p}\}$ such that

$$t_{m(j)}(\tau_0 + 0) = t_j^*, \; j \in J_{(*)},$$

i.e. $t_{m(j)}(\tau)$ is the minimal new zero generated by the old zero t_j^*, $j \in J_{(*)}$. (We have to add the word "minimal" because the old zero t_j^* with $j \in J_{(0)} \subset J_{(*)}$ generates two new zeroes.) For the situation presented in Figure 6, we have

$$J_{(*)} = \{1, 2, 3, 5, 6, 7\}; \; m(1) = 1, \; m(2) = 2, \; m(3) = 4, \; m(5) = 5, \; m(6) = 7, \; m(7) = 8.$$

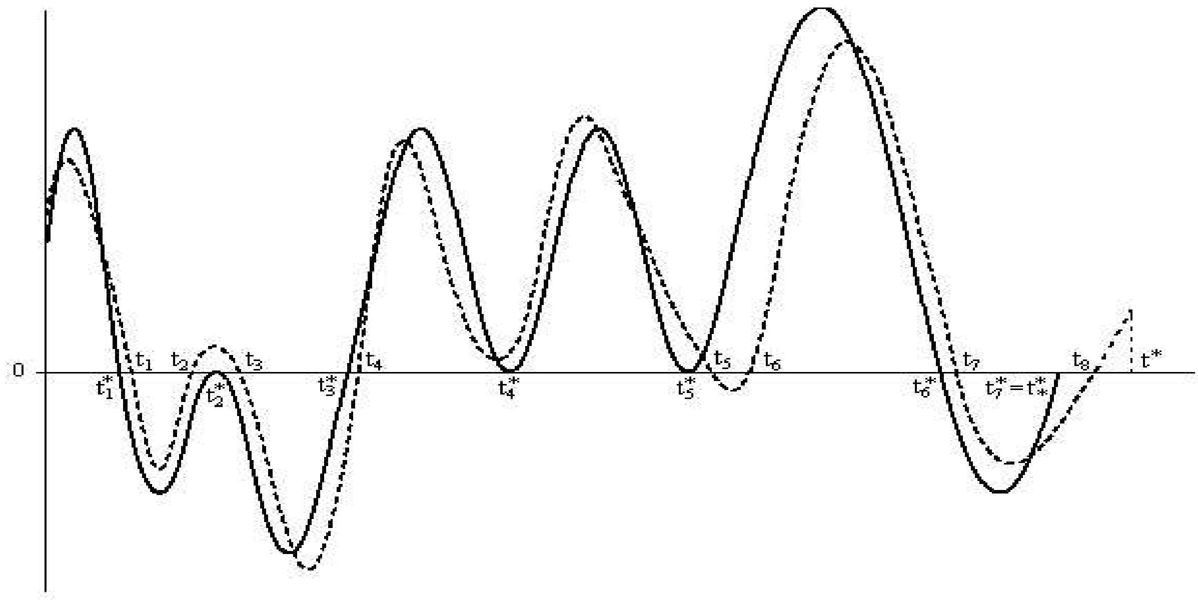

Figure 6: Old and new switching functions.

Having the index correspondence function $m(j), j \in J_{(*)}$, it will be convenient for us to use another variables for denoting the new zeroes besides the notations (54). Namely, let us consider the functions

$$\bar{t}_j(\tau), \; j \in J_{(*)}, \; \Delta t_j(\tau), j \in J_{(0)}, \tag{56}$$

such that

$$\bar{t}_j(\tau_0) = t_j^*, \; j \in J_{(*)}, \; \Delta t_j(\tau_0) = 0, \; j \in J_{(0)}. \tag{57}$$

Then we can use the following presentation of the zeroes (54):

$$t_{m(j)}(\tau) = \bar{t}_j(\tau), \; j \in J_{(*)}; \; t_{m(j)+1}(\tau) = \bar{t}_j(\tau) + \Delta t_j(\tau), \; j \in J_{(0)}. \tag{58}$$

The advantage of this presentation is that it takes into account the correspondence between the old and the new zeroes explicitly. For the example shown in Figure 6, the relations (58) read

$$t_1(\tau) = \bar{t}_1(\tau), \; t_2(\tau) = \bar{t}_2(\tau), \; t_3(\tau) = \bar{t}_2(\tau) + \Delta t_2(\tau),$$
$$t_4(\tau) = \bar{t}_3(\tau), \; t_5(\tau) = \bar{t}_5(\tau), \; t_6(\tau) = \bar{t}_5(\tau) + \Delta t_5(\tau), \; t_7(\tau) = \bar{t}_6(\tau), \; t_8(\tau) = \bar{t}_7(\tau).$$

Further, we will use both presentation (54) and (58) for the zeroes of the new switching function.

3.3 Application of the generalized IFT for determining a new solution structure

Let us check whether our choice (52) of the sets (51) is correct. To reduce a number of subcases under consideration we will suppose that $t_1^* > 0$.

The terminal conditions of the initial problem (35) and the maximum principle conditions (37)–(40) imply that if the choice of the sets (51) is correct then the following equality conditions hold

$$z(\tau_0) + \int_0^{t_*^*} f(t) u_\tau^0(t) dt = 0, \tag{59}$$
$$\sigma(q(\tau), t_j(\tau)) = 0,\ j = 1, \ldots, \bar{p},\ q^T(\tau) z(\tau) + 1 = 0,$$

where $u_\tau^0(t)$, $\sigma(q(\tau), t)$, $f(t)$, $t \in T_\tau$, are defined by (39)–(41).

Introducing an $s-$vector-function $(s = \bar{p} + 1 + n)$

$$y(\tau) = (\bar{t}_j(\tau), j \in J_{(*)}; \Delta t_j(\tau), j \in J_{(0)}; t^*(\tau), q(\tau)),\ \tau \in [\tau_0, \tau_0 + \delta_0], \tag{60}$$

we can rewrite the relations (59) in the form

$$\mathcal{F}(y(\tau), \tau) = 0,\ q^T(\tau) f(\bar{t}_j(\tau)), j \in J_{(*)};\ q^T(\tau) f(\bar{t}_j(\tau) + \Delta t_j(\tau)) = 0, j \in J_{(0)}, \tag{61}$$
$$q^T(\tau) z(\tau) + 1 = 0,$$

where

$$y = (\bar{t}_j, j \in J_{(*)}; \Delta t_j, j \in J_{(0)}; t^*, q)$$

$$\mathcal{F}(y, \tau) = k \sum_{j=1}^{\bar{p}-1} (-1)^j \int_{t_j}^{t_{j+1}} f(t) dt + z(\tau);\quad t_0 = 0,\ t_{\bar{p}+1} = t^*,$$

$$t_{m(j)} = \bar{t}_j,\ j \in J_{(*)};\ t_{m(j)+1} = \bar{t}_j + \Delta t_j,\ j \in J_{(0)},\ k := u^0_{\tau_0+0}(+0).$$

Consequently, if our choice is successful then there is an $s-$vector-function $y(\tau), \tau \in [\tau_0, \tau_0 + \delta_0]$, satisfying the relations (61) and the initial conditions

$$y(\tau_0) = y_0,\ \ y_0 = (t_j^*, j \in J_{(*)}; \Delta t_j = 0, j \in J_{(0)}; t_*^*, q^*). \tag{62}$$

Hence, to check if the choice (52) is correct, first of all it is necessary to check if there exists such a function $y(\tau)$, $\tau \in [\tau_0, \tau_0 + \delta_0]$.

An ordinary way to check this is to use the classical IFT. However, there are several reasons why we are not able to apply the theorem to the system (61). First, due to the presence of the equations

$$q^T f(\bar{t}_j), j \in J_{(*)};\ q^T f(\bar{t}_j + \Delta t_j) = 0, j \in J_{(0)},$$

and the initial conditions (62) , it is evident that, the Jacobi matrix (calculated at the point (y_0, τ_0)) for the system (61), contains $|J_{(0)}|$ pairs of equal rows. Hence, the Jacobian is singular if $J_{(0)} \neq \emptyset$.

To overcome this difficulty, let us consider a new system of the equations

$$\begin{aligned} &\mathcal{F}(y,\tau)=0,\ q^T f(\bar t_j)=0,\ j\in J_{(*)}; \\ q^T(f(\bar t_j+\Delta t_j)-f(\bar t_j))/\Delta t_j=0,\ j\in J_{(0)},&\quad q^T z(\tau)+1=0. \end{aligned} \tag{63}$$

Note that the new system (63) is equivalent to the old one (61), but in the Jacobi matrix (calculated at the point (y_0,τ_0)) for the system (63), there are no equal rows. Hence, we have a chance to apply the classical IFT to the system (63). Indeed, in some situations under some reasonable assumptions we are able to do this (see [7], [8]). However, it often happens in practice that the classical IFT cannot be applied to the system (63) either.

For the problem under consideration such a situation happens if the unperturbed problem $TO(\tau_0)$ is abnormal. This means that the final moment $t^*_* := t^*(\tau_0)$ is a zero of the switching function $\sigma^*(t)=\sigma(q^*,t)$, $t\in T_{\tau_0}=[0,t^*(\tau_0)]$:

$$q^{*T} f(t^*_*)=0. \tag{64}$$

Due to (64), the Jacobian of the system (63) is again singular at the point (y_0,τ_0). Let us show how to apply Theorem 1 for the justification of the existence of a function $y(\tau)$, $\tau\in[\tau_0,\tau_0+\delta_0]$, satisfying (63) and (62). We rewrite the system (63) in the form (34) where

$$\begin{aligned} F^T(y,\tau) \ &= \ \Big(\mathcal{F}(y,\tau)^T,\ q^T f(\bar t_j), j\in J_{(*)}; \\ &\qquad \frac{q^T(f(\bar t_j+\Delta t_j)-f(\bar t_j))}{\Delta t_j}, j\in J_{(0)},\ q^T z(\tau)+1\Big). \end{aligned} \tag{65}$$

Then the matrix $L=\partial F(y_0,\tau_0)/\partial y$ and the vector $b=\partial F(y_0,\tau_0)/\partial\tau$ have the form

$$L=\begin{pmatrix} \bar A_1 & 0 & \bar A_0 & a & 0\\ D_1 & 0 & 0 & 0 & {A^T}_1\\ 0 & 0 & 0 & 0 & {A^T}_0\\ 0 & D_0 & 1/2D_0 & 0 & {A^T}_*\\ 0 & 0 & 0 & 0 & g^T \end{pmatrix}\in R^{s\times s};\ b=\begin{pmatrix} \dot z(\tau_0+0)\\ 0\\ 0\\ 0\\ q^{*T}\dot z(\tau_0+0)\end{pmatrix}\in R^s; \tag{66}$$

$$\begin{aligned} &A_1=\Big(f(t^*_j), j\in J_{(*)}\setminus J_{(0)}\Big),\ A_0=(f(t^*_j), j\in J_{(0)}),\\ &D_1=\mathrm{diag}\Big(q^{*T}\dot f(t^*_j), j\in J_{(*)}\setminus J_{(0)}\Big),\ A_*=\Big(\dot f(t^*_j), j\in J_{(0)}\Big),\\ &D_0=\mathrm{diag}(q^{*T}\ddot f(t^*_j), j\in J_{(0)}),\ \bar A_0=A_0\mathrm{diag}(2\alpha_j, j\in J_{(0)});\\ &\bar A_1=A_1\mathrm{diag}(2\alpha_j, j\in J_{(*)}\setminus J_{(0)}), a=\alpha_{\bar p} f(t^*_*), g=z(\tau_0), \end{aligned} \tag{67}$$

the numbers $\alpha_j, j\in J^*$, are defined according to (47).

To apply Theorem 1 to the system (34) with $F(y,\tau)$ defined by (65), we will need some lemmas, that are formulated below.

Consider a matrix L of the form (66) with the blocks

$$\begin{gathered} A_1\in R^{n\times l}, A_0\in R^{n\times m}, A_*=(A_{*j}, j=1,\dots,m)\in R^{n\times m},\\ A_{*j}\in R^n,\ a\in R^n,\ g\in R^n,\ D_1=\mathrm{diag}(d^1_j, j=1,\dots,l),\ D_0=\mathrm{diag}(d^0_j, j=1,\dots,m),\\ \bar A_0=A_0\mathrm{diag}(\alpha^0_j, j=1,\dots,m);\ \bar A_1=A_1\mathrm{diag}(\alpha^1_j, j=1,\dots,l),\ s=l+2m+1+n. \end{gathered}$$

Suppose that the following conditions are fulfilled

1. $d_j^1\alpha_j^1 > 0, j = 1, \ldots, l;\ d_j^0\alpha_j^0 > 0, j = 1, \ldots, m;$

2. rank $A_0 = m,$ rank $(A_1, A_0, a) =$ rank $(A_1, A_0),$ rank $(A_1, A_0, A_*, g) = n;$

3. $\exists q^* \in R^n$ such that $q^{*T}A_1 = 0,\ q^{*T}A_0 = 0,\ q^{*T}g \neq 0.$

Denote by

$$\varphi(i) \in R^s,\ i = 1, \ldots, k;\ \ \psi(i) \in R^s,\ i = 1, \ldots, k; \tag{68}$$

bases of the spaces Ker L and Ker L^T respectively.

Lemma 1 *Let Conditions 1)-3) be true. Then the vectors (68) can be chosen as*

$$\psi(1) = (q^*, 0, \ldots, 0),\ \psi(i) = (\xi(i), 0, \ldots, 0),\ i = 2, \ldots, k; \tag{69}$$
$$\varphi(i) = (\varphi_j(i),\ 1 \le j \le l + m;\ \Delta\varphi_j(i), 1 \le j \le m; \varphi_*(i), \xi(i)), i = 1, \ldots, k, \tag{70}$$

where

$$\begin{aligned} &\varphi_j(i) = 0, j = 1, \ldots, l;\ \varphi_{l+j}(i) = -\frac{\xi^T(i)A_{*j}}{d_j^0},\ \Delta\varphi_j(i) = 0, j = 1, \ldots, m; \\ &\varphi_*(i) = 0,\ i = 2, \ldots, k, \end{aligned} \tag{71}$$

$\xi(i),\ i = 2, \ldots, k,$ is a basis of the set Ker $(A_1, A_0, g)^T$, and $\varphi(1)$ is some solution of the system $L\varphi(1) = 0$ with the component $\varphi_*(1) = 1$.

Proof is given in Appendix.

Corollary 3 *Let Assumption 1, condition (64) and the following conditions*

$$rank\ \Big(f(t_j^*), j \in J_{(0)}\Big) = |J_{(0)}|,\ rank\ \Big(f(t_j^*), j \in J_{(*)}, \dot f(t_j^*), j \in J_{(0)}, z(\tau_0)\Big) = n \tag{72}$$

be true. Then for the matrix (66) with the data (67), bases of the sets Ker L and Ker L^T can be constructed by the rules (69)-(71).

Proof. It is easy to check that under the assumptions of the Corollary the conditions 1)-3) are fulfilled with $l = |J_{(*)} \setminus J_{(0)}|$, $m = |J_{(0)}|$. $\diamond$

Lemma 2 *Let the conditions of Corollary 3 are fulfilled and $q^{*T}\dot z(\tau_0 + 0) < 0$ (and hence $p^* \in J_{(*)}$) and all components $\Delta\varphi_j(1),\ j \in J_{(0)}$, of the vector $\varphi(1)$ are positive (negative). Then the system (6) has two solutions β^+, β^- ($\beta^+ = -\beta^-$) and*

$$det\ \Phi_*^T S(\beta^+) = -\ det\ \Phi_*^T S(\beta^-) \neq 0. \tag{73}$$

Proof is given in Appendix.

Corollary 4 *If $k = 0$ then the statement of Lemma 2 is true without the additional assumption about positivity (negativity) of all $\Delta\varphi_j(0), j \in J_{(0)}$.*

Theorem 4 *Let the conditions of Lemma 2 be true. Then there exist a number* $\delta > 0$ *and two continuous s-vector-functions*

$$y^+(\tau) = y^+(\tau_0 + \varepsilon^2) = \Big(\bar{t}_j^+(\varepsilon), j \in J_{(*)}; \Delta t_j^+(\varepsilon), j \in J_{(0)}, t^{+*}(\varepsilon), q^+(\varepsilon)\Big), \varepsilon \in [0, \delta];$$
$$y^-(\tau) = y^-(\tau_0 + \varepsilon^2) = \Big(\bar{t}_j^-(\varepsilon), j \in J_{(*)}; \Delta t_j^-(\varepsilon), j \in J_{(0)}, t^{-*}(\varepsilon), q^-(\varepsilon)\Big), \varepsilon \in [0, \delta]; \quad (74)$$

with $\tau = \tau_0 + \varepsilon^2$ *such that*

$$F(y^\pm(\tau), \tau) \equiv 0,\ \tau \in [\tau_0, \tau_0 + \delta^2],\ y^\pm(\tau_0) = y_0 \quad (75)$$

The functions (74) can be represented as follows

$$y^\pm(\tau) = y^\pm(\tau_0 + \varepsilon^2) = y_0 + \varepsilon \sum_{i=0}^{k} \beta_i^\pm(\varepsilon)\varphi(i) + \varepsilon^2 v^\pm(\varepsilon),\ \varepsilon \in [0, \delta], \quad (76)$$

where $\varphi(i) \in R^s$, $i = 1, \dots, k$, *is the basis (70) of the space Ker L,* $\beta^\pm(\varepsilon) = (\beta_i^\pm(\varepsilon),$ $i = 1, \dots, k)$; $v^\pm(\varepsilon)$, $\varepsilon \in [0, \delta]$, *are some continuous functions with* $\beta^+(0) = \beta^0$, $v^+(0) = v_0$ *and* $\beta^-(0) = -\beta^0$, $v^-(0) = v_0$; $(\beta^0,\ v_0)$ *is solution to the system (23), (24),* $\beta_1^0 > 0$.

Proof follows from Corollaries 2 and 3, and Lemma 2. $\diamond$

Theorem 4 yields, that τ_0 is a bifurcation point for the system of equations (65). Due to the Theorem 4 we may conclude that in order to make a right choice (52) of the sets (51) one has to choose the sets that satisfy the assumptions of the theorem. Note, that the conditions of the theorem are only necessary for the sets (52) to be correct. To guarantee that the sets are correct we have to check the following inequality conditions for **one** of the functions (74):

a) For $\tau = \tau_0 + \varepsilon^2 \in \mathcal{E}^+(\tau_0)$, we check the inequalities

$$\Delta t_j^+(\varepsilon) \geq 0, j \in J_{(0)},\ \bar{t}_{p^*}^+(\varepsilon) \leq t_*^+(\varepsilon), \quad (77)$$
$$(\text{or } \Delta t_j^-(\varepsilon) \geq 0, j \in J_{(0)},\ \bar{t}_{p^*}^-(\varepsilon) \leq t_*^-(\varepsilon)). \quad (78)$$

Conditions (77) guarantee, that the function

$$u_\tau^0(t) = (-1)^j k,\ t \in [t_j^+(\tau), t_{j+1}^+(\tau)),\ j = 0, \dots, \bar{p}, \quad (79)$$

where

$$t_{m(j)}^+(\tau) = \bar{t}_j^+(\varepsilon),\ j \in J_{(*)},\ t_{m(j)+1}^+(\tau) = \bar{t}_j^+(\varepsilon) + \Delta t_j^+(\varepsilon),\ j \in J_{(0)}, \quad (80)$$
$$t_{\bar{p}+1}^+(\tau) = t^{+*}(\varepsilon),$$

$\Big($or the function

$$u_\tau^0(t) = (-1)^j k, t \in [t_j^-(\tau), t_{j+1}^-(\tau)), j = 0, \dots, \bar{p}, \quad (81)$$

where

$$t_{m(j)}^-(\tau) = \bar{t}_j^-(\varepsilon),\ j \in J_{(*)},\ t_{m(j)+1}^-(\tau) = \bar{t}_j^-(\varepsilon) + \Delta t_j^-(\varepsilon),\ j \in J_{(0)}, \quad (82)$$
$$t_{\bar{p}+1}^-(\tau) = t^{-*}(\varepsilon),\Big)$$

is an admissible control for the problem $TO(\tau)$ with $\tau = \tau_0 + \varepsilon^2$.

b) For $\tau = \tau_0 + \varepsilon^2 \in \mathcal{E}^+(\tau_0)$, we check the inequalities

$$(-1)^j \sigma(q^+(\varepsilon), t) \geq 0, \ t \in [t_j^+(\tau), t_{j+1}^+(\tau)), \ j = 0, \dots, \bar{p}, \tag{83}$$

$\Big($or

$$(-1)^j \sigma(q^-(\varepsilon), t) \geq 0, \ t \in [t_j^-(\tau), t_{j+1}^-(\tau)), \ j = 0, \dots, \bar{p}\Big). \tag{84}$$

If the inequalities (83) (or (84)) hold then the conditions of maximum principle are fulfilled and hence the admissible control (79) (or (81)) is optimal in the problem $TO(\tau)$, $\tau = \tau_0 + \varepsilon^2 \in \mathcal{E}^+(\tau_0)$.

Having the presentation (76), one can check the inequalities (77), (83) (or (78), (84)) using the vectors (70), (71) and the solution

$$(\beta^0, \ v_0) \ \Big(\text{ or } (-\beta^0, \ v_0)\Big) \text{ with } \beta_1^0 > 0 \tag{85}$$

to the system (23), (24).

Let us give some sufficient conditions for the inequalities (77), (83) to be true.

Theorem 5 *Let conditions of Lemma 2 hold (namely, all components $\Delta\varphi_j(1)$, $j \in J_{(0)}$, of the vector $\varphi(1)$ are positive) and*

$$\varphi_{p^*}(1) < 1, \ \Delta_j^* := \alpha_j \bar{\xi}^{*T} f(t_j^*) < 0, \ j \in J^* \setminus J_{(*)}, \tag{86}$$

where $\bar{\xi}^* = \sum\limits_{i=1}^{k} \beta_i^0 \xi(i)$, $\beta^0 = (\beta_i^0, i = 1, \dots, k)$ is a solution to the system (6) with $\beta_1^0 > 0$, constructed by the vectors (69)-(71) (see also the system (A.23), (A.24) in Appendix). Then the relations (77), (83) hold and the control (79) is optimal in the problem $TO(\tau)$ for $\tau \in \mathcal{E}^+(\tau_0)$. The following expansions are true for the defining elements

$$\begin{aligned} t_{m(j)+1}(\tau_0 + \Delta\tau) \ &- \ t_{m(j)}(\tau_0 + \Delta\tau) = \Delta\varphi_j(1) s^* \sqrt{\Delta\tau} + o(\sqrt{\Delta\tau}) \ \ \textit{if } j \in J_{(0)}; \\ t_{m(j)}(\tau_0 + \Delta\tau) \ &= \ t_j^* + \varphi_j(1) s^* \sqrt{\Delta\tau} + o(\sqrt{\Delta\tau}) \ \ \textit{if } j \in J_{(*)} \setminus J_{(0)}; \\ t^*(\tau_0 + \Delta\tau) \ &= \ t_*^* + \varphi_*(1) s^* \sqrt{\Delta\tau} + o(\sqrt{\Delta\tau}); \\ q(\tau) \ &= \ q^* + \bar{\xi}^* \sqrt{\Delta\tau} + o(\sqrt{\Delta\tau}), \end{aligned} \tag{87}$$

where $s^* = \beta_1^+ > 0$.

Proof. It follows from (70), (71) and (74), (76) that

$$\begin{aligned} &\Delta t_j^+(\varepsilon) = \varepsilon \Delta\varphi_j(1) \beta_1^+(\varepsilon) + \varepsilon^2 \Delta\mu_j(\varepsilon), \ j \in J_{(0)}, \\ &\bar{t}_{p^*}^+(\varepsilon) = t_*^* + \varepsilon \varphi_{p^*}(1) \beta_1^+(\varepsilon) + \varepsilon^2 \mu_{p^*}(\varepsilon), \\ &t_*^+(\varepsilon) = t_*^* + \varepsilon \varphi_*(1) \beta_1^+(\varepsilon) + \varepsilon^2 \mu_*(\varepsilon), \end{aligned} \tag{88}$$

$$q^+(\varepsilon) = q^* + \varepsilon \sum_{i=1}^{k} \beta_i^+(\varepsilon) \xi(i) + \varepsilon^2 \xi^0(\varepsilon), \tag{89}$$

where

$$\beta^+(\varepsilon) = (\beta_i^+(\varepsilon), i = 1, \dots, k),$$
$$v^+(\varepsilon) = (\mu_j(\varepsilon), j \in J_{(*)}, \Delta\mu_j(\varepsilon), j \in J_{(0)}, \mu_*(\varepsilon), \xi^0(\varepsilon)) \tag{90}$$

are the functions as in Theorem 4 with $\beta_1^+(0) = \beta_1^0 > 0$ (see proof of Lemma 2). The relations (77) follow from (88) and the assumption that

$$\Delta\varphi_j(1) > 0,\ j \in J_{(0)},\quad \varphi_{p^*}(1) < 1 = \varphi_*(1).$$

Consequently, the control (79) is admissible in the problem $TO(\tau)$ with $\tau = \tau_0 + \varepsilon^2$.

Let us prove the relations (83). Since Assumption 1 and the latter equations from (63) are true, it is enough to show that the relations

$$\alpha_j\sigma(q^+(\varepsilon), \tilde{t}_j(\varepsilon)) < 0,\ j \in J^* \setminus J_{(*)}, \tag{91}$$

are true, where the functions

$$\tilde{t}_j(\varepsilon),\ \varepsilon \in [0, \delta],\ j \in J^* \setminus J_{(*)}, \tag{92}$$

are given implicitly by

$$\frac{\partial\sigma(q^+(\varepsilon), \tilde{t}_j(\varepsilon))}{\partial t} \equiv 0,\ \varepsilon \in [0, \delta],\ \tilde{t}_j(+0) = t_j^*,\ j \in J^* \setminus J_{(*)}. \tag{93}$$

Indeed, by construction

$$\sigma(q^+(+0), \tilde{t}_j(+0)) = 0,\ \frac{\partial\sigma(q^+(+0), \tilde{t}_j(+0))}{\partial t} = 0,\ \ j \in J^* \setminus J_{(*)}. \tag{94}$$

Taking into account (89), (93), let us calculate

$$\alpha_j \frac{d\sigma(q^+(\varepsilon), \tilde{t}_j(\varepsilon))}{d\varepsilon}\Big|_{\varepsilon=+0} = \alpha_j\Big(\sum_{i=1}^{k}\beta_i^+(0)\xi^T(i)f(t_j^*) + \frac{\partial\sigma(q^+(+0), \tilde{t}_j(+0))}{\partial t}\frac{d\tilde{t}_j(+0)}{d\varepsilon}\Big)$$
$$= \Delta_j^* < 0,\ j \in J^* \setminus J_{(*)}. \tag{95}$$

To finish the proof we note that the relations (91) follow from (94) and (95), and the expansions (87) follow from (70), (71) and (74), (76). ⋄

Theorem 6 *Let Assumption 1 be fulfilled and $q^{*T}f(t_*^*) = 0,\ q^{*T}\dot{z}(\tau_0 + 0) < 0$. Let the sets $J_{(0)}$ and $J_{(*)}$ (see(52)) be such that*

$$rank\Big(f(t_j^*),\ j \in J_{(0)}\Big) = |J_{(0)}|,\ rank\Big(f(t_j^*),\ j \in J_{(*)}, z(\tau_0)\Big) = n, \tag{96}$$

$$\Delta\varphi_j(1) > 0 \ or\ \Delta\varphi_j(1) = 0\ and\ \Delta\mu_j > 0,\ for\ j \in J_{(0)}, \tag{97}$$

$$\Delta_j^* := \alpha_j\xi^T(1)f(t_j^*) < 0\ or\ \Delta_j^* = 0\ and\ \delta_j^* := \alpha_j\xi^{0^T}f(t_j^*) < 0,\ for\ j \in J^* \setminus J_{(*)}, \tag{98}$$

where $\varphi(1) = (\varphi_j(1), j \in J_{(*)}, \Delta\varphi_j(1), j \in J_{(0)}, \varphi_*(1) = 1, \xi(1)) \in Ker\ L$; $\beta^0 = \beta_1^0 > 0$, $v_0 = (\mu_j, j \in J_{(*)}, \Delta\mu_j, j \in J_{(0)}, \mu_*, \xi^0)$ *is a solution to the system (23) (24). Then the relations (77), (83) take place and the control (79) is optimal in the problem* $TO(\tau)$ *for* $\tau \in \mathcal{E}^+(\tau_0)$*. The following expansions are true for the defining elements*

$$\begin{aligned} t_{i(j)+1}(\tau_0 + \Delta\tau) &- t_{i(j)}(\tau_0 + \Delta\tau) = \Delta\varphi_j(1)s^*\sqrt{\Delta\tau} + \Delta\mu_j\Delta\tau + o(\Delta\tau),\ j \in J_{(0)}; \\ t_{i(j)}(\tau_0 + \Delta\tau) &= t_j^* + \varphi_j(1)s^*\sqrt{\Delta\tau} + \mu_j\Delta\tau + o(\Delta\tau),\ j \in J_{(*)} \setminus J_{(0)}; \\ t^*(\tau_0 + \Delta\tau) &= t_*^* + \varphi_*(1)s^*\sqrt{\Delta\tau} + \mu_*\Delta\tau + o(\Delta\tau); \\ q(\tau) &= q^* + s^*\xi(1)\sqrt{\Delta\tau} + \xi^0\Delta\tau + o(\Delta\tau), \end{aligned} \tag{99}$$

where $\Delta\tau = \tau - \tau_0 > 0$, $s^* = \beta_1^0 > 0$.

Proof. First of all let us note that (96) leads to $k :=$ dim Ker $L = 1$. Hence, the relations (77) follow from (88) with $k = 1$ and (97). Consequently, the control (79) is admissible in the problem $TO(\tau)$ with $\tau = \tau_0 + \varepsilon^2$.

Now let us prove the relations (83). Similarly to the proof of Theorem 5, it is enough to show that the relations (91) are true, where the functions $\tilde{t}_j(\varepsilon) < 0$, $\varepsilon \in [0, \delta_0]$, $j \in J^* \setminus J_{(*)}$, are given implicitly by (93).

Indeed, as before, the relations (94) are true and

$$\begin{aligned} &\alpha_j \frac{d\sigma(q^+(\varepsilon), \tilde{t}_j(\varepsilon))}{d\varepsilon}\bigg|_{\varepsilon=+0} = \alpha_j \bigg[\Big(\beta_1^+(\varepsilon)\xi(1) + \varepsilon\dot{\beta}_1^+(\varepsilon)\xi(1) + 2\varepsilon\xi^0(\varepsilon) + \varepsilon^2\dot{\xi}^0(\varepsilon)\Big)^T f(\tilde{t}_j(\varepsilon)) \\ &+ q^+(\varepsilon)\dot{f}(\tilde{t}_j(\varepsilon))\frac{d\tilde{t}_j(\varepsilon)}{d\varepsilon}\bigg]\bigg|_{\varepsilon=+0} = \Delta_j^* \le 0,\ j \in J^* \setminus J_{(*)}. \end{aligned} \tag{100}$$

For $j \in \{j \in J^* \setminus J_{(*)} : \Delta_j^* = 0\}$, let us calculate

$$\begin{aligned} &\alpha_j \frac{d^2\sigma(q^+(\varepsilon), \tilde{t}_j(\varepsilon))}{d\varepsilon^2}\bigg|_{\varepsilon=+0} = \alpha_j \bigg[\Big(2\dot{\beta}_1^+(+0)\xi(1) + 2\xi^0(+0)\Big)^T f(\tilde{t}_j(+0)) \\ &+ \dot{q}^T(+0)\dot{f}(\tilde{t}_j(+0))\frac{d\tilde{t}_j(+0)}{d\varepsilon} + q^T(+0)\ddot{f}(\tilde{t}_j(+0))\Big(\frac{d\tilde{t}_j(+0)}{d\varepsilon}\Big)^2\bigg] \\ &= 2\alpha_j {\xi^0}^T f(t_j^*) = 2\delta_j^* < 0. \end{aligned} \tag{101}$$

Here we have taken into account (89), (94) and the relations

$$\dot{q}^{+T}(\varepsilon)\dot{f}(\tilde{t}_j(\varepsilon)) + q^{+T}(\varepsilon)\ddot{f}(\tilde{t}_j(\varepsilon))\frac{d\tilde{t}_j(\varepsilon)}{d\varepsilon} \equiv 0,\ \varepsilon \in [0, \delta],$$

which follow from (93).

The relations (91) follow from (94), (100) and (101). The expansions (99) are true due to (70), (71) and (74), (76). ⋄

The assumptions made in the beginning of Subsection 3.1 guarantee the existence of a collection of sets (52) satisfying the conditions of Theorem 4 and the inequalities (77), (83) (or (78), (84)). Having the sets (52), one may construct optimal solutions to the problems $TO(\tau)$, $\tau = \tau_0 + \varepsilon^2 \in \mathcal{E}^+(\tau_0)$ by (79), (80) (in case the relations (77), (83) take place) or by (81), (82) (in case the relations (77), (83) take place). Here $\bar{t}_j^{\pm}(\varepsilon)$, $j \in J_{(*)}$, $\Delta t_j^{\pm}(\varepsilon)$, $j \in J_{(0)}$, $t^{*\pm}(\varepsilon)$, $q^{\pm}(\varepsilon)$ are the components of the corresponding $s-$vector-function (74) satisfying (75).

Remark 2 *Above we presented some rules for determining a collection of appropriate sets (52) by enumeration. It is possible to formulate rules which allow us to determine the appropriate sets uniquely without enumeration. For normal time-optimal problems it is done in [8]. For problems under consideration (abnormal problems) the rules are under construction and will be published separately.*

4 Acknowledgments

We would like to thank V. V. Gorokhovik for his interest, support and valuable comments and suggestions.

5 Appendix

Proof of Lemma 1.

Let $\xi(i)$, $i = 2, \ldots, k$, be a basis of the set Ker $(A_1, A_0, g)^T$. Consequently

$$\text{the vectors } \xi(i),\ i = 2, \ldots, k, \text{ are linearly independent,} \tag{A.1}$$

$$\xi^T(i)(A_1, A_0, g) = 0,\ i = 2, \ldots, k;\ \text{rank } (A_1, A_0, g) = n - (k-1). \tag{A.2}$$

It follows from 3) that rank $(A_1, A_0, g) =$ rank $(A_1, A_0) + 1$. Hence,

$$\text{rank } (A_1, A_0, a) = \text{rank } (A_1, A_0, g) - 1 \tag{A.3}$$

and a basis of the set Ker $(A_1, A_0, a)^T$ consists of k vectors. By construction (here we take into account that rank $(A_1, A_0, a) =$ rank (A_1, A_0)) we have

$$\xi^T(i)(A_1, A_0, a) = 0,\ i = 2, \ldots, k;\ q^{*T}(A_1, A_0, a) = 0. \tag{A.4}$$

Let us show that

$$\text{the vectors } \xi(i),\ i = 2, \ldots, k, \text{and } q^* \text{ are linearly independent.} \tag{A.5}$$

Suppose the contrary. Then, taking into account (A.1) we get $q^* = \sum\limits_{i=2}^{k} \xi(i)\mu_i$, with some $\mu_i, i = 2, \ldots, k$. Hence, due to (A.2) we have $q^{*T} g = \sum\limits_{i=2}^{k} \xi(i)^T g \mu_i = 0$. However, this contradicts 3). Consequently, relations (A.5) are true. The relations (A.1)-(A.5) mean, that the vectors $\xi(i)$, $i = 2, \ldots, k$, and q^* form a basis of the set Ker $(A_1, A_0, a)^T$. Thus, we may conclude that the vectors (69) should be included into a basis of the set Ker L^T. Let us show that the vectors (69) form a (complete) basis of the set Ker L^T. Suppose again the contrary. Then there exists a vector

$$(x, y, w, \omega, \alpha),\ (y, w, \omega, \alpha) \neq 0, \tag{A.6}$$
$$x \in R^n, y \in R^l, w \in R^m, \omega \in R^m, \alpha \in R,$$

such that

$$x^T \bar{A}_1 + y^T \bar{D}_1 = 0, \tag{A.7}$$

$$\omega^T D_0 = 0 \tag{A.8}$$

$$x^T \bar{A}_0 + \frac{1}{2}\omega^T D_0 = 0, \tag{A.9}$$

$$x^T a = 0, \tag{A.10}$$

$$y^T \bar{A}_1^T + w^T A_0^T + \omega^T A_*^T + \alpha g^T = 0, \tag{A.11}$$

where the matrix $\bar{D}_1 = \text{diag}(d_j^1 \alpha_j^1, j = 1, \dots, l)$ is positive definite.

The relation (A.8) yields that $\omega = 0$ and (A.9) results in $x^T \bar{A}_0 = x^T A_0 = 0$. Multiplying (A.7) and (A.11) by y and x from the right respectively, and taking into account that $\omega = 0$, $x^T A_0 = 0$, we get

$$x^T \bar{A}_1 + y^T \bar{D}_1 y = 0, \; y^T \bar{A}_1^T x + \alpha g^T x = 0.$$

Consequently,

$$\alpha g^T x = y^T \bar{D}_1 y. \tag{A.12}$$

Multiplying (A.11) by q^* from the right, we get $\alpha g^T q^* = 0$. Due to the condition 3) the latter equality is true only if $\alpha = 0$.

The equalities $\alpha = 0$ and (A.12) lead to $y = 0$. Under the relations $\omega = 0$, $y = 0$, $\alpha = 0$ the equation (A.11) takes the form $w^T A_0^T = 0$. Due to the assumption rank $A_0 = m$, the equation $w^T A_0^T = 0$ yields $w = 0$. Thus, we get $y = 0$, $w = 0$, $\omega = 0$, and $\alpha = 0$. However, this contradicts to (A.6). Consequently, there is no vector (A.6), satisfying the system (A.7)-(A.11). This means that the vectors (69) form a basis of the set Ker L^T.

Let us prove the second part of the Lemma concerning a basis of Ker L. By construction

$$\text{the vectors } \varphi(i), i = 2, \dots, k, \text{are linearly independent,} \tag{A.13}$$
$$L\varphi(i) = 0, i = 2, \dots, k.$$

It has been shown above that dim Ker$L^T = k$, hence, clearly, dim Ker$L = k$ too. Then, the relations (69), (70),(A.13) imply that there exists one more vector $\varphi(1)$ satisfying the properties

$$\text{the vectors } \varphi(i), i = 2, \dots, k, \text{and } \varphi(1) \text{ are linearly independent,} \tag{A.14}$$

$$L\varphi(1) = 0. \tag{A.15}$$

Let the vector $\varphi(1)$ have the form

$$\varphi(1) = (\varphi^1(1), \varphi^0(1), \Delta\varphi(1), \varphi_*(1), \xi(1))$$

with $\varphi^1(1) = (\varphi_j(1), j = 1, \dots, l)$, $\varphi^0(1) = (\varphi_{l+j}(1), j = 1, \dots, m)$, $\Delta\varphi(1) = (\Delta\varphi_j(1), j = 1, \dots, m)$, $\varphi_*(1) \in R$, $\xi(1) \in R^n$. Condition (A.14) yields that $\varphi(1) \neq 0$.

Suppose, that $\varphi_*(1) = 0$. Then relations (A.15) can be rewritten as follows

$$\bar{A}_1\varphi^1(1) + \bar{A}_0\Delta\varphi(1) = 0, \tag{A.16}$$
$$\bar{D}_1\varphi^1(1) + \bar{A}_1^T\xi(1) = 0, \tag{A.17}$$
$$\bar{A}_0^T\xi(1) = 0, \tag{A.18}$$
$$\bar{D}_0\varphi^0(1) + \tfrac{1}{2}\bar{D}_0\Delta\varphi(1) + \bar{A}_*^T\xi(1) = 0, \tag{A.19}$$
$$g^T\xi(1) = 0, \tag{A.20}$$

where $\bar{D}_0 = \text{diag}\,(d_j^0\alpha_j^0, j = 1,\ldots,m)\dot{>}0$, $\bar{A}_* = A_*\text{diag}\,(\alpha_j^0, j = 1,\ldots,m), \bar{A}_0 = A_0\text{diag}\,(\alpha_j^0, j = 1,\ldots,m)$.

Multiplying (A.16) and (A.17) by $\xi^T(1)$ and ${\varphi^1}^T(1)$ from the left respectively and taking into account (A.18) results in ${\varphi^1}^T(1)\bar{D}_1\varphi^1(1) = 0$. This yields $\varphi^1(1) = 0$. Then (A.16) takes the form $\bar{A}_0\Delta\varphi(1) = 0$. Due to the assumption 2) the latter equality leads to $\Delta\varphi(1) = 0$. Hence, the equations (A.17)-(A.20) can be rewritten as

$$\bar{A}_1^T\xi(1) = 0,\ \bar{A}_0^T\xi(1) = 0,\ \bar{D}_0\varphi^0(1) + \bar{A}_*^T\xi(1) = 0,\ g^T\xi(1) = 0. \tag{A.21}$$

It follows from (A.21) that $\xi(1) \in \text{Ker}\,(A_1, A_0, g)^T$ and, hence, $\xi(1) = \sum\limits_{i=2}^{k}\xi(i)\alpha_i$.

Suppose, that $\varphi^0(1) \neq 0$. Consequently,

$$0 \neq \varphi^0(1) = -\bar{D}_0^{-1}A_*^T\xi(1) = -\bar{D}_0^{-1}A_*^T\sum_{i=2}^{k}\xi(i)\alpha_i = \sum_{i=2}^{k}\varphi^0(i)\alpha_i,$$
$$(\alpha_i, i = 2,\ldots,k) \neq 0.$$

Here $\varphi^0(i) = (\varphi_{l+j}(i), j = 1,\ldots,m)$. However, this contradicts (A.14).

Now let us suppose that $\varphi^0(1) = 0$. Then, (A.21) results in

$$\bar{A}_1^T\xi(1) = 0,\ \bar{A}_0^T\xi(1) = 0,\ \bar{A}_*^T\xi(1) = 0,\ g^T\xi(1) = 0. \tag{A.22}$$

Due to the second condition from 2) the relations (A.22) yield $\xi(1) = 0$. Thus, we get

$$\varphi(1) = (\varphi^1(1) = 0, \varphi^0(1) = 0, \Delta\varphi(1) = 0, \varphi_*(1) = 0, \xi(1) = 0).$$

However, this contradicts to the condition $\varphi(1) \neq 0$. Consequently, our assumption that a vector $\varphi(1)$ satisfying relations (A.14), (A.15) has the component $\varphi_*(1)$ equal to zero is wrong.

Summing up, we have proved that there is a vector $\varphi(1)$ satisfying the relations (A.14), (A.15) with the component $\varphi_*(1) \neq 0$. Without loss of generality we may consider that $\varphi_*(1) = 1$. Lemma is proved. ◇

Proof of Lemma 2. Taking into account the properties of the vectors $\psi(i), \varphi(i)$, $i = 1,\ldots,k$, (see Lemma 1 and Corollary 3), we can rewrite the system (6) in the form

$$q^{*T}[(\beta_1)^2 f_* + \beta_1\sum_{j\in J_{(0)}}\sum_{i=2}^{k}2\alpha_j\dot{f}(t_j^*)\Delta\varphi_j(1)\varphi_j(i)\beta_i + \dot{z}(\tau_0+0)] = 0, \tag{A.23}$$

$$\xi^T(i)[(\beta_1)^2 f_* + \beta_1 \sum_{j\in J_{(0)}} \sum_{i=2}^{k} 2\alpha_j \dot f(t_j^*)\Delta\varphi_j(1)\varphi_j(i)\beta_i + \dot z(\tau_0+0)] = 0, \quad (A.24)$$
$$i = 2, \dots, k,$$

where

$$f_* = \sum_{j\in J(*)\setminus J_{(0)}} \alpha_j \dot f(t_j^*)(\varphi_j(1))^2 + \sum_{j\in J_{(0)}} \alpha_j \dot f(t_j^*)(2\varphi_j(1)\Delta\varphi_j(1) + \Delta\varphi_j^2(1)) \quad (A.25)$$
$$+\alpha_{\bar p}\dot f(t_*^*)\varphi_*^2(1)/2.$$

Let us show that $q^{*T} f_* > 0$. Using (A.25) and the fact that by construction

$$q^{*T}\dot f(t_j^*) = 0,\ j\in J_{(0)}, \quad (A.26)$$

we obtain

$$2q^{*T} f_* = \sum_{j\in J_{(*)}\setminus J_{(0)}} 2\alpha_j q^{*T}\dot f(t_j^*)(\varphi_j(1))^2 + \alpha_{\bar p} q^{*T}\dot f(t_*^*). \quad (A.27)$$

Since $L\varphi(1) = 0$, then

$$q^{*T}\dot f(t_j^*)\varphi_j(1) + \xi^T(1) f(t_j^*) = 0, j\in J_{(*)}\setminus J_{(0)}, \quad (A.28)$$
$$\xi^T(1) f(t_j^*) = 0,\ j\in J_{(0)}, \quad (A.29)$$
$$\sum_{j\in J_{(*)}\setminus J_{(0)}} 2\alpha_j f(t_j^*)\varphi_j(1) + \sum_{j\in J_{(0)}} 2\alpha_j f(t_j^*)\Delta\varphi_j(1) + \alpha_{\bar p} f(t_*^*) = 0. \quad (A.30)$$

Consequently, the vector $\tilde\mu = (\tilde\mu_j = \varphi_j(1),\ j\in J_{(*)}\setminus J_{(0)},\ \tilde\mu_j = \Delta\varphi_j(1),\ j\in J_{(0)})$ is a solution of the following problem

$$\min \sum_{j\in J_{(*)}} d_j\mu_j^2, \text{ s.t. } \sum_{j\in J_{(*)}} 2\alpha_j f(t_j^*)\mu_j = -\alpha_{\bar p} f(t_*^*). \quad (A.31)$$

By construction $p^*\in J_{(*)}\setminus J_{(0)}$, $t_{p^*}^* = t_*^*$, consequently, the vector $\bar\mu = (\bar\mu_j = 0,\ j\in J_{(*)}\setminus p^*$, $\bar\mu_{p^*} = 1/2)$ is feasible in problem (A.31). Thus

$$\sum_{j\in J_{(*)}} d_j\tilde\mu_j^2 = \sum_{j\in J_{(*)}\setminus J_{(0)}} \alpha_j q^{*T}\dot f(t_j^*)(\varphi_j(1))^2 \le 1/4 d_{p^*} = \frac{1}{4}\alpha_{\bar p} q^{*T}\dot f(t_*^*). \quad (A.32)$$

It follows from (A.27), (A.32), and the inequality $\alpha_{\bar p} q^{*T}\dot f(t_*^*) > 0$ that

$$q^{*T} f_* > 0. \quad (A.33)$$

We consider now the system (A.23), (A.24). Using (A.26) we can rewrite the equation (A.23)

$$(\beta_1)^2 q^{*T} f_* + q^{*T}\dot z(\tau_0+0) = 0 \quad \text{or} \quad (\beta_1)^2 = -q^{*T}\dot z(\tau_0+0)/q^{*T} f_*.$$

By assumption, $q^{*T}\dot z(\tau_0+0) < 0$, consequently, $\varepsilon_* := -q^{*T}\dot z(\tau_0+0)/q^{*T} f_* > 0$ and the parameter β_1 can take one of two values

$$\beta_1^+ = \sqrt{\varepsilon_*},\ \beta_1^- = -\sqrt{\varepsilon_*}. \quad (A.34)$$

Thus, the parameter β_1 is defined. To determine the other parameters β_i, $i = 2, \ldots, k$, we have the system of $(k-1)$ equations (A.24) with $(k-1)$ unknowns. Using the properties of the vectors $\varphi(i)$, $i = 2, \ldots, k$, we can rewrite the system (A.24) as follows

$$\mathcal{P}^T[(\beta_1)^2 f_* + \dot{z}(\tau_0 + 0) + 2\beta_1 A_* D A^T{}_* \mathcal{P}\bar{\beta}] = 0, \tag{A.35}$$

where

$$\begin{aligned} &\mathcal{P} = (\xi(i), i = 2, \ldots, k), \;\; \bar{\beta} = (\beta_i, i = 2, \ldots, k), \\ &D = \operatorname{diag}\Big(\frac{\alpha_j \Delta\varphi_j(1)}{q^{*T}\ddot{f}(t_j^*)}, j \in J_{(0)}\Big). \end{aligned} \tag{A.36}$$

Under the assumption $\Delta\varphi_j(1) > 0$, $j \in J_{(0)}$, (or $\Delta\varphi_j(1) < 0$, $j \in J_{(0)}$) it is easy to verify that $\mathcal{D}\dot{>}0$ (or $\mathcal{D}\dot{<}0$) and, hence,

$$\det \mathcal{P}^T A_* D A^T{}_* \mathcal{P} \neq 0. \tag{A.37}$$

Thus, for any β_1 the system (A.35) has a unique solution with respect to $\bar{\beta}$. It follows from (A.34), (A.35) that there are two solutions $\beta^+ = (\beta_i^+, i = 1, \ldots, k)$, $\beta^- = (\beta_i^-, i = 1, \ldots, k)$ of system (6), moreover $\beta^+ = -\beta^-$.

Let us prove now the relations (73). Again suppose the contrary: the matrix $\Phi_*^T S(\beta^+)$ is singular. Then there is a vector $\gamma = (\gamma_i,\, i = 1, \ldots, k), \gamma \neq 0$, such that

$$\Phi_*{}^T S(\beta^+)\gamma = 0. \tag{A.38}$$

Using again the features of the vectors $\psi(i)$, $\varphi(i)$, $i = 1, \ldots, k$, we can rewrite the last system in the form

$$q^{*T}\eta = 0, \;\; \mathcal{P}^T\eta = 0, \tag{A.39}$$

where $\mathcal{P} = (\xi(i),\, i = 2, \ldots, k)$, $\bar{\gamma} = (\gamma_i,\, i = 2, \ldots, k)$,

$$\eta = \gamma_1 f_* + \gamma_1 \sum_{j \in J_{(0)}} \sum_{i=2}^{k} 2\alpha_j \dot{f}(t_j^*)\Delta\varphi_j(1)\varphi_j(i)\beta_i^+ - 2\beta_1^+ A_* D A^T{}_* \mathcal{P}\bar{\gamma}.$$

As it has been mentioned before, by construction $q^{*T}\dot{f}(t_j^*) = 0, j \in J_{(0)}$; $q^{*T} f_* > 0$. Hence, the first equation from (A.39) takes the form $\gamma_1 q^{*T} f_* = 0$, implying $\gamma_1 = 0$. Thus, the last equation from (A.39) takes the form

$$2\beta_1^+ \mathcal{P}^T A_* D A^T{}_* \mathcal{P}\bar{\gamma} = 0. \tag{A.40}$$

Relations (A.37), (A.40) yield that $\bar{\gamma} = 0$. Thus, $\gamma = (\gamma_1,\, \bar{\gamma}) = 0$, which contradicts to the assumption that $\gamma \neq 0$. This proves the relation (73) and finishes the proof of the Lemma.

$\diamond$

References

[1] Clarke F.H., On the inverse function theorem, Pacific Journal of Mathematics, 64 (1976), pp. 97–10.

[2] Dontchev A.L. and Hager W.W., *An inverse mapping theorem for set-valued maps,* Proceedings of the American Mathematical Society, 121 (1994), pp. 481–489.

[3] Fiacco A. V., *Sensitivity analysis for nonlinear programming using penalty method,* Mathematical Programming, 10(1976), pp. 287–311.

[4] Gabasov R., Kirillova F.M. and Kostyukova O.I., *Optimally fast position control of linear nonsteady systems,* Phys.-Dokl., 39(1994), pp. 680–682.

[5] Golubitsky M. and Schaeffer D.G., *Singularities and groups in bifurcation theory,* V.1., N.Y. Springer, 1985.

[6] Izmailov A.F. and Trety'akov A.A. *2-Regular Solutions of Nonlinear Problems,* Moscow, Physiko-matematicheskaia Literatura, 1999. (in Russian)

[7] Kostyukova O.I., *Study of structure of co-control for a family of time optimal problems,* Preprint 4(527), Institute of Mathematics, National Academy of Sciences of Belarus, 1997.

[8] Kostyukova O.I. and Kostina E.A., *Analysis of properties of the solutions to parametric time-optimal problems*, Computational Optimization and Applications, 26(2003), pp. 285-326.

[9] Krasovskii N. N., *Theory of Motion Control,* Moscow, Nauka, 1968. (in Russian)

[10] Ledzewicz U. and Schättler H., *A high-order generalization of the Lyusternik theorem,* Nonlinear Analysis, 34 (1994), pp. 793 – 815.

[11] Malanowski K. and Maurer H., *Sensitivity analysis for state constrained optimal control problems. Discrete and Continuous Dynamical Systems*, 4(1998), pp. 241–272.

[12] Malanowski K. and Maurer H., *Sensitivity analysis of optimal control problems subject to higher order state constraints*, Annals of Operations Research., 101(2001) (Optimization with Perturbations II), pp. 43–73.

[13] Pales Z., *Inverse and implicit function theorems for nonsmooth maps in Banach spaces,* Journal of Mathematical Analysis and Applications, 209(1997), pp. 202–220.

[14] L. S. Pontryagin, V. G. Boltyanskij, R. V. Gamkrelidze and E. F. Mishchenko, *Selected Works, Vol. 4, The Mathematical Theory of Optimal Processes,* R. V. Gamkrelidze, ed., Classics of Soviet Mathematics, New York, NY: Gordon and Breach Science Publishers, 1986, (Translation from Russian edition, Nauka, Moscow, 1961)

[15] Pourciau B.H., *Analysis and optimization of Lipschitz continuous mappings,* Journal of Optimization Theory and Applications, 22(1977), pp. 311–351.

[16] Pshenichny B.N., *Implicit function theorem for multi-valued mappings,* Nonsmooth optimization and related topics, Ettore Majorana International Sci. Ser.: Phys. Sci. 43, Plenum Press, pp.371-391.

[17] Robinson S.M., *Perturbed Kuhn-Tucker points and rate of convergence for a class of nonlinear programming algorithms,* Mathematical Programming, 7(1974), pp. 1–16.

[18] Schwartz L., *Analyse Mathematique*, I, Hermann, 1967.

[19] Setharama Gowda M., *Inverse and implicit function theorems for H-differentiable and semismooth functions,* Optimization Methods and Software, 19(2004), pp. 443–461.

[20] Voevodin V.V. and Kuznetsov Yu. A., *Matrices and calculations*, Moscow, Nauka, 1984. (in Russian)