DIFFERENTIAL GAMES WITH INFORMATION LAG

PMM Vol. 34, №5, 1970, pp. 812-819
B. N. SOKOLOV and F. L. CHERNOUS'KO
(Moscow)
(Received May 12, 1969)

Two-player differential games are considered under the assumption that one of the players receives information on the phase coordinates of his opponent with a certain time lag. In practical cases information lag is due to the time required for the reception and treatment of measured data. The presence of information lag distinguishes the problems considered here from ordinary problems of the theory of differential games (e. g. see [1–3]). We show (see also [4]) that under certain general conditions every differential game with information lag is equivalent to a certain differential game with zero lag. This makes possible the use of familiar methods of differential game theory for solving game problems with information lag. Some specific pursuit problems in which one of the players receives information with lag are considered. The problems are solved and analyzed and the conditions under which capture (target locking) can occur are determined.

1. Formulation of the problem. Let us consider a differential game involving two controlled systems called the "players" (or "sides") P and E. We denote the phase coordinate vectors of these systems by $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_s)$ and the controlling function vectors by $u = (u_1, \ldots, u_m)$ and $v = (v_1, \ldots, v_k)$, respectively. The dimensions n, s, m, k of these vectors are arbitrary.

The equations of system motion are of the form

$$dx / dt = f(x, u, t), \qquad dy / dt = g(y, v, t) \tag{1.1}$$

Here t is the time; $f = (f_1, \ldots, f_n)$ and $g = (g_1, \ldots, g_s)$ are given vector functions. We assume that the initial instant t_0 of the game is given and that the instant T at which the game ends is defined by a condition of the form

$$h(x(T), y(T), T) = 0 \tag{1.2}$$

Here h is a given scalar function and T is the smallest root of Eq. (1.2) for which $T > t_0$. Without limiting generality we can assume that $h > 0$ for $t < T$; otherwise we simply multiply the function h by -1. Specifically, if the instant of termination of the game (T_0) is given, then $h = T_0 - t$.

The controlling functions are subject to the restrictions

$$u(t) \in U, \quad v(t) \in V \qquad (t \geqslant t_0) \tag{1.3}$$

Here U, V are closed sets in m- and r-dimensional spaces, respectively. The function (payoff) is given in the form

$$F = F(x(T), y(T), T) \tag{1.4}$$

where $F(x, y, t)$ is a known scalar function. We assume (without limiting generality) that player P strives to minimize and player E to maximize the functional F. In particular, if the payoff is the game duration (in pursuit problems), then $F = t$ or $F = -t$, depending on whether P is the pursuing or the evading player. The conditions and restrictions which we have introduced are standard in the theory of differential games.

Let us also make the following assumption concerning the information made available to the players. At each instant t player P is informed of the instantaneous value of the phase vector of player E at the instant $t - \tau$, i.e. the vector $y(t - \tau)$. Here the constant $\tau > 0$ represents the time lag with which player P receives his information and is equal to the time needed by this player to receive and treat the measured data. It is natural to assume that the lag is sufficiently small (smaller than the duration of the process), i.e. that $\tau < T - t_0$. In addition, player P knows the functions f, g, h, F and the sets U, V in relations (1.1)–(1.4) (i.e. the rules and purpose of the game) prior to the start of the game. We shall solve the problem from the standpoint of player P, i.e. we shall seek the optimal control u for player P as a function of the time t in the interval $[t_0, T]$ and of the instantaneous measured data, namely the vectors $x(t)$ and $y(t - \tau)$. We need make no assumptions as to the information made available to E; we shall merely suppose that E behaves in the worst possible way from the standpoint of P. This will yield the minimum guaranteed value of the functional F.

By the initial instant t_0 we mean the instant when the first information is received; we specify the initial conditions in the form

$$x(t_0) = x^\circ, \qquad y(t_0 - \tau) = y^\circ \tag{1.5}$$

The controlled motion of player P begins at the instant t_0; for $t < t_0$, i.e. prior to the influx of information, the control for player P must be chosen on the basis of some prior considerations (the motion for $t < t_0$ can be considered uncontrolled).

2. Equivalence of games with and without information lag. We introduce the symbolism $\eta(t) = y(t - \tau)$ and rewrite Eqs. (1.1) and initial conditions (1.5) as

$$dx / dt = f(x, u, t), \quad d\eta / dt = g(\eta, v, t - \tau) \tag{2.1}$$

$$x(t_0) = x^0, \qquad \eta(t_0) = y^0$$

Let us denote by $D(y, t)$ the attainability domain for player E in the time τ under the initial condition $y(t) = y$. In other words, $D(y, t)$ is the set of vectors $y(t + \tau)$ obtainable under the condition $y(t) = y$ and for all possible permissible controls $v \in V$ in the interval $[t, t + \tau]$ if the functions y, v are related by the second equation of (1.1). Let us write out the conditions of game termination and the expression for the guaranteed value of the minimized functional F using the symbols introduced above. We shall limit our attention to three important particular cases.

1°. Let the game duration be fixed and equal to T_0; the function h in (1.2) is then equal to $h = T_0 - t$. The guaranteed value of the functional F for given $x(T_0)$, $\eta(T_0)$ can then be determined (for the most adverse circumstances) by maximizing the function F appearing in (1.4) with respect to the permissible values of $y(T_0)$. We obtain

$$F_0 = F_0(x(T_0), \eta(T_0), T_0) = \tag{2.2}$$

$$= \max_y F(x(T_0), y, T_0), \quad y \in D(\eta(T_0), T_0 - \tau)$$

2°. Now let the function h be arbitrary (with the proviso that $h > 0$ for $t < T$) and let the function F appearing in (1.4) be equal to $F = t$. This occurs in a pursuit game in which P is the pursuer. The most adverse game duration from the standpoint of P can then be determined from the condition

$$h_0(x(t), \eta(t), t) = \max_y h(x(t), y, t) = 0$$

$$y \in D(\eta(t), t - \tau) \tag{2.3}$$

3°. Let the function h be arbitrary as above, and let $F = -t$, i.e. let P be the evading player. Then instead of (2.3) we have

$$h_0(x(t),\ \eta(t),\ t) = \min_y h(x(t),\ y,\ t) = 0$$

$$y \in D(\eta(t),\ t-\tau) \qquad (2.4)$$

In all the above cases the condition of game termination and the guaranteed value of the functional are of the form

$$h_0(x(T),\ \eta(T),\ T) = 0$$

$$F_0 = F_0(x(T),\ \eta(T),\ T) \qquad (2.5)$$

Here the function h_0 is defined by the equations $h_0 = T_0 - t$, (2.3), and (2.4), and the function F_0 by Eqs. (2.2), $F_0 = t$ and $F_0 = -t$, respectively, in the three cases considered above. In [4] it is shown that relations (2.5) are also valid in more general cases; the expressions for the functions h_0, F_0 appropriate to these cases are given.

Differential equations and initial conditions (2.1), restrictions (1.3), the condition of process termination, and functional (2.5) clearly define a differential game without information lag. In fact, at each instant t the phase coordinates $x(t)$, $\eta(t)$ of both sides are known to player P. We have thus proved that in order to obtain the guaranteed result in the game with information lag defined by (1.1)–(1.5) we need merely solve differential game (2 1), (1.3), (2.5) without information lag.

Since a differential game with information lag is equivalent to a differential game without lag, games with lag are amenable to all the known approaches and results of the theory of differential games without lag (e.g. see [1–3]). For example, we can construct the Bellman-Isaacs equation for game (2.1), (1.3), (2.5), construct its characteristics, and investigate its solution as is done in [1]. First of all we must calculate the functions h_0, F_0 appearing in (2.5) using formulas (2.2)–(2.4). We note that if the equations of motion of system E are autonomous, i.e. if the function g is independent of t, then equations of motion (2.1) coincide with (1.1).

Pursuit games can be solved by the procedure of [3], which entails determination of the instant of initial absorption. Section 3 of the present paper contains solutions of specific examples on the basis of a self-evident modification of this procedure for games with information lag.

Note 2.1. Having solved differential game (2.1), (1.3), (2.5) without information lag, we can determine the synthesizing controls $u = u(x,\ \eta,\ t)$ and $v = v(x,\ \eta,\ t)$ for both players. The function $u(x,\ \eta,\ t)$ synthesizes the guaranteeing control for player P for the initial problem by way of the instantaneous measurements of $x(t)$ and $\eta(t) = y(t-\tau)$. The function $v(x,\ \eta,\ t)$ represents the control for player E which is least favorable to player P; however, player E cannot use this function, as it defines the control v at the instant $t-\tau$ in terms of a future value of the phase vector $x(t)$. To determine the optimal control for player E we must specify the character of information supply to player E and solve the problem again for player E.

Note 2.2. The above fact of equivalence of games with information lag to games without lag is valid for a broader class of problems than (1.1)–(1.5). For example, it holds for games with integral restrictions on the controls as well as for games with restricted phase coordinates. All that matters is that the phase coordinates and controlling functions of one player not appear in the equations of motion and restrictions of the opposite player.

3. Examples of pursuit games with information lag. Let the motion of two objects (pursuer and evader) be described by the systems of differential equations with restrictions

$$\mathbf{r}_1^{\cdot} = \mathbf{w}_1, \quad \mathbf{w}_1^{\cdot} = \mathbf{a}_1, \quad |\mathbf{a}_1| \leqslant a_1 \qquad (3.1)$$
$$\mathbf{r}_2^{\cdot} = \mathbf{w}_2, \quad \mathbf{w}_2^{\cdot} = \mathbf{a}_2, \quad |\mathbf{a}_2| \leqslant a_2, \quad a_1 > a_2$$

Here $\mathbf{r}_1$, $\mathbf{r}_2$ are the radius vectors of the objects, $\mathbf{w}_1$, $\mathbf{w}_2$ are their velocities, and $\mathbf{a}_1$, $\mathbf{a}_2$ are their accelerations which are the controlling functions restricted in absolute value by the constants a_1, a_2. The subscript 1 refers to the pursuer, the subscript 2 to the evader. All of the vectors in (3.1) are of the arbitrary and equal dimension N. The game is assumed to terminate when the distance between the objects becomes equal to the given number l (i.e. when capture occurs); game termination condition (1.2) then becomes

$$h(\mathbf{r}_1(T_1),\ \mathbf{r}_2(T),\ T) = |\mathbf{r}_1(T) - \mathbf{r}_2(T)| - l = 0 \qquad (3.2)$$

The payoff is defined as the time elapsed from the initial instant of the game until the instant of capture; the pursuer strives to reduce this time and the evader to increase it ($F = T$ in formula (1.4)). We assume that at each instant $t \geqslant 0$ the pursuer knows his phase coordinates $\mathbf{r}_1$, $\mathbf{w}_1$ at this instant and the evader's phase coordinates $\mathbf{r}_2$, $\mathbf{w}_2$ at the instant $t - \tau$.

We solved this problem according to the above scheme. We began by computing the function h_0 and then investigated the resulting game without information lag by the procedure of [1].

We constructed and integrated the equations of the characteristics, constructed the optimal trajectories and controls, and determined the barrier surface in the phase coordinate space. The entire analysis was carried out as in the case of the "isotropic missiles" problem in [1].

Let us describe the solution of the problem by the procedure of [3], which is much more brief by virtue of the simplicity of the attainability domains. We assume that at the initial instant $t = 0$ the pursuer knows his own phase coordinates $\mathbf{r}_1^\circ$, $\mathbf{w}_1^\circ$ and the phase coordinates $\mathbf{r}_2^\circ$, $\mathbf{w}_2^\circ$ of the evader at the instant $t = -\tau$. The attainability domains of the pursuer and the evader at the instant t in the radius vector space are spheres with their centers at the points

$$\mathbf{r}_1^\circ + \mathbf{w}_1^\circ t, \qquad \mathbf{r}_2^\circ + \mathbf{w}_2^\circ (t + \tau)$$

and with the radii

$$a_1 t^2 / 2, \qquad a_2 (t + \tau)^2 / 2$$

respectively. Let us construct the equation which defines the instant of initial absorption of the attainability domain of the evader by the l-neighborhood of the pursuer's attainability domain. We denote the vector connecting the centers of the attainability domains by $R(t)$ and the difference between their radii (with allowance for the l-neighborhood by $Q(t)$ We clearly have

$$R(t) = \mathbf{r}_2^\circ + \mathbf{w}_2^\circ (t + \tau) - \mathbf{r}_1^\circ - \mathbf{w}_1^\circ t \qquad (3.3)$$
$$Q(t) = l + a_1 t^2 / 2 - a_2 (t + \tau)^2 / 2$$

The instant of absorption T is defined as the first positive root of the equation

$$|R(T)| = Q(T) \qquad (3.4)$$

The pursuer's equation which ensures capture at the instant T must effect aiming at that point of its attainability domain which lies at the distance l from the point of tangency of the evader's attainability domain and the l-neighborhood of the pursuer's attainability domain. Hence, the pursuer's control must be taken in the form

$$\mathbf{a}_1 = a_1 R(T) / |R(T)| \tag{3.5}$$

where T is the first positive root of Eq. (3.4). If the pursuer keeps to control (3.5), then the evader can avoid capture until the instant T by aiming towards the same point of tangency, i.e. by employing the control

$$\mathbf{a}_2 = a_2 R(T) / |R(T)| \tag{3.6}$$

Let us assume that

$$Q(0) = l - a_2\tau^2 / 2 > 0 \tag{3.7}$$

This inequality is the necessary condition for the possibility of game termination (if this condition is violated, the evader can always escape from the capture zone of radius l within the lag time τ). In addition, we assume that $|R(0)| > Q(0)$, i.e. capture has not occurred in the initial position. Formulas (3.3) imply that for sufficiently large t and for $a_1 > a_2$ we have $|R(t)| < Q(t)$. This implies that Eq. (3.4) has an odd number of positive roots. Since this equation is reducible with the aid of Eqs. (3.3) to a fourth-degree algebraic equation in T, it follows that it has either one or three positive roots.

Let $Q(t) > 0$ for all positive t. Since the radius of the l-neighborhood of the pursuer's attainability domain is always larger than the radius of the evader's attainability domain, and these domains can touch at one point only. In this case [3, 5] the first positive root T of Eq. (3.4) is the guaranteed time of game termination, and capture can be effected in a finite time under all boundary conditions.

Now let the function $Q(t)$ vanish and let t_1 be its first positive root. If the first positive root T of Eq. (3.4) satisfies the condition $T < t_1$, then the attainability domains can touch at one point only, as above, and T is again the guaranteed time of termination of the game. But for $T > t_1$ there exist instants $t' < T$ such that the radius of the evader's attainability domain can be larger than or equal to the radius of the l-neighborhood of the pursuer's attainability domain. In this case it is generally impossible to guarantee capture in the time T under initial data such that $T > t_1$.

Controlling his motion in accordance with rules (3.4) and (3.5), by the instant T the pursuer ensures a position in which the attainability domain of the pursuer lies entirely in the l-nighborhood of the pursuer's location within the time τ.

This situation satisfies the condition of guaranteed capture for games with information lag (see (2.3)).

Solving the problem by the procedure set out in [1], we arrive at the same Eqs. (3.4), (3.5) for the time T and for the pursuer's optimal equation.

The quantity T and controls (3.5), (3.6) remain constant along each optimal trajectory, so that the optimal trajectories are parabolas. Formulas (3.3)–(3.5) define the synthesis of the pursuer's optimal control if by $\mathbf{r}_1^\circ$, $\mathbf{r}_2^\circ$, $\mathbf{w}_1^\circ$, $\mathbf{w}_2^\circ$ are taken to signify the instantaneous measured data. The time T as a function of the initial phase coordinates has a discontinuity at those points where there is a jump discontinuity in the first root of Eq. (3.4). This corresponds to the merging of the first two positive roots of Eq. (3.4), i.e. to the existence of a multiple root of Eq. (3.4). The multiple-root condition is of the form

$$d|R(T)| / dT = dQ(T) / dT \tag{3.8}$$

Eliminating T from Eqs. (3.4), (3.8), we obtain the equation of the surface (called the "barrier" [1]) in the phase coordinate space at which T experiences a discontinuity. The jump which T experiences at this surface is due to the pursuer's need to execute a turn maneuver [1].

It is possible to prove that the above condition $Q(t) > 0$ is equivalent to the condition of nonpenetration of the barrier [1]. However, if the function $Q(t)$ vanishes for $t > 0$, the two branches constituting the barrier surface intersect and isolate the domain in the phase space where $T < t_1$ and where the game terminates in a finite time. Outside this domain the evader can escape capture (see [1] for a discussion of these matters in the case of the "isotropic missiles" game).

Thus, capture in a finite time T for all initial conditions is guaranteed if $Q(t) > 0$ for all $t \geqslant 0$. Recalling relations (3.3), (3.7), we can write the condition of positive definiteness of the function $Q(t)$ in the form

$$(a_1 - a_2)(l - a_2\tau^2 / 2) > a_2\tau^2 / 2$$

From this we obtain the condition of guaranteed capture in its final form,

$$(a_1 - a_2)a_1^{-1} > a_2\tau^2 l^{-1} / 2 \qquad (3.9)$$

We note that the inequalities $a_1 > a_2$ and (3.7) follow from (3.9). Condition (3.9) shows that the larger the delay time τ, the greater the advantage in acceleration required by the pursuer in order to achieve capture.

Let us consider some further examples of games with information lag which can be solved by the above method.

In all of these examples the payoff is the pursuit time, the game termination condition retains its previous form (3.2), the subscript 1 refers to the pursuer and the subscript 2 to the evader. We solve the problems from the point of view of the pursuer, who receives information on the phase coordinates of the evader after a time lag equal to τ. In all of our examples we denote the players' vector functions by $\mathbf{a}_1$, $\mathbf{a}_2$ as before (but using them to signify either velocity or acceleration, as the case may be). The instant of absorption T is, as before, the first positive root of Eq. (3.4), and the players' optimal controls are defined by formulas (3.5), (3.6). However, the functions $R(T)$, $Q(T)$ in relations (3.4)–(3.6) are of different form than in (3.3). We shall write out these functions for our specific examples.

A. Let both players be subject to velocity control and let their equations of motion and restrictions be of the form

$$\mathbf{r}_1^{\cdot} = \mathbf{a}_1, \quad \mathbf{r}_2^{\cdot} = \mathbf{a}_2, \quad |\mathbf{a}_1| \leqslant a_1, \quad |\mathbf{a}_2| \leqslant a_2$$

Here, as everywhere, a_1 and a_2 are constants. In this case the functions $R(T)$, $Q(T)$ are given by $R(T) = \mathbf{r}_2^\circ - \mathbf{r}_1^\circ$, $Q(T) = l + a_1 T - a_2(T + \tau)$

Capture is guaranteed for all initial situations, provided that the two conditions $a_1 > a_2$ and $l > a_2\tau$ are both fulfilled.

B. Let the pursuer be subject to acceleration control and the evader to velocity control. The equations of motion are

$$\mathbf{r}_1^{\cdot} = \mathbf{w}_1, \quad \mathbf{w}_1^{\cdot} = \mathbf{a}_1, \quad \mathbf{r}_2^{\cdot} = \mathbf{a}_2, \quad |\mathbf{a}_1| \leqslant a_1, \quad |\mathbf{a}_2| \leqslant a_2$$

Here the functions R, Q are given by

$$R(T) = \mathbf{r}_2^\circ - \mathbf{r}_1^\circ - \mathbf{w}_1^\circ T, \quad Q(T) = l + a_1 T^2 / 2 - a_2(T + \tau)$$

Capture is guaranteed for all initial conditions if

$$l > a_2^2 a_1^{-1} / 2 + a_2\tau$$

We note that for $\tau = 0$ this solution becomes the solution of the "isotropic missiles" problem [1].

C. If the pursuer is subject to velocity control and the evader to acceleration control, the equations of motion and the functions R, Q are of the form

$$\mathbf{r}_1^{\cdot} = \mathbf{a}_1, \quad \mathbf{r}_2^{\cdot} = \mathbf{w}_2, \quad \mathbf{w}_2^{\cdot} = \mathbf{a}_2, \quad |\mathbf{a}_1| \leqslant a_1, \quad |\mathbf{a}_2| \leqslant a_2$$

$$R(T) = \mathbf{r}_2^{\circ} + \mathbf{w}_2^{\circ}(T + \tau) - \mathbf{r}_1^{\circ}, \quad Q(T) = l + a_1 T - a_2 (T + \tau)^2 / 2$$

In this case situations such that the evader can escape capture exist for all problem parameters.

BIBLIOGRAPHY

1. Isaacs, R., Differential Games. Moscow, "Mir", 1967.
2. Pontriagin, L. S., On the theory of differential games. Uspekhi Mat. Nauk Vol. 21, №4, 1966.
3. Krasovskii, N. N., Theory of Motion Control. Moscow, "Nauka", 1968.
4. Chernous'ko, F. L., On differential games with information lag. Dokl. Akad. Nauk SSSR Vol. 188, №4, 1969.
5. Krasovskii, N. N. and Subbotin, A. I., The problem of convergence of controlled objects. PMM Vol. 32, №4, 1968.

Translated by A. Y.

ON THE BELLMAN FUNCTION FOR THE TIME-OPTIMAL PROCESS PROBLEM

PMM Vol. 34, №5, 1970, pp. 820-826

N. N. PETROV
(Leningrad)

(Received January 26, 1970)

The necessary and sufficient condition under which the Bellman function for the time-optimal process problem belongs to the class of functions satisfying the Lipschitz condition is developed.

1. Let a controlled process be described by the system of equations

$$dx / dt = f(x, u) \tag{1.1}$$

where x and f are n-dimensional vectors and u is an r-dimensional control vector.

Let us suppose that the set U of permissible values of the controlling functions $u = u(t)$ is a nonempty compact subset of the r-dimensional Euclidean space E_r. As our permissible controlling functions we consider the measurable functions $u = u(t)$ with values in U. In addition, we assume that the vector function $f(x, u)$ is defined and continuous in both its variables on the set $E_n \times U$ and that it satisfies Lipschitz' local condition in x with a constant independent of u. The purpose of control is to bring the system to the position $x = 0$.

Let $G(< T)$ be the set of all points $x_0 \in E_n$ from which it is possible to reach the origin in a time smaller than T. In other words, $x_0 \in G(< T)$ means that there exists a permissible control $u = u(t)$ defined for $t \in [0, \tau]$, $\tau < T$ such that the