ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XII, 3 (1971)

R. GESSING (Gliwice)

TIME-OPTIMAL CONTROL OF PLANTS WHOSE TRANSFER FUNCTIONS CONTAIN ZEROS

1. Introduction. In a few papers, written so far on the time-optimal control of plants with transfer functions containing zeros, the problem does not seem to be ultimatively formulated and solved. Dealing with such plants it is not enough to say that the problem of "time-optimal control" is under consideration but, due to the discontinuity of trajectories in a phase-space Y, it has to be more precisely stated.

To determine the switch curves (which is necessary for a synthesis of time-optimal systems) Chang and Sheldon [2] have introduced new variables x such that in the new phase-space X the trajectories become continuous. Without any justification it is then accepted that the "Bang-Bang" control is the optimum one also in a case of plants whose transfer-functions contain zeros. The problem is rather marginally treated and devoid of precise statement.

Also in [3] the problem has not been precisely formulated. Applying the same variables x, Dmowski tries to justify optimality of the "Bang-Bang" control for the considered case. Without a specification of the initial point x_0 in a new phase-space X and with a wrong definition of the terminal point x_1 the final conclusion has not been achieved.

The precise formulation of the problem has been given by Athans and Falb in [1]. Discussing the problem from the viewpoint of the synthesis of time-optimal systems, authors are mainly concerned with the terminal point of the trajectory. They have shown that in the new phase-space X the terminal point x_1 belongs to a given set. There is, however, no discussion on the determination of the initial point x_0 which is accepted to be an arbitrary point of X.

The author is highly convinced that more attention has to be paid to the initial point x_0 of a trajectory and that the problem has to be precisely formulated in the starting stage of the control. It occurs that (as it will be shown in the sequel) without some additional information, or lacking the precise statement of the problem at the beginning of the control, the starting point x_0 cannot be determined and so the problem has no complete solution.

This will be shown on a simple example. Let us consider a plant which can be described with the differential equation

$$\text{(a)} \qquad \dot{y}+ay = \dot{w}$$

or – which is equivalent – with the transfer function

$$\text{(b)} \qquad \frac{Y(s)}{W(s)} = K(s) = \frac{s}{s+a}.$$

Here w stands for an input (control) of the plant, y for an output, a is a constant, and s is the differential operator. One has to find the control $w(t)$ which, while satisfying the restriction $|w(t)| \leqslant 1$, conducts the coordinate y of the plant from a given value y_0 (which is attained just before the initial moment t_0 of starting the control, more exactly $y(t_0^-) = y_0$) in the shortest possible time to the required value $y_1 = 0$. Having reached this value the coordinate y has to remain constant ($y = y_1 = 0$). Only continuous controls $w(t)$ are admitted or those having first order discontinuity points, i.e. for each discontinuity point τ there exist both the left side limit $w(\tau^-)$ and the right side limit $w(\tau^+)$.

A plant with the given transfer function is shown in Fig. 1a. Quantities w and y are the input voltage (forcing) and the output voltage, respectively, of an electric system RC (for which $a = 1/RC$). From the system RC (or studying solutions of the differential equation (a)) it can be seen that the knowledge of a control $w(t)$ for $t > t_0$, as well as the initial condition $y(t_0^-) = y_0$, is not sufficient for the determination of the course of $y(t)$ for $t > t_0$. In addition one has to know the value of $w(t_0^-)$ which together with $y(t_0^-) = y_0$ is necessary for the determination of the initial condition $x(t_0) = x_0$ (voltage on the capacity C). The value of x is namely the new continuous variable in a new space X.

Two problems may then be brought up:

PROBLEM I. *To find an optimal control $w(t)$, having given the initial point y_0, the terminal point $y_1 = 0$, and the value $w(t_0^-) = w_0$.*

PROBLEM II. *To find an optimal control $w(t)$, as well as the value of $w(t_0^-)$, having given only the initial point y_0 and the terminal point $y_1 = 0$.*

Those two problems will be given the precise formulation in the sequel and solved both in a case when restrictions are imposed upon the signal $w(t)$ itself and in the case when they are imposed upon the k-th derivative of the signal (Appendix 2).

As an example, in Fig. 1b the solution of Problem I is shown for the above system and the selected values: $a = 1$, $y_0 = 2$, $w(t_0^-) = 0$ $y_1 = 0$; and in Fig. 1c the solution of Problem II for $y_0 = 2$, $y_1 = 0$. In the latter case, for an optimal control $w(t)$ with a found value $w(t_0^-) = 1$, the control time is equal to $T = t_1 - t_0 = 0$,

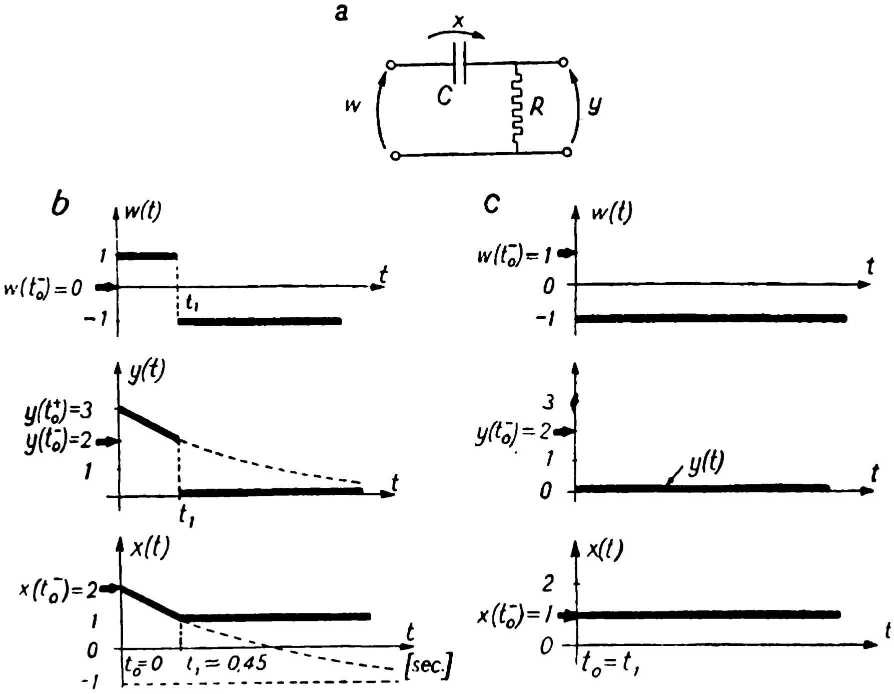


Fig. 1

a — a plant described by a differential equation (a); *b* — optimal runs in the sense of Problem I; *c* — optimal runs in the sense of Problem II

The author wants to express here his sincere gratitude to Professor Stefan Węgrzyn whose valuable remarks and directions were of much help to make more precise the results presented in this paper. Finally, the author takes this opportunity to express his thanks to Professor Czesław Olech for his precious comments concerning solutions ot the differential equation of a plant.

2. Formulation of the problem. We shall discuss here the control of linear plants with a single input w and a single output y which may be described: either by a differential equation of the form(¹)

$$(1) \qquad y^{(n)} + a_1 y^{(n-1)} + \ldots + a_n y = b_{n-m} w^{(m)} + b_{n-m+1} w^{(m-1)} + \ldots + b_n w,$$

(¹) $y^{(i)}$ stand here for the i-th derivatives of the function $y = y(t)$ and $w^{(j)}$ for the j-th derivatives of $w = w(t)$.

where $m \leqslant n$; a_i, b_j $(i = 1, 2, \ldots, n;\ j = n-m, n-m+1, \ldots, n)$ are constant coefficients and $b_{n-m} \neq 0$, or (wich is equivalent) by a transfer function of the form

$$(2) \qquad K(s) = \frac{Y(s)}{W(s)} = \frac{b_{n-m}s^m + b_{n-m+1}s^{m-1} + \ldots + b_n}{s^n + a_1 s^{n-1} + \ldots + a_n}.$$

We assume that the input control $w(t)$ is a piecewise continuous function having at most finite number of first order discontinuity points in any limited interval of time. This means that for each discontinuity point $t = \tau$ there exist both limits

$$(3) \qquad w(\tau^-) = \lim_{\substack{t\to\tau \\ t<\tau}} w(t) \quad \text{and} \quad w(\tau^+) = \lim_{\substack{t\to\tau \\ t>\tau}} w(t)$$

and the Rieman integral $\int_0^t w(\vartheta)\,d\vartheta$ is a continuous function of t at the point τ.

We assume also that the derivatives $w^{(j)}(t)$ $(j = 1, 2, \ldots, m-1)$ of the function $w(t)$ may be discontinuous or even not existing at a finite number of points in the interval of time in question. At each of these points, however, there exist both the left side limits $w^{(j)}(\tau^-)$ and the right side limits $w^{(j)}(\tau^+)$.

Furthermore we assume that the values of $w(t)$ have to belong to the interval

$$(4) \qquad a \leqslant w(t) \leqslant \beta,$$

where a and β are given arbitrary constants.

The value of $w(t)$ at a discontinuity point $t = \tau$ is of no importance for the further discussion. To avoid ambiguity we assume the left side continuity of $w(t)$, i.e. $w(\tau) = w(\tau^-)$. The control $w(t)$ satisfying all above assumptions in the sequel will be called the *admissible control.*

For an admissible control $w(t)$ its derivatives $w^{(j)}(t)$ occuring in equation (1), if understood in the ordinary sense, are also piecewise continuous functions not determined(2) at the discontinuity points of $w(t)$. The solution of equation (1), in an interval where its right-hand side is a continuous function, may be easily found. For the unique determination of the solution $y(t)$ in an arbitrary time interval

$$(5) \qquad t_0' < t < t_1''$$

it is necessary to know the initial conditions at the points belonging to any of the continuity intervals of the right side of equation (1) which make up together the interval (5).

(2) In the sequel we shall understand the derivative in a sense of a distributional operator, so that $w^{(j)}(t)$ and $y^{(i)}(t)$ will be distributions (see [4]).

To make unique the determination of the solution $y(t)$ by initial conditions we will give the following definition of the solution of equation (1).

The solution $y(t)$ of equation (1) *is a function for which the distribution of the left hand side of equation* (1) *is equal to the distribution of the right hand side of that equation.*

Now we shall show that the so defined solution $y(t)$ of equation (1) is uniquely determined by the initial conditions given at an arbitrary continuity point t_0 of the vector-function[3]

$$\boldsymbol{w}(t) = \big(w(t), w^{(1)}(t), \ldots, w^{(m-1)}(t)\big).$$

If we introduce the notation

$$\begin{aligned} \Delta y^{(i)}(\tau) &= y^{(i)}(\tau^+) - y^{(i)}(\tau^-), \\ \Delta w^{(j)}(\tau) &= w^{(j)}(\tau^+) - w^{(j)}(\tau^-), \end{aligned} \tag{6}$$

then, as it is proved in Appendix 1, for the solution $y(t)$ the following relations hold

$$\Delta y^{(i)}(\tau) = \gamma_0 \Delta w^{(i)}(\tau) + \gamma_1 \Delta w^{(i-1)}(\tau) + \ldots + \gamma_i \Delta w(\tau), \qquad i = 1, 2, \ldots, n, \tag{7}$$

where the constants γ_i are defined by a recurrent formula

$$\begin{aligned} \gamma_0 &= b_0, \\ \gamma_i &= b_i - a_1\gamma_{i-1} - a_2\gamma_{i-2} - \ldots - a_i\gamma_0, \qquad i = 1, \ldots, n. \end{aligned} \tag{8}$$

Since coefficients b_i, $i = 0, 1, \ldots, n-m-1$, in equation (1) are equal to zero, then

$$b_i = 0 \quad \text{and} \quad \gamma_i = 0 \quad \text{for } i = 0, 1, \ldots, n-m-1. \tag{9}$$

Formulae (7) and (8) enable us to determine the initial conditions $y^{(i)}(\tau^+)$, $i = 0, 1, \ldots, n-1$, in the next continuity interval of the derivatives $y^{(i)}(t)$, if we know the values of $y^{(i)}(\tau-)$ (from the preceding interval of continuity) and the quantities $\Delta w^{(j)}(\tau)$, $j = 0, 1, \ldots, m-1$ (those are known if the function $w(t)$ is known in the interval (5)). Hence the initial conditions $y^{(i)}(t_0)$, $i = 0, 1, \ldots, n-1$, given at a point t_0 from an arbitrary interval of continuity of the vector-function $\boldsymbol{w}(t)$, determine uniquely the solution $y(t)$ of equation (1) in the whole interval (5) (if only function $w(t)$ is known in this interval).

---

[3] t is a continuity point of the vector-function $\boldsymbol{w}(t)$ if it is a continuity point of all its components $w^{(j)}(t)$, $j = 1, 2, \ldots, m$. Of cource, the knowledge of the function $w(t)$ allows for an easy construction of the vector-function $\boldsymbol{w}(t)$.

From (7)-(9) it follows that in the case $m < n$ the derivatives $y^{(i)}(t)$ $(i = 0, 1, \ldots, n-m-1)$ are continuous ([4]) and only the derivatives of higher order $(i = n-m, n-m+1, \ldots, n)$ have discontinuity points at which their values do not exist.

In an n-dimensional phase-space Y, with coordinates $y, y^{(1)}, \ldots, y^{(n-1)}$, the point corresponding to the solution of equation (1) (for a given piecewise continuous function $w(t)$) sweeps over a phase trajectory which is also a piecewise continuous line. For the discontinuity points $t = \tau_i$ of $\boldsymbol{w}(t)$ the corresponding points of the phase trajectory do not exist. Phase trajectories in space Y are determined by a system of values $y(t)$, $y^{(1)}(t), \ldots, y^{(n-1)}(t)$ and therefore the vector $\boldsymbol{y}(t) = \big(y(t), y^{(1)}(t), \ldots \ldots, y^{(n-1)}(t)\big)$ will be called the *solution of equation* (1). To a solution $\boldsymbol{y}(t)$ there corresponds in space Y a phase trajectory which will be called the *trajectory* $\boldsymbol{y}(t)$.

Consider now two points $\boldsymbol{y}_0$ and $\boldsymbol{y}_1$ in space Y. We shall say that the admissible control $\boldsymbol{w}(t)$ conducts a point in the phase space from the position $\boldsymbol{y}_0$ to the position $\boldsymbol{y}_1$ in the length of time $T = t_1 - t_0 \geqslant 0$ if the solution $\boldsymbol{y}(t)$ (corresponding to the control $w(t)$) of equation (1) with the initial condition $\boldsymbol{y}(t_0^-) = \boldsymbol{y}_0$ satisfies also the relation $\boldsymbol{y}(t_1^+) = \boldsymbol{y}_1$ ([5]).

From (7) it follows that if t_0 is a discontinuity point ([6]) of the vector-function $\boldsymbol{w}(t)$ then the initial conditions $\boldsymbol{y}(t_0^+)$ in a time interval $t_0 < t < t_1$ (for a solution $\boldsymbol{y}(t)$ satisfying $\boldsymbol{y}(t_0^-) = \boldsymbol{y}_0$) depend also on the value of $\boldsymbol{w}(t_0^-)$, i.e. on the terminal value (in the interval of time $t < t_0$) of the control $\boldsymbol{w}(t)$ which has conducted the point in space Y to the position $\boldsymbol{y}_0$ $(\boldsymbol{y}(t_0^-) = \boldsymbol{y}_0)$. Therefore, if we know the function $\boldsymbol{w}(t)$ for $t > t_0$ and the initial condition $\boldsymbol{y}(t_0^-) = \boldsymbol{y}_0$ but the value of $\boldsymbol{w}(t_0^-)$ remains unknown, we cannot find the solution $y(t)$ of equation (1) for the time interval $t > t_0$, satisfying the condition $\boldsymbol{y}(t_0^-) = \boldsymbol{y}_0$.

Bearing in mind what is written above, the problems which will be discussed in the sequel may be formulated as follows:

PROBLEM I. *There are given two points* $\boldsymbol{y}_0 = (y_0, y_0^{(1)}, \ldots, y_0^{(n-1)})$ *and* $\boldsymbol{y}_1 = (y_1, 0, \ldots, 0)$ *in the phase-space* Y *and a vector* $\boldsymbol{w}(t_0^-) = \boldsymbol{w}_0$. *To find an admissible control* $\boldsymbol{w}(t)$, $t > t_0$, *which in the shortest possible time* $T = t_1 - t_0$ *conducts a point in space* Y *from position* $\boldsymbol{y}_0$ *to position* $\boldsymbol{y}_1$.

([4]) Since $y^{(i)}(\tau^-) = y^{(i)}(\tau^+)$ for $i = 0, 1, \ldots, n-m-1$, and at the point τ we may assume $y^{(i)}(\tau) = y^{(i)}(\tau^-)$ as for the function $w(t)$.

([5]) Conditions $\boldsymbol{y}(t_0^-) = \boldsymbol{y}_0$ and $\boldsymbol{y}(t_1^+) = \boldsymbol{y}_1$ have the following meaning: just before the moment t_0 of starting the control the point in space Y tends to position $\boldsymbol{y}_0$ and just after the moment t_1 it "reaches" the position $\boldsymbol{y}_1$. One of the moments t_0 and t_1 may be given in advance.

([6]) τ is a discontinuity point of the vector-function $\boldsymbol{w}(t)$ if it is a discontinuity point of one or more of its components $w^{(i)}(t)$, $i = 0, 1, \ldots, m-1$.

Having reached the terminal position the point has to remain there ($\boldsymbol{y}(t) = \boldsymbol{y}_1$ *for* $t > t_1$).

PROBLEM II. *The only difference from the former one is that the vector* $\boldsymbol{w}(t_0^-)$ *no longer is given in advance. It is also to be found out. Now the control* $\boldsymbol{w}(t)$ *has to be admissible for* $t > t_0 - \varepsilon$ *where* ε *is a sufficiently small positive number.*

It follows from equation (1) that the relation $\boldsymbol{y}(t) = \boldsymbol{y}_1$ (that is $y(t) = y_1$) is fulfilled for $t > t_1$ if and only if the right side of equation (1) is equal to $a_n y_1$ for $t > t_1$. Therefore, if we introduce new variables $y^* = y - y_1$ and $w^* = w - w_1$, where $w_1 = a_n y_1 / b_m$ (7), then the equation (1) for the new variables remains the same. Instead of condition (4) now we get

$$a^* = a - w_1 \leqslant w^* \leqslant \beta - w_1 = \beta^*, \tag{10}$$

which is of the same form (only the values of a^* and β^* are different from those without asterisk). To the point $\boldsymbol{y}_1 = (y_1, 0, \ldots, 0)$ from the space Y there corresponds the point $\boldsymbol{y}_1^* = (0, 0, \ldots, 0)$ in a new space Y^* (with the coordinates $y^*, y^{*(1)}, \ldots, y^{*(n-1)}$).

We may then assume, without any loss in generality, that in both problems $\boldsymbol{y}_1 = 0$ (all its coordinates are equal to zero).

The control $\tilde{w}(t)$, which is to be found in Problem I (or Problem II), will be called the *optimal control*, and the trajectory $\tilde{\boldsymbol{y}}(t)$ (determined by a solution of equation (1) with $w(t) = \tilde{w}(t)$ and such that $\tilde{\boldsymbol{y}}(t_0^-) = \boldsymbol{y}_0$) — the *optimal trajectory in space* Y.

3. Introduction of new variables. Let us introduce new variables $x^1, x^2, \ldots, x^n$ which (as in [1], [2] and [3]) are defined by

$$\begin{aligned}
x^1 &= y, \\
x^2 &= y^{(1)}, \\
&\cdots\cdots\cdots\cdots \\
x^{n-m} &= y^{(n-m-1)}, \\
x^{n-m+1} &= y^{(n-m)} - \gamma_{n-m} w, \\
x^{n-m+2} &= y^{(n-m+1)} - \gamma_{n-m} w^{(1)} - \gamma_{n-m+1} w, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
x^n &= y^{(n-1)} - \gamma_{n-m} w^{(m-1)} - \gamma_{n-m+1} w^{(m-2)} - \ldots - \gamma_{n-1} w.
\end{aligned} \tag{11}$$

(7) Here we assume that $b_m \neq 0$.

In the new system of coordinates the equation (1) may be written in a form of the system of equations

$$(12)\qquad \begin{aligned} \dot{x}^1 &= x^2, \\ &\dots\dots\dots \\ \dot{x}^{n-m-1} &= x^{n-m}, \\ \dot{x}^{n-m} &= x^{n-m+1} + \gamma_{n-m} w, \\ \dot{x}^{n-m+1} &= x^{n-m+2} + \gamma_{n-m+1} w, \\ &\dots\dots\dots\dots\dots\dots \\ \dot{x}^n &= -a_n x^1 - a_{n-1} x^2 - \dots - a_1 x^n + \gamma_n w. \end{aligned}$$

There are no derivatives of $w(t)$ in the system of differential equations (12). Comparing the relations (7) and (11) we may reach the conclusion that, for an admissible control $w(t)$, relations (11) transform the so far considered solutions of equation (1) into the *continuous solutions* $x^i(t)$ $(i = 1, 2, \dots, n)$ of the system of equations (12). Thus, in the sequel we will study the *continuous solutions* of the system of equations (12) and we will write them as a vector $\boldsymbol{x}(t) = \big(x^1(t), x^2(t), \dots, x^n(t)\big)$. Therefore, in space X with the coordinates $x^1, x^2, \dots, x^n$ the phase trajectories, corresponding to solutions $\boldsymbol{x}(t)$ of equations (12) for an arbitrary admissible control $w(t)$, are the continuous lines.

Let us consider now two points $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$ in space X. To find an optimal control $\tilde{w}(t)$, which in the shortest time $T = t_1 - t_0$ conducts the point $\boldsymbol{x}$ in the space X from the position $\boldsymbol{x}_0$ to the position $\boldsymbol{x}_1$ (in a case when the movement is described by equations (12)) may be applied the Maximum Principle of Pontryagin ([8]).

It follows from the application of this principle that the optimal control in this case is a piecewise constant function which alternates taking the two extreme values α and β. In the relevant interval of time $t_0 < t < t_1$ the function $\tilde{w}(t)$ has a finite number of the first order discontinuity points (the so called swiching points).

If we define an m-dimmensional space W with the coordinates $w, w^{(1)}, \dots, w^{(m-1)}$ then equations (11) may be looked upon as relations which transform the pair of points $\boldsymbol{y}$ and $\boldsymbol{w}$ from spaces Y and W into a point $\boldsymbol{x}$ in space X. We shall say that *the point $\boldsymbol{x}$ corresponds, according to relations* (11), *to a pair of points $\boldsymbol{y}$ and $\boldsymbol{w}$.* Similarly, to a pair of points $\boldsymbol{x}$

([8]) See [5], p. 133-136.

and $\boldsymbol{w}$ from spaces X and W there corresponds, according to (11), a point $\boldsymbol{y}$ in space Y ([9]).

If we put into relations (11) the definite functions of time: solution $\boldsymbol{x}(t)$ of the system of equations (12); solution of equation (1) and vector function $\boldsymbol{w}(t)$ corresponding to the control $w(t)$, and afterwards if the relations are fulfilled in a certain interval of time, then in this interval relations (11) assign e.g. to $\boldsymbol{y}(t)$ and $\boldsymbol{w}(t)$ a unique solution $\boldsymbol{x}(t)$ of the system of equations (12). In this case we say that in the definite interval of time *the solution* $\boldsymbol{x}(t)$ (*or the trajectory* $\boldsymbol{x}(t)$) *corresponds, according to relations* (11), *to the solution* $\boldsymbol{y}(t)$ (*or the trajectory* $\boldsymbol{y}(t)$) *and to the function* $\boldsymbol{w}(t)$. For instance, to the optimal trajectory $\tilde{\boldsymbol{y}}(t)$ in space Y and to the optimal control $\tilde{\boldsymbol{w}}(t)$ there corresponds, according to (11), a trajectory $\tilde{\boldsymbol{x}}(t)$ in space X which will be called the optimal trajectory in space X. From (11) it is also easy to see that *for a given control* $w(t)$ (then the vector $\boldsymbol{w}(t)$ is also given) *relations* (11) *determine a one-to-one correspondence between the trajectories* $\boldsymbol{y}(t)$ *in space* Y *and the trajectories* $\boldsymbol{x}(t)$ *in space* X.

Let us denote by $\boldsymbol{x}_0$ the point in space X corresponding, according to (11), to the points $\boldsymbol{y}_0$ and $\boldsymbol{w}_0$ specified in Problem I. If with an arbitrary admissible control $w(t)$, for which $\boldsymbol{w}(t_0^-) = \boldsymbol{w}_0$, we consider the solution $\boldsymbol{y}(t)$ of equation (1), for which $\boldsymbol{y}(t_0) = \boldsymbol{y}_0$, then the solution $\boldsymbol{x}(t)$ corresponding, according to (11), to $\boldsymbol{y}(t)$ and $\boldsymbol{w}(t)$ fulfils the relation $\boldsymbol{x}(t_0^-) = \boldsymbol{x}_0$. Due to the continuity of solutions $\boldsymbol{x}(t)$ (trajectories $\boldsymbol{x}(t)$ in space X) we have $\boldsymbol{x}(t_0^-) = \boldsymbol{x}(t_0^+) = \boldsymbol{x}(t_0)$ and so point $\boldsymbol{x}_0$ is the point passed through by trajectory $\boldsymbol{x}(t)$ at the moment t_0 ($\boldsymbol{x}(t_0) = \boldsymbol{x}_0$).

Thus we have proved the following

LEMMA 1. *Let* $w(t)$ *be an arbitrary admissible control with the given vector* $\boldsymbol{w}(t_0^-) = \boldsymbol{w}_0$ *and* $\boldsymbol{x}(t)$ *be a trajectory in space* X *which corresponds, according to* (11), *to the solution* $\boldsymbol{y}(t)$ *of equation* (1), *for which* $\boldsymbol{y}(t_0^-) = \boldsymbol{y}_0$, *and to the function* $\boldsymbol{w}(t)$ (*determined by the control* $w(t)$). *The trajectory* $\boldsymbol{x}(t)$ *passes through the point* $\boldsymbol{x}_0$ ($\boldsymbol{x}(t_0) = \boldsymbol{x}_0$) *which corresponds, according to* (11), *to the pair of points* $\boldsymbol{y}_0$ *and* $\boldsymbol{w}_0$.

It is necessary to remember that the point $\boldsymbol{x}_0$ is uniquely determined if and only if there is given the vector $\boldsymbol{w}(t_0^-) = \boldsymbol{w}_0$.

4. The control $w(t)$ for $t > t_1$. Let us consider now those of the admissible controls $w(t)$ which are determined for $t > t_1$ and such that there exists a solution $\boldsymbol{y}(t)$ of equation (1) satisfying the condition

$$\boldsymbol{y}(t) = \boldsymbol{y}_1 = 0 \quad \text{for } t > t_1. \tag{13}$$

([9]) Also to a pair of points $\boldsymbol{x}$ and $\boldsymbol{y}$ from spaces X and Y (but only to such points which are equal in the first $n-m$ coordinates) there corresponds a point $\boldsymbol{w}$ in space W. This follows from (11) and from the relation (19) which will be proved in the next paragraph.

Therefrom and from equation (1) it follows that such control has to be chosen among the admissible controls which satisfy the equation

$$b_{n-m}w^{(m)}+b_{n-m+1}w^{(m-1)}+\ldots+b_n w = 0. \tag{14}$$

Let Z_w^+ be the set of all points w_1 in space W such that the solution $w(t)$ of equation (14) with the initial condition w_1 at the moment t_1 ($w(t_1) = w(t_1^+) = w_1$) secure the fulfilment of (4) for $t > t_1$ ([10]).

The construction of set Z_w^+ is a separate problem. For a given case its determination usually is possible although the numerical task increases rapidly with the increasement of the order of equation (14). An example of a set Z_w^+ is shown in Fig. 2a for the case of equation (14) of second order while the corresponding characteristic equation has complex roots with negative real parts. The boundary of the set Z_w^+ in this case consists of the two arcs of phase-trajectories 1 and 2 which are tangent to the straight lines $w = a$ and $w = \beta$ and of the two segments on those lines.

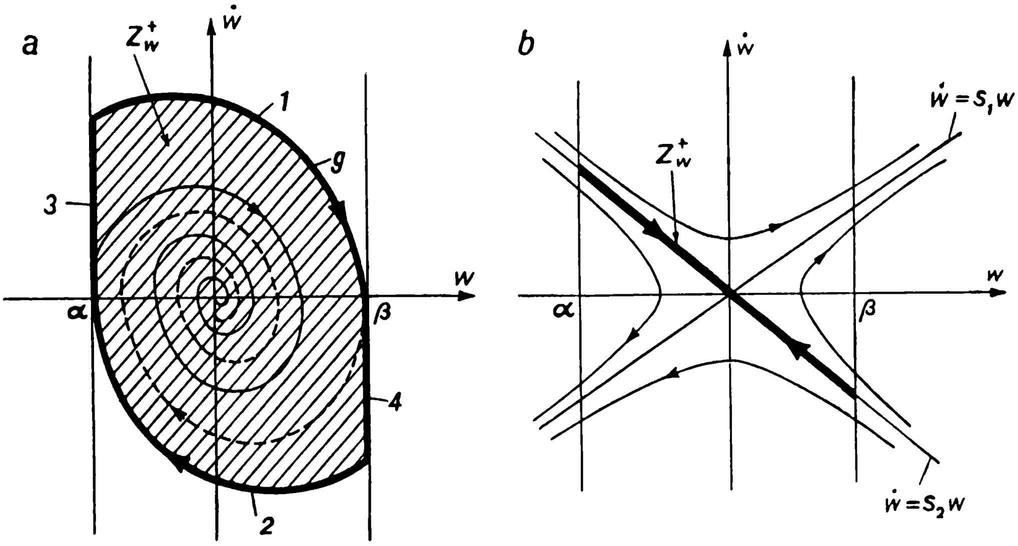


Fig. 2. Examples of the set Z_w^+

a – the case of complex roots of the characteristic equation (15) with negative real parts; *b* – the case of real roots of equation (15) – one of them is positive (s_1) and another negative (s_2)

Another example is shown in Fig. 2b. Here the set Z_w^+ is found for the case of the second order of equation (14) with real roots $s_1 > 0$ and $s_2 < 0$ of the corresponding characteristic equation. Now the set Z_w^+ consists of the points in an interval cut of the straight line $\dot{w} = s_2 w$ by the sector $a \leqslant w \leqslant \beta$.

([10]) Solutions of equation (14) are the continuous functions of time.

It can be noted that *if all the roots of characteristic equation*

(15) $$b_{n-m}s^m + b_{n-m+1}s^{m-1} + \ldots + b_n = 0$$

have positive real part, then Z_w^+ *contains only one point* $\boldsymbol{w}_1 = 0$.

Also it can be seen that if none of the roots of equation (15) equals zero and the values α and β are of the same sign, then Z_w^+ is empty. In this case the problem has no solution since among the admissible controls $w(t)$ does not exist one satisfying the requiurement $\boldsymbol{y}(t) = 0$ for $t > t_1$.

We have to add that Z_w^+ is always a closed set (e.g. in Fig. 2a points on the boundary of Z_w^+ belong to this set).

From the definition of Z_w^+ it follows that the point of a phase-trajectory in space Y can stay in position $\boldsymbol{y}_1 = 0$ for $t > t_1$ under an admissible control $w(t)$ only if the control $w(t)$ for $t > t_1$ is a solution of equation (14) with the initial conditions $\boldsymbol{w}(t_1^+) \in Z_w^+$.

Let us consider now *the set* Z_x^+ *of all such points* $\boldsymbol{x}_1$ *in space* X *which correspond, according to* (11), *to all pairs of points* $\boldsymbol{y}_1 = 0$ *and* $\boldsymbol{w}_1 \in Z_w^+$. Thus, the coordinates of points $\boldsymbol{x}_1$ in Z_x^+ may be obtained from the formulae

(16) $$\begin{aligned} x_1^1 &= 0, \\ &\ldots\ldots\ldots \\ x_1^{n-m} &= 0, \\ x_1^{n-m+1} &= -\gamma_{n-m}w_1, \\ x^{n-m+2} &= -\gamma_{n-m+1}w_1 - \gamma_{n-m}w_1^{(1)}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_1^n &= -\gamma_{n-1}w_1 - \gamma_{n-2}w_1^{(1)} - \ldots - \gamma_{n-m}w_1^{(m-1)}, \end{aligned}$$

where $w_1, w_1^{(1)}, \ldots, w_1^{(m-1)}$ are the coordinates of the corresponding point $\boldsymbol{w}_1 \in Z_w^+$. The last m equations in (16) determine a one-to-one transformation of the space W into an m-dimensional subspace $\mathfrak{X}$ (of the space X) with the coordinates $x^{n-m+1}, x^{n-m+2}, \ldots, x^n$. For, this transformation may be written in the form

(16a) $$\mathfrak{x}_1 = -\Gamma \boldsymbol{w}_1,$$

where

(17) $$\begin{aligned} \boldsymbol{w}_1 &= (w_1, w_1^{(1)}, \ldots, w_1^{(m-1)}), \\ \mathfrak{x}_1 &= (x_1^{n-m+1}, x_1^{n-m+2}, \ldots, x_1^n), \end{aligned}$$

and

(18) $$\Gamma = \begin{bmatrix} \gamma_{n-m} & 0 & \ldots & 0 \\ \gamma_{n-m+1} & \gamma_{n-m} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ \gamma_{n-1} & \gamma_{n-2} & \ldots & \gamma_{n-m} \end{bmatrix}.$$

By $\mathfrak{x}_1$ we denote here a m-dimensional vector with the coordinates equal to the last m coordinates of the vector $\boldsymbol{x}_1$. From (16) it follows that if a vector $\boldsymbol{x}_1$ belongs to Z_x^+ then its first $n-m$ coordinates are equal to zero. Thus, if $\boldsymbol{x}_1 \in Z_x^+$ then those coordinates which do not occur in $\mathfrak{x}_1$ are all zeros ([11]). Let us notice that

$$\det\Gamma = \gamma_{n-m}^{m} = b_{n-m}^{m} \neq 0 \tag{19}$$

since, according to previous assumption, $b_{n-m} \neq 0$. Thus, to any point of Z_w^+ there corresponds a unique point of Z_x^+ and vice versa.

Now we may formulate the following

LEMMA 2. *For an arbitrary admissible control $w(t)$, for which there exists a solution $\boldsymbol{y}(t)$ of equation (1) satisfying the relation $\boldsymbol{y}(t) = \boldsymbol{y}_1 = 0$ for $t > t_1$, the phase-trajectory $\boldsymbol{x}(t)$ in space X which corresponds, according to (11), to the solution $\boldsymbol{y}(t)$ and to the function $\boldsymbol{w}(t)$ passes at the moment $t = t_1$ through one of the points of the set Z_x^+ ($\boldsymbol{x}(t_1) \in Z_x^+$).*

To prove the lemma we shall remember that if there exists a solution $\boldsymbol{y}(t)$ of equation (1) such that $\boldsymbol{y}(t) = \boldsymbol{y}_1 = 0$ for $t > t_1$, then the admissible control $w(t)$ for $t > t_1$ must be a solution of equation (14) with the initial conditions $\boldsymbol{w}(t_1^+) = \boldsymbol{w}_1 \in Z_w^+$. The coordinates of the point of trajectory $\boldsymbol{x}(t)$ (which is the subject of the lemma) at "the moment" t_1^+, that is the coordinates of the vector $\boldsymbol{x}(t_1^+)$, may be calculated from formulae (11). For $\boldsymbol{y}(t_1^+) = 0$ and $\boldsymbol{w}(t_1^+) = \boldsymbol{w}_1 \in Z_w^+$ formulae (11) reduce themself to (16). Thus, we have $\boldsymbol{x}(t_1^+) \in Z_x^+$ which follows from the definition of Z_x^+. Due to the continuity of trajectories in space X we get $\boldsymbol{x}(t_1^+) = \boldsymbol{x}(t_1^-) = \boldsymbol{x}(t_1)$ and finaly $\boldsymbol{x}(t_1) \in Z_x^+$.

From lemma 2 it follows that, for the control $w(t)$ which we are looking for, the terminal point $\boldsymbol{x}_1$ reached at the moment t_1 by the phase-trajectory $\boldsymbol{x}(t)$ in space X ($\boldsymbol{x}(t_1) = \boldsymbol{x}_1$) is one of the points of the set Z_x^+ ($\boldsymbol{x}_1 \in Z_x^+$).

5. The solution of problem I. Let us assume now that at the moment t_1 the trajectory in space X passes through the point $\boldsymbol{x}_1 = (0, \ldots, 0, x_1^{n-m+1}, x_1^{n-m+2}, \ldots, x_1^n)$ and $\boldsymbol{x}_1 \in Z_x^+$. To the point $\boldsymbol{x}_1 \in Z_x^+$ there corresponds in Z_w^+ a unique point $\boldsymbol{w}_1$ determined by the one-to-one transformation (16a). Thus, we get

$$\boldsymbol{w}_1 = -\Gamma^{-1}\mathfrak{x}_1, \tag{20}$$

([11]) As it was with the set Z_w^+ also Z_x^+ is the closed set (see [1], lemma 7.2). If all the roots of equation (15) have positive real part then Z_x^+ contains only one point $\boldsymbol{x}_1 = 0$. (Lemma 7.3 in [1] which says that "if one of the roots is positive or having a positive real part then Z_x^+ contains only one point $\boldsymbol{x}_1 = 0$" is obviously false and its proof is wrong — see the counterexample in Fig. 2b).

where

$$\mathfrak{x}_1 = (x_1^{n-m+1}, x_1^{n-m+2}, \dots, x_1^n).$$

Now we may formulate the following

LEMMA 3. *Let us assume that the phase-trajectory $\boldsymbol{x}(t)$ in space X at the moment t_1 passes through the point $\boldsymbol{x}_1 \in Z_x^+$ ($\boldsymbol{x}(t_1) = \boldsymbol{x}_1$), and the control $w(t)$ for $t > t_1$ is a solution of equation* (14) *with the initial conditions $\boldsymbol{w}(t_1^+) = \boldsymbol{w}_1$ obtained from relation* (20). *Then and only then the phase-trajectory $\boldsymbol{y}(t)$ in space Y, which corresponds according to* (11) *to the trajectory $\boldsymbol{x}(t)$ and to the function $\boldsymbol{w}(t)$, remains for $t > t_1$ at the point $\boldsymbol{y}_1 = 0$ (i.e. $\boldsymbol{y}(t) = \boldsymbol{y}_1 = 0$ for $t > t_1$) and the control $w(t)$ for $t > t_1$ is an admissible one.*

To prove lemma 3 note first the consequence of the assumption that the control $w(t)$ for $t > t_1$ is a solution of (14). From this it follows that for $t > t_1$ the movement of a point in space Y is described by a differential equation of the form

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_n y = 0. \tag{21}$$

Phase-trajectories in space Y (for $t > t_1$) are determined by the solutions of this equation and their course depends on the initial conditions $\boldsymbol{y}(t_1^+)$. From the assumption of the lemma we have also $\boldsymbol{x}(t_1) = \boldsymbol{x}_1 \in Z_x^+$ and, due to the continuity of trajectories in space X, $\boldsymbol{x}(t_1^+) = \boldsymbol{x}(t_1) = \boldsymbol{x}_1 = (0, \dots, 0, x_1^{n-m+1}, x_1^{n-m+2}, \dots, x_1^n)$. According to the assumption there is also $\boldsymbol{w}(t_1^+) = \boldsymbol{w}_1$, and the quantities $\boldsymbol{w}_1$ and $\mathfrak{x}_1 = (x_1^{n-m+1}, x_1^{n-m+2}, \dots \dots, x_1^n)$ fulfil relation (20). Therefore, the coordinates of vectors $\boldsymbol{x}_1$ and $\boldsymbol{w}_1$ fulfil the equation (16). Point $\boldsymbol{y}(t_1^+)$ of the trajectory $\boldsymbol{y}(t)$ (being the subject of the lemma) corresponds, according to (11) to the pair of points $\boldsymbol{x}(t_1^+) = \boldsymbol{x}_1$ and $\boldsymbol{w}(t_1^+) = \boldsymbol{w}_1$. Hence, in virtue of (16) we get $\boldsymbol{y}(t_1^+) = 0$. The solution of equation (21) for $t > t_1$ with the initial conditions $\boldsymbol{y}(t_1^+) = 0$ is of the form $\boldsymbol{y}(t) = 0$, and so there is also $\boldsymbol{y}(t) = \boldsymbol{y}_1 = 0$ for $t > t_1$. Since $\boldsymbol{x}_1 \in Z_x^+$ and transformation (20) is a one-to-one transformation of Z_x^+ onto Z_w^+, then $\boldsymbol{w}(t_1^+) = \boldsymbol{w}_1 \in Z_w^+$. According to the definition of Z_w^+, the solution $w(t)$ of equation (14) (for $t > t_1$) with the initial conditions $\boldsymbol{w}(t_1^+) = \boldsymbol{w}_1$ fulfils the condition (4), hence $w(t)$ is an admissible control for $t > t_1$.

On the other hand, when $\boldsymbol{y}(t) = 0$ for $t > t_1$, and a control $w(t)$ is a admissible one then it is easy to proove that $w(t)$ is a solution of equation (14) with the initial conditions fulfils (20), where $\boldsymbol{x}_1$ is a point for which $\boldsymbol{x}_1 = \boldsymbol{x}(t_1)$.

Directly from lemmas 1 and 3 we get the following

COROLLARY 1. *Assume that the point $\boldsymbol{x}_0$, which according to* (11) *corresponds to the pair of points $\boldsymbol{y}_0$ and $\boldsymbol{w}(t_0^-) = \boldsymbol{w}_0$, belongs to Z_x^+. Let us denote by $\tilde{w}(t)$ the control which for $t > t_0$ is a solution of equation* (14) *with the initial conditions* $\tilde{\boldsymbol{w}}(t_0^+) = -\Gamma^{-1}\mathfrak{x}_0$. *Under the control* $\tilde{w}(t)$ *the solution* $\boldsymbol{y}(t)$ *of equation* (1), *if* $\boldsymbol{y}(t_0^-) = \boldsymbol{y}_0$, *fulfils the relation* $\boldsymbol{y}(t) = \boldsymbol{y}_1 = 0$ *for* $t > t_0$. *Hence we have* $t_1 = t_0$ *and the point in space* Y *may be conducted under the control* $\tilde{w}(t)$ (*which is the admissible one for* $t > t_0$) *from position* $\boldsymbol{y}_0$ *into position* $\boldsymbol{y}_1 = 0$ *in the time* $T = t_1 - t_0 = 0$.

Now we may interpret Z_x^+ as the set of points in space X which, according to (11), correspond to all pairs of points $\boldsymbol{w}(t_0^-) = \boldsymbol{w}_0$ and $\boldsymbol{y}_0^*$; where $\boldsymbol{y}_0^*$ is an arbitrary point of space Y which may be conducted in time $T = t_1 - t_0 = 0$ (under an admissible control $w(t)$ with the given vector $\boldsymbol{w}(t_0^-) = \boldsymbol{w}_0$) to the position $\boldsymbol{y}_1 = 0$ in such a way that $\boldsymbol{y}(t) = \boldsymbol{y}_1 = 0$ for $t > t_1 = t_0$.

Hence, *for the case of* $\boldsymbol{x}_0 \in Z_x^+$ *the control* $\tilde{w}(t)$ *specified in Corollary* 1 *is the optimal control required in Problem* I.

If a point $\boldsymbol{x}_0$ does not belong to Z_x^+ ($\boldsymbol{x}_0 \notin Z_x^+$) then it has to be $t_1 > t_0$. Indeed, solutions $\boldsymbol{x}(t)$ of the system of equations (12) are unique and continuous([12]) and $\boldsymbol{x}_0 \notin Z_x^+$ yields $\boldsymbol{x}_0 \neq \boldsymbol{x}_1$ for any $\boldsymbol{x}_1 \in Z_x^+$. From lemmas 1 and 2 follows that $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$ and $\boldsymbol{x}(t_1) = \boldsymbol{x}_1$ where $\boldsymbol{x}_1 \in Z_x^+$. Hence, $\boldsymbol{x}(t_0) \neq \boldsymbol{x}(t_1)$ and $t_0 \neq t_1$. From the assumption $t_1 \geqslant t_0$ we finaly get $t_1 > t_0$.

Let us denote by $\tilde{w}(t)$ the optimal control in the interval of time $t_0 < t < t_1$ which conducts in the shortest possible time $T = t_1 - t_0$ the point in space X from the initial position $\boldsymbol{x}_0$ to an arbitrary point of Z_x^+.

The trajectory in space X, which under the optimal control $\tilde{w}(t)$ ($t_0 < t < t_1$) passes through $\boldsymbol{x}_0$ at t_0, later at t_1 passes through a point which we will denote by $\bar{\boldsymbol{x}}_1$ (obviously $\bar{\boldsymbol{x}}_1 \in Z_x^+$). In other words, if $w(t) = \overline{w}(t)$ ($t_0 < t < t_1$) and $\boldsymbol{x}(t)$ is a solution of the system of differetial equations (12) with the initial condition $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$, then $\boldsymbol{x}(t_1) = \bar{\boldsymbol{x}}_1$. Hence, $\bar{\boldsymbol{x}}_1$ is the selected point of Z_x^+ which is available in the shortest time from the position $\boldsymbol{x}_0$ (under an admissible control). Both $\overline{w}(t)$ and $\bar{\boldsymbol{x}}_1$ may be obtained by solving the problem of time-optimal control in space X from the given initial point $\boldsymbol{x}_0$ to the target set Z_x^+. Selection of the point $\bar{\boldsymbol{x}}_1$ in Z_x^+ can be made easier by the *proper application of the transversality conditions*([13]). If we know the coordinates of $\bar{\boldsymbol{x}}_1$, e.g. $\bar{\boldsymbol{x}}_1 = (0, \ldots, 0, \bar{x}_1^{n-m+1}, \ldots, \bar{x}_1^n)$, then by a transformation inverse to (16a) we may find the point $\boldsymbol{w}_1^* \in Z_w^+$ corresponding to $\bar{\boldsymbol{x}}_1$

$$\boldsymbol{w}_1^* = -\Gamma^{-1}\bar{\mathfrak{x}}_1, \tag{22}$$

where

$$\bar{\mathfrak{x}}_1 = (\bar{x}_1^{n-m+1}\,\bar{x}_1^{n-m+2}, \ldots, \bar{x}_1^n).$$

([12]) They are unique, that is for an arbitrary moment t' there is only one point $\boldsymbol{x}(t')$ in space X and if $\boldsymbol{x}(t') \neq \boldsymbol{x}(t'')$ then $t' \neq t''$, and the continuity means $\boldsymbol{x}(t'^-) = \boldsymbol{x}(t'^+) = \boldsymbol{x}(t')$.

([13]) See [5], p. 59 and further remarks in this paper.

Now, denote by $w^*(t)$ the control (for $t > t_1$) which is the solution of differential equation (14) with the initial conditions $\boldsymbol{w}^*(t_1^+) = \boldsymbol{w}_1^*$.

Next, we shall prove the following

THEOREM 1. *The time optimal control $\tilde{w}(t)$ for $t > t_0$ (in the sense of Problem* I), *which in the shortest time $T = t_1 - t_0$ conducts a point in space Y from position $\boldsymbol{y}_0$ to position $\boldsymbol{y}_1 = 0$ and makes it stay there for $t > t_1$ (i.e. $\boldsymbol{y}(t) = \boldsymbol{y}_1 = 0$ for $t > t_1$), is a function of time which:*

1. *in the interval of time $t_0 < t < t_1$ coincides with a piecewise constant function $\overline{w}(t)$ which is the solution of the problem of time-optimal control from the given point $\boldsymbol{x}_0$ to the set Z_x^+;*

2. *for $t > t_1$ coincides with the function $w^*(t)$ which is the solution of differential equation* (14) *with the initial conditions $\boldsymbol{w}^*(t_1^+) = \boldsymbol{w}_1^* = -\Gamma^{-1}\overline{\boldsymbol{x}}_1$.*

The point $\boldsymbol{x}_0$ is determined by $\boldsymbol{w}_0$ and $\boldsymbol{y}_0$ (which are given in Problem I). *The point $\overline{\boldsymbol{x}}_1 \epsilon Z_x^+$ and the moment t_1 result from the solution of time-optimal control from the point $\boldsymbol{x}_0$ to the set Z_x^+* ([14]).

Proof of theorem 1. It follows from lemmas 1 and 2 that if a control $w(t)$ $(t > t_0)$ is admissible, has a given vector $\boldsymbol{w}(t_0^-)$ and conducts the point in space Y from position $\boldsymbol{y}_0$ to position $\boldsymbol{y}_1 = 0$ in such a way that $\boldsymbol{y}(t) = \boldsymbol{y}_1 = 0$ for $t > t_1$, then the trajectory $\boldsymbol{x}(t)$ in space X, which according to (11) corresponds to the solution $\boldsymbol{y}(t)$ $(\boldsymbol{y}(t_0^-) = \boldsymbol{y}_0)$ and to the function $\boldsymbol{w}(t)$, passes the points $\boldsymbol{x}_0$ and $\boldsymbol{x}_1 \epsilon Z_x^+$ (i.e. $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$ and $\boldsymbol{x}(t_1) = \boldsymbol{x}_1 \epsilon Z_x^+$). The control $\overline{w}(t)$ $(t_0 < t < t_1)$, according to the definition, conducts the point in space X from position $\boldsymbol{x}_0$ to position $\boldsymbol{x}_1 \epsilon Z_x^+$ in the shortest time $T = t_1 - t_0$, hence for $w(t) = \overline{w}(t)$ $(t_0 < t < t_1)$ there is $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$, $\boldsymbol{x}(t_1) = \overline{\boldsymbol{x}}_1 \epsilon Z_x^+$ and $T = T_{\min}$. According to the maximum principle of Pontryagin, as we have mentioned previously, the optimal control $\overline{w}(t)$ $(t_0 < t < t_1)$ is a piecewise constant function taking only the extreme values α and β.

From $\boldsymbol{x}(t_1) = \overline{\boldsymbol{x}}_1 \epsilon Z_x^+$ and from lemma 3 follows that under the control $w^*(t)$, $t > t_1$, (which is admissible) a point in the space Y remains in position $\boldsymbol{y}_1 = 0$ for $t > t_1$, that is $\tilde{\boldsymbol{y}}(t) = \boldsymbol{y}_1 = 0$ for $t > t_1$, where $\tilde{\boldsymbol{y}}(t)$ is a solution satisfying the condition $\tilde{\boldsymbol{y}}(t_0^-) = \boldsymbol{y}_0$.

The control $\tilde{w}(t)$ which is equal to $\overline{w}(t)$ for $t_0 < t < t_1$ and to $w^*(t)$ for $t > t_1$ is an admissible control for $t > t_0$. Indeed the control $\overline{w}(t)$ is admissible in the interval of time $t_0 < t < t_1$ and $w^*(t)$ admissible for $t > t_1$ whereas at the point t_1, which may be a discontinuity point of $\tilde{\boldsymbol{w}}(t)$, there exist the values $\tilde{\boldsymbol{w}}(t_1^-) = \overline{\boldsymbol{w}}(t_1^-)$ and $\tilde{\boldsymbol{w}}(t_1^+) = \boldsymbol{w}^*(t_1^+)$. Therefore $\tilde{w}(t)$ is an admissible control for $t > t_0$ which (for a given vector $\boldsymbol{w}(t_0^-)$)

([14]) We remind here that the moment t_0 is either given in advance or it may be arbitrarily chosen.

conducts a point in the space Y from the position $\boldsymbol{y}_0$ to $\boldsymbol{y}_1 = 0$ and $\tilde{\boldsymbol{y}}(t) = \boldsymbol{y}_1 = 0$ for $t > t_1$.

Now we shall prove that the time $T = t_1 - t_0 = T_{\min}$ for the control $\tilde{w}(t)$ is the shortest one of all times required by admissible controls $w(t)$. To show this let us assume that there exists an admissible control $w'(t) \neq \tilde{w}(t)$ for which $T' = t_1' - t_0' < T_{\min}$. From lemmas 1 and 2 follows that for the control $w'(t)$ the corresponding solution $\boldsymbol{x}'(t)$ of (12) has to fulfil the conditions $\boldsymbol{x}'(t_0) = \boldsymbol{x}_0$ and $\boldsymbol{x}'(t_1) = \boldsymbol{x}_1 \in Z_x^+$. Therefore $w'(t)$ for $t_0 < t < t_1$ is a control which in time $T' = t_1' - t_0'$, shorter than that for the control $\tilde{w}(t) = \overline{w}(t)$ for $t_0 < t < t_1$, conducts a point of the space X from the position $\boldsymbol{x}_0$ to the position $\boldsymbol{x}_1 \in Z_x^+$. This contradicts the definition of $\overline{w}(t)$, $t_0 < t < t_1$, as a control which in the shortest time conducts a point of the space X from the position $\boldsymbol{x}_0$ to the position $\boldsymbol{x}_1 \in Z_x^+$.

6. The solution of Problem II. Let $w(t)$ be an admissible control for $t > t_0$. We shall extend it for $t < t_0$ in such a way that $\boldsymbol{w}(t_0^-) = \boldsymbol{w}_0$. *We denote by Z_w^- the set of all points $\boldsymbol{w}_0$ of the m-dimentional space W (with coordinates $w, w^{(1)}, \ldots, w^{(m-1)}$) for which the so extended control is admissible for $t > t_0 - \varepsilon$ (here ε is an arbitrarily small positive number).*

One can notice that Z_w^- consists of the points $\boldsymbol{w}_0$ of space W for which

$$\begin{aligned} &\alpha < w_0 \leqslant \beta \quad \text{if } w_0^{(1)} > 0, \\ &\alpha \leqslant w_0 < \beta \quad \text{if } w_0^{(1)} < 0. \end{aligned} \tag{23}$$

An example of the set Z_w^- for $m = 2$ is presented in Fig. 3. In this figure the points of dashed halflines do not belong to Z_w^-.

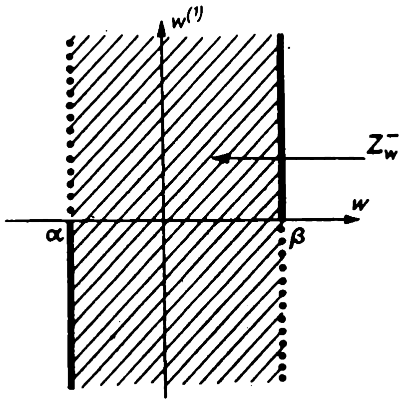


Fig. 3. An example of the set Z_w^- for $m = 2$

Now let us denote by $Z_x^-(\boldsymbol{y}_0)$ the set of points $\boldsymbol{x}_0$ in X which correspond, according to (11), *to all pairs of points $\boldsymbol{y}_0$ and $\boldsymbol{w}_0$, where $\boldsymbol{w}_0$ is an arbitrary point of Z_w^-.* The coordinates of the points $\boldsymbol{x}_0 \in Z_x^-(\boldsymbol{y}_0)$ may be calculated as follows

$$
(24)\qquad
\begin{aligned}
x_0^1 &= y_0, \\
&\dots\dots\dots \\
x_0^{n-m} &= y_0^{(n-m-1)}, \\
x_0^{n-m+1} &= y_0^{(n-m)} - \gamma_{n-m} w_0, \\
x_0^{n-m+2} &= y_0^{(n-m+1)} - \gamma_{n-m+1} w_0 - \gamma_{n-m} w_0^{(1)}, \\
&\dots\dots\dots\dots\dots\dots\dots\dots \\
x_0^{n} &= y_0^{(n-1)} - \gamma_{n-1} w_0 - \gamma_{n-2} w_0^{(1)} - \dots - \gamma_{n-m} w_0^{(m-1)},
\end{aligned}
$$

where $\boldsymbol{y}_0 = (y_0, y_0^{(1)}, \dots, y_0^{(m-1)})$ and $w_0, w_0^{(1)}, \dots, w_0^{(m-1)}$ are the coordinates of an arbitrary point $\boldsymbol{w}_0 \in Z_w^-$. One can easily notice that having given a point $\boldsymbol{y}_0$ formulae (24) determine a one-to-one correspondence between the points $\boldsymbol{w}_0$ of Z_w^- and the points $\boldsymbol{x}_0$ of the set $Z_x^-(\boldsymbol{y}_0)$ and for any point $\boldsymbol{x}_0 \in Z_x^-(\boldsymbol{y}_0)$ the first $n-m$ of its coordinates are equal to the corresponding coordinates of $\boldsymbol{y}_0$.

Since it may be difficult in practice to realize the vector $\boldsymbol{w}(t_0^-)$ if some of its coordinates are of large values one can also impose some bounds upon the coordinates of $\boldsymbol{w}(t_0^-)$. This will induce the appropriate changes in the sets Z_w^- and $Z_x^-(y_0)$.

Now we may formulate the following

LEMMA 4. *Let $w(t)$ be an arbitrary admissible control for $t > t_0 - \varepsilon$ and $\boldsymbol{x}(t)$ a trajectory in space X which according to* (11) *corresponds to the solution $\boldsymbol{y}(t)$ of equation* (1) *(for which $\boldsymbol{y}(t_0^-) = \boldsymbol{y}_0$) and to the function $\boldsymbol{w}(t)$. The trajectory $\boldsymbol{x}(t)$ at the moment t_0 passes one of the points of $Z_x^-(\boldsymbol{y}_0)$, that is $\boldsymbol{x}(t_0) \in Z_x^-(y_0)$.*

The proof of lemma 4 follows immediately from the definition of sets Z_w^- and $Z_x^-(\boldsymbol{y}_0)$ as well as from the continuity of the trajectory in space X.

Let us assume now that for an admissible control $w(t)$, $t > t_0$, to a trajectory $\boldsymbol{x}(t)$, $t > t_0$, in space X which at t_0 leaves an arbitrary point $\boldsymbol{x}_0 \in Z_x^-(\boldsymbol{y}_0)$ (i.e. $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$) and to a function $\boldsymbol{w}(t)$ there corresponds according to (11) a trajectory $\boldsymbol{y}(t)$, $t > t_0$, in space Y. We want to find a vector $\boldsymbol{w}(t_0^-)$ for the function $w(t)$ such that the above mentioned trajectory $\boldsymbol{y}(t)$ passes at t_0^- the point $\boldsymbol{y}_0$ (i.e. $\boldsymbol{y}(t_0^-) = \boldsymbol{y}_0$). Since both trajectories $\boldsymbol{y}(t)$ and $\boldsymbol{x}(t)$ as well as the function $\boldsymbol{w}(t)$ are related according to (11) then the values of $\boldsymbol{y}(t_0^-) = \boldsymbol{y}_0$, $\boldsymbol{x}(t_0^-) = \boldsymbol{x}(t_0) = \boldsymbol{x}_0$ and $\boldsymbol{w}(t_0^-)$ have to fulfil also the same relations. From the last m equations in (11) we thus obtain

$$
(25)\qquad \mathfrak{x}_0 = \mathfrak{y}_0 - \Gamma \boldsymbol{w}(t_0^-),
$$

where $\mathfrak{x}_0$ and $\mathfrak{y}_0$ are m-dimensional vectors with the coordinates equal to the last m coordinates of the vectors $\boldsymbol{x}_0$ and $\boldsymbol{y}_0$, respectively.

Taking into account (19) we may calculate from equation (25) the value

$$w(t_0^-) = \Gamma^{-1}(\mathfrak{y}_0 - \mathfrak{x}_0). \tag{26}$$

If $x_0 \in Z_x^-(y_0)$ then $w(t_0^-) \in Z_w^-$. From the definition of Z_w^- and from the assumption that $w(t)$ is an admissible control for $t > t_0$ follows then that the same control $w(t)$ extended for $t < t_0$ so as to fulfil (26) is an admissible control for $t > t_0 - \varepsilon$ (point t_0 may be a discontinuity point of the function $w(t)$ however there exist both side limits $w(t_0^-)$ and $w(t_0^+)$).

Thus we have proved the following

LEMMA 5. *Let for an admissible control $w(t)$, $t > t_0$, for a trajectory $x(t)$ in space X which at t_0 leaves an arbitrary point $x_0 \in Z_x^-(y_0)$ and for a function $w(t)$ there corresponds, according to* (11), *a trajectory $y(t)$ $(t > t_0)$ in space Y. If we extend the control $w(t)$ for $t < t_0$ so as to fulfil the equation $w(t_0^-) = \Gamma^{-1}(\mathfrak{y}_0 - \mathfrak{x}_0)$, then the extended control is an admissible one for $t > t_0 - \varepsilon$ and the trajectory $y(t)$ passes at t_0^- the point y_0, i.e. $y(t_0^-) = y_0$.*

From lemmas 3 4 and 5 it follows

COROLLARY 2. *Let us assume that the set $Z_x^* = Z_x^-(y_0) \cap Z_x^+$ is not empty and let x_0' be an arbitrary point of Z_x^*. By $\tilde{w}(t)$ we denote the control which for $t > t_0$ is a solution of equation* (14) *with the initial conditions $\tilde{w}(t_0^+) = -\Gamma^{-1}\mathfrak{x}_0'$ and which fulfils the condition $\tilde{w}(t_0^-) = \Gamma^{-1}(\mathfrak{y}_0 - \mathfrak{x}_0')$. Under the control $\tilde{w}(t)$ the solution $y(t)$ of equation* (1) *for which $y(t_0^-) = y_0$ fulfils the relation $y(t) = y_1 = 0$ for $t > t_0$. Therefore $t_1 = t_0$ and a point in space Y under the control $\tilde{w}(t)$ may be conducted from the position y_0 to y_1 in the time $T = t_1 - t_0 = 0$. Besides $\tilde{w}(t)$ is an admissible control for $t > t_0 - \varepsilon$.*

It is easy to see that *in the case of a non-empty set Z_x^* the control $\tilde{w}(t)$ discussed in corollary 2 is the optimal control which is asked for in problem* II.

Point x_0' is an arbitrary point of the set Z_x^* so if Z_x^* contains more then one point there exists also more then one admissible control $w(t)$, $t > t_0 - \varepsilon$, which may conduct a point in space Y from *the* position y_0 to y_1 in time $T = t_1 - t_0 = 0$ (there may be even an infinite number of such controls).

In a general case we may formulate the following theorem specifying the control $\tilde{w}(t)$ which is asked for in problem II.

THEOREM 2. *The time-optimal control $\tilde{w}(t)$ for $t > t_0 - \varepsilon$ (in the sense of problem* II), *which in the shortest time $T = t_1 - t_0$ conducts a point in space Y from position y_0 to position $y_1 = 0$ and makes it stay there for $t > t_1$ (i.e. $y(t) = y_1 = 0$ for $t > t_1$), is a function of time which:*

1. *fulfils the equation $\tilde{w}(t_0^-) = \Gamma^{-1}(\mathfrak{y}_0 - \bar{\mathfrak{x}}_0)$;*

2. *in the interval of time $t_0 < t < t_1$ coincides with a piecewise constant function $\overline{w}(t)$ which is the solution of the problem of time-optimal control from the set $Z_x^-(y_0)$ to the set Z_x^+;*

3. *for $t > t_1$ coincides with the function $w^*(t)$ which is the solution of differential equation* (14) *with the initial conditions $w^*(t_1^+) = w_1^* = -\Gamma^{-1}\bar{x}_1$.*

The points $\bar{x}_0 \in Z_x^-(y_0)$ and $\bar{x}_1 \in Z_x^+$ as well as the moment t_1 result also from the solution of time-optimal control from the set $Z_x^-(y_0)$ to the set Z_x^+.

The proof of theorem 2 is analogous with that of theorem 1 with the modification that instead of referring to lemma 1 now we have to refer to lemma 4. In addition we have to ensure that the trajectory $\tilde{y}(t)$ $(t > t_0)$, which for $t > t_1$ coincides with the point $y_1 = 0$ of space Y and which corresponds (according to (11)) to the trajectory $x(t)$ (for which $x(t_0) = \bar{x}_0$ and $x(t_1) = \bar{x}_1$), at the moment t_0^- passes the point y_0 (i.e. $y(t_0^-) = y_0$). On behalf of lemma 5 this condition holds if the optimal control $w(t)$ contains the vector $\tilde{w}(t_0^-) = \Gamma^{-1}(\mathfrak{y}_0 - \bar{\mathfrak{x}}_0)$.

Remarks on the application of the transversality conditions. For the determination of points $\bar{x}_0$ and $\bar{x}_1$ the transversality conditions may be applied. Those conditions were proved in [5] on the assumption that both the starting set and the target set consist of points lying on smooth hypersurfaces. Let S_0 be an r_0-dimensional hypersurface $(r_0 < n)$ containing the starting set and S_1 a r_1-dimensional one $(r_1 < n)$ containig the target set. If we carefuly study the proof of the maximum principle and in particular that of the transversality conditions (published in [5]) we can notice that the above assumption may be partially released. Thus the transversality conditions hold also for all points $\bar{x}_0$ and $\bar{x}_1$ having neighbourhoods in space X inside which the hypersurfices S_0 and S_1 are smooth. The whole construction of the proof, for ε sufficiently small to restrict the movement to the given neighbourhoods, remain valid also for the points $\bar{x}_0$ and $\bar{x}_1$ which fulfil the released assumption. So the hypersurfices S_0 and S_1 do not need to be totally smooth, they may contain edges and the sets of starting points x_0 and targets x_1 may contain borders consisting of points which lay on hypersurfices S_0 and S_1. Those edges and borders are in general also hypersurfices of dimension lower than r_0 and r_1, respectively. With respect to the points $\bar{x}_0$ and $\bar{x}_1$ on the edges and borders one can also apply the transversality conditions, taking into account the lower dimension of hypersurfices (edges or borders), if the above condition holds at the points in question.

These remarks are of importance for problems discussed in the paper since the geometrical objects corresponding to the sets Z_x^+ and $Z_x^-(y_0)$ are often not totally smooth.

7. Examples. Let us consider a plant for which the differential equation (1) is

$$\ddot{y} = \ddot{w} + w \tag{27}$$

and the admissible control $w(t)$ is a piecewise continuous function which obeys the restriction

$$-1 \leqslant w(t) \leqslant +1. \tag{28}$$

So we have $n = m = 2$, $a_1 = 0$, $a_2 = 0$, $b_1 = 0$, $b_2 = 1$ and from (8) we get $\gamma_0 = 1$, $\gamma_1 = 0$ and $\gamma_2 = 1$. According to (11) we introduce the new variables

$$\begin{aligned} x^1 &= y - w, \\ x^2 &= \dot{y} - \dot{w}. \end{aligned} \tag{29}$$

The system of equations (12) which describes the movement of a point in two-dimensional space X, i.e. in the plane X with coordinates x^1, x^2, now takes the form

$$\begin{aligned} \dot{x}^1 &= x^2, \\ \dot{x}^2 &= w. \end{aligned} \tag{30}$$

In the plane W with coordinates $w, \dot{w}$ the trajectories $w(t)$ determined by the solutions of equation

$$\ddot{w} + w = 0 \tag{31}$$

are (as it is well known) circles with the common centre in the origin of the system of coordinates. Thus, taking into account the definition of Z_w^+, one can easily notice that the set Z_w^+ consists of the points w_1 of the plane W which lay within and on the border of a circle with the centre in the point $w = 0$ and the radius equal to 1 (see Fig. 4a).

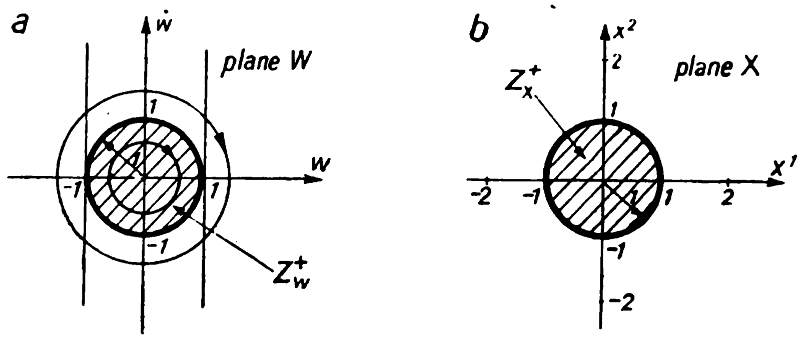


Fig. 4
a – the set Z_w^+, and b – the set Z_x^+ for the discussed example

From (16) we get

$$\begin{aligned} x_1^1 &= -w_1, \\ x_1^2 &= -\dot{w}_1, \end{aligned} \tag{32}$$

so that Z_x^+ is the set of points $x_1 = (x_1^1, x_1^2)$ of the plane X which lay inside and on the border of the unitary circle g with the centre in the point $x = 0$ (see Fig. 4b).

The solution of Problem I. Let us begin with solving the problem of optimal control in plane X with a given starting point x_0 (determined by the points w_0 and y_0 in Problem I) and with a free target x_1 belonging to the above determined set Z_x^+. To this effect, following the notation used in [5], we introduce the function [15]

$$H(\psi, x, w) = \psi_1 x^2 + \psi_2 w, \tag{33}$$

where ψ_1, ψ_2 are the coordinates of the vector ψ which fulfil the system of equations

$$\dot{\psi}_1 = 0,$$
$$\dot{\psi}_2 = -\psi_1.$$

This yields

$$\begin{aligned} \psi_1 &= c_1, \\ \psi_2 &= c_2 - c_1 t, \end{aligned} \tag{34}$$

where c_1 and c_2 are constants.

It follows from the maximum principle that the optimal control $\overline{w}(t)$ at any moment has to fulfil the condition

$$H(\psi, x, \overline{w}) = \sup_{-1 \leqslant w \leqslant +1} H(\psi, x, w).$$

On the ground of (33) we obtain

$$\overline{w}(t) = \operatorname{sign} \psi_2(t) = \operatorname{sign}(c_2 - c_1 t). \tag{35}$$

Thus, the optimal control $\overline{w}(t)$ is a piecewise constant function having only the values $+1$ or -1 and no more then one switch (from the value $+1$ to -1 or vice versa).

In the case $w = +1$ or $w = -1$ it is easy to determine the phase-trajectories in plane X applying the system of equations (30). They are paraboles of the form

$$\begin{aligned} (x^2)^2 &= 2x^1 + C \qquad \text{for } w = +1, \\ (x^2)^2 &= -2x^1 + C \quad \text{for } w = -1, \end{aligned} \tag{36}$$

where C is a constant.

Let us assume now that the point x_0 does not belong to Z_x^+ (i.e. x_0 lies outside of the circle g)[16]. Since the trajectory is continuous in the plane X we may then restrict ourselves to the case when the target x_1 lies on the border g of the set Z_x^+.

---

[15] See Pontryagin et al. [5], p. 23-26. As it can be seen from (33) the function H is an inner product of two vectors ψ and $\dot{x}$.

[16] If $x_0 \in Z_x^+$ we may, applying Corollary 1, to get a control $\tilde{w}(t)$ for which $T = t_1 - t_0 = 0$.

The transversality condition for the target x_1 of a trajectory demands that the vector $\psi(t_1)$ be orthogonal to the border g at the point x_1. On the other hand, from the condition

$$H\big(\psi(t_1), x(t_1), \overline{w}(t_1)\big) \geqslant 0$$

which appears in the maximum principle, and remembering that the function H is an inner product of the vectors ψ and $\dot{x}$, we obtain an additional information saying that $\psi(t_1)$ should be directed inside the circle. Thus, the coordinates of $\psi(t_1)$ have to fulfil the equation

$$\frac{\psi_2(t_1)}{\psi_1(t_1)} = \frac{x_1^2}{x_1^1},$$

wherefrom on the ground of (34) we obtain

$$\varrho - t_1 = x_1^2/x_1^1, \tag{37}$$

where $\varrho = c_2/c_1$ is constant.

To simplify further calculation we will assume that $t_0 = 0$.

From (37) we may deduce some conclusions on the course of optimal trajectories in space X which terminate at different points of the circle g.

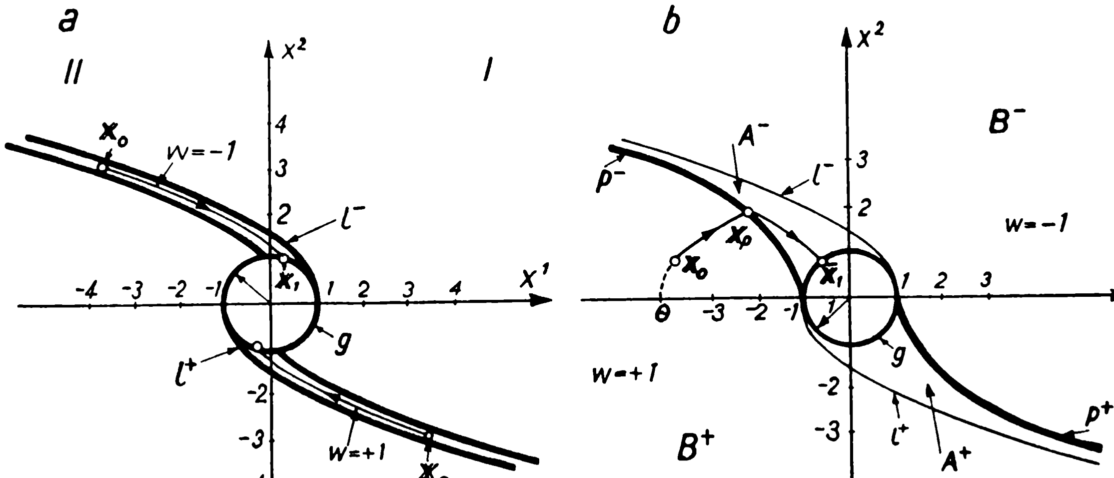


Fig. 5

a — the course of optimal trajectories in plane X which terminate at the points x_1 of the circle g laying in the first and the third quadrant; *b* — the course of optimal trajectories which terminate at the points x_1 of the circle g in the second quadrant

Thus, for the trajectories terminating at the points $x_1 \epsilon g$ which lay in the first and the third quadrant of the plane X (see Fig. 5a) we have $x_1^2/x_1^1 > 0$ and therefore $\varrho - t_1 > 0$. Since $t_1 > 0$, then it is also $\varrho > 0$ and in the interval of time $t_0 = 0 < t < t_1$ there is $\psi_2(t)/\psi_1(t) > 0$. Now, from $\psi_1(t) = c_1$ follows that $\psi_2(t)$ is of a constant sign in the above interval.

Thus, the optimal control $\overline{w}(t)$, according to (35), for the whole interval $t_0 < t < t_1$ retains constant value $+1$ or -1 (there are no swiching points). It is easy to notice that $\overline{w}(t) = -1$ for the trajectories terminating at points $\boldsymbol{x}_1$ of the circle g which belong to the first quadrant, and $\overline{w}(t) = +1$ for those in third quadrant. In order that the vector $\boldsymbol{\psi}(t_1)$ be orthogonal to g at point $\boldsymbol{x}_1$ and directed inside the circle, for the points $\boldsymbol{x}_1 \epsilon g$ in the second and the third quadrant it has to be

$$\psi_1(t_1) = \psi_1(t) = c_1 > 0, \tag{38}$$

while in the first and the fourth quadrant it has to be

$$\psi_1(t) = c_1 < 0.$$

Let us consider now the trajectories terminating at those points $\boldsymbol{x}_1$ of the circle g which belong to the second quadrant of the plane X.

It is known that from conditions of the maximum principle the vector $\boldsymbol{\psi}$ may be determined with preciseness to the constant coefficient (only the argument but not the absolute value of $\boldsymbol{\psi}$ may be determined). Hence because of (38) for the second quadrant we may accept

$$\begin{aligned} \psi_1(t) &= 1, \\ \psi_2(t) &= \varrho - t, \end{aligned} \tag{39}$$

where $\varrho = c_2/c_1$ is a constant.

From (35) it follows that the swich occurs at the moment t_p, $t_0 < t_p < t_1$, at which the function $\psi_2(t)$ equals zero, so $\psi_2(t_p) = 0$. Since $t_0 = 0$ then $t_p > 0$ and $\psi_2(t_p) = \varrho - t_p = 0$ hence

$$\varrho = t_p. \tag{40}$$

Thus, for the existence of a swich it is necessary that $\varrho > 0$; if $\varrho \leqslant 0$ no swich occurs.

Taking into account (40) from (37) we obtain

$$t_p - t_1 = x_1^2/x_1^1. \tag{41}$$

In the interval of time $t_p < t < t_1$ there is $\overline{w}(t) = -1$. Substituting this value into the second equation of (30) and integrating it over t from t_p to t_1 we get

$$x^2(t_1) - x^2(t_p) = t_p - t_1.$$

If we denote $x^1(t_p) = x_p^1$, $x^2(t_p) = x_p^2$ and remember that $x^1(t_1) = x_1^1$, $x^2(t_1) = x_1^2$ then

$$x_1^2 - x_p^2 = t_p - t_1. \tag{42}$$

Combining formulae (41) and (42) we obtain

$$x_1^2 - x_p^2 = x_1^2/x_1^1. \tag{43}$$

For the control $\overline{w}(t) = +1$ $(t_0 < t < t_p)$ the corresponding trajectory is a parabola passing the points x_0, x_p and on the strength of (36) we have

$$(x_p^2)^2 - (x_0^2)^2 = 2(x_p^1 - x_0^1). \tag{44}$$

For $\overline{w}(t) = -1$ $(t_p < t < t_1)$ the parabola passes the points x_p and $x_1 \epsilon g$ so we have

$$(x_1^2)^2 - (x_p^2)^2 = -2(x_1^1 - x_p^1). \tag{45}$$

Since the point x_1 lies on the circle g its coordinates obey the equation

$$(x_1^1)^2 + (x_1^2)^2 = 1. \tag{46}$$

Thus we have obtained the system of four equations (43)-(46) wherefrom one can get the unknown coordinates of the points x_p and $\overline{x}_1$ (the coordinates of x_0 are given in advance).

From this system we will calculate the equation of the swich curve (let us remember that now we consider the case of trajectories terminating at points x_1 of the circle g which lie in the second quadrant). From (43) we obtain

$$x_p^2 = x_1^2\left(1 - \frac{1}{x_1^1}\right). \tag{47}$$

Substituting this into (45) and solving it with respect to x_p^1 yields

$$x_p^1 = \frac{1}{x_1^1} - \frac{1}{2}\frac{(x_1^2)^2}{(x_1^1)^2}$$

or, applying (46),

$$\begin{aligned} x_p^1 &= \frac{1}{x_1^1} - \frac{1}{2}\left(\frac{1}{x_1^1}\right)^2 + \frac{1}{2}, \\ x_p^2 &= \sqrt{1 - \frac{1}{(1/x_1^1)^2}}\left(1 - \frac{1}{x_1^1}\right). \end{aligned} \tag{48}$$

Equations (48) are the parametric equations of the switch curve and in the second quadrant the parameter x_1^1 runs over the interval $-1 \leqslant x_1^1 < 0$ or the parameter $1/x_1^1$ over the interval $-\infty < 1/x_1^1 \leqslant -1$. Remembering this, from the first one of equations (48) we obtain

$$\frac{1}{x_1^1} = 1 - \sqrt{2 - 2x_p^1}. \tag{49}$$

Substituting (49) into the second equation of (48) we obtain the following equation of the swich curve p^-

(50) $$x_p^2 = \sqrt{1-1/(1-\sqrt{2-2x_p^1})^2}\sqrt{2-2x_p^1}.$$

It is easy to notice that the equation of the swich curve p^+ for trajectories terminating at points x_1 of the circle g which belong to the fourth quadrant may be obtained from (50) by substituting $-x_p^1$ and $-x_p^2$ for x_p^1 and x_p^2, respectively. Thus, for the fourth quadrant we have

(51) $$x_p^2 = -\sqrt{1-1/(1-\sqrt{2+2x_p^1})^2}\sqrt{2+2x_p^1}.$$

The knowledge of the course of swich curves p^- and p^+ enables the determination of the optimal control $\overline{w}(x^1, x^2)$ which is a function of the coordinates of the point x in plane X. The swich curves p^-, p^+ and the circle g divide the plane X into parts in each of them the control $\overline{w}(t)$ is determined in a different way. So, to the left from the line consisting of the curve p^-, a part of the circle g and of the curve p^+ including points on p^+ there is $\overline{w}(x^1, x^2) = +1$. In the remaining part of the plane X with the exception of the set Z_x^+ there is $\overline{w}(x^1, x^2) = -1$. Inside and on the circle g (i.e. for the points $x \in Z_x^+$) the control $\overline{w}(x^1, x^2)$ at the point $x = (x^1, x^2)$ has to fulfil the relation $\overline{w}(x^1, x^2) = -x^1$. This follows from the definition of Z_x^+ and from the first equation of (29) in which one has to remember that for $t > t_1$ there is $w(t) \in Z_w^+$ and $y(t) = 0$.

The knowledge of $\overline{w}(x^1, x^2)$ may help us to construct the system which realizes the optimal control in the sense of Problem I.

Depending on to which part of the plane X belongs the starting point x_0 the control $\overline{w}(t)$, $t_0 < t < t_1$, differs in its course. For example, at points x_0 belonging to the region A^- bounded by the curve p^-, a part of the circle g and the parabola l^- (see Fig. 5b) there is $\overline{w}(t) = -1$, $t_0 < t < t_1$, while at the points x_0 in the region B^+ bounded by the curve p^-, a part of the circle g and the curve p^+ there is $\overline{w}(t) = +1$, $t_0 < t < t_p$ and $\overline{w}(t) = -1$, $t_p < t < t_1$.

If x_0 belongs to A^-, then for a trajectory passing through the point x_0 we may find coordinates of the terminal point $\overline{x}_1$ applying the equations

(52) $$(x_1^2)^2-(x_0^2)^2 = -2(x_1^1-x_0^1)$$

and (46). Thus we obtain

(53) $$\overline{x}_1^1 = 1-\sqrt{2-2x_0^1-(x_0^2)^2},$$

and $\overline{x}_1^2$ may be calculated from (46).

If x_0 belongs to B^+, then the control $\overline{w}(t)$ has a switch. Point x_p in which the switching occurs is an intersection point of the corresponding parabola passing through the point x_0 and the switch curve p^- and, therefore, its coordinates may be calculated from the system of equations

$$(x_p^2)^2-(x_0^2)^2 = 2(x_p^1-x_0^1),$$

$$x_p^2 = \sqrt{1-1/(1-\sqrt{2-2x_p^1})^2}\sqrt{2-2x_p^1},$$

where x_0^1 and x_0^2 are known. Eliminating x_p^2 and introducing a new variable

(54) $$z = 1-\sqrt{2-2x_p^1} = 1/x_1^1$$

we obtain an algebraic equation of the fourth degree

(55) $$2z^4-4z^3-[(x_0^2)^2-2x_0^1+1]z^2+2z-1 = 0$$

which may be solved, e.g. applying the Cardano formulae. If $\bar{z}$ is a root of the equation (55) and $\bar{z} < -1$, then the coordinates x_p^1 and $\bar{x}_1^1$ may be calculated as follows

(56) $$x_p^1 = 1-\tfrac{1}{2}(1-\bar{z})^2, \qquad \bar{x}_1^1 = 1/\bar{z}.$$

Since the calculation of $\bar{z}$ from equation (55) may require a lot of work while it is quite easy to calculate $\theta = x_0^1-\frac{1}{2}(x_0^2)^2$ for a given value of $\bar{z}$ we may draw a diagram of the function $\theta = \Phi(\bar{z})$ (see Fig. 6). This

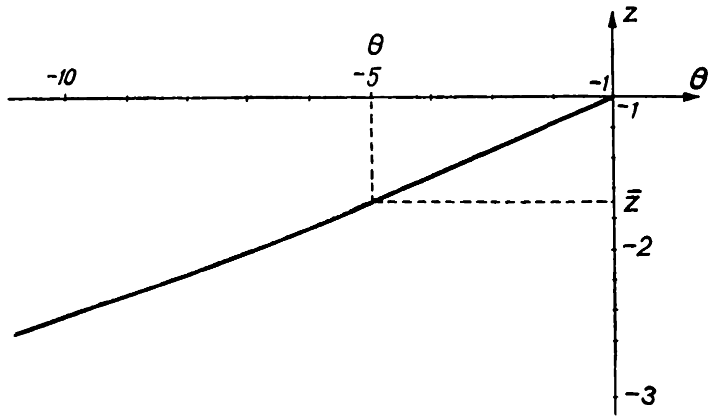


Fig. 6. A diagram for the approximate solution of equation (55)

leads to a graphical method of determination of $\bar{z}$ for the given values of x_0^1 and x_0^2. Applying (56) we then find x_p^1 and $\bar{x}_1^1$ to obtain finally from (50) and (46) the values of x_p^2 and $\bar{x}_1^2$. The meaning of θ is shown in Fig. 5b.

To determine the time interval for which $\bar{w}(t) = +1$ we may use the formula

(57) $$t_p-t_0 = x_p^2-x_0^2$$

which follows directly from the second equation of (30). The time interval for which $\bar{w}(t) = -1$ is given by formula (42).

To determine coordinates of the point $\bar{x}_1 \in g$ in the case when x_0 belongs to A^+ or to B^- one should apply the formulae following from (53) or (54) and (55), (56) by the substitution of $-x_0^1$, $-\bar{x}_1^1$, $-\bar{z}$, $-x_p^1$ instead of $x_0^1, \bar{x}_1^1, \bar{z}, x_p^1$.

If we know the coordinates of the terminal point x_1 then, applying Theorem 1, we may determine the control $w^*(t)$ for $t > t_1$. The solutions of equation (31) are of the from

$$w^*(t) = A\sin(t+\varphi), \tag{58}$$

where $A > 0$ and φ are constant.

Initial conditions for the control $w^*(t)$ at the moment t_1 may be obtained from relation (32). This yields

$$\begin{aligned} w^*(t_1) &= -\bar{x}_1^1, \\ \dot{w}^*(t_1) &= -\bar{x}_1^2. \end{aligned}$$

Therefore it has to be

$$\begin{aligned} A\sin(t_1+\varphi) &= -\bar{x}_1^1, \\ A\cos(t_1+\varphi) &= -\bar{x}_1^2. \end{aligned}$$

Taking into account (46) we obtain for constants A and φ the following values:

$$A = 1, \quad \varphi = \operatorname{arc\,tg}\frac{\bar{x}_1^1}{\bar{x}_1^2} - t_1. \tag{59}$$

The solution of Problem II. According to Theorem 2 first we have to solve in plane X the problem of time-optimal control from the set $Z_x^-(y_0)$ to Z_x^+. The set Z_x^+ has been already determined. To determine the set $Z_x^-(y_0)$ we have to know the set Z_w^- in plane W. In the case $m = 2$ the set Z_w^- is shown in Fig. 3 though we have to remember that in our case $\alpha = -1$ and $\beta = +1$. According to the definition of $Z_x^-(y_0)$ (for a given point y_0) the coordinates of points $x_0 \in Z_x^-(y_0)$ may be obtained from the formulae

$$\begin{aligned} x_0^1 &= y_0 - w_0, \\ x_0^2 &= \dot{y}_0 - \dot{w}_0, \end{aligned} \tag{60}$$

where $y_0, \dot{y}_0$ are the coordinates of y_0 and $(w_0, \dot{w}_0) \in Z_w^-$. The set $Z_x^-(y_0)$ is shown in Fig. 7 where points on the borders γ^- and γ^+ marked by a continuous line belong to $Z_x^-(y_0)$ while those marked by a dashed line do not belong to $Z_x^-(y_0)$. Since the trajectories in plane X are continuous, then we may restrict ourselves, as we did formerly, to the case when the starting point x_0 lays on straight lines γ^- or γ^+ and the terminal point x_1 lays on the circle g. Though we have to remember that if solving the problem of time-optimal control in plane X we get a starting point x_0 in dashed parts of straight lines γ^- or γ^+ then the problem in fact has no solution (since x_0 does not belong to $Z_x^-(y_0)$). In this case the obtained solution determines only the lower limit for the control time $T = t_1 - t_0$. We may arbitraryly aproach this limit (choosing the starting point $x_0' \in Z_x^-(y_0)$ sufficiently close to $\bar{x}_0$) but it cannot be reached by any control which is admissible in the time interval $t > t_0 - \varepsilon$.

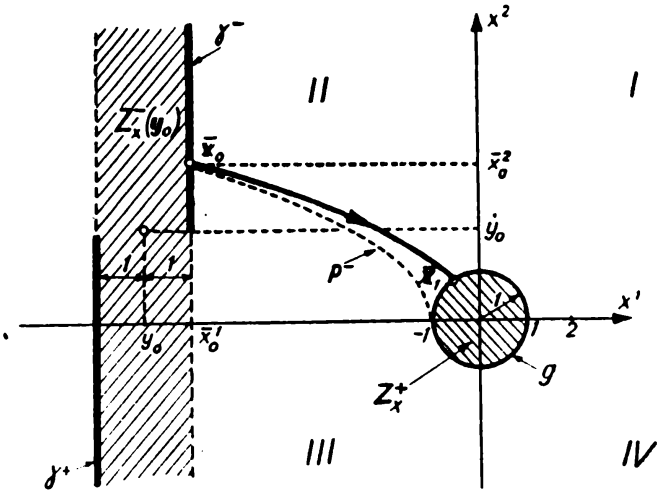


Fig. 7. Optimal trajectory in the plane X in the case of Problem II

From the transversality condition in the starting point x_0 of a trajectory we get

$$\psi_2(t_0) = c_2 - c_1 t_0 = 0$$

or

(61)
$$\psi_2(t) = c_1(t_0 - t),$$

so the function $\psi_2(t)$ for $t > t_0$ does not alter its sign. Taking into account the relation (35), which holds as well in the present case, we conclude that the control $\bar{w}(t)$ takes either the value $+1$ or the value -1 in the whole time interval $t_0 < t < t_1$ (there are no switch points). The transversality condition in the terminal point x_1 is of the form

$$\frac{\psi_2(t_1)}{\psi_1(t_1)} = \frac{c_1(t_0 - t_1)}{c_1} = \frac{x_1^2}{x_1^1}$$

or

(62)
$$t_0 - t_1 = \frac{x_1^2}{x_1^1}.$$

Since in the case when $\overline{w}(t) = -1$ for $t_0 < t < t_1$ there is

$$t_0 - t_1 = x_1^2 - x_0^2$$

then we have

$$x_1^2 - x_0^2 = \frac{x_1^2}{x_1^1}. \tag{63}$$

This case obviously takes place if sets $Z_x^-(y_0)$ and Z_x^+ have no points in common and $Z_x^-(y_0)$ lays in plane X to the left of the set Z_x^+, i.e. if $y_0 < -2$. To determine the coordinates of the points $\overline{x}_0$ and $\overline{x}_1$ besides (63) we have the following equations

$$(x_1^2)^2 - (x_0^2)^2 = -2(x_1^1 - x_0^1), \tag{64}$$

$$(x_1^1)^2 + (x_1^2)^2 = 1, \tag{65}$$

$$\overline{x}_0^1 = y_0 + 1. \tag{66}$$

Let us notice that equations (63), (64) and (65) are similar to those in (43), (45) and (46), with the only difference that in (63) and (64) there are coordinates of the point x_0 while in (43) and (45) there are coordinates of x_p. Therefore some of the previous results may be used now if we substitute the coordinates of x_0 for those of x_p. Thus, formula (49) yields

$$\overline{x}_1^1 = \frac{1}{2\overline{x}_0^1 - 1}\left(1 + \sqrt{2 + 2\overline{x}_0^1}\right) \tag{67}$$

and from (50) we obtain

$$\overline{x}_0^2 = \sqrt{1 - 1/\left(1 - \sqrt{2 - 2\overline{x}_0^1}\right)^2}\sqrt{2 - 2\overline{x}_0^1}. \tag{68}$$

Since $\overline{x}_0^1$ can be directly calculated from (66) so using formulae (67), (68) and equation (63) we can calculate the remaining coordinates of $\overline{x}_0$ and $\overline{x}_1$.

Comparing formulae (50) and (68) we find out that the point $\overline{x}_0$ lies on the curve p^-. It can be obtained as the point of intersection of the straight line $\overline{x}_0^1 = y_0 + 1$ (which is paraller to the axis x^2) and the switch curve p^- (if $y_0 < -2$) which was determined in Problem I.

If $y_0 > 2$, i.e. in the case when $Z_x^-(y_0)$ lies to the right of the set Z_x^+, then the control $\overline{w}(t)$, $t_0 < t < t_1$ takes the value $+1$ and the appropriate formulae for coordinates of $\overline{x}_0$ and $\overline{x}_1$ can be obtained by a substitution of $-\overline{x}_0^1, -\overline{x}_1^1, -y_0, -\overline{x}_0^2$ for $\overline{x}_0^1, \overline{x}_1^1, y_0, \overline{x}_0^2$ in the formulae discussed above.

If $-2 < y_0 < 2$, then $Z_x^-(y_0)$ and Z_x^+ have some points in common and by Corollary 2 we can find a control for which $T = t_1 - t_0 = 0$.

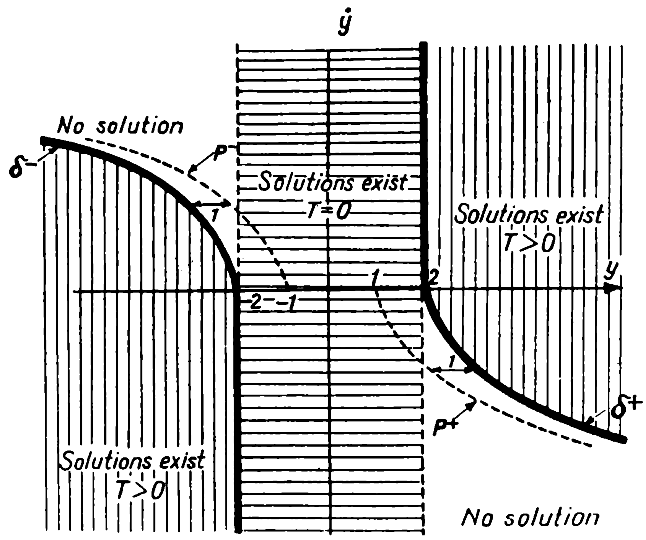
According to the Theorem 2 the knowledge of coordinates of both the starting point $\overline{x}_0$ and the terminal point $\overline{x}_1$ makes it possible:

(a) to determine the optimal control $w^*(t)$ for $t > t_1$ using formulae (58) and (59);

(b) to calculate the coordinates of the vector $\boldsymbol{w}(t_0^-)$ by the formulae

$$w(t_0^-) = y_0 - \bar{x}_0^1, \qquad \dot{w}(t_0^-) = \dot{y}_0 - \bar{x}_0^2. \tag{69}$$

As we have mentioned, the optimal control does not exist if $\bar{\boldsymbol{x}}_0$ lies on the dashed parts of the straight lines γ^- or γ^+. Taking into account that $\bar{\boldsymbol{x}}_0$ is the point of intersection of the straight line γ^- (or γ^+) with the curve p^- (or p^+) and that the solution of Problem II does not exist if $\bar{\boldsymbol{x}}_0$ lies on the dashed part of γ^- or γ^+, we can now determine a region of points $\boldsymbol{y}_0$ (in space Y) for which there exists a solution of Problem II.



Fif. 8. The region of points $\boldsymbol{y}_0$ in plane Y for which there exist solutions of Problem II

This region is shown in Fig. 8. It is easy to notice that the line δ^- may be obtained from curve p^- by shifting it parallelly to the axis y one unit to the left and the line δ^+ from curve p^+ by shifting it one unit to the right.

Formulae derived above were used to obtain a solution of Problem I in the case when $\boldsymbol{y}_0 = (-5, 0)$, $\boldsymbol{w}_0 = (0, 0)$, $\boldsymbol{y}_1 = (0, 0)$. The results are presented graphically: in Fig. 9a the optimal trajectory in space X; in Fig. 9b the optimal trajectory in space Y; in Fig. 9c the optimal control $\tilde{w}(t)$. For a comparison in Fig. 10 there are presented results of solving Problem II for $\boldsymbol{y}_0 = (-5,0)$ and $\boldsymbol{y}_1 = (0,0)$; the optimal trajectory in space X (Fig. 10a); the optimal trajectory in space Y (Fig. 10b); the optimal control $\tilde{w}(t)$ with the vector $\boldsymbol{w}_0 \approx (-1, -2{,}8)$ (Fig. 10c). From this example we can see that for the same points $\boldsymbol{y}_0$ and $\boldsymbol{y}_1$ time $T = t_1 - t_0$ under the optimal control in the sense of Problem II is shorter than that under the optimal control in the sense of Problem I.

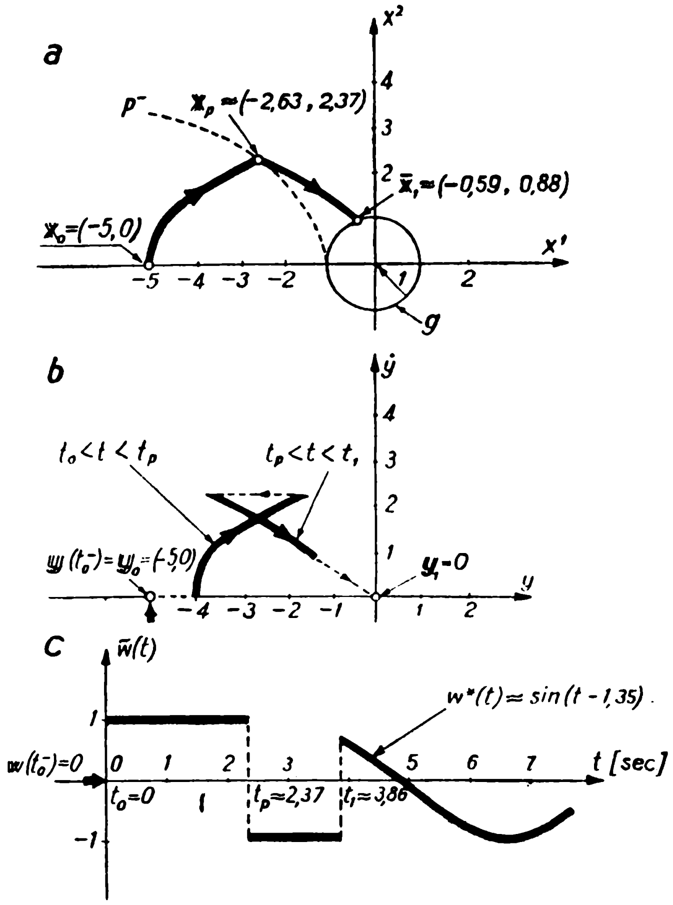


Fig. 9. Optimal solution of Problem I for the following data: $y_0 = (-5, 0)$, $w_0 = 0$, $y_1 = 0$

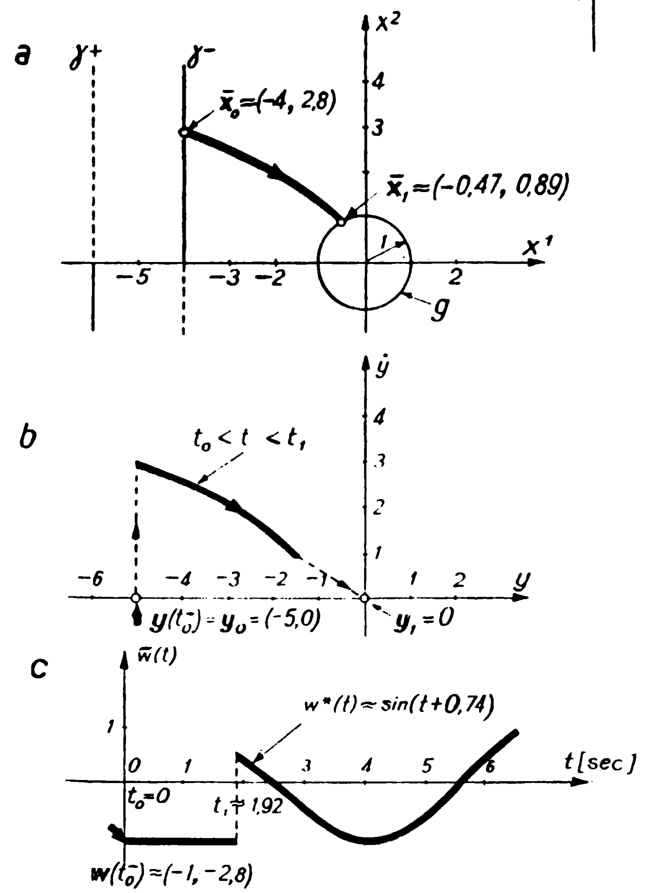


Fig. 10. Optimal solution of Problem II for the following data: $y_0 = (-5, 0)$, $y_1 = 0$

8. Closing remarks. The discussion presented above shows that the time-optimal control of plants whose transfer functions contain zeros considerably differs from that of plants with transfer functions having no zeros (or more general from the time-optimal control of plants with continuous solutions — under admissible controls — of the equations describing the plant). In the last case the time-optimal control does not depend on the way in which coordinates of the plant "have reached the location of the point y_0". In other words it does not depend on what happened to the plant since the initial moment t_0 of the control (i.e. on the history of the plant) depending only on the starting point y_0 (having given the terminal point y_1). In the former case the time-optimal control depends not only on the starting point y_0 but also on the vector $w(t_0^-)$, i.e. on the "position of the rudder" at which coordinates of the plant "reached the starting point y_0" and so it depends on what happened to the plant just before the initial moment t_0 of the control.

Therefore in the paper two problems were stated. With the first one we have to deal if we cannot influence the control $w(t)$ before the initial moment t_0 accepting the fact that coordinates of the plant were brought to the position of the point y_0 with the given "position of the rudder", that is with the given value of $w(t_0^-)$. In this case, using the new variable x and the function $\overline{w} = \overline{w}(x)$ which expresses the dependence of the optimal value $\overline{w}$ on the point x determined by new coordinates of the plant [2], we may perform the synthesis of a system giving the optimal control (in the sense of Problem I) of the plant in the interval of time $t > t_0$. This control depends on the starting point y_0 as well as on the "position of the rudder" $w(t_0^-)$ (and in the space X on the point x_0 determined by the points y_0 and $w(t_0^-) = w_0$). We have to emphasize that such a system may realize the optimal control immediately after the moment of receiving information on the initial point y_0 and the "position of the rudder" $w(t_0^-)$.

It is obvious that in the case of Problem II the realization of an optimal control has to be based on different rules. In this case the sulution of the problem has to be attained before the initial moment t_0 of the control, to provide the calculated optimal "position of the rudder" $w(t_0^-)$ when coordinates of the plant reach the position of point y_0. To fulfil that requirement a certain interval of time is necessary before the initial moment t_0 of the control. If no restriction is imposed, then coordinates of the plant may be brought to the position of point y_0 at the moment t_0 in an arbitrary way which only guarantees the calculated "position of the rudder" $w(t_0^-)$ just before the moment t_0.

Thus the realization of the optimal control in the sense of Problem II consists of two steps. The first step is a preparatory one and from the point of view of Problem II only the final effect of this step is relevant: to reach the best conditions, under existing restrictions, for the time-optimal control during the principal second step. In the second step the optimal control in the sense of Problem I is then performed. Since now the optimal control is carried on in the best initial conditions achieved in step one so the time $T = t_1 - t_0$ for the optimal control in the sense of Problem II in general is smaller then that for the optimal control in the sense of Problem I.

The arguments of this paper do not cover all the considered problems but they may be used to find similar answers to many other questions concerning the discussed plants.

APPENDIX 1. PROOF OF FORMULAE (7) AND (8)

From the assumption that the derivatives $w^{(j)}(t)$ and $y^{(i)}(t)$ in (1) are understood in the distributional sense follows(17) that

$$\int_{t'}^{t''} y^{(i)}(t)\,dt = y^{(i-1)}(t'') - y^{(i-1)}(t'), \qquad i = 0, 1, \ldots, n, \tag{1'}$$

$$\int_{t'}^{t''} w^{(j)}(t)\,dt = w^{(j-1)}(t'') - w^{(j-1)}(t'), \qquad j = 1, 2, \ldots, m, \tag{2'}$$

for any pair of moments t' and t'' at which the derivatives $y^{(i-1)}(t)$ and $w^{(j-1)}(t)$ are determined.

Let us assume now that $b_i = 0$ for $i = 0, 1, \ldots, n-m-1$, so that equation (1) may be written in the form

$$y^{(n)} + a_1 y^{(n-1)} + \ldots + a_n y = b_0 w^{(n)} + b_1 w^{(n-1)} + \ldots + b_n w. \tag{3'}$$

Let τ be a discontinuity point of the vector-function $w(t)$. From (2') it follows

$$\lim_{\Delta t \to 0} \int_{\tau-\Delta t}^{\tau+\Delta t} dt_1^* \int_{\tau-\Delta t}^{t_1^*} dt_2^* \ldots \int_{\tau-\Delta t}^{t_{i-1}^*} w^{(j)}(t_i^*)\,dt_i^* = \begin{cases} \Delta w^{(j-i)}(t) & \text{for } i \leqslant j, \\ 0 & \text{for } i > j, \end{cases} \tag{4'}$$

where, according to (6), there is

$$\Delta w^{(j-i)}(\tau) = w^{(j-i)}(\tau^+) - w^{(j-i)}(\tau^-). \tag{5'}$$

(17) See [4], p. 62.

The operation specified on the left-hand side of equation (4′) will be called in the sequel an *i times integration around the point* $t = \tau$.

In a similar way we have

$$\lim_{\Delta t\to 0}\int\limits_{\tau-\Delta t}^{\tau+\Delta t} dt_1^*\int\limits_{\tau-\Delta t}^{t_1^*} dt_2^*\ldots\int\limits_{\tau-\Delta t}^{t_{i-1}^*} y^{(j)}(t_i^*)\,dt_i^* = \Delta y^{(j-i)}(\tau), \tag{6′}$$

where

$$\Delta y^{(j-i)}(\tau) = y^{(j-i)}(\tau^+)-y^{(j-i)}(\tau^-) \tag{7′}$$

and $y^{(j-i)}(t)$, if $j-i<0$, stands for the indefinite integration $i-j$ times of the function $y(t)$.

It is easy to notice that solution $y(t)$ of equation (3′) is at least piecewise continuous function with only first order discontinuity points and a continuous function if $m<n$. To prove this let us suppose that $t=\tau$ was a first order discontinuity point of the function $\psi(t) = \int y(t)dt$. Then, integrating $n+1$ times around the point $t=\tau$ both sides of equation (3′) we would get zero on the right-hand side and $\Delta\psi(\tau)\neq 0$ on the left-hand side, so that the equation would not be fulfilled. Therefore, it has to be $\Delta\psi(\tau) = 0$ and $t=\tau$ is at most a first order discontinuity point of the function $y(t)$. Having this in mind we get

$$\Delta y^{(j-i)}(\tau) = 0 \quad \text{for } j<i. \tag{8′}$$

To prove folmulae (7) and (8) we shall proceed by induction. First we shall integrate n times around the point $t=\tau$ both sides of equation (3′). Because of (4′), (6′) and (8′) we obtain

$$\Delta y(\tau) = b_0\Delta w(\tau) = \gamma_0\Delta w(\tau), \tag{9′}$$

where $\gamma_0 = b_0$.

In a similar way, integrating $n-1$ times around the point $t=\tau$ both sides of equation (3′), we get

$$\Delta y^{(1)}(\tau)+a_1\Delta y(\tau) = b_0\Delta w^{(1)}(\tau)+b_1\Delta w(\tau).$$

Applying (9′), we obtain

$$\Delta y^{(1)}(\tau) = b_0\Delta w^{(1)}(\tau)+(b_1-a_1\gamma_0)\Delta w(\tau) = \gamma_0\Delta w^{(1)}(\tau)+\gamma_1\Delta w(\tau),$$

where $\gamma_0 = b_0$, $\gamma_1 = b_1-a_1\gamma_0$, and so formulae (7) and (8) are valid for $i=0$ and 1. We shall now assume their validity for $i = 0, 1, \ldots, j$, where $j < n$. Integrating $n-j-1$ times around the point $t = \tau$ both sides of equation (3'), we obtain

$$\Delta y^{(j+1)}(\tau)+a_1\Delta y^{(j)}(\tau)+\ldots+a_{j+1}\Delta y(\tau)$$
$$= b_0\Delta w^{(j+1)}(\tau)+b_1\Delta w^{(j)}(\tau)+\ldots+b_{j+1}\Delta w(\tau).$$

Now, applying formulae (7) and (8) for $i = 0, 1, \ldots, j$, we get

$$\Delta y^{(j+1)} = b_0\Delta w^{(j+1)}(\tau)+(b_1-a_1\gamma_0)\Delta w^{(j)}(\tau)+(b_2-a_1\gamma_1-a_2\gamma_0)\Delta w^{(j-1)}(\tau)+$$
$$+\ldots+(b_{j+1}-a_1\gamma_j-a_2\gamma_{j-1}-\ldots-a_{j+1}\gamma_0)\Delta w(\tau)$$

or

$$\Delta y^{(j+1)}(\tau) = \gamma_0\Delta w^{(j+1)}(\tau)+\gamma_1\Delta w^{(j)}(\tau)+\ldots+\gamma_{j+1}\Delta w(\tau),$$

where

$$\gamma_{j+1} = b_{j+1}-a_1\gamma_j-a_2\gamma_{j-1}-\ldots-a_{j+1}\gamma_0.$$

This completes the proof of formulae (7) and (8).

APPENDIX 2. THE CASE OF LIMITED VALUE OF THE k-TH DERIVATIVE OF THE CONTROL INPUT $w(t)$

Up to now we have discussed the case where the control input $w(t)$ is a piecewise continuous function having at most first order discontinuity points and obeying the restriction (4). Such a function we have called the admissible control.

Here we shall consider the case where some bounds are imposed upon the value of k-th derivative of the control input.

Thus, we assume now that the admissible control $w(t)$ is a function of time for which the derivative $w^{(k)}(t)$ ($k < m$ is a given positive integer) is a piecewise continuous function having at most finite number of first order discontinuity points in any limited interval of time and taking values which belong to the fixed interval

$$\alpha \leqslant w^{(k)}(t) \leqslant \beta. \tag{4*}$$

From this assumption it follows of course that the admissible control $w(t)$ and all its derivatives $w^{(i)}(t)$ of lower order $i = 1, 2, \ldots, k-1$ are continuous functions of time. With respect to the derivatives of higher order $i = k+1, \ldots, m-1$ we accept similar requirements to those which were formerly imposed upon the derivatives of $w(t)$. We may substitute

$$\begin{aligned} w^{(k)} &= u, \\ w^{(k+1)} &= u^{(1)}, \\ &\cdots\cdots \\ w^{(m)} &= u^{(m-k)} \end{aligned} \tag{4*a}$$

considering $u(t)$ to be a new control input. The restrictions imposed upon $u(t)$ are now of the same kind as those formerly imposed upon $w(t)$. Let us introduce now the new variables:

$$
\begin{aligned}
x^1 &= y,\\
x^2 &= y^{(1)},\\
&\dots\dots\dots\\
x^{n-m+k} &= y^{(n-m+k-1)},\\
x^{n-m+k+1} &= y^{(n-m+k)} - \gamma_{n-m} w^{(k)},\\
x^{n-m+k+2} &= y^{(n-m+k+1)} - \gamma_{n-m} w^{(k+1)} - \gamma_{n-m+1} w^{(k)},\\
&\dots\dots\dots\\
x^n &= y^{(n-1)} - \gamma_{n-m} w^{(m-1)} - \gamma_{n-m+1} w^{(m-2)} - \dots - \gamma_{n-k-1} w^{(k)},\\
x^{n+1} &= w,\\
x^{n+2} &= w^{(1)},\\
&\dots\dots\dots\\
x^{n+k} &= w^{(k-1)},
\end{aligned}
\tag{11*}
$$

where the constants γ_i are now defined by a recurrent formula

$$
\begin{aligned}
\gamma_{n-m} &= b_{n-m},\\
\gamma_i &= b_i - a_1\gamma_{i-1} - a_2\gamma_{i-2} - \dots - a_{i-n+m}\gamma_{n-m}\\
&\qquad \text{for } i = n-m+1,\ n-m+2, \dots, n-k.
\end{aligned}
\tag{8*}
$$

Applying formulae (4*a) and the last k of equations (11*) to equation (1) we obtain

$$
\begin{aligned}
y^{(n)} + a_1 y^{(n-1)} + \dots + a_n y &= b_{n-m} u^{(m-k)} + b_{n-m+1} u^{(m-k-1)} +\\
&+ \dots + b_{n-k} u + (b_{n-k+1} x^{n+k} + b_{n-k+2} x^{n+k-1} + \dots + b_n x^{n+1}).
\end{aligned}
\tag{1*}
$$

Now, from the first n equations (11*) we may calculate $y^{(i)}, i = 0, 1, \dots, n$. The substitution of these to (1*) and the use of (8*) yields

$$
\begin{aligned}
&\dot{x}^n + a_1 x^n + a_2 x^{n-1} + \dots + a_n x^1\\
&\qquad = \gamma_{n-m} u + (b_{n-k+1} x^{n+k} + b_{n-k+2} x^{n+k-1} + \dots + b_n x^{n+1}).
\end{aligned}
\tag{1*a}
$$

Taking into account (11*) and (1*a) we get for $n+k$ variables $x^1, x^2, \dots, x^{n+k}$ the following system of $n+k$ differential equations describing the movement of a point in the $(n+k)$-dimensional space X:

$$
(12^*)\qquad
\begin{aligned}
\dot{x}^1 &= x^2,\\
\dot{x}^2 &= x^3,\\
&\dots\dots\dots\dots\\
\dot{x}^{n-m+k-1} &= x^{n-m+k},\\
\dot{x}^{n-m+k} &= x^{n-m+k+1}+\gamma_{n-m}u,\\
\dot{x}^{n-m+k+1} &= x^{n-m+k+2}+\gamma_{n-m+1}u,\\
&\dots\dots\dots\dots\dots\dots\\
\dot{x}^{n-1} &= x^n+\gamma_{n-k-1}u,\\
\dot{x}^{n} &= -a_nx^1-a_{n-1}x^2-\dots-a_1x^n+b_nx^{n+1}+b_{n-1}x^{n+2}+\\
&\qquad +\dots+b_{n-k+1}x^{n+k}+\gamma_{n-k}u,\\
\dot{x}^{n+1} &= x^{n+2},\\
\dot{x}^{n+2} &= x^{n+3},\\
&\dots\dots\dots\dots\\
\dot{x}^{n+k} &= u.
\end{aligned}
$$

In a similar way as it has been done before, it is easy to show that to the defined in section 2 solutions $y(t)$ of equation (1) (with the new definition of admissible controls) there correspond, according to (11*), continuous solutions $x(t) = \big(x^1(t), x^2(t), \dots, x^{n+k}(t)\big)$ of the system of equations (12*).

Equations (11*) may be looked upon (just as it formerly was with equations (11)) as relations which transform a pair of points y and w (y belongs to n-dimensional space Y and w to m-dimensional space W) into a point x in $(n+k)$-dimensional space X. We shall say that the point x corresponds, according to relations (11*), to the pair of points y and w. To a pair of points $x\in X$ and $w\in W$ (but only in the case where the last k coordinates of x are equal to those of w) there corresponds, according to (11*) a unique point y in the space Y. Similar remarks remain valid in the case where x, y and w in (11*) denote the functions of time $x(t)$, $y(t)$ and $w(t)$, respectively.

With some adjustments, which we shall discuss in the sequel, all previous lemmas, theorems and their proofs remain valid also in the circumstances described here. To start with, all the formulae occurring in the formulation and the proofs of all lemmas and theorems in the main part of the paper now should be replaced by the new ones which are similar in form (especially in the matrix notation) and have the same numbers with only an asterisk added. Thus, for example, in a former statement: "the point x (or the trajectory $x(t)$) corresponds, according to (11), to the pair of points y and w (or to the trajectory $y(t)$ and the function $w(t)$)" now we have to replace the phrase "according to (11)" by the new form of it "according to (11*)".

The previously used term "admissible control $w(t)$" now has to be understood in its new meaning defined here (with bounds imposed upon $w^{(k)}(t)$).

The set Z_w^+ is now a set of all points $\mathbf{w}_1$ of the m-dimensional space W which, taken for the initial condition for solutions $w(t)$ of equation (14) at the moment t_1 ($\mathbf{w}(t_1) = \mathbf{w}(t_1^+) = \mathbf{w}_1$), guarantee the fulfilment of condition (4*) for $t > t_1$.

The set Z_x^+ is now defined as a set of points $\mathbf{x}_1$ of the $(n+k)$-dimensional space X which correspond, according to (11*), to all pairs of points $\mathbf{y}_1 = 0$ and $\mathbf{w}_1 \in Z_w^+$. Thus, the coefficients $x_1^1, x_1^2, \ldots, x_1^{n+k}$ of points $\mathbf{x}_1 \in Z_x^+$ may be calculated from the formulae

$$
\begin{aligned}
x_1^1 &= 0,\\
&\cdots\cdots\cdots\\
x_1^{n-m+k} &= 0,\\
x_1^{n-m+k+1} &= -\gamma_{n-m} w_1^{(k)},\\
x_1^{n-m+k+2} &= -\gamma_{n-m+1} w_1^{(k)} - \gamma_{n-m} w_1^{(k+1)},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
x_1^{n} &= -\gamma_{n-k-1} w_1^{(k)} - \gamma_{n-k-2} w_1^{(k+1)} - \ldots - \gamma_{n-m} w_1^{(m-1)},\\
x_1^{n+1} &= w_1,\\
x_1^{n+2} &= w_1^{(1)},\\
&\cdots\cdots\cdots\\
x_1^{n+k} &= w_1^{(k-1)},
\end{aligned}
\tag{16*}
$$

where $w_1, w_1^{(1)}, \ldots, w_1^{(m-1)}$ are the coordinates of an arbitrary point $\mathbf{w}_1 \in Z_w^+$. The last m formulae of (16*) may be written in the form

$$\mathfrak{x}_1 = -\Gamma^* \mathbf{w}_1, \tag{16*a}$$

where

$$\mathbf{w}_1 = (w_1, w_1^{(1)}, \ldots, w_1^{(m-1)}),$$

$$\mathfrak{x}_1 = (x_1^{n-m+k+1}, x_1^{n-m+k+2}, \ldots, x_1^{n+k}), \tag{17*}$$

and

$$
\Gamma^* = \begin{bmatrix}
0 & 0 & \ldots & 0 & \gamma_{n-m} & 0 & 0 & \ldots & 0\\
0 & 0 & \ldots & 0 & \gamma_{n-m+1} & \gamma_{n-m} & 0 & \ldots & 0\\
\cdot & \cdot & \ldots & \cdot & \cdot & \cdot & \cdot & \ldots & \cdot\\
\cdot & \cdot & \ldots & \cdot & \cdot & \cdot & \cdot & \ldots & \cdot\\
0 & 0 & \ldots & 0 & \gamma_{n-k+1} & \gamma_{n-k+2} & \gamma_{n-k+2} & \cdots & \gamma_{n-m}\\
1 & 0 & \ldots & 0 & 0 & 0 & 0 & \ldots & 0\\
0 & 1 & \ldots & 0 & 0 & 0 & 0 & \ldots & 0\\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\
0 & 0 & \ldots & 1 & 0 & 0 & 0 & \ldots & 0
\end{bmatrix}. \tag{18*}
$$

Due to the assumption $b_{n-m} \neq 0$ the matrix Γ^* is nonsingular

$$(19^*) \qquad \det \Gamma^* = (-1)^{k(m-k)} b_{n-m}^{m-k} \neq 0$$

and there exists an inverse matrix Γ^{*-1} so that

$$(20^*) \qquad \boldsymbol{w}_1 = -\Gamma^{*-1} \mathfrak{x}_1.$$

Thus, formula (16*a) or (20*) determines an one-to-one correspondance between the points of spaces W and $\mathfrak{X}$.

From (16*) it follows that any point $\boldsymbol{x}_1 \in Z_x^+$ has the first $n-m+k$ coordinates equal to zero

$$\boldsymbol{x}_1 = (0, 0, \ldots, 0, x_1^{n-m+k+1}, x_1^{n-m+k+2}, \ldots, x_1^{n+k})$$

and the corresponding m-dimensional vector $\mathfrak{x}_1$ contains (so as it was before) all the remaining coordinates of $\boldsymbol{x}_1$ which may be different from zero.

The definition of the set Z_k^- remains the same with the only change that now we admit only those controls $w(t)$ for which the k-th derivative fulfils the specified conditions. The set Z_w^- now consists of such points $\boldsymbol{w}_0$ for which

$$(23^*) \qquad \begin{aligned} &\alpha < w_0^{(k)} \leqslant \beta \quad \text{if } w_0^{(k+1)} > 0, \\ &\alpha \leqslant w_0^{(k)} < \beta \quad \text{if } w_0^{(k+1)} < 0. \end{aligned}$$

To $Z_x^-(\boldsymbol{y}_0)$ now belong those points $\boldsymbol{x}_0 \in X$ which correspond, according to (11*), to all pairs of points $\boldsymbol{y}_0$ and $\boldsymbol{w}_0$, where $\boldsymbol{w}_0$ is an arbitrary point of Z_w^-.

Equation (25) takes now the form

$$(25^*) \qquad \mathfrak{x}_0 = \mathfrak{y}_0 - \Gamma^* \boldsymbol{w}(t_0^-),$$

where

$$\mathfrak{x}_0 = (x_0^{n-m+k+1}, x_0^{n-m+k+2}, \ldots, x_0^{n+k}),$$

$$\mathfrak{y}_0 = (y_0^{(n-m+k)}, y_0^{(n-m+k+1)}, \ldots, y_0^{(n-1)}, 0, \ldots, 0)$$

are the corresponding m-dimensional vectors.

Taking into account (19*), we may obtain from (25*)

$$(26^*) \qquad \boldsymbol{w}(t_0^-) = \Gamma^{*-1}(\mathfrak{y}_0 - \mathfrak{x}_0).$$

There are no difficulties to write further formulae with asterisks which now correspond to the former ones occurring in all lemmas, theorems an proofs of the main part of the paper.

References

[1] M. Athans and P. L. Falb, *Optimal control*, McGraw Hill, New York 1966.
[2] S. L. Chang, *Synthesis of optimum control*, McGraw Hill, New York 1961.
[3] R. Dmowski, *Sterowanie optymalne dla obiektów o przepustowości zawierającej zera*, Archiwum Automatyki i Telemechaniki 9 (1964), p. 357-383.
[4] J. Mikusiński i R. Sikorski, *Elementarna teoria dystrybucji*, PWN, Wrocław 1961.
[5] Л. С. Понтрягин, В. Г. Болтянский, Р. В. Гамкеридзе и Е. Ф. Мищенко, *Математическая теория оптимальных процессов*, Физматгиз, Москва 1961.

KATEDRA KOMPLEKSOWYCH SYSTEMÓW STEROWANIA
POLITECHNIKA ŚLĄSKA, GLIWICE

Received on 7. 11. 1967;
revised version on 27. 1. 1971

R. GESSING (Gliwice)

STEROWANIE CZASOWO-OPTYMALNE DLA OBIEKTÓW, KTÓRYCH FUNKCJE PRZEJŚCIA POSIADAJĄ ZERA

STRESZCZENIE

Prawa strona równań różniczkowych rozpatrywanych obiektów jest liniową formą sygnału sterującego $w(t)$ i jego pochodnych $w^{(1)}(t), \ldots, w^{(m)}(t)$. Rozpatruje się przy tym funkcje $w(t)$, mogące posiadać punkty nieciągłości. Dla takich sygnałów $w(t)$ wprowadzono w niniejszej pracy określenie rozwiązania $y(t)$ danych równań różniczkowych. Pokazano, że rozwiązania $\boldsymbol{y}(t) = (y(t), y^{(1)}(t), \ldots, y^{(n-1)}(t))$ mogą posiadać punkty nieciągłości. W związku z tym, do wyznaczenia rozwiązania $y(t)$ z warunkiem początkowym $\boldsymbol{y}(t_0^-) = \boldsymbol{y}_0$ potrzebna jest znajomość wielkości $\boldsymbol{w}(t_0^-) = (w(t_0^-), w^{(1)}(t_0^-), \ldots, w^{(m-1)}(t_0^-))$. Również do wyznaczenia sterowania optymalnego potrzebna jest znajomość nie tylko punktów — początkowego $\boldsymbol{y}_0$ i końcowego $\boldsymbol{y}_1$ — ale także wielkości $\boldsymbol{w}(t_0^-)$, czyli „ustawienia steru" bezpośrednio przed chwilą t_0 rozpoczęcia sterowania.

Biorąc to pod uwagę, w pracy postawiono dwa następujące problemy: pierwszy — gdy dane są punkty $\boldsymbol{y}_0$ i $\boldsymbol{y}_1$ oraz wielkość $\boldsymbol{w}(t_0^-)$, a należy znaleźć sterowanie czasowo-optymalne dla $t > t_0$ — oraz drugi — gdy dane są punkty $\boldsymbol{y}_0$ i $\boldsymbol{y}_1$, a wielkość $\boldsymbol{w}(t_0^-)$ należy również znaleźć, przy czym poszukiwane sterowanie optymalne powinno być dopuszczalne w przedziale czasu $t > t_0 - \varepsilon$ ($\varepsilon > 0$). W obu problemach stawia się przy tym — podobnie jak w [1] — wymaganie, aby punkt po osiągnięciu położenia $\boldsymbol{y}_1$ (czyli po chwili t_1) pozostawał nadal w tym położeniu.

Wprowadzając, podobnie jak w [1]-[3], nowe zmienne $\boldsymbol{x}$, które przy rozpatrywanych sygnałach sterujących są ciągłymi funkcjami czasu, sformułowano i udo-

wodniono twierdzenia dotyczące obydwu problemów. Z twierdzeń tych wynika, że sterowanie optymalne dla $t > t_1$ jest rozwiązaniem równania różniczkowego, powstającego z przyrównania do zera prawej strony równania obiektu. W przedziale czasu $t_0 < t < t_1$ sterowanie optymalne — w przypadku problemu pierwszego — można otrzymać, rozwiązując zagadnienie sterowania z zadanego punktu początkowego do zbioru punktów końcowych. W przypadku problemu drugiego, sterowanie optymalne dla $t_0 < t < t_1$ wynika z rozwiązania zagadnienia sterowania ze zbioru punktów początkowych do zbioru punktów końcowych. Do rozwiązania obu tych zagadnień mogą być pomocne odpowiednio zastosowane warunki transwersalności.

Okazuje się, że w przypadku problemu drugiego, gdy mamy możność wyboru „ustawienia steru" bezpośrednio przed chwilę t_0 rozpoczęcia sterowania (czyli gdy możemy „przygotować obiekt" do optymalnego sterowania), możemy zazwyczaj znacznie skrócić całkowity czas sterowania $T = t_1 - t_0$. Widać to wyraźnie w rozpatrzonym szczegółowo przykładzie sterowania obiektem o równaniu $\ddot{y} = \ddot{w} + w$.

W pracy rozpatruje się zasadniczo przypadek, gdy ograniczenia nałożone są na sam sygnał sterujący $w(t)$. W dodatku 2 pokazano, jednak, że oba twierdzenia i inne rozważania zamieszczone w pracy są również prawdziwe, gdy ograniczenia nałożone są na k-tą pochodną sygnału sterującego (gdzie $k < m$).