MAXIMUM PRINCIPLE AND ITS EXTENSION FOR BOUNDED CONTROL PROBLEMS WITH BOUNDARY CONDITIONS

OLGA VASILIEVA

This note is focused on a bounded control problem with boundary conditions. The control domain need not be convex. First-order necessary condition for optimality is obtained in the customary form of the maximum principle, and second-order necessary condition for optimality of singular controls is derived on the basis of second-order increment formula using the method of increments along with linearization approach.

1. Introduction. The role of the Pontryagin maximum principle is crucial for any research related to optimal processes that have control constraints. The simplicity of the principle's formulation together with its meaningful and beneficial directness has become an extraordinary attraction and one of the major causes for the appearance of new tendencies in mathematical sciences. The maximum principle is by nature a *necessary* first-order condition for optimality since it was born as an extension of Euler-Lagrange and Weierstrass necessary conditions of variational calculus. That is why the application of the Pontryagin maximum principle provides information on *extremal* controls, called sometimes *Pontryagin extremals*, among which an optimal control is to be searched. On the other hand, there is always a fair chance that such extremals may cease to be optimal, and it should be emphasized that in this note, we will discuss only necessary conditions for optimal control. Moreover, they will be derived under an essential qualitative assumption that optimal control does exist, that is, we will not discuss the problem of existence itself. The core results related to existence of optimal control can be found, for example, in [2, 11].

Proved initially for time-optimal linear systems, the maximum principle later on was generalized and extended for many other classes of control problems. The circle of problems solvable by means of the maximum principle is broadening every year. In this note, we demonstrate the maximum principle for nonlinear control problem with boundary conditions (Section 3). A distinctive feature of this particular problem is related to the system of differential equations that governs the dynamic process. Namely, it is supposed that one group of the state variables is specified in the initial time while the other group is specified in the terminal time and both groups take share and share alike in the boundary condition assigned to the dynamic system. That means that there are no phase or terminal constraints of any type. In other words, the boundary conditions are linked to the ODE system (instead of customary-used initial conditions) and mostly serve for determining its particular solution. On the other hand, a mere

replacement of initial conditions by the boundary ones is not acceptable. Therefore, our starting point will be to extract first-order dominant term in increment of objective functional and then to formulate the necessary condition for optimality in the customary form of the maximum principle.

Thus, the novelty of our approach consists in considering the nonlinear boundary conditions together with the dynamic system of ODEs, that is, as an entire object that for a specified admissible control produces a bunch of state profiles. The customary approach described in the past tends to separate these conditions by initial and terminal ones and proposes to unify the initial conditions with the dynamic system and to treat the terminal conditions as constraints.

It should be clarified that our target problem is not the same as the problem considered in [16] where *all* state profiles $x_k(t)$, $k = 1, 2, \ldots, n$, are subject to both initial and boundary conditions, and the initial states of the system are treated as controls. Neither is it similar to the problem considered in [9] where all system trajectories $x_k(t)$ are supposed to satisfy both initial and boundary conditions and the last ones are treated as terminal constraints.

It is not our intention to depreciate the maximum principle, but it would be fair to say that the maximum principle may cease working in some situations. An illustration of such a "misbehavior" is offered by a typical instance of *singular controls*. This issue was disclosed by Rosonoèr [10] in 1959, right after the announcement of the maximum principle was made, and then discussed by many other researches (see, e.g., [1, 3, 4, 5, 6, 7, 8, 9]). In fact, the maximum principle (as first-order necessary condition for optimality) provides no information about singular extremal controls. In this situation, one must seek an *extension of the maximum principle*, that is, higher-order necessary condition for optimality of singular controls. The proposed task can be accomplished using diverse techniques, such as the generalized Legendre-Clebsch condition [8, 9], the concept of a bundle of variations of [6, 7], the concept of Legendre bundles of disturbances used for time-optimal control problems [1], and the concept of sub-Riemannian "abnormal" minimizers that appear in [4]. This note is strongly motivated by the approach of Gabasov and Kirillova [6, 7] destined for free endpoint problems. The basic strategy for the proof (Section 4) is exactly the same as that of the classical argument given in [6, 7]. However, since a free endpoint problem is a special case of boundary value problem (BVP), it is clear that our version of necessary condition for optimality of singular controls also includes the classical version [6, 7] as a special case.

2. Problem statement and preliminaries. We start by posing the problem to minimize the objective functional

$$J(\mathbf{u}) = \varphi(\mathbf{x}(t_0), \mathbf{x}(t_1)) + \int_{t_0}^{t_1} F(\mathbf{x}, \mathbf{u}, t) dt \longrightarrow \min \tag{2.1}$$

which is defined on the solution set of the nonlinear BVP

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t), \quad \mathbf{x}(t) \in \mathbb{R}^n, \quad t \in T = [t_0, t_1], \tag{2.2}$$

$$\mathbf{\Phi}(\mathbf{x}(t_0), \mathbf{x}(t_1)) = \mathbf{0}, \quad \mathbf{\Phi} : \mathbb{R}^{2n} \longrightarrow \mathbb{R}^n, \tag{2.3}$$

where vector function $\mathbf{x} = \mathbf{x}(t)$, $t \in T$, describes the state of dynamic process (2.2)-(2.3), while $\mathbf{u} = \mathbf{u}(t)$, $\mathbf{u}(t) \in \mathbb{R}^r$, $t \in T$, represents the control vector.

Here we must outline the distinctive feature of the BVP (2.2)-(2.3). Namely, we suppose that one group of the state variables $x_i(t)$, $i \in \mathcal{I} \subset \{1,2,\dots,n\}$, is specified in the initial time, while the other group of $x_j(t)$, $j \in \{1,2,\dots,n\} \setminus \mathcal{I}$, is specified in the terminal time so that the BVP defined this way is well posed.

REMARK 2.1. A simple example of $\Phi(\mathbf{x}(t_0),\mathbf{x}(t_1)) = \mathbf{0}$ could be given by separated two-point boundary conditions: $x_i(t_0) - a_i = 0$, $a_i \in \mathbb{R}$, $i = 1,\dots,k$, $k < n$, and $x_j(t_1) - b_j = 0$, $b_j \in \mathbb{R}$, $j = k+1,\dots,n$. That means that there is a group of state profiles x_j for which the initial conditions are not specified, and another group x_i for which the terminal conditions are not specified either. Thus, (2.3) provides just enough conditions in order to define n arbitrary constants and to obtain the particular solution of ODE system (2.2) for a specific control vector $\mathbf{u}(t)$. Under such approach, the boundary conditions are "linked" to the ODE system and we do not treat them as terminal state constraints. The majority of research papers dedicated to the high-order maximum principle are focused on initial value systems without state constraints (e.g., [3, 6, 7]) or on initial value systems with additional endpoint constraint (see, e.g., [9, 16]). It should be clarified that this note provides an analysis of necessary optimality conditions of optimal control for BVP without using the standard procedure for dealing with state or phase constraints (such as Lagrange multipliers, transversality conditions, etc.).

Thus, the problem under consideration does not contain any state constraints and we will refer to the class of *admissible controls* as a set of measurable vector functions $\mathbf{u}(\cdot) \in \mathbb{L}_\infty^r(T)$ with direct inclusion constraint

$$\mathbf{u}(t) \in U, \quad t \in T, \tag{2.4}$$

where U is some compact nonempty set in $\mathbb{R}^r$. It should be emphasized that the control domain U need not be convex. A pair of functions $\{\mathbf{u},\mathbf{x}\}$, where $\mathbf{u} = \mathbf{u}(t)$ is an admissible control and $\mathbf{x} = \mathbf{x}(t,\mathbf{u})$ is the corresponding state of dynamic process calculated in $\mathbf{u}(t)$, will be called *admissible process*. Control function $\mathbf{u}^* = \mathbf{u}^*(t)$ which provide the minimum value to the objective functional (2.1) will be called *optimal control*, and the process $\{\mathbf{u}^*,\mathbf{x}^*\}$ will be referred to as *optimal process*.

Now we will impose the following hypotheses on the entries of problem (2.1), (2.2), (2.3), and (2.4) formulated above:

(1) vector function $\mathbf{f} = (f_1,\dots,f_n)$ is continuous in $(\mathbf{x},\mathbf{u},t)$ together with $\partial \mathbf{f}/\partial \mathbf{x}$ and satisfies the Lipschitz condition with respect to $\mathbf{x}$,

$$\left\| \mathbf{f}(\mathbf{x} + \Delta\mathbf{x},\mathbf{u},t) - \mathbf{f}(\mathbf{x},\mathbf{u},t) \right\| \leq \mathcal{L} \left\| \Delta\mathbf{x}(t) \right\|, \tag{2.5}$$

in any closed domain of variables $(\mathbf{x},\mathbf{u},t)$ for the same constant $\mathcal{L} > 0$;

(2) scalar functions φ and F are continuous in their arguments together with partial derivatives $\partial\varphi/\partial\mathbf{x}(t_0)$, $\partial\varphi/\partial\mathbf{x}(t_1)$, $\partial F/\partial\mathbf{x}$;

(3) vector function $\mathbf{\Phi} = (\Phi_1, \ldots, \Phi_n)$ is continuous in $\mathbf{x}(t_0), \mathbf{x}(t_1)$ together with partial derivatives $\partial\mathbf{\Phi}/\partial\mathbf{x}(t_0)$, $\partial\mathbf{\Phi}/\partial\mathbf{x}(t_1)$, and

$$\mathrm{rank}\left[\frac{\partial\mathbf{\Phi}}{\partial\mathbf{x}(t_0)}\,\frac{\partial\mathbf{\Phi}}{\partial\mathbf{x}(t_1)}\right] = n. \tag{2.6}$$

The latter implies that BVP (2.2)-(2.3) is well posed.

REMARK 2.2. For $\mathbf{\Phi} = \mathbf{\Phi}(\mathbf{x}(t_0))$, problem (2.1), (2.2), (2.3), and (2.4) turns into familiar free endpoint problem. In this case, the question of existence and uniqueness of solution to the initial value problem in any admissible control is answered straightforwardly. Namely, the fact that $\mathbf{f}$ is Lipschitzian clearly indicates it. The same may be said about BVP (2.2)-(2.3) *only* if system (2.2) is linear in the state variable (see, e.g., [13, 15]).

Since the existence of BVP solution is a different issue, we do not incorporate it in this note and simply introduce the following assumption.

ASSUMPTION 2.3. Suppose that BVP (2.2)-(2.3) is resolvable in any admissible control $\mathbf{u} = \mathbf{u}(t)$, $\mathbf{u}(t) \in U$, and that the set formed by all admissible pairs $\{\mathbf{u}, \mathbf{x} = \mathbf{x}(t, \mathbf{u})\}$ is closed, that is, in the point-set topological sense, that set contains all of its limit points.

The purpose of Assumption 2.3 is to give a simplified justification to the existence of finite infimum of $J(\mathbf{u})$. For more formal details, see [11, pages 500–518] that provides thorough analysis of Weierstrass theorem in functional spaces.

3. First-order optimality condition. In order to obtain the necessary conditions for optimality, we will use the standard procedure (see, e.g., [12]). Namely, we should analyze the response of objective functional caused by some control impulse or, in other words, we must derive the increment formula that originates from the Taylor's series expansion. A suitable definition of the conjugate system will facilitate the extraction of dominant term that is destined to determine the necessary condition for optimality. For the sake of simplicity, it will be reasonable to construct a linearized model of nonlinear system (2.2)-(2.3) in some small vicinity.

3.1. Increment formula. For two admissible processes, the basic one $\{\mathbf{u}, \mathbf{x} = \mathbf{x}(t, \mathbf{u})\}$ and the varied one $\{\tilde{\mathbf{u}} = \mathbf{u} + \Delta\mathbf{u},\ \tilde{\mathbf{x}} = \mathbf{x} + \Delta\mathbf{x} = \mathbf{x}(t, \tilde{\mathbf{u}})\}$, we can define the BVP (2.2)-(2.3) in increments:

$$\Delta\dot{\mathbf{x}} = \Delta\mathbf{f}(\mathbf{x}, \mathbf{u}, t), \qquad \Delta\mathbf{\Phi}(\mathbf{x}(t_0), \mathbf{x}(t_1)) = \mathbf{0}, \tag{3.1}$$

where

$$\Delta\mathbf{f}(\mathbf{x}, \mathbf{u}, t) = \mathbf{f}(\tilde{\mathbf{x}}, \tilde{\mathbf{u}}, t) - \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \tag{3.2}$$

denotes the total increment in contrast to partial increments to be used later on:

$$\Delta_{\tilde{u}}\mathbf{f}(\mathbf{x}, \mathbf{u}, t) = \mathbf{f}(\mathbf{x}, \tilde{\mathbf{u}}, t) - \mathbf{f}(\mathbf{x}, \mathbf{u}, t). \tag{3.3}$$

We introduce some nontrivial differentiable vector function $\boldsymbol{\psi} = \boldsymbol{\psi}(t)$, $\boldsymbol{\psi}(t) \in \mathbb{R}^n$, and a numerical vector $\boldsymbol{\lambda} \in \mathbb{R}^n$, both undefined for now, that will be used later to define the conjugate system. Then the increment of objective functional (2.1) may be represented as

$$\begin{aligned}
\Delta J(\mathbf{u}) = {} & \Delta \varphi(\mathbf{x}(t_0), \mathbf{x}(t_1)) + \int_T \Delta F(\mathbf{x}, \mathbf{u}, t)\, dt \\
& + \int_T \langle \boldsymbol{\psi}(t), \Delta \dot{\mathbf{x}}(t) - \Delta \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \rangle\, dt \\
& + \langle \boldsymbol{\lambda}, \Delta \boldsymbol{\Phi}(\mathbf{x}(t_0), \mathbf{x}(t_1)) \rangle,
\end{aligned} \tag{3.4}$$

where $\langle \cdot, \cdot \rangle$ stands for inner product in $\mathbb{R}^n$.

REMARK 3.1. It should be noted that an arbitrary numerical vector $\boldsymbol{\lambda}$ does not play the role of Lagrange multiplier here since the formulated problem (2.1)–(2.4) does not include additional phase constraints, while the boundary condition (2.3) is adhered to the dynamic system (2.2). It is essential that $\boldsymbol{\lambda}$ be arbitrary and $\boldsymbol{\lambda} \neq \mathbf{0}$ in order to fulfill the boundary condition (2.3).

Now we will perform a few necessary transformations, namely,
(i) introduce the Hamiltonian function

$$H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t) = \langle \boldsymbol{\psi}(t), \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \rangle - F(\mathbf{x}, \mathbf{u}, t); \tag{3.5}$$

(ii) expand $\Delta \varphi$ and $\Delta \boldsymbol{\Phi}$ in Taylor series up to the first-order term:

$$\begin{aligned}
\Delta \varphi(\mathbf{x}(t_0), \mathbf{x}(t_1)) = {} & \left\langle \frac{\partial \varphi}{\partial \mathbf{x}(t_0)}, \Delta \mathbf{x}(t_0) \right\rangle + \left\langle \frac{\partial \varphi}{\partial \mathbf{x}(t_1)}, \Delta \mathbf{x}(t_1) \right\rangle \\
& + o_\varphi(\|\Delta \mathbf{x}(t_0)\|, \|\Delta \mathbf{x}(t_1)\|),
\end{aligned} \tag{3.6}$$

$$\begin{aligned}
\Delta \boldsymbol{\Phi}(\mathbf{x}(t_0), \mathbf{x}(t_1)) = {} & \frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_0)} \Delta \mathbf{x}(t_0) + \frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_1)} \Delta \mathbf{x}(t_1) \\
& + \mathbf{o}_{\boldsymbol{\Phi}}(\|\Delta \mathbf{x}(t_0)\|, \|\Delta \mathbf{x}(t_1)\|);
\end{aligned} \tag{3.7}$$

(iii) perform the integration by parts:

$$\begin{aligned}
\int_T \langle \boldsymbol{\psi}(t), \Delta \dot{\mathbf{x}}(t) \rangle\, dt = {} & \langle \boldsymbol{\psi}(t_1), \Delta \mathbf{x}(t_1) \rangle - \langle \boldsymbol{\psi}(t_0), \Delta \mathbf{x}(t_0) \rangle \\
& - \int_T \langle \dot{\boldsymbol{\psi}}(t), \Delta \mathbf{x}(t) \rangle\, dt.
\end{aligned} \tag{3.8}$$

It should be mentioned that in previous deductions, o_φ and $\mathbf{o}_{\boldsymbol{\Phi}}$ are the remainder terms of the Taylor-series expansions with respect to $\mathbf{x}(t_0)$, $\mathbf{x}(t_1)$ of the scalar function φ and vector function $\boldsymbol{\Phi}$, respectively, and that $o(\alpha)/\alpha \to 0$, $\alpha \to 0$.

Thus, the extended version of (3.4) will be the following:

$$
\begin{aligned}
\Delta J(\mathbf{u}) = &-\int_T \langle \dot{\boldsymbol{\psi}}(t), \Delta \mathbf{x}(t) \rangle dt - \int_T \Delta_{\tilde{x}\tilde{u}} H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t) dt \\
&+ \left\langle \left[\frac{\partial \varphi}{\partial \mathbf{x}(t_0)} - \boldsymbol{\psi}(t_0) + \left(\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_0)} \right)' \boldsymbol{\lambda} \right], \Delta \mathbf{x}(t_0) \right\rangle \\
&+ \left\langle \left[\frac{\partial \varphi}{\partial \mathbf{x}(t_1)} + \boldsymbol{\psi}(t_1) + \left(\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_1)} \right)' \boldsymbol{\lambda} \right], \Delta \mathbf{x}(t_1) \right\rangle \\
&+ o_\varphi(\|\Delta \mathbf{x}(t_0)\|, \|\Delta \mathbf{x}(t_1)\|) + \langle \boldsymbol{\lambda}, \mathbf{o}_\Phi(\|\Delta \mathbf{x}(t_0)\|, \|\Delta \mathbf{x}(t_1)\|) \rangle,
\end{aligned}
\tag{3.9}
$$

where the prime $'$ denotes the transpose of the matrix and

$$
\Delta_{\tilde{x}\tilde{u}} H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t) = H(\boldsymbol{\psi}, \tilde{\mathbf{x}}, \tilde{\mathbf{u}}, t) - H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t).
\tag{3.10}
$$

It is easy to see that the right-hand side of formula (3.9) consists of three parts—the integral part (IP) formed by the first row, the terminal part (TP) formed by the second and third rows, and the remainder part (RP) formed by the fourth row.

In the IP of (3.9), having performed the following rearrangements

$$
\begin{aligned}
\Delta_{\tilde{x}\tilde{u}} H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t) &= \Delta_{\tilde{x}} H(\boldsymbol{\psi}, \mathbf{x}, \tilde{\mathbf{u}}, t) + \Delta_{\tilde{u}} H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t), \\
\Delta_{\tilde{x}} H(\boldsymbol{\psi}, \mathbf{x}, \tilde{\mathbf{u}}, t) &= \left\langle \frac{\partial H(\boldsymbol{\psi}, \mathbf{x}, \tilde{\mathbf{u}}, t)}{\partial \mathbf{x}}, \Delta \mathbf{x}(t) \right\rangle + o_H(\|\Delta \mathbf{x}(t)\|), \\
\frac{\partial H(\boldsymbol{\psi}, \mathbf{x}, \tilde{\mathbf{u}}, t)}{\partial \mathbf{x}} &= \Delta_{\tilde{u}} \frac{\partial H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}} + \frac{\partial H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}},
\end{aligned}
\tag{3.11}
$$

we obtain

$$
\begin{aligned}
\text{IP} = &-\int_T \langle \dot{\boldsymbol{\psi}}(t), \Delta \mathbf{x}(t) \rangle dt - \int_T \Delta_{\tilde{x}\tilde{u}} H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t) dt \\
= &-\int_T \Delta_{\tilde{u}} H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t) dt - \int_T o_H(\|\Delta \mathbf{x}(t)\|) dt \\
&-\int_T \left\langle \Delta_{\tilde{u}} \frac{\partial H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}} + \frac{\partial H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}} + \dot{\boldsymbol{\psi}}(t), \Delta \mathbf{x}(t) \right\rangle dt.
\end{aligned}
\tag{3.12}
$$

In order to simplify this expression, we will use the arbitrariness of differentiable function $\boldsymbol{\psi}(t)$ to define the *conjugate system*:

$$
\dot{\boldsymbol{\psi}} = -\frac{\partial H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}}.
\tag{3.13}
$$

Thus, we arrive at the expression

$$
\begin{aligned}
\text{IP} = &-\int_T \Delta_{\tilde{u}} H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t) dt - \int_T \left\langle \Delta_{\tilde{u}} \frac{\partial H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}}, \Delta \mathbf{x}(t) \right\rangle dt \\
&-\int_T o_H(\|\Delta \mathbf{x}(t)\|) dt.
\end{aligned}
\tag{3.14}
$$

In the TP of (3.9), we note the presence of arbitrary vector $\boldsymbol{\lambda}$. It would be useful to eliminate the parameter $\boldsymbol{\lambda}$ in some way in order to guarantee the fulfillment of both

boundary conditions (3.1) and (2.3). To do so, we will equate the TP of (3.9) to zero:

$$
\begin{aligned}
\text{TP} = &\left\langle \left[\frac{\partial \varphi}{\partial \mathbf{x}(t_0)} - \boldsymbol{\psi}(t_0) + \left(\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_0)} \right)' \boldsymbol{\lambda} \right], \Delta \mathbf{x}(t_0) \right\rangle \\
&+ \left\langle \left[\frac{\partial \varphi}{\partial \mathbf{x}(t_1)} + \boldsymbol{\psi}(t_1) + \left(\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_1)} \right)' \boldsymbol{\lambda} \right], \Delta \mathbf{x}(t_1) \right\rangle = 0.
\end{aligned}
\tag{3.15}
$$

This means that we can also equate to zero each term involving $\boldsymbol{\lambda}$:

$$
\begin{aligned}
\frac{\partial \varphi}{\partial \mathbf{x}(t_0)} - \boldsymbol{\psi}(t_0) + \left(\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_0)} \right)' \boldsymbol{\lambda} = \mathbf{0}, \\
\frac{\partial \varphi}{\partial \mathbf{x}(t_1)} + \boldsymbol{\psi}(t_1) + \left(\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_1)} \right)' \boldsymbol{\lambda} = \mathbf{0}.
\end{aligned}
\tag{3.16}
$$

If for some arbitrary nontrivial $(n \times n)$ matrices $\mathbf{B}_0$, $\mathbf{B}_1$,

$$
\begin{aligned}
\mathbf{B}_0 \left[\frac{\partial \varphi}{\partial \mathbf{x}(t_0)} - \boldsymbol{\psi}(t_0) + \left(\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_0)} \right)' \boldsymbol{\lambda} \right] = \mathbf{0}, \\
\mathbf{B}_1 \left[\frac{\partial \varphi}{\partial \mathbf{x}(t_1)} + \boldsymbol{\psi}(t_1) + \left(\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_1)} \right)' \boldsymbol{\lambda} \right] = \mathbf{0},
\end{aligned}
\tag{3.17}
$$

then their sum is also equal to zero:

$$
\begin{aligned}
\mathbf{B}_0 \frac{\partial \varphi}{\partial \mathbf{x}(t_0)} + \mathbf{B}_1 \frac{\partial \varphi}{\partial \mathbf{x}(t_1)} - \mathbf{B}_0 \boldsymbol{\psi}(t_0) \\
+ \mathbf{B}_1 \boldsymbol{\psi}(t_1) + \left[\mathbf{B}_0 \left(\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_0)} \right)' + \mathbf{B}_1 \left(\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_1)} \right)' \right] \boldsymbol{\lambda} = \mathbf{0}.
\end{aligned}
\tag{3.18}
$$

So, the matrices $\mathbf{B}_0$, $\mathbf{B}_1$ must be chosen according to the condition

$$
\mathbf{B}_0 \left(\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_0)} \right)' + \mathbf{B}_1 \left(\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_1)} \right)' = \mathbf{0}.
\tag{3.19}
$$

In this case, we can write the boundary conditions for the conjugate system (3.13) as

$$
\mathbf{B}_0 \frac{\partial \varphi}{\partial \mathbf{x}(t_0)} + \mathbf{B}_1 \frac{\partial \varphi}{\partial \mathbf{x}(t_1)} - \mathbf{B}_0 \boldsymbol{\psi}(t_0) + \mathbf{B}_1 \boldsymbol{\psi}(t_1) = \mathbf{0}.
\tag{3.20}
$$

REMARK 3.2. The selection of $\mathbf{B}_0$ and $\mathbf{B}_1$ according to (3.19) is not a difficult task since system (3.19) consists of n^2 linear algebraic equations in $2n^2$ unknowns. The only thing which remains unusual here is that the boundary conditions for the conjugate system are defined in nonunique manner.

On closer examination, it becomes clear that the conjugate system (3.13) is linear with respect to $\boldsymbol{\psi}$ and that the conjugate boundary conditions (3.20) for $\mathbf{B}_0$, $\mathbf{B}_1$ defined by equality (3.19) are also linear with respect to $\boldsymbol{\psi}(t_0)$, $\boldsymbol{\psi}(t_1)$ for any admissible process. Thus, for the original *nonlinear* BVP (2.2)-(2.3), we have obtained the corresponding *linear* conjugate BVP (3.13), (3.19), and (3.20). It should be also noted that if boundary conditions (2.3) for $\mathbf{x}$ involve $x_i(t_0)$, $i \in \mathscr{I} \subset \{1, 2, \ldots, n\}$, and $x_j(t_1)$, $j \in \{1, 2, \ldots, n\} \setminus \mathscr{I}$, then boundary conditions (3.20) for $\boldsymbol{\psi}$ will involve the opposite $\psi_j(t_0)$ and $\psi_i(t_1)$.

Finally, we can write the increment formula (3.9) in the simplified form:

$$\Delta J(\mathbf{u}) = -\int_T \Delta_{\tilde{u}} H(\boldsymbol{\psi},\mathbf{x},\mathbf{u},t)dt - \int_T \left\langle \Delta_{\tilde{u}} \frac{\partial H(\boldsymbol{\psi},\mathbf{x},\mathbf{u},t)}{\partial \mathbf{x}}, \Delta \mathbf{x}(t) \right\rangle dt + \eta_{\tilde{u}}, \qquad (3.21)$$

$$\eta_{\tilde{u}} = o_\varphi(||\Delta\mathbf{x}(t_0)||,||\Delta\mathbf{x}(t_1)||) + \langle \boldsymbol{\lambda}, \mathbf{o}_\Phi(||\Delta\mathbf{x}(t_0)||,||\Delta\mathbf{x}(t_1)||)\rangle - \int_T o_H(||\Delta\mathbf{x}(t)||)dt. \qquad (3.22)$$

REMARK 3.3. If instead of boundary conditions (2.3) for the original system (2.2), we have the initial conditions, that is, $\boldsymbol{\Phi} = \boldsymbol{\Phi}(\mathbf{x}(t_0))$, then for $\mathbf{B}_0 = \mathbf{0}$, $\mathbf{B}_1 = \mathbf{I}$, the conjugate BVP (3.13), (3.19), (3.20) turns into familiar terminal value problem for conjugate function $\boldsymbol{\psi}$:

$$\dot{\boldsymbol{\psi}} = -\frac{\partial H(\boldsymbol{\psi},\mathbf{x},\mathbf{u},t)}{\partial \mathbf{x}}, \qquad \boldsymbol{\psi}(t_1) = -\frac{\partial \varphi}{\partial \mathbf{x}(t_1)}. \qquad (3.23)$$

Thus, $\boldsymbol{\psi}(t_1)$ is an opposite match against $\mathbf{x}(t_0)$.

3.2. The maximum principle. A customary form of the necessary optimality condition—the maximum principle—will follow from the increment formula (3.21) if we will succeed to show that on the needle-shaped (spiky) variation

$$\tilde{\mathbf{u}}(t) = \mathbf{u}_\varepsilon(t) = \begin{cases} \mathbf{v} \in U, & t \in (\tau - \varepsilon, \tau] \subset T, \ \varepsilon > 0, \\ \mathbf{u}(t), & t \in T \setminus (\tau - \varepsilon, \tau], \end{cases} \qquad (3.24)$$

the state increment $\Delta_\varepsilon \mathbf{x}(t)$ has the order ε. In the control theory for initial value problems, that result follows from the natural assumption on the validity of the Lipschitz condition with respect to $\mathbf{x}$ for $\mathbf{f}(\mathbf{x},\mathbf{u},t)$ and from the Gronwall lemma. In the case of BVP, the substantiation of this result essentially depends upon the qualitative Assumption 2.3 on resolvability of the original BVP (2.2)-(2.3).

In order to demonstrate this, we represent the BVP in increments (3.1) in the form

$$\Delta\dot{\mathbf{x}} = \Delta_{\tilde{x}}\mathbf{f}(\mathbf{x},\tilde{\mathbf{u}},t) + \Delta_{\tilde{u}}\mathbf{f}(\mathbf{x},\mathbf{u},t), \qquad (3.25)$$

$$\Delta_{\tilde{x}(t_0)}\boldsymbol{\Phi}(\mathbf{x}(t_0),\tilde{\mathbf{x}}(t_1)) + \Delta_{\tilde{x}(t_1)}\boldsymbol{\Phi}(\mathbf{x}(t_0),\mathbf{x}(t_1)) = \mathbf{0}. \qquad (3.26)$$

On the two indicated admissible processes $\{\mathbf{u},\mathbf{x}\}$, $\{\tilde{\mathbf{u}},\tilde{\mathbf{x}}\}$, we define a mean weighted $(n \times n)$ matrix function

$$\mathbf{A}[t] = \int_0^1 \frac{\partial \mathbf{f}(\mathbf{x}(t) + \theta\Delta\mathbf{x}(t),\tilde{\mathbf{u}},t)}{\partial \mathbf{x}}d\theta. \qquad (3.27)$$

REMARK 3.4. This representation results from the Lagrange theorem on finite increments (see, e.g., [11])

$$\Delta_{\tilde{x}}\mathbf{G}(\mathbf{x},\cdot) = \frac{\partial \mathbf{G}(\mathbf{x}+\alpha\Delta\mathbf{x},\cdot)}{\partial \mathbf{x}}\Delta\mathbf{x}, \quad 0 \le \alpha \le 1, \qquad (3.28)$$

where the partial derivative is represented using the mean value theorem

$$\int_0^1 \frac{\partial \mathbf{G}(\mathbf{x}+\theta\Delta\mathbf{x},\cdot)}{\partial \mathbf{x}}d\theta = \frac{\partial \mathbf{G}(\mathbf{x}+\alpha\Delta\mathbf{x},\cdot)}{\partial \mathbf{x}} \cdot [1-0] \qquad (3.29)$$

so that

$$\Delta_{\tilde{x}}\mathbf{G}(\mathbf{x},\cdot) = \left[\int_0^1 \frac{\partial \mathbf{G}(\mathbf{x}+\theta\Delta\mathbf{x},\cdot)}{\partial \mathbf{x}}d\theta\right]\cdot\Delta\mathbf{x}. \tag{3.30}$$

This enables us to represent the partial increment by the formula

$$\Delta_{\tilde{x}}\mathbf{f}(\mathbf{x},\tilde{\mathbf{u}},t) = \mathbf{A}[t]\Delta\mathbf{x}(t). \tag{3.31}$$

Analogously, we can define

$$\begin{aligned}
\mathbf{L}_0 &= \int_0^1 \frac{\partial \mathbf{\Phi}(\mathbf{x}(t_0)+\theta\Delta\mathbf{x}(t_0),\tilde{\mathbf{x}}(t_1))}{\partial \mathbf{x}(t_0)}d\theta, \\
\mathbf{L}_1 &= \int_0^1 \frac{\partial \mathbf{\Phi}(\mathbf{x}(t_0),\mathbf{x}(t_1)+\theta\Delta\mathbf{x}(t_1))}{\partial \mathbf{x}(t_1)}d\theta,
\end{aligned} \tag{3.32}$$

whence

$$\begin{aligned}
\Delta_{\tilde{x}(t_0)}\mathbf{\Phi}(\mathbf{x}(t_0),\tilde{\mathbf{x}}(t_1)) &= \mathbf{L}_0\Delta\mathbf{x}(t_0), \\
\Delta_{\tilde{x}(t_1)}\mathbf{\Phi}(\mathbf{x}(t_0),\mathbf{x}(t_1)) &= \mathbf{L}_1\Delta\mathbf{x}(t_1).
\end{aligned} \tag{3.33}$$

In terms of the matrices introduced above, we construct a linear BVP relating certain state $\mathbf{z}(t) \in \mathbb{R}^n$ with the admissible controls $\mathbf{v} = \mathbf{v}(t)$, $\mathbf{v}(t) \in U$:

$$\dot{\mathbf{z}} = \mathbf{A}[t]\mathbf{z}+\Delta_v\mathbf{f}(\mathbf{x},\mathbf{u},t), \qquad \mathbf{L}_0\mathbf{z}(t_0)+\mathbf{L}_1\mathbf{z}(t_1) = \mathbf{0}. \tag{3.34}$$

BVP (3.34) is referred to as *associated linear BVP* to the original nonlinear BVP (2.2)-(2.3).

LEMMA 3.5. *For* $\mathbf{v}(t) = \mathbf{u}(t)$, *the solution of associated linear BVP (3.34) is trivial* $\mathbf{z}(t) \equiv \mathbf{0}$, *and for* $\mathbf{v}(t) = \tilde{\mathbf{u}}(t)$, *its solution is* $\mathbf{z}(t) \equiv \Delta\mathbf{x}(t)$.

PROOF. First statement is obvious since $\Delta_u\mathbf{f}(\mathbf{x},\mathbf{u},t) \equiv \mathbf{0}$ and the linear homogeneous BVP

$$\dot{\mathbf{z}} = \mathbf{A}[t]\mathbf{z}, \qquad \mathbf{L}_0\mathbf{z}(t_0)+\mathbf{L}_1\mathbf{z}(t_1) = \mathbf{0} \tag{3.35}$$

has only trivial solution $\mathbf{z}(t) \equiv \mathbf{0}$. On the other hand, for $\mathbf{v}(t) = \tilde{\mathbf{u}}(t)$, the associated linear BVP (3.34) coincides with the BVP in increments (3.25)–(3.26). Therefore, $\mathbf{z}(t) \equiv \Delta\mathbf{x}(t)$. □

Lemma 3.5 displays the connection between the solutions of two BVPs—the associated linear BVP (3.34) and the BVP in increments (3.1) or (3.25)–(3.26). The solution of the latter is also linked to the solution of the original nonlinear BVP (2.2)-(2.3). Thus, under the qualitative Assumption 2.3 on resolvability of the original nonlinear BVP (2.2)-(2.3), the associated linear BVP (3.34) is always resolvable for any admissible control $\mathbf{v} = \mathbf{v}(t)$.

This fact follows from the *resolvability condition* and the *analogue of Cauchy formula* for representation of the solution to linear BVP (see [13, 14] for more details) which we reproduce here:

$$\mathbf{z}(t) = \int_{t_0}^{t_1} \mathscr{F}(t,t_1,\tau)\Delta_v\mathbf{f}(\mathbf{x},\mathbf{u},\tau)d\tau + \int_{t_0}^{t} \mathbf{X}(t)\mathbf{X}^{-1}(\tau)\Delta_v\mathbf{f}(\mathbf{x},\mathbf{u},\tau)d\tau, \tag{3.36}$$

where

$$\mathscr{F}(t, t_1, \tau) = -\mathbf{X}(t) \left[\mathbf{L}_0 + \mathbf{L}_1 \mathbf{X}(t_1) \right]^{-1} \mathbf{L}_1 \mathbf{X}(t_1) \mathbf{X}^{-1}(\tau),$$
$$\dot{\mathbf{X}} = \mathbf{A}[t] \mathbf{X}, \qquad \mathbf{X}(t_0) = \mathbf{I}, \tag{3.37}$$

and the resolvability condition of linear BVP (3.34) is

$$\det \left[\mathbf{L}_0 + \mathbf{L}_1 \mathbf{X}(t_1) \right] \neq 0. \tag{3.38}$$

Formula (3.36) yields obviously the estimate

$$\left\| \mathbf{z}(t) \right\| \leq \mathcal{M} \int_T \left\| \Delta_v \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \right\| dt, \quad \mathcal{M} = \mathrm{const} > 0. \tag{3.39}$$

The latter for $\mathbf{v}(t) = \tilde{\mathbf{u}}(t)$ turns into

$$\left\| \Delta \mathbf{x}(t) \right\| \leq \mathcal{M} \int_T \left\| \Delta_{\tilde{u}} \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \right\| dt, \quad \mathcal{M} = \mathrm{const} > 0, \tag{3.40}$$

what proves our hypothesis on response of the state increment caused by the needle-shaped variation $\tilde{\mathbf{u}}(t) = \mathbf{u}_\varepsilon(t)$ given by (3.24):

$$\left\| \Delta_\varepsilon \mathbf{x}(t) \right\| \leq \mathcal{K} \cdot \varepsilon, \quad t \in T, \ \mathcal{K} = \mathrm{const} > 0. \tag{3.41}$$

This also implies that for $\tilde{\mathbf{u}}(t) = \mathbf{u}_\varepsilon(t)$,

$$\int_{\tau - \varepsilon}^{\tau} \left\langle \Delta_v \frac{\partial H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}}, \Delta_\varepsilon \mathbf{x}(t) \right\rangle dt + \eta_{u_\varepsilon} \left(\left\| \Delta_\varepsilon \mathbf{x} \right\| \right) \sim \varepsilon, \tag{3.42}$$

where

$$\Delta_\varepsilon \mathbf{x}(t) = \mathbf{x}(t, \mathbf{u}_\varepsilon) - \mathbf{x}(t, \mathbf{u}) \sim \varepsilon. \tag{3.43}$$

Therefore, the response of objective functional caused by the needle-shaped variation (3.24) can be represented according to (3.21) as

$$\Delta_\varepsilon J(\mathbf{u}) = J(\mathbf{u}_\varepsilon) - J(\mathbf{u}) = -\Delta_v H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, \tau) \cdot \varepsilon + o(\varepsilon), \quad \mathbf{v} \in U, \ \tau \in T. \tag{3.44}$$

It should be noted that in the last expression, we have made use of the mean value theorem.

For the needle-shaped variation of optimal process $\{\mathbf{u}^*, \mathbf{x}^* = \mathbf{x}(t, \mathbf{u}^*)\}$, the increment formula (3.44) with regard to the estimate (3.41) implies the necessary optimality condition in the form of the maximum principle.

THEOREM 3.6 (maximum principle). *Suppose that admissible process $\{\mathbf{u}^*, \mathbf{x}^*\}$ is optimal for problem (2.1)–(2.4), and that $\boldsymbol{\psi}^*$ is the solution to conjugate BVP (3.13),(3.19),*

(3.20) calculated on the optimal process. Then, for all $\tau \in T$, the following inequality holds:

$$\Delta_v H(\boldsymbol{\psi}^*, \mathbf{x}^*, \mathbf{u}^*, \tau) \leq 0, \quad \forall \mathbf{v} \in U. \tag{3.45}$$

REMARK 3.7. If $\boldsymbol{\Phi}, \mathbf{f}$ are linear and φ, F are convex in $\mathbf{x}(t_0)$, $\mathbf{x}(t_1)$, and $\mathbf{x}$, respectively, then the maximum principle (3.45) is both necessary and sufficient condition for optimality. This fact follows from the increment formula

$$\Delta J(\mathbf{u}) = -\int_T \Delta_{\tilde{u}} H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t) dt + o_\varphi (\|\Delta \mathbf{x}(t_0)\|, \|\Delta \mathbf{x}(t_1)\|)$$
$$+ \int_T o_F(\|\Delta \mathbf{x}(t)\|) dt, \tag{3.46}$$

where $o_{\varphi_0} \geq 0$, $o_F \geq 0$.

4. Second-order optimality condition. The significance of the maximum principle as a necessary condition of optimality for optimal control problems is commonly recognized and accepted. On the other hand, in certain cases, the maximum principle has a tendency to "degenerate," that is, to be fulfilled trivially on a series of admissible controllers, and therefore it cannot serve neither as a verifying condition for optimality nor as a basis for construction of optimal design algorithms. This situation, called in the theory of optimal control "singular," is not a rare exception. On the contrary, it is rather a regularity typical for dynamic processes with complicated structure.

We will recall that an admissible control $\mathbf{u} = \mathbf{u}(t)$, $t \in T$, is called *singular* within some range $\Omega \subseteq T$ of positive measure, that is, $\mathrm{mes}\,\Omega > 0$, if

$$\Delta_{\tilde{u}} H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t) \equiv 0 \tag{4.1}$$

at any $t \in \Omega$ and for all $\tilde{\mathbf{u}}(t) \in U$.

For example, if optimal control $\mathbf{u}^*(t)$ is singular within some range $\Omega \subseteq T$, $\mathrm{mes}\,\Omega > 0$, then function $\Delta_v H(\boldsymbol{\psi}^*, \mathbf{x}^*, \mathbf{u}^*, t)$ does not depend upon control variable $\mathbf{v}$ on the direct product $U \times \Omega$. Therefore, within Ω, the maximum principle (3.45) formulated in the form of Theorem 3.6 becomes completely useless. In other words, condition (4.1) conveys the degeneracy (or triviality) of the maximum principle within $\Omega \subseteq T$ and indicates the need of another optimality condition. One way to obtain such a condition is to try to extract the second-order term in the increment formula (3.44) corresponding to the needle-shaped variation of the optimal process. In order to use the Taylor series expansions up to second order, we have to assume that scalar functions $\varphi(\mathbf{x}(t_0), \mathbf{x}(t_1))$, $F(\mathbf{x}, \mathbf{u}, t)$ and vector functions $\boldsymbol{\Phi}(\mathbf{x}(t_0), \mathbf{x}(t_1))$, $\mathbf{f}(\mathbf{x}, \mathbf{u}, t)$ admit second-order partial derivatives with respect to $\mathbf{x}(t_0)$, $\mathbf{x}(t_1)$, and $\mathbf{x}$, respectively.

4.1. Second-order increment formula. As we did before, for two admissible processes, the basic one $\{\mathbf{u}, \mathbf{x} = \mathbf{x}(t, \mathbf{u})\}$ and the varied one $\{\tilde{\mathbf{u}} = \mathbf{u} + \Delta \mathbf{u}, \tilde{\mathbf{x}} = \mathbf{x} + \Delta \mathbf{x} = \mathbf{x}(t, \tilde{\mathbf{u}})\}$, we define the incremental BVP (3.1) and make use of the conjugate vector function $\boldsymbol{\psi} = \boldsymbol{\psi}(t)$ and the arbitrary numerical vector $\boldsymbol{\lambda} \in \mathbb{R}^n$ introduced in Section 3. To extract the second-order term in the increment formula (3.4), we will need to introduce additionally some differentiable matrix function $\boldsymbol{\Psi} = \boldsymbol{\Psi}(t)$, $\boldsymbol{\Psi}(t) \in \mathbb{R}^{n \times n}$ and two

arbitrary numerical $(n \times n)$ matrices Λ_0, Λ_1 undefined for now. Then the increment formula (3.4) may be also written as

$$\Delta J(\mathbf{u}) = \Delta \varphi\left(\mathbf{x}(t_0), \mathbf{x}(t_1)\right) + \int_T \Delta F(\mathbf{x}, \mathbf{u}, t)\, dt$$
$$+ \int_T \left\langle \boldsymbol{\psi}(t) + \frac{1}{2}\boldsymbol{\Psi}(t)\Delta \mathbf{x}(t), \Delta \dot{\mathbf{x}}(t) - \Delta \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \right\rangle dt \qquad (4.2)$$
$$+ \left\langle \boldsymbol{\lambda} + \frac{1}{2}\Lambda_0 \Delta \mathbf{x}(t_0) + \frac{1}{2}\Lambda_1 \Delta \mathbf{x}(t_1), \Delta \boldsymbol{\Phi}\left(\mathbf{x}(t_0), \mathbf{x}(t_1)\right) \right\rangle.$$

To proceed, we extract the second-order terms in the Taylor series expansions (3.6) and (3.7):

$$\Delta \varphi\left(\mathbf{x}(t_0), \mathbf{x}(t_1)\right) = \left\langle \frac{\partial \varphi}{\partial \mathbf{x}(t_0)}, \Delta \mathbf{x}(t_0) \right\rangle + \left\langle \frac{\partial \varphi}{\partial \mathbf{x}(t_1)}, \Delta \mathbf{x}(t_1) \right\rangle$$
$$+ \frac{1}{2}\left\langle \frac{\partial^2 \varphi}{\partial \mathbf{x}(t_0)^2}\Delta \mathbf{x}(t_0) + \frac{\partial^2 \varphi}{\partial \mathbf{x}(t_1)\partial \mathbf{x}(t_0)}\Delta \mathbf{x}(t_1), \Delta \mathbf{x}(t_0) \right\rangle$$
$$+ \frac{1}{2}\left\langle \frac{\partial^2 \varphi}{\partial \mathbf{x}(t_1)^2}\Delta \mathbf{x}(t_1) + \frac{\partial^2 \varphi}{\partial \mathbf{x}(t_0)\partial \mathbf{x}(t_1)}\Delta \mathbf{x}(t_0), \Delta \mathbf{x}(t_1) \right\rangle \qquad (4.3)$$
$$+ o_\varphi\left(\|\Delta \mathbf{x}(t_0)\|^2, \|\Delta \mathbf{x}(t_1)\|^2\right);$$

$$\Delta \Phi_i\left(\mathbf{x}(t_0), \mathbf{x}(t_1)\right) = \left\langle \frac{\partial \Phi_i}{\partial \mathbf{x}(t_0)}, \Delta \mathbf{x}(t_0) \right\rangle + \left\langle \frac{\partial \Phi_i}{\partial \mathbf{x}(t_1)}, \Delta \mathbf{x}(t_1) \right\rangle$$
$$+ \frac{1}{2}\left\langle \frac{\partial^2 \Phi_i}{\partial \mathbf{x}(t_0)^2}\Delta \mathbf{x}(t_0) + \frac{\partial^2 \Phi_i}{\partial \mathbf{x}(t_1)\partial \mathbf{x}(t_0)}\Delta \mathbf{x}(t_1), \Delta \mathbf{x}(t_0) \right\rangle$$
$$+ \frac{1}{2}\left\langle \frac{\partial^2 \Phi_i}{\partial \mathbf{x}(t_1)^2}\Delta \mathbf{x}(t_1) + \frac{\partial^2 \Phi_i}{\partial \mathbf{x}(t_0)\partial \mathbf{x}(t_1)}\Delta \mathbf{x}(t_0), \Delta \mathbf{x}(t_1) \right\rangle \qquad (4.4)$$
$$+ o_{\Phi_i}\left(\|\Delta \mathbf{x}(t_0)\|^2, \|\Delta \mathbf{x}(t_1)\|^2\right), \quad i = 1, \dots, n.$$

It should be noted that $\Delta \boldsymbol{\Phi} = (\Delta \Phi_1, \dots, \Delta \Phi_n)$ and that $\mathbf{o}_\Phi = (o_{\Phi_1}, \dots, o_{\Phi_n})$. In the partial increment of H given by formula (3.11), we will also extract the second-order term

$$\Delta_{\tilde{x}\tilde{u}}H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t) = \Delta_{\tilde{u}}H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t) + \Delta_{\tilde{x}}H(\boldsymbol{\psi}, \mathbf{x}, \tilde{\mathbf{u}}, t)$$
$$= \Delta_{\tilde{u}}H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t) + \left\langle \frac{\partial H(\boldsymbol{\psi}, \mathbf{x}, \tilde{\mathbf{u}}, t)}{\partial \mathbf{x}}, \Delta \mathbf{x}(t) \right\rangle \qquad (4.5)$$
$$+ \frac{1}{2}\left\langle \frac{\partial^2 H(\boldsymbol{\psi}, \mathbf{x}, \tilde{\mathbf{u}}, t)}{\partial \mathbf{x}^2}\Delta \mathbf{x}(t), \Delta \mathbf{x}(t) \right\rangle + o_H\left(\|\Delta \mathbf{x}(t)\|^2\right)$$

and represent each summand involving $\tilde{\mathbf{u}}$ as

$$\frac{\partial H(\boldsymbol{\psi}, \mathbf{x}, \tilde{\mathbf{u}}, t)}{\partial \mathbf{x}} = \frac{\partial H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}} + \Delta_{\tilde{u}}\frac{\partial H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}},$$
$$\frac{\partial^2 H(\boldsymbol{\psi}, \mathbf{x}, \tilde{\mathbf{u}}, t)}{\partial \mathbf{x}^2} = \frac{\partial^2 H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}^2} + \Delta_{\tilde{u}}\frac{\partial^2 H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}^2}. \qquad (4.6)$$

It will be helpful to make some transformations and expansions for the right-hand side of the incremental system $\Delta\dot{\mathbf{x}} = \Delta\mathbf{f}(\mathbf{x}, \mathbf{u}, t)$:

$$\Delta\mathbf{f}(\mathbf{x}, \mathbf{u}, t) = \Delta_{\tilde{x}}\mathbf{f}(\mathbf{x}, \tilde{\mathbf{u}}, t) + \Delta_{\tilde{u}}\mathbf{f}(\mathbf{x}, \mathbf{u}, t),$$

$$\Delta_{\tilde{x}}\mathbf{f}(\mathbf{x}, \tilde{\mathbf{u}}, t) = \frac{\partial\mathbf{f}(\mathbf{x}, \tilde{\mathbf{u}}, t)}{\partial\mathbf{x}}\Delta\mathbf{x}(t) + \mathbf{o}_f(\|\Delta\mathbf{x}(t)\|), \tag{4.7}$$

$$\frac{\partial\mathbf{f}(\mathbf{x}, \tilde{\mathbf{u}}, t)}{\partial\mathbf{x}} = \frac{\partial\mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial\mathbf{x}} + \Delta_{\tilde{u}}\frac{\partial\mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial\mathbf{x}}.$$

In addition to the integration by parts (3.8), we must perform another one:

$$\int_T \langle \mathbf{\Psi}(t)\Delta\mathbf{x}(t), \Delta\dot{\mathbf{x}}(t) \rangle dt = \langle \mathbf{\Psi}(t_1)\Delta\mathbf{x}(t_1), \Delta\mathbf{x}(t_1) \rangle - \langle \mathbf{\Psi}(t_0)\Delta\mathbf{x}(t_0), \Delta\mathbf{x}(t_0) \rangle$$
$$- \int_T \langle \dot{\mathbf{\Psi}}(t)\Delta\mathbf{x}(t), \Delta\mathbf{x}(t) \rangle dt - \int_T \langle \mathbf{\Psi}(t)\Delta\dot{\mathbf{x}}(t), \Delta\mathbf{x}(t) \rangle dt, \tag{4.8}$$

where

$$\int_T \langle \mathbf{\Psi}(t)\Delta\dot{\mathbf{x}}(t), \Delta\mathbf{x}(t) \rangle dt$$
$$= \int_T \langle \mathbf{\Psi}(t)\Delta\mathbf{f}(\mathbf{x}, \mathbf{u}, t), \Delta\mathbf{x}(t) \rangle dt$$
$$= \int_T \left\langle \mathbf{\Psi}(t)\left[\Delta_{\tilde{u}}\mathbf{f}(\mathbf{x}, \mathbf{u}, t) + \frac{\partial\mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial\mathbf{x}}\Delta\mathbf{x}(t) \right.\right. \tag{4.9}$$
$$\left.\left. + \Delta_{\tilde{u}}\frac{\partial\mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial\mathbf{x}}\Delta\mathbf{x}(t) + \mathbf{o}_f(\|\Delta\mathbf{x}(t)\|) \right], \Delta\mathbf{x}(t) \right\rangle dt.$$

The second entry of the right-hand side in the last formula can be also written as

$$\int_T \left\langle \mathbf{\Psi}(t)\frac{\partial\mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial\mathbf{x}}\Delta\mathbf{x}(t), \Delta\mathbf{x}(t) \right\rangle dt$$
$$= \frac{1}{2}\int_T \left[\left\langle \mathbf{\Psi}(t)\frac{\partial\mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial\mathbf{x}}\Delta\mathbf{x}(t), \Delta\mathbf{x}(t) \right\rangle \right. \tag{4.10}$$
$$\left. + \left\langle \left(\frac{\partial\mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial\mathbf{x}} \right)' \mathbf{\Psi}(t)'\Delta\mathbf{x}(t), \Delta\mathbf{x}(t) \right\rangle \right] dt.$$

Now we have to substitute all the expressions processed above into the increment formula (4.2). In the IP of (4.2), the conjugate vector function $\boldsymbol{\psi} = \boldsymbol{\psi}(t)$ is subject to system (3.13) as it was accomplished in the previous section. On the other hand, we must also make use of the arbitrary differentiable $(n \times n)$ matrix function $\mathbf{\Psi} = \mathbf{\Psi}(t)$ in order to reorganize and cancel some terms. This can be done by defining the *conjugate matrix system*:

$$\dot{\mathbf{\Psi}} = -\left(\frac{\partial\mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial\mathbf{x}} \right)' \mathbf{\Psi}' - \mathbf{\Psi}\frac{\partial\mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial\mathbf{x}} - \frac{\partial^2 H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t)}{\partial\mathbf{x}^2}. \tag{4.11}$$

REMARK 4.1. The idea to use the conjugate matrix function $\boldsymbol{\Psi}$ and the respective technique has originally emerged in [6] and then was described in detail in [7]. In the mentioned sources, the concept of $\boldsymbol{\Psi}$ was utilized in order to derive the necessary condition for optimality of singular controls that may appear in free endpoint problems of optimal control. It should be noted however that conjugate matrix function $\boldsymbol{\Psi}$ that appears in [6, 7] is *symmetric* by definition. The same cannot be affirmed for $\boldsymbol{\Psi}$ used in this note and defined by system (4.11).

In order to obtain the boundary conditions for both conjugate functions $\boldsymbol{\psi}$ and $\boldsymbol{\Psi}$, we will use a procedure similar to that of the previous section. Namely, we must substitute the expansions of $\Delta\boldsymbol{\Phi}(\mathbf{x}(t_0),\mathbf{x}(t_1))$ given by (4.4) into the last term of the increment formula (4.2), then rearrange all the terminal terms in the resultant expression and try to reorganize the TP of (4.2) by summands of the form

$$
\begin{aligned}
&\langle(\cdot),\Delta\mathbf{x}(t_0)\rangle, \quad &&\langle(\cdot),\Delta\mathbf{x}(t_1)\rangle, \\
&\langle[\cdot]\Delta\mathbf{x}(t_0),\Delta\mathbf{x}(t_0)\rangle, \quad &&\langle[\cdot]\Delta\mathbf{x}(t_1),\Delta\mathbf{x}(t_0)\rangle, \\
&\langle[\cdot]\Delta\mathbf{x}(t_0),\Delta\mathbf{x}(t_1)\rangle, \quad &&\langle[\cdot]\Delta\mathbf{x}(t_1),\Delta\mathbf{x}(t_1)\rangle.
\end{aligned}
\tag{4.12}
$$

After equating to zero the vector quantities corresponding to $(\cdot)$ in the upper row of (4.12), we can manage to eliminate the arbitrary nontrivial vector $\boldsymbol{\lambda}$ by choosing two $(n\times n)$ numerical matrices $\mathbf{B}_0$, $\mathbf{B}_1$ according to the relationship (3.19). This was accomplished in the previous section and yielded to the boundary conditions (3.20). In a similar manner, we can equate to zero the matrix quantities corresponding to $[\cdot]$ in the two lower rows of (4.12) and then make use of the same matrices $\mathbf{B}_0$, $\mathbf{B}_1$ in order to try to eliminate the undefined parameters $\boldsymbol{\lambda}, \Lambda_0, \Lambda_1$. This routine involves very extensive intermediate computations; therefore, we refrain from reproducing them here due to space limitations. Unfortunately, by using such a technique, we only can eliminate the arbitrary matrices Λ_0 and Λ_1 and the arbitrary vector $\boldsymbol{\lambda}$ will be still present in the final form of the boundary conditions for the matrix conjugate system (4.11):

$$
\begin{aligned}
&\mathbf{B}_0\left\{\frac{\partial^2\varphi}{\partial\mathbf{x}(t_0)^2}+\frac{\partial^2\varphi}{\partial\mathbf{x}(t_1)\partial\mathbf{x}(t_0)}+\left[\frac{\partial^2\boldsymbol{\Phi}}{\partial\mathbf{x}(t_0)^2}\boldsymbol{\lambda}\right]'+\left[\frac{\partial^2\boldsymbol{\Phi}}{\partial\mathbf{x}(t_1)\partial\mathbf{x}(t_0)}\boldsymbol{\lambda}\right]'\right\} \\
&\quad+\mathbf{B}_1\left\{\frac{\partial^2\varphi}{\partial\mathbf{x}(t_1)^2}+\frac{\partial^2\varphi}{\partial\mathbf{x}(t_0)\partial\mathbf{x}(t_1)}+\left[\frac{\partial^2\boldsymbol{\Phi}}{\partial\mathbf{x}(t_1)^2}\boldsymbol{\lambda}\right]'+\left[\frac{\partial^2\boldsymbol{\Phi}}{\partial\mathbf{x}(t_0)\partial\mathbf{x}(t_1)}\boldsymbol{\lambda}\right]'\right\} \\
&\qquad\qquad\qquad-\mathbf{B}_0\boldsymbol{\Psi}(t_0)+\mathbf{B}_1\boldsymbol{\Psi}(t_1)=\mathbf{0}.
\end{aligned}
\tag{4.13}
$$

It should be clarified that

$$
\left[\frac{\partial^2\boldsymbol{\Phi}}{\partial\mathbf{x}(t_0)^2}\boldsymbol{\lambda}\right]', \left[\frac{\partial^2\boldsymbol{\Phi}}{\partial\mathbf{x}(t_1)\partial\mathbf{x}(t_0)}\boldsymbol{\lambda}\right]', \left[\frac{\partial^2\boldsymbol{\Phi}}{\partial\mathbf{x}(t_1)^2}\boldsymbol{\lambda}\right]', \left[\frac{\partial^2\boldsymbol{\Phi}}{\partial\mathbf{x}(t_0)\partial\mathbf{x}(t_1)}\boldsymbol{\lambda}\right]'
\tag{4.14}
$$

are the matrices of dimension $(n \times n)$ calculated according to the rules

$$
\begin{aligned}
\left[\frac{\partial^2 \mathbf{\Phi}}{\partial \mathbf{x}(t_0)^2} \boldsymbol{\lambda}\right]' &\stackrel{\text{def}}{=} \frac{\partial}{\partial \mathbf{x}(t_0)}\left[\left(\frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}(t_0)}\right)' \boldsymbol{\lambda}\right], \\
\left[\frac{\partial^2 \mathbf{\Phi}}{\partial \mathbf{x}(t_1)^2} \boldsymbol{\lambda}\right] &\stackrel{\text{def}}{=} \frac{\partial}{\partial \mathbf{x}(t_1)}\left[\left(\frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}(t_1)}\right)' \boldsymbol{\lambda}\right], \\
\left[\frac{\partial^2 \mathbf{\Phi}}{\partial \mathbf{x}(t_1)\partial \mathbf{x}(t_0)} \boldsymbol{\lambda}\right] &\stackrel{\text{def}}{=} \frac{\partial}{\partial \mathbf{x}(t_1)}\left[\left(\frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}(t_0)}\right)' \boldsymbol{\lambda}\right], \\
\left[\frac{\partial^2 \mathbf{\Phi}}{\partial \mathbf{x}(t_0)\partial \mathbf{x}(t_1)} \boldsymbol{\lambda}\right] &\stackrel{\text{def}}{=} \frac{\partial}{\partial \mathbf{x}(t_0)}\left[\left(\frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}(t_1)}\right)' \boldsymbol{\lambda}\right].
\end{aligned}
\tag{4.15}
$$

It also should be noted that instead of writing $\mathbf{\Psi}(t)$ for a solution of conjugate matrix BVP (4.11), (4.13), we will write $\mathbf{\Psi}(\boldsymbol{\lambda}, t)$ since each solution of (4.11), (4.13), that is, solutions obtained for diverse admissible choices of $\mathbf{B}_0$, $\mathbf{B}_1$, will depend on $\boldsymbol{\lambda}$ via the boundary condition (4.13) in the general case, although we must point out two particular cases where such a dependence can be avoided.

REMARK 4.2. If instead of boundary conditions (2.3) for the original system (2.2), we have the initial conditions, that is, $\mathbf{\Phi} = \mathbf{\Phi}(\mathbf{x}(t_0))$, then for $\mathbf{B}_0 = \mathbf{0}$, $\mathbf{B}_1 = \mathbf{I}$, the boundary condition for conjugate matrix function $\mathbf{\Psi}$ (4.13) turns into the terminal value condition

$$
\mathbf{\Psi}(t_1) = -\frac{\partial^2 \varphi}{\partial \mathbf{x}(t_1)^2}
\tag{4.16}
$$

that was defined for symmetric $\mathbf{\Psi}$ in [6, 7].

REMARK 4.3. In case of linear boundary conditions (2.3), that is, $\mathbf{\Phi}(\mathbf{x}(t_0), \mathbf{x}(t_1)) = \mathbf{D}_0 \mathbf{x}(t_0) + \mathbf{D}_1 \mathbf{x}(t_1) - \mathbf{d} = \mathbf{0}$, $\text{rank}[\mathbf{D}_0 \mathbf{D}_1] = n$, $\mathbf{D}_0, \mathbf{D}_1 \in \mathbb{R}^{n \times n}$, $\mathbf{d} \in \mathbb{R}^n$, we do not need to make use of Taylor series expansions (4.4) since the boundary conditions for incremental BVP (3.1) are also linear:

$$
\Delta\mathbf{\Phi}(\mathbf{x}(t_0), \mathbf{x}(t_1)) = \mathbf{D}_0 \Delta\mathbf{x}(t_0) + \mathbf{D}_1 \Delta\mathbf{x}(t_1) = \mathbf{0}
\tag{4.17}
$$

and all second partial derivatives of $\mathbf{\Phi}$ are equal to zero. This implies that the boundary conditions (4.13) for conjugate matrix system (4.11) do not involve arbitrary parameter $\boldsymbol{\lambda}$:

$$
\begin{aligned}
-\mathbf{B}_0 \mathbf{\Psi}(t_0) &+ \mathbf{B}_1 \mathbf{\Psi}(t_1) + \mathbf{B}_0 \frac{\partial^2 \varphi}{\partial \mathbf{x}(t_0)^2} + \mathbf{B}_1 \frac{\partial^2 \varphi}{\partial \mathbf{x}(t_1)^2} \\
&+ \mathbf{B}_0 \frac{\partial^2 \varphi}{\partial \mathbf{x}(t_1)\partial \mathbf{x}(t_0)} + \mathbf{B}_1 \frac{\partial^2 \varphi}{\partial \mathbf{x}(t_0)\partial \mathbf{x}(t_1)} = \mathbf{0}.
\end{aligned}
\tag{4.18}
$$

More details about this particular case can be found in [15].

Thus, having defined two conjugate BVPs (3.13), (3.19), (3.20) and (4.11), (4.13), we can rewrite the second-order increment formula (4.2) in the following form:

$$
\begin{aligned}
\Delta J(\mathbf{u}) = & -\int_T \Delta_{\tilde{u}} H(\boldsymbol{\psi},\mathbf{x},\mathbf{u},t)\,dt \\
& -\int_T \left\langle \Delta_{\tilde{u}} \frac{\partial H(\boldsymbol{\psi},\mathbf{x},\mathbf{u},t)}{\partial \mathbf{x}} + \boldsymbol{\Psi}(\boldsymbol{\lambda},t)\Delta_{\tilde{u}}\mathbf{f}(\mathbf{x},\mathbf{u},t), \Delta\mathbf{x}(t) \right\rangle dt + \eta_1 - \eta_2,
\end{aligned}
\tag{4.19}
$$

where the remainders η_1 and η_2 are given by

$$
\begin{aligned}
\eta_1 = & -\int_T o_H\left(\|\Delta\mathbf{x}(t)\|^2\right)dt - \int_T \left\langle \boldsymbol{\Psi}(\boldsymbol{\lambda},t)\Delta\mathbf{x}(t), \mathbf{o}_f(\|\Delta\mathbf{x}(t)\|) \right\rangle dt \\
& + \left\langle \boldsymbol{\lambda} + \frac{1}{2}\boldsymbol{\Lambda}_0\Delta\mathbf{x}(t_0) + \frac{1}{2}\boldsymbol{\Lambda}_1\Delta\mathbf{x}(t_1), \mathbf{o}_\Phi\left(\|\Delta\mathbf{x}(t_0)\|^2, \|\Delta\mathbf{x}(t_1)\|^2\right) \right\rangle \\
& + o_\varphi\left(\|\Delta\mathbf{x}(t_0)\|^2, \|\Delta\mathbf{x}(t_1)\|^2\right), \\
\eta_2 = & \frac{1}{2}\int_T \left\langle \Delta_{\tilde{u}}\left[\frac{\partial^2 H(\boldsymbol{\psi},\mathbf{x},\mathbf{u},t)}{\partial \mathbf{x}^2} + 2\boldsymbol{\Psi}(\boldsymbol{\lambda},t)\frac{\partial \mathbf{f}(\mathbf{x},\mathbf{u},t)}{\partial \mathbf{x}}\right]\Delta\mathbf{x}(t), \Delta\mathbf{x}(t) \right\rangle dt.
\end{aligned}
\tag{4.20}
$$

4.2. Needle-shaped variation. At first, we recall the state increment estimate (3.40) that was obtained in Section 3. Now if we replace an arbitrary admissible control $\tilde{\mathbf{u}}(t) \in U$ by the needle-shaped variation (3.24), then by virtue of estimate (3.41), the increment of state $\Delta_\varepsilon\mathbf{x}(t)$ caused by the needle-shaped variation (3.24) will be of order ε and the corresponding increment of the objective functional will be represented according to (4.19) as

$$
\begin{aligned}
\Delta_\varepsilon J(\mathbf{u}) = & -\int_{\tau-\varepsilon}^{\tau} \Delta_v H(\boldsymbol{\psi},\mathbf{x},\mathbf{u},t)\,dt \\
& -\int_{\tau-\varepsilon}^{\tau} \left\langle \Delta_v \frac{\partial H(\boldsymbol{\psi},\mathbf{x},\mathbf{u},t)}{\partial \mathbf{x}} + \boldsymbol{\Psi}(\boldsymbol{\lambda},t)\Delta_v\mathbf{f}(\mathbf{x},\mathbf{u},t), \Delta_\varepsilon\mathbf{x}(t) \right\rangle dt + o(\varepsilon^2),
\end{aligned}
\tag{4.21}
$$

where $o(\varepsilon^2)/\varepsilon^2 \to 0$, $\varepsilon \to 0$.

REMARK 4.4. Generally speaking, the necessary condition for optimality in the form of the maximum principle (3.45) also results from formula (4.21). Factually, on the needle-shaped variation of the optimal process $\{\mathbf{u}^*, \mathbf{x}^*\}$, formula (4.21) does not contradict (3.44) that yields the maximum principle.

On the other hand, if optimal control $\mathbf{u}^*$ is singular within some range $\Omega \subseteq T$, the dominant first-order term in (3.44) vanishes, that is,

$$
\Delta_v H(\boldsymbol{\psi}^*,\mathbf{x}^*,\mathbf{u}^*,t) \equiv 0, \quad \forall \mathbf{v} \in U,\ t \in \Omega,
\tag{4.22}
$$

and the maximum principle degenerates. Thus, in order to deal with singular controls, we ought to extract an explicit coefficient of ε^2 in second-order variational increment formula (4.21). To cope with this task, it will be sufficient to extract a coefficient of ε in the correspondent state increment $\Delta_\varepsilon\mathbf{x}(\tau)$.

One way to approach this task is to examine the incremental BVP (3.1) on the needle-shaped variation (3.24). It is clear that

$$
\begin{aligned}
\Delta \dot{\mathbf{x}} &= \Delta_{\tilde{x}} \mathbf{f}(\mathbf{x}, \tilde{\mathbf{u}}, t) + \Delta_{\tilde{u}} \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \\
&= \left[\frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}} + \Delta_{\tilde{u}} \mathbf{f}(\mathbf{x}, \mathbf{u}, t) + \Delta_{\tilde{u}} \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}} \right] \Delta \mathbf{x}(t) + \mathbf{o}_f(\|\Delta \mathbf{x}(t)\|)
\end{aligned}
\tag{4.23}
$$

or in the integral form,

$$
\begin{aligned}
\Delta \mathbf{x}(t) = \Delta \mathbf{x}(t_0) + \int_{t_0}^{t} \Bigg[&\frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u}, \xi)}{\partial \mathbf{x}} \Delta \mathbf{x}(\xi) + \Delta_{\tilde{u}} \mathbf{f}(\mathbf{x}, \mathbf{u}, \xi) \\
&+ \Delta_{\tilde{u}} \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}} + \mathbf{o}_f(\|\Delta \mathbf{x}(t)\|) \Bigg] d\xi.
\end{aligned}
\tag{4.24}
$$

The latter on the needle-shaped variation (3.24) turns into

$$
\Delta_\varepsilon \mathbf{x}(t) = \Delta_\varepsilon \mathbf{x}(t_0) + \int_{t_0}^{t} \left[\frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u}, \xi)}{\partial \mathbf{x}} \Delta_\varepsilon \mathbf{x}(\xi) + \Delta_v \mathbf{f}(\mathbf{x}, \mathbf{u}, \xi) \right] d\xi + \mathbf{o}(\varepsilon).
\tag{4.25}
$$

Thus, the BVP in increments corresponding to the needle-shaped variation will be given by

$$
\Delta_\varepsilon \dot{\mathbf{x}} = \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}} \Delta_\varepsilon \mathbf{x} + \Delta_v \mathbf{f}(\mathbf{x}, \mathbf{u}, t) + \mathbf{o}(\varepsilon),
\tag{4.26}
$$

$$
\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_0)} \Delta_\varepsilon \mathbf{x}(t_0) + \frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_1)} \Delta_\varepsilon \mathbf{x}(t_1) + \mathbf{o}(\varepsilon) = \mathbf{0}.
\tag{4.27}
$$

We can represent the solution of *linear* BVP (4.26), (4.27) using the Cauchy formula analogue (3.36). Let $\mathbf{X} = \mathbf{X}(t)$ be the fundamental $(n \times n)$ matrix function of the homogeneous system (4.26):

$$
\dot{\mathbf{X}} = \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}} \mathbf{X}, \qquad \mathbf{X}(t_0) = \mathbf{I}.
\tag{4.28}
$$

Suppose that BVP (4.26), (4.27) has unique solution for any admissible process $\{\mathbf{u}, \mathbf{x}\}$; in other words, suppose that resolvability condition (3.38) holds, that is,

$$
\det \left[\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_0)} + \frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_1)} \mathbf{X}(t_1) \right] \neq 0.
\tag{4.29}
$$

Then

$$
\begin{aligned}
\Delta_\varepsilon \mathbf{x}(t) = &-\int_{\tau-\varepsilon}^{\tau} \mathbf{X}(t) \mathbf{P}(t_1) \mathbf{X}^{-1}(\xi) \Delta_v \mathbf{f}(\mathbf{x}, \mathbf{u}, \xi) d\xi \\
&+ \int_{t_0}^{t} \mathbf{X}(t) \mathbf{X}^{-1}(\xi) \begin{cases} 0, & t \in [t_0, \tau-\varepsilon] \\ \Delta_v \mathbf{f}(\mathbf{x}, \mathbf{u}, \xi), & t \in [\tau-\varepsilon, \tau) \\ 0, & t \in [\tau, t_1] \end{cases} d\xi + \mathbf{o}(\varepsilon),
\end{aligned}
\tag{4.30}
$$

where $\mathbf{P}(t_1)$ is $(n \times n)$ numerical matrix given by

$$
\mathbf{P}(t_1) = \left[\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_0)} + \frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_1)} \mathbf{X}(t_1) \right]^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_1)} \mathbf{X}(t_1).
\tag{4.31}
$$

Whence it follows that

$$\Delta_\varepsilon \mathbf{x}(t) = -\mathbf{X}(t)\mathbf{P}(t_1)\mathbf{X}^{-1}(\tau)\Delta_v\mathbf{f}(\mathbf{x},\mathbf{u},\tau)\cdot\varepsilon$$

$$+ \begin{cases} 0, & t \in [t_0, \tau-\varepsilon] \\ \mathbf{X}(t)\mathbf{X}^{-1}(\tau)\Delta_v\mathbf{f}(\mathbf{x},\mathbf{u},\tau)(t-\tau+\varepsilon), & t \in [\tau-\varepsilon, \tau) \\ \mathbf{X}(t)\mathbf{X}^{-1}(\tau)\Delta_v\mathbf{f}(\mathbf{x},\mathbf{u},\tau)\cdot\varepsilon, & t \in [\tau, t_1] \end{cases} + \mathbf{o}(\varepsilon) \tag{4.32}$$

and therefore

$$\Delta_\varepsilon \mathbf{x}(\tau) = \big[\mathbf{I} - \mathbf{X}(\tau)\mathbf{P}(t_1)\mathbf{X}^{-1}(\tau)\big]\Delta_v\mathbf{f}(\mathbf{x},\mathbf{u},\tau)\cdot\varepsilon + \mathbf{o}(\varepsilon). \tag{4.33}$$

It should be noted that inverse matrix $\mathbf{X}^{-1}(t)$ satisfies the following initial value problem:

$$\dot{\mathbf{X}}^{-1} = -\mathbf{X}^{-1}\cdot\frac{\partial\mathbf{f}(\mathbf{x},\mathbf{u},t)}{\partial\mathbf{x}}, \qquad \mathbf{X}^{-1}(t_0) = \mathbf{I}. \tag{4.34}$$

Apparently,

$$\mathbf{X}^{-1}(t)\mathbf{X}(t) = \mathbf{I}, \qquad \dot{\mathbf{X}}^{-1}(t)\mathbf{X}(t) + \mathbf{X}^{-1}(t)\dot{\mathbf{X}}(t) = \mathbf{0} \tag{4.35}$$

and with regard to (4.28),

$$\left[\dot{\mathbf{X}}^{-1}(t) + \mathbf{X}^{-1}(t)\frac{\partial\mathbf{f}(\mathbf{x},\mathbf{u},t)}{\partial\mathbf{x}}\right]\mathbf{X}(t) = \mathbf{0} \tag{4.36}$$

so that (4.34) becomes obvious.

Now we will introduce a differentiable $(n \times n)$ matrix function $\mathbf{Y} = \mathbf{Y}(t)$ according to

$$\mathbf{Y}(t) = \mathbf{X}(t)\mathbf{P}(t_1)\mathbf{X}^{-1}(t), \tag{4.37}$$

which for $t = t_0$ yields $\mathbf{Y}(t_0) = \mathbf{X}(t_0)\mathbf{P}(t_1)\mathbf{X}^{-1}(t_0) = \mathbf{P}(t_1)$. Then with regard to (4.28) and (4.34), we can write

$$\dot{\mathbf{Y}}(t) = \dot{\mathbf{X}}(t)\mathbf{P}(t_1)\mathbf{X}^{-1}(t) + \mathbf{X}(t)\mathbf{P}(t_1)\dot{\mathbf{X}}^{-1}(t)$$

$$= \frac{\partial\mathbf{f}(\mathbf{x},\mathbf{u},t)}{\partial\mathbf{x}}\mathbf{X}(t)\mathbf{P}(t_1)\mathbf{X}^{-1}(t) - \mathbf{X}(t)\mathbf{P}(t_1)\mathbf{X}^{-1}(t)\frac{\partial\mathbf{f}(\mathbf{x},\mathbf{u},t)}{\partial\mathbf{x}} \tag{4.38}$$

$$= \frac{\partial\mathbf{f}(\mathbf{x},\mathbf{u},t)}{\partial\mathbf{x}}\mathbf{Y}(t) - \mathbf{Y}(t)\frac{\partial\mathbf{f}(\mathbf{x},\mathbf{u},t)}{\partial\mathbf{x}}.$$

Using $\mathbf{Y}$, we can rewrite (4.33) as

$$\Delta_\varepsilon\mathbf{x}(\tau) = \big[\mathbf{I} - \mathbf{Y}(\tau)\big]\Delta_v\mathbf{f}(\mathbf{x},\mathbf{u},\tau)\varepsilon + \mathbf{o}(\varepsilon), \tag{4.39}$$

where

$$\dot{\mathbf{Y}} = \frac{\partial\mathbf{f}(\mathbf{x},\mathbf{u},t)}{\partial\mathbf{x}}\mathbf{Y} - \mathbf{Y}\frac{\partial\mathbf{f}(\mathbf{x},\mathbf{u},t)}{\partial\mathbf{x}}, \tag{4.40}$$

$$\mathbf{Y}(t_0) = \left[\frac{\partial\mathbf{\Phi}}{\partial\mathbf{x}(t_0)} + \frac{\partial\mathbf{\Phi}}{\partial\mathbf{x}(t_1)}\mathbf{X}(t_1)\right]^{-1}\frac{\partial\mathbf{\Phi}}{\partial\mathbf{x}(t_1)}\mathbf{X}(t_1). \tag{4.41}$$

REMARK 4.5. It should be noted that the fulfillment of the resolvability condition (4.29) for linear BVP (4.26)-(4.27) will also provide the uniqueness of its solution. The same condition may also stipulate the existence of (at least) approximate solution of the nonlinear BVP (2.2)-(2.3) but cannot provide its uniqueness.

REMARK 4.6. It may occur that for some admissible process $\{\mathbf{u}, \mathbf{x}(t, \mathbf{u})\}$,

$$\frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}(t_0)} = \mathbf{0}, \qquad \frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}(t_1)} = \mathbf{0}. \tag{4.42}$$

In this case, instead of boundary condition (4.27) for linearized system (4.26), we must assign the initial condition

$$\Delta_\varepsilon \mathbf{x}(t_0) = \mathbf{0} \tag{4.43}$$

so that to fulfill the original boundary condition (4.27) at least within ε and to prevent the failure of the resolvability condition (4.29). The final result in this particular case will be $\mathbf{Y}(t) \equiv \mathbf{0}$ and

$$\Delta_\varepsilon \mathbf{x}(\tau) = \Delta_v \mathbf{f}(\mathbf{x}, \mathbf{u}, \tau) \cdot \varepsilon + \mathbf{o}(\varepsilon). \tag{4.44}$$

Thus, in order to calculate the coefficient of $\Delta_v \mathbf{f}(\mathbf{x}, \mathbf{u}, \tau)\varepsilon$ in the expression of the state increment $\Delta_\varepsilon \mathbf{x}(\tau)$, $\tau \in T$, we should

(a) determine the fundamental matrix of solutions $\mathbf{X} = \mathbf{X}(t)$ of the homogeneous system (4.28);

(b) calculate the numerical matrix $\mathbf{Y}(t_0)$ by formula (4.41);

(c) solve the auxiliary initial value problem (4.40)-(4.41).

4.3. Extension of the maximum principle. Having calculated the explicit coefficient of ε in the correspondent state increment $\Delta_\varepsilon \mathbf{x}(\tau)$ according to (4.39), we can represent the response of objective functional to the needle-shaped variation of control function by the following formula:

$$
\begin{aligned}
\Delta_\varepsilon J(\mathbf{u}) = &-\int_{\tau-\varepsilon}^{\tau} \Delta_v H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, t)\,dt \\
&- \left\langle \Delta_v \frac{\partial H(\boldsymbol{\psi}, \mathbf{x}, \mathbf{u}, \tau)}{\partial \mathbf{x}} + \mathbf{\Psi}(\boldsymbol{\lambda}, \tau)\Delta_v \mathbf{f}(\mathbf{x}, \mathbf{u}, \tau), [\mathbf{I} - \mathbf{Y}(\tau)]\Delta_v \mathbf{f}(\mathbf{x}, \mathbf{u}, \tau) \right\rangle \cdot \varepsilon^2 + o(\varepsilon^2).
\end{aligned}
\tag{4.45}
$$

Suppose that in the original problem (2.1)–(2.4), the optimal control $\mathbf{u}^* = \mathbf{u}^*(t)$ is singular within some range $\Omega \subseteq T$, $\operatorname{mes}\Omega > 0$. That means that $\Delta J(\mathbf{u}^*) \geq 0$ for all $\mathbf{v} \in U$ and that the first term in (4.45) vanishes. In this case, the second term in (4.45) must be nonpositive for all $\boldsymbol{\lambda} \in \mathbb{R}^n$, $\boldsymbol{\lambda} \neq \mathbf{0}$, and for the corresponding $\boldsymbol{\psi}^*$, $\mathbf{\Psi}^*$, $\mathbf{Y}^*$ associated with the optimal process $\{\mathbf{u}^*, \mathbf{x}^*\}$, that is,

$$\left\langle \Delta_v \frac{\partial H(\boldsymbol{\psi}^*, \mathbf{x}^*, \mathbf{u}^*, \tau)}{\partial \mathbf{x}} + \mathbf{\Psi}^*(\boldsymbol{\lambda}, \tau)\Delta_v \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*, \tau), [\mathbf{I} - \mathbf{Y}^*(\tau)]\Delta_v \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*, \tau) \right\rangle \leq 0. \tag{4.46}$$

The foregoing deductions result in the necessary condition for optimality of singular controls.

THEOREM 4.7 (extension of the maximum principle). *Suppose that $\mathbf{u}^* = \mathbf{u}^*(t)$ is optimal in the original problem (2.1)–(2.4) and singular within some range $\Omega \subseteq T$, $\mathrm{mes}\,\Omega > 0$. Then it is necessary that two conditions be fulfilled:*

(1) *the maximum principle (first-order condition) with respect to $\mathbf{u}^*$ for the Hamiltonian function*

$$\Delta_v H(\boldsymbol{\psi}^*, \mathbf{x}^*, \mathbf{u}^*, t) \le 0, \quad \forall \mathbf{v} \in U \tag{4.47}$$

almost everywhere on $T \setminus \Omega$;

(2) *the necessary condition for optimality of singular controls (second-order condition) in the form of inequality*

$$\left\langle \Delta_v \frac{\partial H(\boldsymbol{\psi}^*, \mathbf{x}^*, \mathbf{u}^*, t)}{\partial \mathbf{x}} + \boldsymbol{\Psi}^*(\boldsymbol{\lambda}, t) \Delta_v \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*, t), [\mathbf{I} - \mathbf{Y}^*(t)] \Delta_v \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*, t) \right\rangle \le 0$$
$$\forall \mathbf{v} \in U, \ \forall \boldsymbol{\lambda} \in \mathbb{R}^n, \ \boldsymbol{\lambda} \ne \mathbf{0}, \tag{4.48}$$

almost everywhere on Ω for the correspondent solutions $\boldsymbol{\psi}^*$, $\boldsymbol{\Psi}^*$ to conjugate BVPs (3.13), (3.19), (3.20) and (4.11), (4.13), as well as for $\mathbf{X}^*$, $\mathbf{Y}^*$ defined by the auxiliary initial value problems (4.28) and (4.40)-(4.41), respectively.

It should be emphasized that the second condition of Theorem 4.7 must be fulfilled for all nontrivial $\boldsymbol{\lambda} \in \mathbb{R}^n$. In other words, if there exists even one single $\boldsymbol{\lambda} \ne \mathbf{0}$ altering the sign of (4.48), then the control function which is being tested does not definitely satisfy the necessary condition for optimality of singular controls. Obviously, the presence of such a $\boldsymbol{\lambda}$ in (4.13) and (4.48) is quite an obstacle here. Nevertheless, the necessary condition for optimality of singular controls can be sometimes useful in practice. We will illustrate it by means of the following example.

EXAMPLE 4.8. Consider a nonlinear bidimensional dynamic system with nonlinear boundary conditions and with a scalar control function defined on $T = [0, 1]$:

$$\begin{aligned}
\dot{x}_1 &= u, & \Phi_1 &= [x_2(0) - 1]^2 + x_1(1) = 0, \\
\dot{x}_2 &= -x_1^2, & \Phi_2 &= x_1^2(0) + x_2(1) - 1 = 0.
\end{aligned} \tag{4.49}$$

The objective functional (2.1) and the conventional inclusion constraint (2.4) will be

$$J(u) = x_2(1) \longrightarrow \min, \qquad |u(t)| \le 1. \tag{4.50}$$

It should be noted that the resolvability condition (4.29) holds for system (4.49) and it is easy to find that

$$\det\left[\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_0)} + \frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}(t_1)} \mathbf{X}(t_1) \right] = 1. \tag{4.51}$$

We check if the candidate control $u^*(t) \equiv 0$, $t \in [0, 1]$, is singular within some range $\Omega \subseteq [0, 1]$. This control produces

$$\begin{aligned}
\dot{x}_1 &= 0, & \mathbf{x}^*(t) &= \begin{cases} x_1^*(t) = 0, \\ x_2^*(t) = 1. \end{cases}
\end{aligned} \tag{4.52}$$

It is clear that $H(\boldsymbol{\psi}, \mathbf{x}, u, t) = u \cdot \psi_1 - x_1^2 \cdot \psi_2$ so that the conjugate system (3.13) is

$$\dot{\boldsymbol{\psi}} = \begin{pmatrix} \dot{\psi}_1 \\ \dot{\psi}_2 \end{pmatrix}, \quad \dot{\psi}_1 = 2x_1 \psi_2, \ \dot{\psi}_2 = 0. \tag{4.53}$$

In order to define the boundary conditions of the form (3.20) for the conjugate vector system (4.53), we must calculate

$$\frac{\partial \varphi}{\partial \mathbf{x}(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \frac{\partial \varphi}{\partial \mathbf{x}(1)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.54}$$

and choose two (2×2) matrices $\mathbf{B}_0$, $\mathbf{B}_1$ satisfying the relationship (3.19). Let

$$\mathbf{B}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{B}_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathrm{rank}\,[\mathbf{B}_0\mathbf{B}_1] = 2 \tag{4.55}$$

so that

$$-\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \boldsymbol{\psi}(0) + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{\psi}(1) + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \mathbf{0} \tag{4.56}$$

or

$$\psi_1(0) = 0, \qquad \psi_2(1) = -1. \tag{4.57}$$

It is obvious that the solution $\boldsymbol{\psi}(t)$ to conjugate vector BVP (4.53), (4.57) corresponding to $\{u^*, \mathbf{x}^*\}$ will be

$$\boldsymbol{\psi}(t) = \begin{cases} \psi_1^*(t) = 0, \\ \psi_2^*(t) = -1, \end{cases} \quad t \in [0,1]. \tag{4.58}$$

Since $\psi_1^*(t) \equiv 0$, then the partial increment of H with respect to any admissible control is identically equal to zero for all $t \in [0,1]$:

$$\Delta_v H(\boldsymbol{\psi}^*, \mathbf{x}^*, u^*, t) = \psi_1^*(v - u^*(t)) \equiv 0, \quad \forall v(t) : |v(t)| \le 1, \ t \in [0,1]. \tag{4.59}$$

Therefore, the candidate control $u^*(t) \equiv 0$ is *singular* on the entire $T = [0,1]$ and $\Omega = [0,1]$.

In order to define the conjugate matrix system (4.11), we must calculate $\partial \mathbf{f}/\partial \mathbf{x}$ and $\partial^2 H/\partial \mathbf{x}^2$ corresponding to $\{u^*, \mathbf{x}^*\}$:

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} 0 & 0 \\ -2x_1 & 0 \end{bmatrix}, \qquad \frac{\partial^2 H}{\partial \mathbf{x}^2} = \begin{bmatrix} -2\psi_2 & 0 \\ 0 & 0 \end{bmatrix}. \tag{4.60}$$

By replacing $x_1 = x_1^*(t) = 0$ and $\psi_2 = \psi_2^*(t) = -1$, we arrive at

$$\begin{bmatrix} \dot{\Psi}_{11} & \dot{\Psi}_{12} \\ \dot{\Psi}_{21} & \dot{\Psi}_{22} \end{bmatrix} = \dot{\boldsymbol{\Psi}} = -\frac{\partial^2 H(\boldsymbol{\psi}^*, \mathbf{x}^*, u^*, t)}{\partial \mathbf{x}^2} = \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix}. \tag{4.61}$$

In order to determine the boundary conditions of the form (4.13) for conjugate matrix function $\mathbf{\Psi}$, we must calculate first

$$\frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}(0)} = \begin{bmatrix} 0 & 2[x_2(0)-1] \\ 2x_1(0) & 0 \end{bmatrix}, \qquad \frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{4.62}$$

then take some arbitrary vector $\boldsymbol{\lambda} \neq \mathbf{0}$, $\boldsymbol{\lambda} \in \mathbb{R}^2$, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)'$, and calculate

$$\begin{aligned}
\left[\frac{\partial^2 \mathbf{\Phi}}{\partial \mathbf{x}(t_0)^2} \boldsymbol{\lambda} \right]' &= \frac{\partial}{\partial \mathbf{x}(t_0)} \left[\left(\frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}(t_0)} \right)' \boldsymbol{\lambda} \right] = \begin{bmatrix} 2\lambda_2 & 0 \\ 0 & 2\lambda_1 \end{bmatrix}, \\
\left[\frac{\partial^2 \mathbf{\Phi}}{\partial \mathbf{x}(t_1)^2} \boldsymbol{\lambda} \right]' &= \frac{\partial}{\partial \mathbf{x}(t_1)} \left[\left(\frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}(t_1)} \right)' \boldsymbol{\lambda} \right] = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \\
\left[\frac{\partial^2 \mathbf{\Phi}}{\partial \mathbf{x}(t_1) \partial \mathbf{x}(t_0)} \boldsymbol{\lambda} \right]' &= \frac{\partial}{\partial \mathbf{x}(t_1)} \left[\left(\frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}(t_0)} \right)' \boldsymbol{\lambda} \right]' = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \\
\left[\frac{\partial^2 \mathbf{\Phi}}{\partial \mathbf{x}(t_0) \partial \mathbf{x}(t_1)} \boldsymbol{\lambda} \right]' &= \frac{\partial}{\partial \mathbf{x}(t_0)} \left[\left(\frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}(t_1)} \right)' \boldsymbol{\lambda} \right] = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.
\end{aligned} \tag{4.63}$$

Then according to (4.13), we obtain

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2\lambda_2 & 0 \\ 0 & 2\lambda_1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \mathbf{\Psi}(0) + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{\Psi}(1) = \mathbf{0} \tag{4.64}$$

so that the boundary conditions for (4.61) are given by

$$\begin{aligned}
-\Psi_{11}(0) + 2\lambda_2 &= 0, & -\Psi_{12}(0) &= 0, \\
\Psi_{21}(1) &= 0, & \Psi_{22}(1) &= 0.
\end{aligned} \tag{4.65}$$

Therefore, the solution to conjugate matrix BVP (4.61), (4.65) is

$$\mathbf{\Psi}^*(\boldsymbol{\lambda}, t) = \begin{bmatrix} \Psi_{11}^* & \Psi_{12}^* \\ \Psi_{21}^* & \Psi_{22}^* \end{bmatrix} = \begin{bmatrix} -2t + 2\lambda_2 & 0 \\ 0 & 0 \end{bmatrix}. \tag{4.66}$$

It is easy to see that for $\{u^*(t), \mathbf{x}^*(t) = \mathbf{x}(t, u^*(t))\}$, formulas (4.62) yield

$$\frac{\partial \mathbf{\Phi}^*}{\partial \mathbf{x}(0)} = \frac{\partial \mathbf{\Phi}^*}{\partial \mathbf{x}(1)} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \tag{4.67}$$

and according to Remark 4.6, $\mathbf{Y}(t) = \mathbf{0}$ so that the necessary condition for optimality of singular controls (4.48) is given by

$$\langle \mathbf{\Psi}^*(\boldsymbol{\lambda}, t) \Delta_v \mathbf{f}(\mathbf{x}^*, u^*, t), \Delta_v \mathbf{f}(\mathbf{x}^*, u^*, t) \rangle \leq 0, \quad \forall \boldsymbol{\lambda} \in \mathbb{R}^2, \ |v(t)| \leq 1, \ t \in [0,1] \tag{4.68}$$

that results in

$$\left\langle \begin{bmatrix} 2(\lambda_2 - t) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ 0 \end{bmatrix}, \begin{bmatrix} v \\ 0 \end{bmatrix} \right\rangle \leq 0 \tag{4.69}$$

or

$$2(\lambda_2 - t)v^2 \leq 0, \quad \forall (\lambda_1, \lambda_2)' \in \mathbb{R}^2, \ |v(t)| \leq 1, \ t \in [0,1]. \tag{4.70}$$

Now it is obvious that for $\lambda_2 > 1$, inequality (4.70) does not hold. Therefore, the candidate singular control $u^*(t) \equiv 0$ cannot be optimal since it does not satisfy the necessary condition for optimality of singular controls (4.48).

We also keep in mind that for the candidate control $u^*(t) \equiv 0$, the value of objective functional is $J(u^*) = x_2^*(t) = 1$ and then see if there is another control function with a better value of J. For example, we calculate the value of J for admissible control $\bar{u}(t) \equiv -1$. For this control function, the general solution to the ODE system (4.49) is

$$\bar{x}_1(t) = c_1 - t,$$
$$\bar{x}_2(t) = c_2 - \frac{t^3}{3} + c_1 t^2 - c_1^2 t, \tag{4.71}$$

where the arbitrary constants c_1, c_2 are to be defined according to the boundary conditions of (4.49). Actually, the boundary conditions lead to the algebraic system,

$$c_2^2 - 3c_2 = -\frac{4}{3}, \qquad c_1 + c_2 = \frac{4}{3}, \tag{4.72}$$

that has two solution sets: $\{\bar{c}_1, \bar{c}_2\} = \{-1.12409, 2.45743\}$ and $\{\hat{c}_1, \hat{c}_2\} = \{0.79076, 0.54257\}$. These are approximate values rounded off to five decimals.

The resolvability condition (4.29) takes the form of

$$\det\left[\frac{\partial \Phi}{\partial \mathbf{x}(t_0)} + \frac{\partial \Phi}{\partial \mathbf{x}(t_1)} \mathbf{X}(t_1) \right] = 3 - 2c_2 \tag{4.73}$$

and obviously holds for both $\bar{c}_2 = 2.45743$ and $\hat{c}_2 = 0.54257$. That justifies the existence of solution to BVP (4.49) without its uniqueness (see Remark 4.5). Consequently, there are two sets of the system profiles $\{\bar{x}_1(t), \bar{x}_2(t)\}$ and $\{\hat{x}_1(t), \hat{x}_2(t)\}$ that correspond to the admissible control function $\bar{u}(t) \equiv -1$, $t \in [0,1]$. For each of these two sets, we can calculate the corresponding values of objective functional:

$$J(\bar{u}) = \begin{cases} \bar{x}_2(1) = -0.263572, \\ \hat{x}_2(1) = 0.374695. \end{cases} \tag{4.74}$$
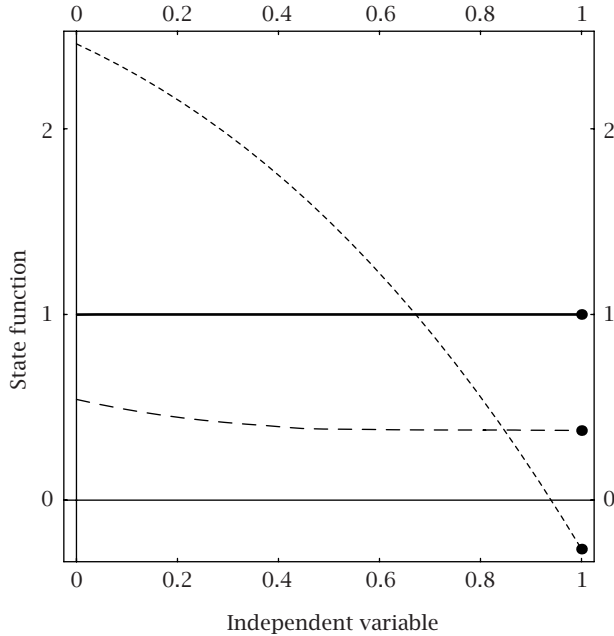
FIGURE 4.1. State profiles $x_2^*(t)$, $\bar{x}_2(t)$, and $\hat{x}_2(t)$.

Figure 4.1 shows three profiles where $x_2^*(t)$ is given by a thick solid line, $\bar{x}_2(t)$ is drawn by a dashed line, and $\hat{x}_2(t)$ is represented by a dotted line. Thick point in the end of each curve indicates the value of $J(u)$. Finally, having compared these values, it becomes perfectly clear that $u^*(t) \equiv 0$, $t \in [0,1]$, is not optimal.

References

[1]   A. A. Agračev and R. V. Gamkrelidze, *The principle of second order optimality for time-optimal problems*, Mat. Sb. (N.S.) **100(142)** (1976), no. 4, 610–643, 648 (Russian).

[2]   N. U. Ahmed, *Elements of Finite-Dimensional Systems and Control Theory*, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 37, John Wiley & Sons, New York, 1988.

[3]   D. J. Bell and D. H. Jacobson, *Singular Optimal Control Problems*, Academic Press, New York, 1975.

[4]   A. Bellaïche and J.-J. Risler (eds.), *Sub-Riemannian Geometry*, Progress in Mathematics, vol. 144, Birkhäuser Verlag, Basel, 1996.

[5]   G. Fraser-Andrews, *Finding candidate singular optimal controls: a state of the art survey*, J. Optim. Theory Appl. **60** (1989), no. 2, 173–190.

[6]   R. F. Gabasov, *Necessary conditions for optimality of singular control*, Izv. Akad. Nauk SSSR Tehn. Kibernet. **1968** (1968), no. 5, 34–43 (Russian), translated in Engrg. Cybernetics **1968** (1968), no. 5, 28–37.

[7]   R. F. Gabasov and F. M. Kirillova, *Singular Optimal Controls*, Izdat. "Nauka", Moscow, 1973.

[8]   H. J. Kelley, R. E. Kopp, and H. G. Moyer, *Singular extremals*, Topics in Optimization (G. Leitmann, ed.), Academic Press, New York, 1967, pp. 63–101.

[9]   A. J. Krener, *The high order maximal principle and its application to singular extremals*, SIAM J. Control Optim. **15** (1977), no. 2, 256–293.

[10]   L. I. Rozonoèr, *L. S. Pontryagin maximum principle in the theory of optimum systems. I, II, III*, Avtomat. i Telemekh. **20** (1959), 1320-1334, 1441-1458, 1561-1578 (Russian), translated in Autom. Remote Control **20** (1959), 1288-1302, 1405-1421, 1517-1532.

[11]   F. P. Vasil'ev, *Optimization Methods*, Factorial Press, Moscow, 2002.

[12]   O. V. Vasiliev, *Optimization Methods*, Advanced Series in Mathematical Science and Engineering, vol. 5, World Federation Publishers Company, Georgia, 1996.

[13]   O. O. Vasilieva, *Two-parameter algorithm for optimal control problems with boundary conditions*, Saitama Math. J. **20** (2002), 45-62.

[14]   O. O. Vasilieva and K. Mizukami, *Optimal control of a boundary value problem*, Izv. Vyssh. Uchebn. Zaved. Mat. **38** (1994), no. 12, 33-41 (Russian), translation in Russian Math. (Iz. VUZ) **38** (1994), no. 12, 31-39.

[15]   ———, *Optimality criterion for singular controllers: linear boundary conditions*, J. Math. Anal. Appl. **213** (1997), no. 2, 620-641.

[16]   W. Xu, *A maximum principle for optimal control for a class of controlled systems*, J. Austral. Math. Soc. Ser. B **38** (1996), no. 2, 172-181.

Olga Vasilieva: Department of Mathematics, University of Valle, Cali, Colombia

E-mail address: olgavas@mafalda.univalle.edu.co

Advances in
Operations Research

Advances in
Decision Sciences

Journal of
Applied Mathematics

Algebra

Journal of
Probability and Statistics

The Scientific
World Journal

International Journal of
Differential Equations

International Journal of
Combinatorics

Hindawi

Submit your manuscripts at
http://www.hindawi.com

Advances in
Mathematical Physics

Journal of
Complex Analysis

Journal of
Mathematics

Mathematical Problems
in Engineering

Abstract and
Applied Analysis

Discrete Dynamics in
Nature and Society

International
Journal of
Mathematics and
Mathematical
Sciences

Journal of
Discrete Mathematics

Journal of
Function Spaces

International Journal of
Stochastic Analysis

Journal of
Optimization