On the Maximum Principle for Systems with Delays

A. V. Kim, V. M. Kormyshev and A. V. Ivanov

Abstract In this article we present the Pontryagin maximum principle of a time-optimal control problem for general form of functional-differential equations. The obtained results are the direct generalization of the case for ordinary differential equations: if the delay disappear then the results turn into the classic Pontryagin maximum principle for finite dimensional systems. In this work we apply the methodology and constructions of the i-Smooth analysis.

1 Introduction

The delay phenomenon plays an important role in the study of processes arising in natural science, technology and society. First of all, this is due to the fact that the future development of many processes depends not only on their present state but is essentially influenced by their previous history. Such processes can be described mathematically using the functional-differential equations (hereinafter FDE). At present FDE theory is the well developed branch of the differential equations and offenly uses in description and modeling of automatic control processes with aftereffect, mechanics, technology, economics, medicine and other areas of human activity [6, 10].

This work is devoted to establishing the necessary optimality conditions in the form of Pontryagin's maximum principle for general FDEs. The discovery of the

A. V. Kim (✉) · V. M. Kormyshev · A. V. Ivanov
Ural Federal University, 620002 19 Mira Street, Ekaterinburg, Russia
e-mail: avkim@imm.uran.ru

V. M. Kormyshev
e-mail: vkormyshev@gmail.com

A. V. Ivanov
e-mail: avi@imm.uran.ru

famous Pontryagin maximum principle [12] started the development of the mathematical theory of optimal processes. This classic fundamental book already included a variant of the maximum principle for systems with discrete delays. The origin of the development of the theory of delayed optimal processes goes back to [7], where an analog of the Pontryagin maximum principle was proved for optimal systems with constant delays in state coordinates. The maximum principle was later proved for some classes of systems with distributed delays ([1, 2, 5, 11, 13]). However, there is no principle maximum variant for general form FDE, that is systems without a priory specification of delay types. In this work we apply i-Smooth analysis [8, 9] to obtain the Pontryagin maximum principle for general form FDEs. i-Smooth analysis allows to obtain results by using methods and arguments similar to ordinary differential equations. In our article we apply an analog of the methodology developed in [3] for deriving the Pontryagin maximum principle for finite-dimensional systems.

This article is organized as follows. In the second section, we obtain special conditions of optimality in the form of the Bellman functional by applying the i-smooth analysis. In the third section we use these relations to obtain the maximum principle for general form of FDEs.

2   Problem Statement and Preliminaries

In the article we consider a control system with delays

$$\dot{x} = f(x(t), x(t + s), u(t)), \tag{1}$$

where $x(t) = (x_1(t), x_2(t), \ldots, x_n(t))^T \in R^n$, $x(t + \cdot) = \{x(t + s), -\tau \leq s < 0\}$, $f(x, y(\cdot), u) : R^n \times Q[-\tau, 0) \times P \to R^n$; $Q[-\tau, 0)$ is the space of piecewise continuous n-dimensional functions $x(\cdot)$ on $[-\tau, 0)$ (right continuous at points of discontinuity) with the norm $\|x(\cdot)\|_Q = \sup_{-\tau \leq t < 0} \|x(t)\|$, $P \subseteq R^r$ is a control region; $h(x, y(\cdot)) \in H = R^n \times Q[-\tau, 0)$, $x_t = \{x(t), x(t + \cdot)\} \in H$.

The problem is to find a control which transfers the system (1) from a phase (functional) state (position) $h(x, y(\cdot)) \in H$ into a given point $x^* \in R^n$. Herewith as an initial position h we will consider various points of the phase space H.

We assume that further the following condition is valid

Assumption 1. For every position $h(x, y(\cdot)) \in H$ there is the time-optimal transition process from the position h into the point x^*.

We denote by $T[x, y(\cdot)]$ the optimal transition time from the position $h(x, y(\cdot)) \in H$ into a given point x^*. For the convenience we consider the functional

$$W[x, y(\cdot)] = -T[h], \tag{2}$$

which depends on $2n$ variables

$$W[x, y(\cdot)] = W[x^1, x^2, \ldots, x^n, y^1(\cdot), y^2(\cdot), \ldots, y^n(\cdot)].$$

We also assume that for the considered problem the following condition is also valid

Assumption 2. The functional $W[x, y(\cdot)]$ has the following partial and invariant derivatives

$$\frac{\partial W}{\partial x_1}, \frac{\partial W}{\partial x_2}, \ldots, \frac{\partial W}{\partial x_n}, \quad \partial W_{y^1}, \partial W_{y^2}, \ldots, \partial W_{y^n}.$$

which are invariantly continuous in domains.

Let $h(x_0, y_0(\cdot))$ be an arbitrary point of the phase space H, and $u_o \in P$ is an arbitrary point of the control region.

Consider a process which starts at a moment t_0 from the position h_0 under the constant control $u = u_0$. Therefore the phase trajectory of the process $x(t) = (x_1(t), x_2(t), \ldots, x_n(t))$ satisfies the following functional differential equation

$$\dot{x} = f(x(t), x(t + \cdot), u_0), \quad \text{for} \quad t > t_0 \tag{3}$$

and the initial condition

$$x_{t_0} = h_0. \tag{4}$$

It takes time $t - t_0$ to move along this trajectory from the point x_0 to the point $x(t)$. Applying from the moment t an optimal control we move from x_t into the terminal point x^* during the time $T[x_t]$.

Such movement from the point x_0 into the terminal point x^* takes time $(t - t_0) + T[x_t]$. Taking into account that optimal (minimal) time from the position (point) $h_0(x^0)$ is equal to $T[h_0] = T[x_{t_0}]$ we obtain the following inequality

$$T[x_{t_0}] \leq (t - t_0) + T[x_t],$$

from which (see (2)) we have

$$-W[x_{t_0}] \leq (t - t_0) - W[x_t].$$

Therefore

$$W[x_t] - W[x_{t_0}] \leq t - t_0,$$

$$\frac{W[x_t] - W[x_{t_0}]}{t - t_0} \leq 1.$$

Proceeding in the last inequality to limit as $t \to t_0$ we obtain

$$\frac{d}{dt} W[x_t]|_{t=t_0} \leq 1. \tag{5}$$

The left-hand side of the inequality (5) can be expressed in terms of the partial and the invariant derivatives, then (5) can be presented in the form

$$\frac{\partial W[x_0, y_0(\cdot)]}{\partial x} \cdot f(x_0, y_0(\cdot), u_0) + \partial W[x_0, y_0(\cdot)] \leq 1.$$

$h = \{x_0, y_0\}$ and u_0 are arbitrary elements, therefore for any position $h = \{x_0, y(\cdot)\} \in H$ and every point $u \in P$ the following relation is valid

$$\frac{\partial W[x, y(\cdot)]}{\partial x} \cdot f(x, y(\cdot), u) + \partial W[x, y(\cdot)] \leq 1. \tag{6}$$

Let $\{x(\cdot), y(\cdot)\}$ be the time-optimal process of transferring the system from the position h_0 into the point x^*, and $[t_0, t_1]$ is the corresponding time interval, therefore: $x_{t_0} = h_0$, $x_{t_1} = x_1$ and $t_1 = t_1 + T[h_0]$.

The process satisfies the equation

$$\dot{x}(t) = f(x_t, u(t)), \quad t_0 \leq t \leq t_1. \tag{7}$$

Movement along the optimal trajectory from the position $h_0(x_0, y_0(\cdot))$ to a point $x(t)$ takes $t - t_0$, and from the point $x(t)$ to the terminal point x^* the system moves during $t_1 - t$, then $T[h_0] - (t - t_0)$ is the minimal time of transferring the system from the state x_t into the point x^*, that is

$$T[x_t] = T[h_0] - (t - t_0).$$

By virtue of $T[h] = -W[h]$ we obtain

$$W[x_t] = -W[h_0] + (t - t_0),$$

$$W[x(t), x(t + \cdot)] = -W[h_0] + (t - t_0).$$

Differentiating this equality by t we obtain

$$\sum_{i=1}^{n} \frac{\partial W[x_t]}{\partial x^i} \cdot \dot{x}^i(t) + \partial_{y^0} W[x(t)] = 1.$$

Taking into account (7) we have

$$\sum_{i=1}^{n} \frac{\partial W[x_t]}{\partial x^i} \cdot f^i(x_t, u(t)) + \partial_y W[x(t)] = 1, \quad t_0 \leq t \leq t_1. \tag{8}$$

Thus for every optimal process the equality (8) is valid during the process. Consider the functional

$$B[x, y(\cdot), u] = \sum_{i=1}^{n} \frac{\partial W[h]}{\partial x^i} \cdot f^i(h, u) \tag{9}$$

then relations (6), (8) can be presented in the following form

$$B[h, u] \leq 1, \text{ for every } h \in H \text{ and } u \in P. \tag{10}$$

$$B[h, u] = 1, \text{ along any optimal process } (x(\cdot), y(\cdot)). \tag{11}$$

Thus the following theorem is proved

Theorem 1 *If the assumptions for the control system (1) and a fixed terminal point x^* are valid, then the relations (10) and (11) take place.*

This theorem presents the essence of the dynamic programming method for systems with delays. Its main mathematical relation can be expressed in other form.

From (11) with $t = t_0$ we have $B[h_0, u(t_0)] = 1$. Taking into account (10) we obtain relation

$$\max_{u \in P} B[h, u] = 1, \quad \forall h \in H,$$

or equivalently

$$\max_{u \in P} \sum_{i=1}^{n} \frac{\partial W[h]}{\partial x^i} \cdot f^i(x, y(\cdot), u) + \partial_y W[x(t)] = 1, \quad \forall h \in H. \tag{12}$$

3 Maximum Principle

Further along with the assumptions 1,2 we suppose that the following conditions are satisfied.
  Assumption 1.

- The functional $W[x, y(\cdot)]$ has invariantly continuous derivatives with respect to $x^i, i = 1, \ldots, n$, up to the second order, that is functionals

$$\frac{\partial W[h]}{\partial x^i}, \frac{\partial^2 W[h]}{\partial x^i \partial x^j}, \quad i, j = 1, \ldots, n.$$

  are invariantly continuous.
- Functionals $f^i(x, y(\cdot), u), i = 1, \ldots, n$ have invariantly continuous partial derivatives

$$\frac{\partial f^i(h, u)}{\partial x^j}, \quad i, j = 1, \ldots, n.$$

Let $(x(t), u(t)), t_0 \leq t \leq t_1$ be the time-optimal process transferring the system (1) from the position h_0 into the terminal point x^*.
  Fix a moment $t \in [t_0, t_1)$ and consider the functional $B(x, y(\cdot), u(t))$ of variables $x, y(\cdot)$.

From the definition of the functional B (see. 9) and the hypothesis 3 it follows that the functional $B(x, y(\cdot), u(t))$ has the invariantly continuous derivatives with respect to variables $x^1, x^2, \ldots, x^n$:

$$\frac{\partial B(x, y(\cdot), u(t))}{\partial x^k} = \sum_{i=1}^{n} \frac{\partial^2 W[h]}{\partial x^i \partial x^k} \cdot f^i(h, u(t)) + \sum_{i=1}^{n} \frac{\partial W[h]}{\partial x^i} \cdot \frac{\partial f^i(x, y(\cdot), u(t))}{\partial x^k}, \quad k = 1, \ldots, n.$$

(13)

By virtue of (10), (11) we have

$$B[h, u(t)] \leq 1, \quad \forall h \in H;$$

$$B[h, u(t)] = 1, \quad \forall h = x_t.$$

These two relations mean that the functional achieves the maximum at the element $h = x_t$.

Therefore, if we fix $x(t + \cdot)$ and $u(t)$ in the functional $B[x, x(t + \cdot), u(t)]$, and consider it as the function of x, then this function has the maximum at the point $x = x(t)$. Hence its partial derivatives with respect to $x^1, x^2, ldots, x^n$ are equals to zero at this point:

$$\sum_{i=1}^{n} \frac{\partial^2 W[x_t]}{\partial x^i \partial x^k} \cdot f^i(h, u(t)) + \sum_{i=1}^{n} \frac{\partial W[x_t]}{\partial x^i} \cdot \frac{\partial f^i(x, y(\cdot), u(t))}{\partial x^k} = 0, \quad k = 1, \ldots, n.$$

(14)

(see (13)).

Differentiating the function $\frac{\partial W[x_t]}{\partial x^k}$ with respect to t and taking into account (7), we find

$$\frac{d}{dt}\left(\frac{\partial W[x_t]}{\partial x}\right) = \sum_{i=1}^{n} \frac{\partial W[x_t]}{\partial x^k \partial x^i} \dot{x}^i(t) = \sum_{i=1}^{n} \frac{\partial W[x_t]}{\partial x^k \partial x^i} f^i(x_t, u(t)), \quad k = 1, \ldots, n.$$

(15)

Then relation (15) can be presented in the following form:

$$\frac{d}{dt}\left(\frac{\partial W[x_t]}{\partial x}\right) = \sum_{i=1}^{n} \frac{\partial W[x_t]}{\partial x^k \partial x^i} \dot{x}^i(t) = \sum_{i=1}^{n} \frac{\partial W[x_t]}{\partial x^k \partial x^i} f^i(x_t, u(t)), \quad k = 1, \ldots, n.$$

(16)

(note, that $\frac{\partial^2 W}{\partial x^k \partial x^i} = \frac{\partial^2 W}{\partial x^i \partial x^k}$ due to continuity of the second derivatives).

Formulas (10)–(12), and (16) do not include the functional W, but only its partial derivatives with respect to $x^1, \ldots, x^n$: $\frac{\partial W}{\partial x^1}, \ldots, \frac{\partial W}{\partial x^n}$, so, for the convenience, we will use the following notation:

$$\frac{\partial W[x_t]}{\partial x^1} = \psi_1[t], \frac{\partial W[x_t]}{\partial x^2} = \psi_2[t], \ldots, \frac{\partial W[x_t]}{\partial x^n} = \psi_n[t]. \tag{17}$$

Then the functional B (see (9)) can be presented in the form:

$$B[x_t, y(\cdot), u(t)] = \sum_{i=1}^{n} \psi_i[t] \cdot f^i(x_t, u(t))$$

and the relation (11) becomes

$$\sum_{i=1}^{n} \psi_i[t] \cdot f^i(x_t, u(t)) \equiv 1 \text{ for optimal process } (x(t), u(t)), \ t_0 \le t \le t_1. \quad (18)$$

Besides, according to (10)

$$\sum_{i=1}^{n} \psi_i[t] \cdot f^i(x_t, u(t)) \le 1 \text{ for every point } u \in P \text{ and all } t_0 \le t \le t_1. \quad (19)$$

Finally, relations (15) can be presented in the following form:

$$\dot{\psi}_k[t] + \sum_{i=1}^{n} \psi_i[t] \cdot \frac{f^i(x_t, u(t))}{\partial x^k} = 0, \quad k = 1, \ldots, n. \quad (20)$$

In summary, if $(x(t), u(t))$, $t_0 \le t \le t_1$ is the optimal process, then there exist functionals $\psi_1[t], \psi_2[t], \ldots, \psi_n[t]$ (defined by (16)), such that the relations are valid.

The form of the left-hand sides of (17), (18) lead us to consideration of the functional

$$H[\psi, x, y(\cdot), u] = \sum_{i=1}^{n} \psi_i \cdot f^i(x, u) = \psi_1 \cdot f^1(x, u) + \cdots + \psi_n f^n(x, u), \quad (21)$$

depending on $2n + r$ variables $\psi_1, \ldots, \psi_n, x^1, \ldots, x^n, u^1, \ldots, u^r$. In terms of this functional relations (17), (18) can be presented in the form of two following relations:

$$H[\psi[t], x_t, y(\cdot), u(t)] \equiv 1 \text{ for optimal process } (x(t), u(t)), \ t_0 \le t \le t_1, \quad (22)$$

where $\psi[t] = (\psi_1[t], \ldots, \psi_n[t])$ is defined by (16).

$$H[\psi[t], x_t, y(\cdot), u(t)] \le 1 \text{ for every point } u \in P \text{ and all } t_0 \le t \le t_1. \quad (23)$$

Relations (22) and (23) can be unified in a compact form

$$\max_{u \in P} H[\psi[t], x(t), u(t)] = H[\psi[t], x_t, u(t)], \quad t_0 \le t \le t_1. \quad (24)$$

Additionally, the relation (19) can be presented in the form:

$$\dot{\psi}_k[t] = -\frac{\partial H[\psi[t], x_t, y(\cdot), u(t)]}{\partial x^k}, \quad k = 1, \ldots, n. \tag{25}$$

Thus, if $(x(t), u(t))$, $t_0 \le t \le t_1$ is the optimal process, then a function $\psi[t] = (\psi_1[t], \ldots, \psi_n[t])$ exists and the relations (22), (24), (25) are valid, in which the functional H is defined by (21).

Formulas (21), (22), (24), (25) do not contain explicitly the functional $W[x, y(\cdot)]$, so equalities (17), representing the functions $\psi_1[t], \ldots, \psi_n[t]$ by the functional W, do not give us additional information and will be out of our consideration. Relation (25) is the system of equations which satisfy these functions. Note that the functions $\psi_1[t], \ldots, \psi_n[t]$ are nontrivial solutions of this system (that is the functions do not equal to zero at the same time); indeed, if at some moment t we have $\psi_1[t] = \ldots = \psi_n[t] = 0$, then from (21) we obtain $H[\psi[t], x_t, u(t)] = 0$ that contradicts to equality (22). Thus we obtain the following theorem in the form of **the maximum principle**.

Theorem 2 *Let for the control system*

$$\dot{x}(t) = f(x(t), x(t+s), u(t)), \quad u \in P, \tag{26}$$

and a terminal point x^*, assumptions 1, 2 and 3 are valid, and let $(x(t), u(t))$, $t_0 \le t \le t_1$ be a process transferring the system from an initial state $h_0 \in H$ into the final point x_1. Consider a functional depending on variables $x^1, \ldots, x^n, u^1, \ldots, u^r$ and auxiliary variables $\psi_1, \ldots, \psi_n$ (cf. (21)):

$$H[\psi, x, y(\cdot), u] = \sum_{i=1}^{n} \psi_i f^i(x, y(\cdot), u). \tag{27}$$

Consider for the auxiliary variables the system of differential equations

$$\dot{\psi}_k[t] = -\frac{\partial H[\psi[t], x_t, u(t)]}{\partial x^k}, \quad k = 1, \ldots, n, \tag{28}$$

where $(x(t), u(t))$ is the process under consideration (cf. (25)). Then, if $(x(t), u(t))$, $t_0 \le t \le t_1$ is the time-optimal process, then there exists nontrivial solution $\psi_1[t], \ldots, \psi_n[t]$, $t_0 \le t \le t_1$ of the system (28) such that for every moment $t_0 \le t \le t_1$ the following maximum condition

$$H[\psi[t], x_t, u(t)] = \max_{u \in P} H[\psi[t], x(t), y(\cdot), u]. \tag{29}$$

(cf. (24)) and the equality (cf. (22)

$$H[\psi[t], x_t, u(t)] = 1$$

are valid.

The Theorem 2 presents necessary conditions for optimality of systems with delays in the form of the maximum principle.

References

1. Banks, H.T.: Necessary conditions for control problems with variable time lags. SIAM J. Control (1968). https://doi.org/10.1137/0306002
2. Bokov, G.V.: Pontryagin's maximum principle of optimal control problems with time-delay. J. Math. Sci. (2011). https://doi.org/10.1007/s10958-011-0208-y
3. Boltyanskii, V.G.: Mathematical Methods of Optimal Control. Nauka, Moscow (1968)
4. Fleming, W.H., Rishel, R.W.: Deterministic and Stochastic Optimal Control. Springer, New York (1975)
5. Göllmann, L., Kern, D., Maurer, H.: Optimal control problems with delays in state and control variables subject to mixed controlstate constraints. Optim. Control Appl. Meth. (2008). https://doi.org/10.1002/oca.843
6. Hale, J.K.: Theory of Functional Differential Equations. Springer, New York (1977)
7. Kharatishvili, G.L.: Maximum principle in the theory of optimal processes with delays Dokl. Akad. Nauk SSSR. **136**(1), 39–42 (1961)
8. Kim, A.V.: Functional Differential Equations. Application of i-smooth Analysis. Kluwer Academic Publishers, Netherlands (1999)
9. Kim, A.V.: i-Smooth Analysis. Theory and Applications (Wiley, 2015)
10. Kolmanovskii, V.B., Myshkis, A.D.: Introduction to the Theory and Applications of Functional Differential Equations. Kluwer Academic Publishers, Netherlands (1999)
11. Lalwani, C.S., Desai R.C.: The maximum principle for systems with time-delay. Int. J. Control (1973). https://doi.org/10.1080/00207177308932508
12. Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., Mishchenko, E.F.: The Mathematical Theory of Optimal Processes. Gordon and Breach, New York (1962)
13. Teo, K.L., Moore, E.J.: Necessary conditions for optimality for control problems with time delays appearing in both state and control variables. J. Optim. Theory Appl. (1977). https://doi.org/10.1007/BF00933450