JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **213,** 620–641 (1997)
ARTICLE NO. AY975565

Optimality Criterion for Singular Controllers: Linear Boundary Conditions

O. Vasilieva and K. Mizukami

Faculty of Integrated Arts and Sciences, Hiroshima University, 1-7-1 Kagamiyama

Submitted by Harold L. Stalford

Received June 24, 1996

This paper is concerned with the dynamic process governed by the boundary value problem. We examine the situation with the appearance of singular controllers and formulate a second-order optimality criterion on the basis of the increment formula. We also show how to apply this criterion as a verifying condition for optimality and how to perform an optimal design for a singular controller. © 1997 Academic Press

1. INTRODUCTION

In the theory of optimal control, great importance is fairly attached to the maximum principle of L. S. Pontryagin. On the other hand, in certain cases the maximum principle has a tendency to "degenerate," i.e., to be fulfilled trivially on a series of admissible controllers, and therefore it cannot serve as a verifying condition for optimality nor as a basis for the 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 which is typical for complicated dynamic processes.

A number of research works have been devoted to the investigation of singularities in dynamic processes governed by the Cauchy problem (i.e., the initial value problem), see, e.g., [1, 2]. One of the directions in this research was to obtain an optimality criterion for a singular controller on the basis of second-order increment formula (if the classic increment formula whence the maximum principle follows is regarded as the first-

0022-247X/97 $25.00

order increment formula). For the dynamic process governed by the Cauchy problem, the second-order increment formula and optimality criterion for a singular controller expressed in terms of an adjoint matrix system have been discovered in [1, 2]. For a long time this criterion has been used only as a verifying condition. Later in [3] this criterion has been integrated into the general scheme of an optimal design algorithm.

This paper deals with the dynamic processes governed by the boundary value problem which have various applications in the simulation of physical, mechanical, chemical, and other systems (and also includes the processes considered in [1–3] as a special case). For this type of dynamic processes the authors have recently obtained an optimality criterion in the form of the maximum principle and set up an optimal design algorithm [5, 6]. To complete this research it is necessary to examine the situation with the appearance of singularities and formulate the optimality criterion for a singular controller.

This is exactly the purpose of our paper. However, it should be noted that here we consider linear boundary conditions in contrast to the generalized formulation of the problem given in [5, 6]. This restriction is not caused by the essence of the research technique. The complexity of deductions forces us to deal with linear boundary conditions in order to make all the descriptions and explanations as understandable and clear as possible under the limited length of the paper.

2. PROBLEM FORMULATION AND PRELIMINARIES

Let us start by posing the problem to minimize the performance index

$$J(u) = \varphi(x(t_0), x(t_1)) + \int_T F(x, u, t)\, dt \to \min \tag{1}$$

which is defined on the solution set of the boundary value problem (henceforward BVP)

$$\dot{x} = f(x, u, t), \qquad t \in T = [t_0, t_1] \tag{2}$$

$$L_0 x(t_0) + L_1 x(t_1) - b = 0. \tag{3}$$

Here $x = x(t)$, $x(t) \in R^n$ describes the state of dynamic process (2), (3); $u = u(t)$, $u(t) \in R^r$ represents the controller; vector-function $f = (f_1, \dots, f_n)$ and scalar functions φ, F are continuous in (x, u, t) together with their partial derivatives up to second order; L_0, L_1, and b are specified numerical $(n \times n), (n \times n), (n \times 1)$ matrices, and

$$\operatorname{rank}[L_0 L_1] = n.$$

We shall refer to the class of admissible controllers as a set of measurable vector-functions $u(\cdot) \in L_\infty^r(T)$ with direct constraint

$$u(t) \in U, \qquad t \in T, \tag{4}$$

where U is a compact set in R^r.

Remark 2.1. For $L_0 = I$ (identity matrix), $L_1 = 0$, problem (1)–(4) turns into the familiar free end-point problem. In this case, the question of existence and uniqueness of solution to the Cauchy problem in any admissible controller is answered straightforwardly. The same may be said about BVP (2)–(3) only if system (2) is linear in the state variable (see, e,g., [5]). Therefore, from now on we ought to introduce

Assumption 2.1. Suppose that BVP (2)–(3) is resolvable in any admissible controller, and that the set formed by all admissible pairs $\{u, x = x(t, u)\}$ is closed.

3. SECOND-ORDER INCREMENT FORMULA

For two admissible processes—basic $\{u, x = x(t, u)\}$ and varying $\{\tilde{u} = u + \Delta u, \tilde{x} = x + \Delta x = x(t, \tilde{u})\}$—we can define an incremental BVP

$$\Delta \dot{x} = \Delta f(x, u, t), \qquad L_0 \Delta x(t_0) + L_1 \Delta x(t_1) = 0, \tag{5}$$

where

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

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

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

Let us introduce some non-trivial vector-functions $\psi = \psi(t)$, $\psi(t) \in R^n$, the matrix function $\Psi = \Psi(t)$, $\Psi(t) \in R^{n \times n}$, numerical vector $\lambda \in R^n$, and numerical matrices Λ_0, Λ_1 of dimension $(n \times n)$. Then the increment of performance index (1) may be represented as

$$\Delta J(u) = \Delta\varphi(x(t_0), x(t_1)) + \int_T \Delta F(x,u,t)\,dt + \int_T \left\langle \psi(t) + \frac{1}{2}\Psi(t)\Delta x, \Delta\dot{x}(t) - \Delta f(x,u,t) \right\rangle dt + \left\langle \lambda + \frac{1}{2}\Lambda_0\Delta x(t_0) + \frac{1}{2}\Lambda_1\Delta x(t_1), L_0\Delta x(t_0) + L_1\Delta x(t_1) \right\rangle, \tag{6}$$

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

Now we are about to perform a few necessary transformations, viz.,

—introduce the Hamiltonian function

$$H(\psi, x, u, t) = \langle \psi(t), f(x,u,t)\rangle - F(x,u,t);$$

—expand $\Delta\varphi, \Delta_{\tilde{x}}H(\psi, x, \tilde{u}, t)$ in Taylor series up to the second-order term

$$\begin{aligned}\Delta\varphi(x(t_0), x(t_1)) &= \left\langle \frac{\partial\varphi}{\partial x(t_0)}, \Delta x(t_0) \right\rangle + \left\langle \frac{\partial\varphi}{\partial x(t_1)}, \Delta x(t_1) \right\rangle \\ &+ \frac{1}{2}\left\langle \frac{\partial^2\varphi}{\partial x(t_0)^2}\Delta x(t_0), \Delta x(t_0) \right\rangle + \frac{1}{2}\left\langle \frac{\partial^2\varphi}{\partial x(t_1)^2}\Delta x(t_1), \Delta x(t_1) \right\rangle \\ &+ \frac{1}{2}\left\langle \frac{\partial^2\varphi}{\partial x(t_0)\,\partial x(t_1)}\Delta x(t_0), \Delta x(t_1) \right\rangle \\ &+ \frac{1}{2}\left\langle \frac{\partial^2\varphi}{\partial x(t_1)\,\partial x(t_0)}\Delta x(t_1), \Delta x(t_0) \right\rangle \\ &+ o_\varphi\left(\|\Delta x(t_0)\|^2, \|\Delta x(t_1)\|^2\right)\end{aligned}$$

$$\begin{aligned}\Delta_{\tilde{x}\tilde{u}}H(\psi, x, u, t) &= \Delta_{\tilde{u}}H(\psi, x, u, t) + \Delta_{\tilde{x}}H(\psi, x, \tilde{u}, t) \\ &= \Delta_{\tilde{u}}H(\psi, x, u, t) + \left\langle \frac{\partial H(\psi, x, \tilde{u}, t)}{\partial x}, \Delta x(t) \right\rangle \\ &\quad + \frac{1}{2}\left\langle \frac{\partial^2 H(\psi, x, \tilde{u}, t)}{\partial x^2}\Delta x(t), \Delta x(t) \right\rangle + o_H\left(\|\Delta x(t)\|^2\right);\end{aligned}$$

—in the last equation represent each summand of the form $K(\tilde{u})$ as

$$K(\tilde{u}) = K(u) + \Delta_{\tilde{u}} K(u);$$

—perform the integration by parts

$$\int_T \langle \psi(t), \Delta \dot{x}(t) \rangle \, dt = \langle \psi(t_1), \Delta x(t_1) \rangle - \langle \psi(t_0), \Delta x(t_0) \rangle - \int_T \langle \dot{\psi}(t), \Delta x(t) \rangle \, dt,$$

$$\begin{aligned} \int_T \langle \Psi(t) \Delta x(t), \Delta \dot{x}(t) \rangle \, dt &= \langle \Psi(t_1) \Delta x(t_1), \Delta x(t_1) \rangle - \langle \Psi(t_0) \Delta x(t_0), \Delta x(t_0) \rangle \\ &\quad - \int_T \langle \dot{\Psi}(t) \Delta x(t), \Delta x(t) \rangle \, dt - \int_T \langle \Psi(t) \Delta \dot{x}(t), \Delta x(t) \rangle \, dt, \end{aligned}$$

where

$$\begin{aligned} \int_T \langle \Psi(t) \Delta \dot{x}(t), \Delta x(t) \rangle \, dt &= \int_T \langle \Psi(t) \Delta f(x, u, t), \Delta x(t) \rangle \, dt \\ &= \int_T \left\langle \Psi(t) \left[\Delta_{\tilde{u}} f(x, u, t) + \frac{\partial f(x, u, t)}{\partial x} \Delta x(t) \right. \right. \\ &\qquad \left. \left. + \Delta_{\tilde{u}} \frac{\partial f(x, u, t)}{\partial x} \Delta x(t) + \hat{o}_f(\| \Delta x(t) \|) \right], \Delta x(t) \right\rangle dt; \end{aligned}$$

—transform the second entry in the last row

$$\begin{aligned} \int_T \left\langle \Psi(t) \frac{\partial f(x, u, t)}{\partial x} \Delta x(t), \Delta x(t) \right\rangle dt &= \frac{1}{2} \int_T \left\langle \Psi(t) \frac{\partial f(x, u, t)}{\partial x} \Delta x(t), \Delta x(t) \right\rangle dt \\ &\quad + \frac{1}{2} \int_T \left\langle \frac{\partial f(x, u, t)'}{\partial x} \Psi(t)' \Delta x(t), \Delta x(t) \right\rangle dt; \end{aligned}$$

—and convert the last summand in (6) into

$$\left\langle \lambda + \frac{1}{2}\Lambda_0 \Delta x(t_0) + \frac{1}{2}\Lambda_1 \Delta x(t_1), L_0 \Delta x(t_0) + L_1 \Delta x(t_1) \right\rangle$$

$$= \langle L_0' \lambda, \Delta x(t_0)\rangle + \langle L_1' \lambda, \Delta x(t_1)\rangle + \frac{1}{2}\langle L_0' \Lambda_0 \Delta x(t_0), \Delta x(t_0)\rangle$$

$$+ \frac{1}{2}\langle L_1' \Lambda_1 \Delta x(t_1), \Delta x(t_1)\rangle + \frac{1}{2}\langle L_1' \Lambda_0 \Delta x(t_0), \Delta x(t_1)\rangle$$

$$+ \frac{1}{2}\langle L_0' \Lambda_1 \Delta x(t_1), \Delta x(t_0)\rangle.$$

Then, having substituted all these processed expressions into the increment formula (6), one ought to separate the principal summands

$$-\int_T \Delta_{\tilde{u}} H \, dt - \int_T \left\langle \Delta_{\tilde{u}} \frac{\partial H}{\partial x} + \Psi \Delta_{\tilde{u}} f, \Delta x \right\rangle dt$$

from the reminder $\eta = \eta_1 - \eta_2$ where

$$\eta_1 = o_\varphi\left(\| \Delta x(t_0) \|^2, \| \Delta x(t_1) \|^2\right) - \int_T o_H\left(\| \Delta x(t) \|^2\right) dt$$

$$- \int_T \langle \Psi(t) \Delta x(t), \hat{o}_f(\| \Delta x(t) \|)\rangle \, dt, \tag{7}$$

$$\eta_2 = \frac{1}{2}\int_T \left\langle \Delta_{\tilde{u}}\left[\frac{\partial^2 H(\psi, x, u, t)}{\partial x^2} + 2\Psi(t)\frac{\partial f(x, u, t)}{\partial x}\right]\Delta x(t), \Delta x(t)\right\rangle dt \tag{8}$$

and by using the arbitrary choice of $\psi, \Psi, \lambda, \Lambda_0, \Lambda_1$ make all the corresponding coefficients vanish. Thus, we obtain an increment formula of second order,

$$\Delta J(u) = -\int_T \Delta_{\tilde{u}} H(\psi, x, u, t) \, dt$$

$$- \int_T \left\langle \Delta_{\tilde{u}} \frac{\partial H(\psi, x, u, t)}{\partial x} + \Psi(t) \Delta_{\tilde{u}} f(x, u, t), \Delta x(t) \right\rangle dt$$

$$+ \eta_1 - \eta_2, \tag{9}$$

where the vector-function $\psi = \psi(t)$ is subordinated to the vector adjoint system

$$\dot{\psi} = -\frac{\partial H(\psi, x, u, t)}{\partial x} \tag{10}$$

and the matrix function $\Psi = \Psi(t)$ is subordinated to the matrix adjoint system

$$\dot{\Psi} = -\frac{\partial f(x, u, t)'}{\partial x}\Psi' - \Psi\frac{\partial f(x, u, t)}{\partial x} - \frac{\partial^2 H(\psi, x, u, t)}{\partial x^2}. \tag{11}$$

Boundary conditions for both system (10) and (11) are defined by the equalities

$$\left.\begin{aligned} \frac{\partial \varphi}{\partial x(t_0)} + L_0'\lambda - \psi(t_0) = 0 \\ \frac{\partial \varphi}{\partial x(t_1)} + L_1'\lambda + \psi(t_1) = 0 \end{aligned}\right\} \tag{12}$$

$$\left.\begin{aligned} \frac{\partial^2 \varphi}{\partial x(t_0)^2} + L_0'\Lambda_0 - \Psi(t_0) = 0 \\ \frac{\partial^2 \varphi}{\partial x(t_1)^2} + L_1'\Lambda_1 + \Psi(t_1) = 0 \\ \frac{\partial^2 \varphi}{\partial x(t_0)\,\partial x(t_1)} + L_1'\Lambda_0 = 0 \\ \frac{\partial^2 \varphi}{\partial x(t_1)\,\partial x(t_0)} + L_0'\Lambda_1 = 0. \end{aligned}\right\} \tag{13}$$

Remark 3.1. In contrast to the free end-point problem (see, e.g., [1, 2]), the matrix function $\Psi = \Psi(t)$ here is asymmetric in the general sense.

Further, to eliminate vector λ and matrices Λ_0, Λ_1 out of (12), (13), we select such matrices B_0, B_1 distinct from zero that

$$B_0 L_0' + B_1 L_1' = 0. \tag{14}$$

Then boundary conditions (12), (13) become equivalent to linear boundary conditions for vector and matrix adjoint systems (10) and (11), respectively,

$$-B_0\psi(t_0) + B_1\psi(t_1) + B_0\frac{\partial\varphi}{\partial x(t_0)} + B_1\frac{\partial\varphi}{\partial x(t_1)} = 0 \tag{15}$$

$$-B_0\Psi(t_0) + B_1\Psi(t_1) + B_0\frac{\partial^2\varphi}{\partial x(t_0)^2} + B_1\frac{\partial^2\varphi}{\partial x(t_1)^2}$$
$$+ B_0\frac{\partial^2\varphi}{\partial x(t_1)\,\partial x(t_0)} + B_1\frac{\partial^2\varphi}{\partial x(t_0)\,\partial x(t_1)} = 0. \tag{16}$$

To complete this section it is useful to survey some examples. First of all, let us consider the Cauchy problem as a special case of BVP. In other words, let $L_0 = I$, $L_1 = 0$. Then $\varphi = \varphi(x(t_1))$. It is obvious that equality (14) holds for $B_0 = 0$, $B_1 = I$, and Eqs. (15), (16) turn into

$$\psi(t_1) = -\frac{\partial\varphi}{\partial x(t_1)} \qquad \text{and} \qquad \Psi(t_1) = -\frac{\partial^2\varphi}{\partial x(t_1)^2}.$$

In this case the increment formula of the form (9) has been obtained in [1, 2] though in another way.

Now let us examine another type of BVP generally known as the two-point BVP. For the system (2) it is defined by the boundary conditions

$$x^{(1)}(t_0) = b^{(1)}, \qquad x^{(2)}(t_1) = b^{(2)} \tag{17}$$

where $x = (x^{(1)}, x^{(2)})$, $x^{(1)} \in R^m$, $x^{(2)} \in R^{(n-m)}$. Boundary conditions (17) come out of (3) when

$$L_0 = \begin{bmatrix} I_{m\times m} & 0_{m\times(n-m)} \\ 0_{(n-m)\times m} & 0_{(n-m)\times(n-m)} \end{bmatrix}, \qquad L_1 = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}.$$

There is no difficulty to prove that the matrix equality (14) holds for

$$B_0 = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}, \qquad B_1 = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}.$$

Then boundary conditions (15) for the vector adjoint system (10) are given by

$$\psi^{(1)}(t_1) = -\frac{\partial\varphi}{\partial x^{(1)}(t_1)}, \qquad \psi^{(2)}(t_0) = +\frac{\partial\varphi}{\partial x^{(2)}(t_0)},$$

where, of course, $\varphi = \varphi(x^{(1)}(t_1), x^{(2)}(t_0))$, and boundary conditions (16) for the matrix adjoint system (11) will be written as

$$\Psi_{11}(t_1) = -\frac{\partial^2\varphi}{\partial x^{(1)}(t_1)^2}, \qquad \Psi_{12}(t_1) = 0$$

$$\Psi_{21}(t_0) = +\frac{\partial^2\varphi}{\partial x^{(1)}(t_1)\,\partial x^{(2)}(t_0)}, \qquad \Psi_{22}(t_0) = +\frac{\partial^2\varphi}{\partial x^{(2)}(t_0)^2}. \quad (18)$$

Here

$$\Psi_{(n\times n)} = \begin{bmatrix} \Psi_{11m\times m} & \Psi_{12m\times(n-m)} \\ \Psi_{21(n-m)\times m} & \Psi_{22(n-m)\times(n-m)} \end{bmatrix}.$$

A great many dynamic processes in physics and mechanics are usually described by BVP of the form

$$\ddot{y} = f(y, \dot{y}, u, t), \qquad y(t) \in R^n$$
$$y(t_0) = b^0, \qquad y(t_1) = b^1.$$

It becomes clear that this type also belongs to the class of two-point BVP when we denote $x = (x^{(1)}, x^{(2)})$, $x \in R^{2n}$, $x^{(1)} \in R^n$, $x^{(2)} \in R^n$, $x^{(1)} = y$, $x^{(2)} = \dot{y}$. Boundary conditions $x^{(1)}(t_0) = b^0$, $x^{(2)}(t_1) = b^1$ follow from the boundary conditions (3) for

$$L_0 = \begin{bmatrix} I_{n\times n} & 0_{n\times n} \\ 0_{n\times n} & 0_{n\times n} \end{bmatrix}, \qquad L_1 = \begin{bmatrix} 0 & 0 \\ I & 0 \end{bmatrix}.$$

It is obvious that $\varphi = \varphi(x^{(2)}(t_0), x^{(2)}(t_1))$ and the matrix equality (14) holds for

$$B_0 = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix}, \qquad B_1 = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}.$$

Then boundary conditions (15) in this case are given by

$$\psi^{(2)}(t_0) = +\frac{\partial \varphi}{\partial x^{(2)}(t_0)}, \qquad \psi^{(2)}(t_1) = -\frac{\partial \varphi}{\partial x^{(2)}(t_1)}$$

and boundary conditions (16) will be written as

$$\Psi_{21}(t_0) = 0, \qquad \Psi_{21}(t_1) = 0$$

$$\Psi_{22}(t_0) = \frac{\partial^2 \varphi}{\partial x^{(2)}(t_0)^2} + \frac{\partial^2 \varphi}{\partial x^{(2)}(t_0)\, \partial x^{(2)}(t_1)},$$

$$\Psi_{22}(t_1) = -\frac{\partial^2 \varphi}{\partial x^{(2)}(t_1)\, \partial x^{(2)}(t_0)}.$$

4. NEEDLE-SHAPED VARIATION

First, let us recall the state-increment estimate caused by control variation, which has been obtained in [5]

$$\|\Delta x(t)\| \leq K \int_T \|\Delta_{\tilde{u}} f(x, u, t)\|\, dt, \qquad K = \text{const} > 0. \tag{19}$$

Now if we replace an arbitrary admissible controller $\tilde{u}(t) \in U$ by needle-shaped variation, i.e., set $\tilde{u}(t) = u(t) + \Delta_\varepsilon u(t)$ where

$$\Delta_\varepsilon u(t) = v - u(t), \qquad v \in U, t \in (\tau - \varepsilon, \tau] \subset T, \varepsilon > 0$$

$$\Delta_\varepsilon u(t) = 0, \qquad t \in T \setminus (\tau - \varepsilon, \tau] \tag{20}$$

then by virtue of estimate (19), the increment of state $\Delta_\varepsilon x(t)$ caused by needle-shaped variation (20) will be of order ε,

$$\|\Delta x(t)\| \leq K_1 \cdot \varepsilon, \qquad K_1 = \text{const} > 0$$

and increment formula (7)–(9) turns into

$$\Delta_\varepsilon J(u) = -\int_{\tau-\varepsilon}^{\tau} \Delta_v H(\psi, x, u, t)\, dt$$
$$- \int_{\tau-\varepsilon}^{\tau} \left\langle \Delta_v \frac{\partial H(\psi, x, u, t)}{\partial x} + \Psi(t) \Delta_v f(x, u, t), \Delta_\varepsilon x(t) \right\rangle dt$$
$$+ o(\varepsilon^2), \tag{21}$$

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

Generally speaking, the necessary condition for optimality (i.e., maximum principle) which has been obtained in [3–5] results from the formula (21). Factually, it follows from (21) that

$$\Delta_\varepsilon J(u) = -\Delta_v H(\psi, x, u, \tau) \cdot \varepsilon + o(\varepsilon), \qquad \tau \in T, v \in U$$

which implies that for the optimal process $\{u^*, x^*\}$ and associated solution ψ^* of adjoint BVP (10), (14), (15)

$$\Delta_\varepsilon J(u^*) \geq 0, \qquad \tau \in T, v \in U$$

and indeed

$$\Delta_v H(\psi^*, x^*, u^*, \tau) \leq 0, \qquad \tau \in T, v \in U. \tag{22}$$

Further, we ought to extract an explicit coefficient of ε^2 in the variational increment formula (21). To cope with this task, it is sufficient to extract a coefficient of ε in $\Delta_\varepsilon x(\tau)$.

One should approach this task by examining the incremental BVP (5) on the needle-shaped variation (20). It is clear that

$$\begin{aligned}\Delta \dot{x} &= \Delta_{\tilde{x}} f(x, \tilde{u}, t) + \Delta_{\tilde{u}} f(x, u, t) \\ &= \frac{\partial f(x, u, t)}{\partial x} \Delta x(t) + \Delta_{\tilde{u}} f(x, u, t) \Delta x(t) \\ &\quad + \Delta_{\tilde{u}} \frac{\partial f(x, u, t)}{\partial x} \Delta x(t) + \hat{o}_f(\| \Delta x(t) \|)\end{aligned}$$

or in integral form

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

Then, after carry-over onto needle-shaped variation (20)

$$\begin{aligned}\Delta_\varepsilon x(t) &= \Delta_\varepsilon x(t_0) \\ &\quad + \int_{t_0}^{t} \left[\frac{\partial f(x, u, \xi)}{\partial x} \Delta_\varepsilon x(\xi) + \Delta_v f(x, u, \xi) \right] d\xi + o(\varepsilon).\end{aligned}$$

Thus, incremental BVP (5) may be rewritten as

$$\Delta_\varepsilon \dot{x} = \frac{\partial f(x,u,t)}{\partial x}\Delta_\varepsilon x + \Delta_v f(x,u,t) + o(\varepsilon) \tag{23}$$
$$L_0 \Delta_\varepsilon x(t_0) + L_1 \Delta_\varepsilon x(t_1) = 0.$$

The solution to the linear BVP (23) is found (to within $o(\varepsilon)$) by the Cauchy formula analogy derived by the authors in [5].

Let $X = X(t)$ be a fundamental $(n \times n)$ matrix function of homogeneous system (23):

$$\dot{X} = \frac{\partial f(x,u,t)}{\partial x} X, \qquad X(t_0) = I. \tag{24}$$

Suppose that BVP (23) has a unique solution in any admissible process $\{u, x\}$, i.e.,

$$\det\left[L_0 + L_1 X(t_1)\right] \neq 0. \tag{25}$$

Then

$$\begin{aligned} \Delta_\varepsilon x(t) &= -\int_{\tau-\varepsilon}^{\tau} X(t)\Phi(t_1)X^{-1}(\xi)\Delta_v f(x,u,\xi)\,d\xi \\ &\quad + \int_{t_0}^{t} X(t)X^{-1}(\xi)\left\{\begin{matrix} 0, & t \in [t_0, \tau-\varepsilon) \\ \Delta_v f(x,u,\xi), & t \in [\tau-\varepsilon, \tau) \\ 0, & t \in [\tau, t_1] \end{matrix}\right\} d\xi \\ &\quad + o(\varepsilon), \end{aligned}$$

where

$$\Phi(t_1) = \left[L_0 + L_1 X(t_1)\right]^{-1} L_1 \cdot X(t_1). \tag{26}$$

Whence

$$\begin{aligned} \Delta_\varepsilon x(t) = {} & -X(t)\Phi(t_1)X^{-1}(\tau)\Delta_v f(x,u,\tau)\cdot\varepsilon \\ & + \left\{\begin{matrix} 0, & t \in [t_0, \tau-\varepsilon) \\ X(t)X^{-1}(\tau)\Delta_v f(x,u,\tau)(t-\tau+\varepsilon), & t \in [\tau-\varepsilon, \tau) \\ X(t)X^{-1}(\tau)\Delta_v f(x,u,\tau)\cdot\varepsilon, & t \in [\tau, t_1] \end{matrix}\right\} \\ & + o(\varepsilon) \end{aligned}$$

and therefore

$$\Delta_\varepsilon x(\tau) = \left[I - X(\tau)\Phi(t_1)X^{-1}(\tau)\right]\Delta_v f(x,u,\tau)\cdot\varepsilon + o(\varepsilon). \quad (27)$$

It should be noted that

$$\dot{X}^{-1} = -X^{-1}\cdot\frac{\partial f(x,u,t)}{\partial x}, \qquad X^{-1}(t_0) = I. \quad (28)$$

It is obvious that

$$X^{-1}(t)X(t) = I \qquad \text{and} \qquad \dot{X}^{-1}(t)X(t) + X^{-1}(t)\dot{X}(t) = 0$$

and with regard to (24)

$$\left[\dot{X}^{-1}(t) + X^{-1}(t)\right]X(t) = 0$$

whence Eqs. (28) immediately follow.

If we set

$$Y(t) = X(t)\Phi(t_1)X^{-1}(t), \qquad Y(t_0) = \Phi(t_1)$$

then with regard to (24) and (27),

$$\begin{aligned}\dot{Y}(t) &= \dot{X}(t)\Phi(t_1)X^{-1}(t) + X(t)\Phi(t_1)\dot{X}^{-1}(t)\\ &= \frac{\partial f(x,u,t)}{\partial x}X(t)\Phi(t_1)X^{-1}(t) - X(t)\Phi(t_1)X^{-1}(t)\frac{\partial f(x,u,t)}{\partial x}\\ &= \frac{\partial f(x,u,t)}{\partial x}Y(t) - Y(t)\frac{\partial f(x,u,t)}{\partial x}.\end{aligned}$$

So far,

$$\Delta_\varepsilon x(\tau) = \left[I - Y(\tau)\right]\Delta_v f(x,u,\tau)\varepsilon + o(\varepsilon), \quad (29)$$

where

$$\dot{Y} = \frac{\partial f(x,u,t)}{\partial x}Y - Y\frac{\partial f(x,u,t)}{\partial x} \quad (30)$$

$$Y(t_0) = \left[L_0 + L_1X(t_1)\right]^{-1}L_1X(t_1). \quad (31)$$

Thus, in order to calculate the coefficient of $\Delta_v f(x,u,\tau)\varepsilon$ in $\Delta_\varepsilon x(\tau)$, $\tau \in T$ one should

(a) solve the matrix Cauchy problem (24);

(b) compute the matrix $Y(t_0)$ by the formula (31);

(c) solve the matrix Cauchy problem (30), (31).

5. OPTIMALITY CRITERION FOR SINGULAR CONTROLLERS

Taking into account (29), the increment formula (21) will take on its final representation

$$\Delta_\varepsilon J(u) = -\int_{\tau-\varepsilon}^{\tau} \Delta_v H(\psi, x, u, t)\, dt - \left\langle \Delta_v \frac{\partial H(\psi, x, u, \tau)}{\partial x} + \Psi(\tau)\Delta_v f(x,u,\tau), [I - Y(\tau)]\Delta_v f(x,u,\tau) \right\rangle \cdot \varepsilon^2 + o(\varepsilon^2). \tag{32}$$

DEFINITION 5.1. The admissible controller $u = u(t)$ is called singular on the set $\Omega \subset T$ of positive measure if

$$\Delta_{\tilde{u}} H(\psi, x, u, t) \equiv 0 \tag{33}$$

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

For example, if the optimal controller $u^*(t)$ is singular on $\Omega \subset T$, mes $\Omega > 0$ then the function $H(\psi^*, x^*, u^*, t)$ does not depend upon the control variable u on the direct product $U \times \Omega$. Therefore, at all $t \in \Omega$ the maximum principle (22) becomes useless as a verifying condition for optimality in the first place. Secondly, the singular controller $u^k = u^k(t)$ may appear on some kth step of the iterative decision process which would stop the process even if u^k is not optimal yet.

In other words, condition (33) conveys the degeneracy (or triviality) of the maximum principle within $\Omega \in T$ and indicates the need of another optimality criterion which would involve a deeper analysis of the primary problem (1)–(4). The second-order increment formula (32) just allows us to formulate such a criterion if the variation procedure is carried out within the range $(\tau - \varepsilon, \tau] \in \Omega$ where the maximum principle loses its significance.

In this case, having calculated ψ^*, Ψ^*, and Y^* which are associated with the optimal process $\{u^*, x^*\}$ it becomes clear that

$$\Delta_\varepsilon J(u^*) = -\left\langle \Delta_v \frac{\partial H(\psi^*, x^*, u^*, \tau)}{\partial x} + \Psi^*(\tau)\Delta_v f(x^*, u^*, \tau), \right.$$

$$\left. [I - Y^*(\tau)]\Delta_v f(x^*, u^*, \tau) \right\rangle \varepsilon^2 + o(\varepsilon^2) \geq 0$$

$$\tau \in \Omega, \qquad v \in U.$$

Thus, all the foregoing calculations result in the second-order optimality criterion being formulated in the form of

THEOREM 5.1. *In order that an admissible controller $u^* = u^*(t)$ singular on $\Omega \subset T$ be optimal in the primary problem* (1)–(4) *it is necessary that two conditions hold*:

(1) *The maximum condition with respect to u^* for the Hamiltonian function*

$$\Delta_v H(\psi^*, x^*, u^*, t) \leq 0, \qquad v \in U \tag{34}$$

almost everywhere on $T \setminus \Omega$;

(2) *The second-order condition in the form of inequality*

$$\left\langle \Delta_v \frac{\partial H(\psi^*, x^*, u^*, t)}{\partial x} + \Psi^*(t)\Delta_v f(x^*, u^*, t), \right.$$

$$\left. [I - Y^*(t)]\Delta_v f(x^*, u^*, t) \right\rangle \leq 0, \qquad v \in U \tag{35}$$

almost everywhere on Ω and along the solutions ψ^*, Ψ^*, X^*, Y^* to adjoint vector (10), (14), (15) *and matrix* (11), (14), (16) *BVP and auxiliary initial value problems* (24), (30), (31).

6. APPLICATION OF THE OPTIMALITY CRITERION

First of all, the optimality criterion for singular controller (35) may be adopted as the verifying condition. In other words, by means of this criterion one can easily exclude a singular non-optimal process out of consideration and narrow the class of admissible controllers which are suspected to be optimal.

Moreover, the necessary condition (35) may be highly useful in computations. For example, when the decision process for the primary problem

(1)–(4) is carried out by some successive approximation algorithm of maximum principle there might appear some locality $\Omega \in T$ where the maximum principle degenerates (i.e., identity (33) holds) and the decision process would be forced to stop. In this situation we can introduce an additional calculation procedure which gives us a chance to improve an admissible controller within such a locality Ω. This procedure (henceforward called the improving procedure) will complete the main body of iterative method of the maximum principle set forth in [3–6].

Before the description of the improving procedure, we shall introduce a scalar function

$$Q(u,v,t) = \left\langle \Delta_v \frac{\partial H(\psi,x,u,t)}{\partial x} + \Psi(t)\Delta_v f(x,u,t), [I - Y(t)]\Delta_v f(x,u,t) \right\rangle$$

$$Q(u,u,t) \equiv 0 \tag{36}$$

defined on some admissible process $\{u, x\}$ and associated solutions ψ, Ψ, X, Y.

Assumption 6.1. Assume the resolvability of the maximum condition for function (36) in $v \in U$ for all $t \in T$

$$\hat{u}(t) = \arg\max_{v \in U} Q(u,v,t). \tag{37}$$

Now let us proceed to the improving procedure. Suppose that on some kth iteration of the algorithm of the maximum principle like [3–6]

$$W(u^k, \tau_k) = \max_{t \in T} \max_{v \in U} W(u^k, v, t) = 0 \tag{38}$$

$$W(u^k, v, t) = \Delta_v H(\psi^k, x^k, u^k, t)$$

if only within the limits of adequate accuracy which implies that either the controller $u^k = u^k(t)$ satisfies the maximum principle or the algorithm has exhausted all its resources because the value of the performance index (1) does not decrease any more.

Thereupon we form $Q(u^k, v, t)$ by formula (36), find $\hat{u}^k(t)$ according to (37), and calculate

$$\hat{Q}_k(t) = Q(u^k, \hat{u}^k, t) \geq 0.$$

Then we construct the set

$$\Omega_k = \{t \in T : W(u^k, \hat{u}^k, t) \equiv 0\}. \tag{39}$$

If mes $\Omega_k = 0$ then there are no singularities and the decision process for the primary problem (1)–(4) is terminated on the non-singular controller $u^k = u^k(t)$. Otherwise, if mes $\Omega_k > 0$ then we look for a non-isolated point

$$\tau_k = \arg\max_{t \in \Omega_k} \hat{Q}_k(t) \tag{40}$$

and construct a set $T_k(\varepsilon) \subset \Omega_k$, $\varepsilon \in [0, 1]$ according to the rule

$$T_k(\varepsilon) = \left[\tau_k - \varepsilon\left(\tau_k - t_0^k\right), \tau_k + \varepsilon\left(t_1^k - \tau_k\right)\right], \qquad \left[t_0^k, t_1^k\right] \subseteq \Omega_k. \tag{41}$$

If $\hat{Q}_k(\tau_k) = 0$ then the singular controller $u^k = u^k(t)$ satisfies the second-order optimality criterion (35) and the decision process is terminated on the singular control $u^k = u^k(t)$. Otherwise, if $\hat{Q}_k(\tau_k) > 0$ then we compose a one-parameter family of controllers as

$$u_\varepsilon^k(t) = \begin{cases} \hat{u}^k(t), & t \in T_k(\varepsilon) \subset \Omega_k,\ \varepsilon \in [0, 1] \\ u^k(t), & t \notin T_k(\varepsilon), \end{cases} \tag{42}$$

calculate the best value of the parameter

$$\varepsilon_k = \arg\min_{\varepsilon \in [0, 1]} J\left(u_\varepsilon^k\right), \tag{43}$$

and set

$$u^{k+1}(t) = u_{\varepsilon_k}^k(t). \tag{44}$$

Having completed all these operations, we should return to the main body of the algorithm of the maximum principle [3–6] since this new controller $u^{k+1}(t)$ may not satisfy the maximum principle much longer.

It is obvious that for the computational scheme (38)–(44)

$$J\left(u_\varepsilon^k\right) - J(u^k) = -\hat{Q}_k(\tau_k) \cdot \varepsilon^2 + o_k(\varepsilon^2), \qquad \varepsilon \in [0, 1].$$

Whence the existence of solutions to problem (43) and the relaxation property $J(u^{k+1}) < J(u^k)$ follows.

Improving procedure (38)–(44) yields quite good results only if u^k contains singular sections defined by (39). Thereupon, the improving procedure will either find a singular optimal controller or just jump out of some deep local minimum back on the main solution algorithm. This is an advantage of the improving procedure.

Theoretically, the improving procedure might have been applied without waiting for the situation when (38) holds, but it would be inexpedient since this procedure involves a great deal of additional calculations.

In conclusion let us consider two examples which illustrate how to make use of the optimality criterion (35) as a verifying condition.

EXAMPLE 6.1.

$$\dot{x}_1 = u, \qquad x_1(0) = 1, \qquad |u(t)| \le 1, T = [0,2]$$
$$\dot{x}_2 = \frac{1}{2}x_1^2, \qquad x_2(2) = 0,$$

$$J(u) = \frac{1}{2}x_1^2(2) - x_2(0) \to \min.$$

As a controller which is suspected to be optimal let us take

$$u^*(t) = \begin{cases} -1, & t \in [0,1) \\ 0, & t \in [1,2] \end{cases}.$$

On this controller

$$x_1^*(t) = \begin{cases} 1-t, & t \in [0,1) \\ 0, & t \in [1,2], \end{cases} \qquad x_2^*(t) = \begin{cases} -\frac{1}{6}(1-t)^3, & t \in [0,1) \\ 0, & t \in [1,2]. \end{cases}$$

The vector adjoint BVP (10), (14), (15) has the form

$$\dot{\psi}_1 = -x_1^*\psi_2, \; \dot{\psi}_2 = 0; \qquad \psi_1(2) = -x_1^*(2), \; \psi_2(0) = -1.$$

Whence

$$\psi_1^*(t) = \begin{cases} -\frac{1}{2}(1-t^2), & t \in [0,1) \\ 0, & t \in [1,2], \end{cases} \qquad \psi_2^*(t) = -1.$$

Then

$$\Delta_v H(\psi^*, x^*, u^*, t) = \psi_1^*(v - u^*)$$
$$= \begin{cases} -\frac{1}{2}(1-t)^2(v+1), & t \in [0,1) \\ 0 \cdot v, & t \in [1,2] \end{cases}$$

which means that the controller $u^*(t)$ satisfies the maximum principle on $[0,1)$ and that $[1,2] = \Omega$ is a locality where the maximum principle degenerates, i.e., $u^*(t)$ is singular within $[1,2]$.

Let us write down all the constructions which are needed in order to apply the optimality criterion for the singular controller (35). The matrix adjoint BVP (11), (18) takes on the form

$$\begin{bmatrix} \dot{\Psi}_{11} & \dot{\Psi}_{12} \\ \dot{\Psi}_{21} & \dot{\Psi}_{22} \end{bmatrix} = -\begin{bmatrix} 0 & x_1^* \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{bmatrix} - \begin{bmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{bmatrix}\begin{bmatrix} 0 & 0 \\ x_1^* & 0 \end{bmatrix} - \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \tag{45}$$

$$\begin{bmatrix} \Psi_{11}(2) & \Psi_{12}(2) \\ \Psi_{21}(0) & \Psi_{22}(0) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \tag{46}$$

and after integration

$$\Psi_{11}^*(t) = t - 3, \qquad \Psi_{12}^*(t) = \Psi_{21}^*(t) = \Psi_{22}^*(t) \equiv 0$$

$$\Psi^*(t)\Delta_v f(x^*, u^*, t) = \begin{bmatrix} \Psi_{11}^*(t)(v - u^*(t)) \\ 0 \end{bmatrix}.$$

It should be noted that

$$\Delta_v \frac{\partial H(\psi^*, x^*, u^*, t)}{\partial x} = 0.$$

Then we solve the Cauchy problem

$$\begin{bmatrix} \dot{X}_{11} & \dot{X}_{12} \\ \dot{X}_{21} & \dot{X}_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ x_1^* & 0 \end{bmatrix}\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}, \qquad X(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and obtain

$$X_{11}^*(t) = 1, \qquad X_{12}^*(t) = 0,$$

$$X_{21}^*(t) = \begin{cases} t - \dfrac{1}{2}t^2, & t \in [0,1) \\ \dfrac{1}{2}, & t \in [1,2] \end{cases}, \qquad X_{22}^*(t) = 1.$$

Calculate matrix

$$L_0 + L_1 X^*(2) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 0 \end{bmatrix}$$

and its inverse

$$[L_0 + L_1 X^*(2)]^{-1} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 1 \end{bmatrix}$$

in order to compute

$$\Phi^*(2) = \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & 1 \end{bmatrix}$$

by formula (26) which is needed to write down an additional Cauchy problem (30), (31)

$$\begin{bmatrix} \dot{Y}_{11} & \dot{Y}_{12} \\ \dot{Y}_{21} & \dot{Y}_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ x_1^* & 0 \end{bmatrix} \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} - \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ x_1^* & 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_{11}(0) & Y_{12}(0) \\ Y_{21}(0) & Y_{22}(0) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & 1 \end{bmatrix}.$$

Finally, after integration

$$Y_{11}^*(t) = 0, \qquad Y_{12}^*(t) = 0,$$

$$Y_{21}^*(t) = \begin{cases} \frac{1}{2} - t + \frac{1}{2}t^2, & t \in [0, 1) \\ 0, & t \in [1, 2] \end{cases}, \qquad Y_{22}^*(t) = 1$$

and

$$[I - Y^*(t)] \Delta_v f(x^*, u^*, t)$$
$$= \begin{bmatrix} 1 & 0 \\ -Y_{21}^* & 0 \end{bmatrix} \begin{bmatrix} v - u^* \\ 0 \end{bmatrix} = \begin{bmatrix} v - u^* \\ -Y_{21}^*(v - u^*) \end{bmatrix}.$$

Thereby, the optimality criterion for the singular controller (35), (36) may be represented as

$$\begin{aligned} Q(u^*, v, t) &= \left\langle \Psi^*(t)\Delta_v f(x^*, u^*, t), [I - Y^*(t)]\Delta_v f(x^*, u^*, t)\right\rangle \\ &= \Psi^*_{11}(t)(v - u^*(t))^2 \\ &= (t-3)(v - u^*(t))^2 \le 0 \qquad \text{since } (t-3) < 0, t \in [0,2] \end{aligned}$$

which means that the singular controller on $\Omega = [1,2]$ and the optimal controller $u^* = u^*(t)$ satisfy both the maximum principle and the optimality criterion for the singular controller, i.e., both first and second-order necessary conditions for optimality hold in $u^* = u^*(t)$ whose optimality immediately follows from geometric interpretation of the problem as well.

EXAMPLE 6.2. In Example 6.1 we define the performance index in a different way

$$J(u) = -\frac{1}{2}x_1^2(2) - x_2(0)$$

and verify whether the same controller $u^* = u^*(t)$ of Example 6.1 is optimal. In this case

$$\dot{\psi}_1 = -x_1^*\psi_2, \qquad \psi_1(2) = +x_1^*(2), \qquad \text{and}$$

$$\psi_1^*(t) = \begin{cases} -\frac{1}{2}(1-t)^2, & t \in [0,1) \\ 0, & t \in [1,2] \end{cases}$$

$$\Delta_v H(\psi^*, x^*, u^*, t) = \psi_1^*(t)(v - u^*) = \begin{cases} \le 0, & t \in [0,1) \\ \equiv 0, & t \in [1,2]. \end{cases}$$

Obviously, u^* here also satisfies the maximum principle on $[0,2]$ and the definition of the singular controller on $\Omega = [1,2]$. In the matrix adjoint BVP (45), (46) $\Psi_{11}(2) = +1$ and therefore $\Psi_{11}^*(t) = t - 1$. The rest of the Ψ_{ij}^* are the same as in Example 6.1. Then

$$Q(u^*, v, t) = (t-1)(v - u^*(t))$$

which implies that u^* satisfies the second-order optimality criterion on $[0,1)$ but does not on $\Omega = [1,2]$, i.e., $u^* = u^*(t)$ cannot be an optimal controller.

Generally speaking, when one is facing some practical problem of the form (1)–(4) and has already found a suitable controller $u = u(t)$ which satisfies the maximum principle, we would suggest to verify this controller by using the optimality criterion (35) and apply the improving procedure (38)–(44) if needed.

REFERENCES

1. R. F. Gabasov, About optimality conditions for singular controllers, *Soviet Izv. Ser. Tech. Cybernetics* **5** (1968), 34–43.
2. R. F. Gabasov and F. M. Kirillova, "Singular Optimal Controls," Nauka, Moscow, 1973. [In Russian]
3. O. V. Vasiliev and A. I. Tyatyushkin, About one solution method for optimal control problems based on the maximum principle, *U.S.S.R. Comput. Math. and Math. Phys.* **21** (1981), 1376–1384.
4. O. Vasilieva and K. Mizukami, Convergence of improved successive approximation method for optimal control problems, *in* "Proceedings, Asian Control Conference, Tokyo, 1994," Vol. 2, pp. 901–904.
5. O. Vasilieva and K. Mizukami, Optimal control in boundary value problems, *Russian Izv. Ser. Math.* **12** (1994), 33–41.
6. O. Vasilieva and K. Mizukami, Dynamic process governed by the boundary value problem: Necessary conditions for optimality and method of solution, *Russian Tech. Cybernetics*, in press.