PMM U.S.S.R.,Vol.54,No.3,pp.321-325,1990
Printed in Great Britain

0021-8928/90 $10.00+0.00


# COOPERATIVE SOLUTIONS IN A MANY-PERSON POSITIONAL DIFFERENTIAL GAME WITH CONTINUOUS PAYMENT FUNCTIONS*

A.F. KLEIMENOV

Continuing previous research /1-2/, we consider a non-zero-sum positional differential game which allows deviations of individual players and also groups of players from a coordinated solution**. (**Kleimenov A.F., One approach to the analysis of cooperative positional differential games, Sverdlovsk, 1989. Unpublished manuscript, VINITI 11.07.89, 4609-V89.) The proposed formalization of a non-zero-sum differential game is based on formalization and results of the theory of zero-sum positional differential games /3, 4/. We assume that in the course of the game each player may make continuous payments to other players. The payments are measured in units of the criterion of the player making the payments, and at any moment in the game they may not exceed a given fraction of the increment in that player's guaranteed outcome. Each player evaluates individually the payments received from other players. This approach covers, in particular, the case of transferable /5/ rewards. The proposed cooperative solution of the game is such that none of the admissible coalitions will be better off by deviating from this solution in the course of the game. The cooperative solutions provide for a certain penalty strategy /6/. An approach to the analysis of cooperative differential games with transferable rewards using the notation of dynamic stability of optimality principles was studied in /7/. The notion of a strong equilibrium solution stable in relation to the deviations of coalitions was considered in /8/.

1. Assume that the dynamics of a controlled system is described by the equation

$$x^{\cdot} = f(t, x, u_1, \ldots, u_m), \quad x \in R^n, \quad u_i \in P_i \in \operatorname{comp} R^{p_i}, \quad x[t_0] = x_0 \tag{1.1}$$

where the function $f$ satisfies the standard restrictions accepted in the theory of zero-sum positional differential games /3, 4/. Player $i$ $(i = 1, \ldots, m)$ equipped with a set of controls $u_i$ attempts to maximize his reward $\sigma_i(x[\vartheta])$. Here $\sigma_i$ are continuous functions and $\vartheta$ is the given termination time of the game.

We make the following assumptions 1°-3°.

1°. The system phase vector $x[t]$ is known to each player at the current time $t$.

A coalition is any non-empty subset $K$ of the set of players $I = \{1, \ldots, m\}$. Let $\mathbf{K}$ be the set of possible coalitions that may form in the course of the game. We have $1 \leqslant |K| \leqslant m - 1$.

2°. One of the possible coalitions (it is not known which in advance) may deviate from the cooperative solution at any moment $\tau \in [t_0, \vartheta)$.

The notation of a "cooperative solution" (CS) will be formalized below (see Definition 3). We assume that the existence of a deviation from the CS and the identity of the deviating coalition become known to all players as soon as a deviation occurs. This information is conveniently formalized as follows.

3°. If the coalition $K \in \mathbf{K}$ deviates from the CS at time $\tau \in [t_0, \vartheta)$, then the current values of the function $\alpha_K^{\tau}[t] = \{\varnothing$ for $t_0 \leqslant t < \tau$; $K$ for $\tau \leqslant t \leqslant \vartheta\}$ are announced to all players; if no deviation has occurred during the entire game, then the current values of the function $\alpha^{\circ}[t] = \varnothing$ are announced to the players for $t_0 \leqslant t \leqslant \vartheta$.

On the basis of assumptions 1°-3°, we will define the strategies of the players and the motions generated by these strategies using the concepts of strategies and motions from /3, 4/. For simplicity, we consider only pure positional strategies. Mixed strategies and counter-strategies can be considered similarly.

A pure positional strategy (simply a strategy) of player $i$ is identified with the pair $U_i - \{u_i(t, x, \alpha, \varepsilon), \beta_i(\varepsilon)\}$, where $u_i(\cdot)$ is an arbitrary function of the position $(t, x)$, the parameter $\alpha \in \mathbf{K} \cup \varnothing$, and the accuracy parameter /4/ $\varepsilon > 0$ with values in the set $P_i$, and the continuous monotone function $\beta_i$: $(0, \infty) \mapsto (0, \infty)$ satisfies the condition $\beta_i(\varepsilon) \to 0$ as

*Prikl.Matem.Mekhan.,54,3,389-394,1990

$\varepsilon \to 0$. Two types of motion are considered as the motions generated by the collection of strategies $U = (U_1, \ldots, U_m)$: approximation motions (Euler polygonal lines) and ideal (limiting) motions. Note that if the functions $\alpha_K{}^\tau[\cdot]$ is realized in the course of the game, then starting at a time $\tau$ the members of coalition $K$ switch to some collection of strategies $U_K{}^\tau = (U_j{}^\tau, j \in K)$, which we call the collection of deviation strategies of the coalition $K$.

The control law of player $i$ corresponding to strategy $U_i$ is determined by three components: the function $u_i(\cdot)$ occurring in the strategy $U_i$, the value of the accuracy parameter $\varepsilon_i$, and the partition $\Delta_i = \{t_j{}^i\}$ of the interval $[t_0, \vartheta]$ with a step not exceeding $\beta_i(\varepsilon_i)$ (see /4/, and also /1, 2/). We assume that $\varepsilon_1 = \ldots = \varepsilon_m = \varepsilon$ if there are no deviations from the cooperative solution. After the deviation of the coalition $K$ from the cooperative solution, the values of $\varepsilon_j, j \in K$, may be arbitrary. The Euler polygonal lines $x[t, t_*, x_*, U_i, \varepsilon, \Delta_i, \alpha^\circ[\cdot]]$ and $x[t, t_*, x_*, U_i, \varepsilon, \Delta_i, \alpha_K{}^\tau[\cdot], U_K{}^\tau]$, generated from the initial position $(t_*, x_*)$ by the collection of control laws $(U_i, \varepsilon, \Delta_i), i \in I$, for the realizations $\alpha^\circ[\cdot]$ and $\alpha_K{}^\tau[\cdot]$ respectively are defined as in /3, 4/ (see also /1, 2/). The limiting motions $x[t, t_*, x_*, U, \alpha^\circ[\cdot]]$ and $x[t, t_*, x_*, U, \alpha_K{}^\tau[\cdot], U_K{}^\tau]$ generated from the initial position $(t_*, x_*)$ by the collection of strategies $U$ for the realizations $\alpha^\circ[\cdot]$ and $\alpha_K{}^\tau[\cdot]$ are defined as the uniform limits for $\varepsilon \to 0$ of the convergent sequences of Euler polygonal lines. The sets of limiting motions, denoted by $X(t_*, x_*, U, \alpha^\circ[\cdot])$ and $X(t_*, x_*, U, \alpha_K{}^\tau[\cdot], U_K{}^\tau)$ respectively, are non-empty and compact in the metric of the space $C[t_0, \vartheta]$.

In what follows, we only consider those collections of strategies $U$ for which $X(t_0, x_0, U, \alpha^\circ[\cdot])$ is a one-element set. The solution of Problem 1 stated below is achieved on precisely such collections, and the restriction of the set of admissible collections therefore does not affect the result.

Let us now give a formal description of payments that can be made between players in the course of the game. Consider zero-sum positional differential games $\Gamma_i, i \in I$, in which player $i$ attempts to maximize his reward $\sigma_i(x[\vartheta])$ and the other players $I \setminus \{i\}$ try to counteract him. The player strategies in the games $\Gamma_i$ are defined, according to /3, 4/, as $U_i \div u_i(t, x, \varepsilon)$. It follows from the general theory /3, 4/ that the games $\Gamma_i$ have continuous value functions $\gamma_i(\cdot, \cdot)$. The value $\gamma_i(t, x)$ is the guaranteed outcome of player $i$ in game $\Gamma_i$ when the game starts from position $(t, x)$.

Assume that the trajectory $x^*[\cdot] = \{x^*[t], t_0 \leqslant t \leqslant \vartheta\}$ is realized in our non-zero-sum game. If the increment in $[t_0, t]$ of the guaranteed (in game $\Gamma_i$) outcome of player $i$, $\Delta\gamma_i = \gamma_i(t, x^*[t]) - \gamma_i(t_0, x_0)$ is positive, then this increment may be interpreted as the "accumulated" reward of player $i$ at time $t$, measured in units of the criterion $\sigma_i$. We assume that a proportion $\lambda_i(\gamma_i(t, x^*[t]) - \gamma_i(t_0, x_0))$, of the accumulated reward, where $0 < \lambda_i \leqslant 1$, may be spent by player $i$ on payments to other players, and that for the purpose of making these payments player $i$ has a certain "reserve" of units of his criterion. Denote by $\mu_{ij}(t)$ the payment made by player $i$ to player $j$ by the time $t$. We assume that $\mu_{ij}(\cdot), i, j \in I, i \neq j$, are non-negative continuous monotone non-decreasing functions (monotonicity implies that payments, once made, cannot be returned). From the above we have

$$\sum_{j \in I \setminus \{i\}} \mu_{ij}(t) \leqslant \lambda_i(\gamma_i(t, x[t]) - \gamma_i(t_0, x_0)), \quad i \in I \tag{1.2}$$

Assume that each player evaluates individually the payments received from other players. Denote by $g_{ij}, i, j \in I, i \neq j$, the value in units of the criterion $\sigma_i$ attached by player $i$ to one unit of the criterion of player $j$. If all $g_{ij} = 1$, we have transferability of rewards /5/.

Assume that the players choose a collection of strategies $U$ that generates a unique trajectory $x[\cdot]$ and a collection of payments $\mu(U) = (\mu_{ij}(\cdot), i, j \in I, i \neq j)$ that satisfies condition (1.2). Then the reward of each player is given by

$$\sigma_i{}^* = \sigma_i(x[\vartheta]) + \sum_{j \in I \setminus \{i\}} (g_{ij}\mu_{ji}(\vartheta) - \mu_{ij}(\vartheta)) \tag{1.3}$$

The collection of strategies $U$ in general corresponds to a whole set of collections of payments $\mu(U)$ that satisfy condition (1.2). The choice from this set will be restricted by the following assumption.

4°. As the collection of payments $\mu(U)$ corresponding to a fixed collection of strategies, the players choose only those collections that are Pareto-unimprovable relative to the criteria (1.3); the reward of each player allowing for payments should not be less than the reward without payments earned on a given trajectory.

The set of payment collections corresponding to the collection of strategies $U$ that satisfy the conditions of assumption 4° will be denoted by $M(U)$.

For fixed $U$ and $\mu(U) \in M(U)$, define the set

$$T(U, \mu(U)) \overset{\Delta}{=} \{\tau \in [t_0, \vartheta): \exists K \in \mathbf{K} \exists \mathbf{U}_K^\tau \neq \varnothing \forall U_K^\tau \in \mathbf{U}_K^\tau: \sigma_i(x[\vartheta, t_0, x_0, U, \alpha^\circ[\cdot]]) + \sum_{j \in I \setminus \{i\}} (g_{ij}\mu_{ji}(\vartheta) - \mu_{ij}(\vartheta)) \leqslant \min_{x[\cdot]} [\sigma_i(x[\vartheta, t_0, x_0, U, \alpha_{K^\tau}[\cdot], U_K^\tau]) \sum_{j \in I \setminus K} (g_{ij}\mu_{ji}(\tau) - \mu_{ij}(\tau)) + \sum_{j \in K \setminus \{i\}} (g_{ij}\mu_{ji}^*(\vartheta) - \mu_{ij}^*(\vartheta))],\ i \in K\} \tag{1.4}$$

where $\mathbf{U}_K^\tau$ is the set of deviation strategies $U_K^\tau$ of coalition $K$; $\mu_K^\tau(U, U_K^\tau) = (\mu_{ij}^*(\cdot), i, j \in K, i \neq j)$ is a collection of payments that satisfy the condition $\mu_{ij}^*(t) = \mu_{ij}(t)$ for $t_0 \leqslant t \leqslant \tau$ and condition (1.2) for $\tau < t \leqslant \vartheta$, where the trajectory $x[\cdot]$ is identified with that motion from the bundle $X(t_0, x_0, U, \alpha_K^\tau[\cdot], U_K^\tau)$ on which the minimum in (1.4) is attained.

By definition, the set $T(U, \mu(U))$ consists of the instants of time $\tau \in [t_0, \vartheta)$ with the following property: there exists a coalition $K$ which, starting at time $\tau$, may switch from the strategies in $U$ to the collection of deviation strategies $U_K^\tau$, simultaneously replacing the payment functions from $\mu(U)$ with the collection of payment functions $\mu_K^\tau(U, U_K^\tau)$, so that all the coalition members benefit as a result of the change; it is assumed that as a result of the deviation of the coalition $K$ from the strategies $U$, players who are not members in $K$ stop exchanging payments with members of the coalition $K$ as of the time of their deviation. (The case when a coalition deviates from the collection of payments $\mu(U) \in M(U)$ while preserving the collection of strategies $U$ is not considered here).

For fixed $U$, $\mu(U) \in M(U)$ put

$$s = \inf T(U, \mu(U)) \tag{1.5}$$

*Definition 1.* The functions

$$\rho_i(U, \mu(U)) = \{\sigma_i(x[\vartheta, t_0, x_0, U, \alpha^\circ[\cdot]]) + \sum_{j \in I \setminus \{i\}} (g_{ij}\mu_{ji}(\vartheta) - \mu_{ij}(\vartheta)),\ \text{if}\ T(U, \mu(U)) = \varnothing;\ \inf_{U_K^s \in \mathbf{U}_K^s} \min_{x[\cdot]} [\sigma_i(x[\vartheta, t_0, x_0, U, \alpha_K^s[\cdot], U_K^s]) + \varphi_i(\mu(U), \mu_K^s(U, U_K^s))]\ \text{otherwise}\} \tag{1.6}$$

$$\varphi_i(\mu(U), \mu_K^s(U, U_K^s)) = \{\sum_{j \in I \setminus K} (g_{ij}\mu_{ji}(s) - \mu_{ij}(s)) + \sum_{j \in K \setminus \{i\}} (g_{ij}\mu_{ji}^*(\vartheta) - \mu_{ij}^*(\vartheta))\ \text{for}\ i \in K;\ \sum_{j \in I \setminus \{i\}} (g_{ij}\mu_{ji}(s) - \mu_{ij}(s))\ \text{for}\ i \notin K\} \tag{1.7}$$

defined on the couples $(U, \mu(U) \in M(U))$ will be called the guaranteed outcome functions of the players.

This definition is motivated by the following behaviour of the players. If for one of the coalitions, coalition $K$ say, it is profitable to switch at some time $\tau$ from the collection of strategies $U$ to the collection of deviation strategies $U_K^\tau$ with the collection of payments $\mu_K^\tau(U, U_K^\tau)$, then this coalition will make the switch at time $s$ directly preceding the time $\tau$.

*Definition 2.* Let assumptions 1°-4° hold. The differential game with guaranteed outcome functions $\rho_i$ (1.6), (1.7) will be called a differential game with continuous payment functions.

The main problem in this game is formulated as follows.

*Problem 1.* Find the couple $(U^\circ, \mu^\circ(U^\circ))$, where $\mu^\circ(U^\circ) \in M(U^\circ)$, which is Pareto-unimprovable relative to the guaranteed outcomes $\rho_i$ (1.6), (1.7).

*Definition 3.* The couple $(U^\circ, \mu^\circ(U^\circ))$, that solves Problem 1 will be called a cooperative solution (CS) of the game with continuous payment functions.

A cooperative version of a differential game without payments has been previously studied*; (*Kleimenov A.F., One approach to the analysis of cooperative positional differential games, Sverdlovsk, 1989. Unpublished manuscript, VINITI 11.07.89, 4609-V89.) the problem in this game can be obtained from Problem 1 by setting equal to zero the payments $\mu(U)$ and $\mu_K^\tau(U, U_K^\tau)$ in formulas (1.3)-(1.7). In what follows, the game with continuous payment functions will be called game 1 and the cooperative version of the game without payments game 2. It has been shown that the solution set $\mathbf{U}^2$ of game 2 consists of the collections of strategies $U_i = \{u_i(t, x, \alpha, \varepsilon), \beta_i(\varepsilon)\}$, having the following structure:

$$u_i(t, x, \alpha, \varepsilon) = \begin{cases} u_i^*(t) & \text{for } \alpha = \varnothing \\ u_i^\circ(t, x, \varepsilon \mid \tau, K, x^*(\cdot)) & \text{for } \alpha = K, i \notin K \\ \text{arbitrary} & \text{for } \alpha = K, i \in K \end{cases} \tag{1.8}$$

where $u_i^\circ(t, x, \varepsilon \mid \tau, K, x^*(\cdot))$ are the optimal strategies of coalition $I \setminus K$ in the pursuit-and-evasion game at time $\vartheta$ with a specially constructed set $N$; in the structure (1.8), these strategies play the role of penalty strategies /6/ in relation to players of the deviating coalition $K$.

*Lemma.* If $(U^\circ, \mu^\circ(U^\circ))$ is a solution of game 1, then $U^\circ$ is a solution of game 2.

Indeed, if $U^\circ$ is not a solution of game 2, then there exists a collection of strategies $U^*$ in game 2 on which the players' rewards are not smaller (and for at least one player actually larger). Then the couple $(U^*, \mu^*(U^*))$ constructed from $U^*$ will produce rewards that are not smaller (and at least for one player actually larger) than the couple $(U^\circ, \mu^\circ(U^\circ))$, which is a contradiction.

For fixed $U \in \mathbf{U}^2$, construct the set $L(U)$ of couples $(U, \mu(U))$, where $\mu(U) \in M(U)$. Put $L^* = \bigcup_{U \in \mathbf{U}^2} L(U)$. We denote by $P(L^*)$ the set of couples $(U, \mu(U))$ from the set $L^*$ that are Pareto-unimprovable relative to the criteria $\rho_i$ (1.6), (1.7). Our lemma leads to the following theorem.

*Theorem.* $P(L^*)$ is the solution set of game 1.

When the cooperative solution is realized, the players construct Euler polygonal lines, which suggests the following procedure for detecting the deviation of a coalition from the CS.

We introduce two additional assumptions 5° and 6°.

5°. Alongside the choice of the couple $(U, \mu(U))$, the players define the family $H(\varepsilon)$, $\varepsilon > 0$, of sets that are bounded and open in $(t_0, \vartheta) \times R^n$ and satisfy the following conditions:

$$\bigcap_{\varepsilon > 0} H(\varepsilon) = \{(t, x) : t_0 < t < \vartheta, x = x[t, t_0, x_0, U, \alpha^\circ[\cdot]]\} \tag{1.9}$$

$$H(\varepsilon_1) \subset H(\varepsilon_2), \quad \text{if} \quad \varepsilon_2 > \varepsilon_1 \tag{1.10}$$

$$(t, \psi[t]) \in H(\varepsilon), \psi[\cdot] \in X(t_0, x_0, U, \varepsilon, \Delta_i, \alpha^\circ[\cdot]), \quad t_0 < t < \vartheta \tag{1.11}$$

The deviation time $\tau$ identified with the least $t$ when the inclusion (1.11) does not hold, and the deviating coalition is the one whose actions have led to the violation of (1.11). Conditions (1.9), (1.10) of assumption 5° are satisfied by the sets $H(\varepsilon)$ which are open (in the space $t, x$) $\varepsilon$-neighbourhoods of the limiting motion generated by the collection of strategies $U$.

6°. Simultaneously with the beginning of the game, the players choose a common value of the accuracy parameter $\varepsilon$.

When assumptions 5° and 6° are satisfied, the players can realize the CS by Euler polygonal lines and avoid deviation from the cooperative solution.

2. Assume that the motion of the controlled system is described by an equation of the form

$$\xi^{\cdot\cdot} = u + v, \ \xi, u, v \in R^2; \ \|u\| = (u_1^2 + u_2^2)^{1/2} \leqslant 1, \ \|v\| \leqslant 1 \tag{2.1}$$

The vector $u$ is chosen by the first player and the vector $v$ by the second player. The initial conditions $\xi[t_0] = \xi_0, \xi^\cdot[t_0] = \xi_0^\cdot$ and the game termination time $\vartheta$ are given. The criteria maximized by the players have the form

$$\sigma_i(\xi[\vartheta]) = (c^i, \xi[\vartheta]) = \sum_{j=1}^{2} c_j^i \xi_j[\vartheta], \quad i = 1, 2 \tag{2.2}$$

where $c^i$ are given vectors.

Eq.(2.1) can be interpreted as the equation of motion of a material point of unit mass in the plane $(\xi_1, \xi_2)$ under the action of a force generated by the two players.

Setting $y_1 = \xi_1, y_2 = \xi_2, y_3 = \xi_1^\cdot, y_4 = \xi_2^\cdot$ in (2.2) and making the change of variables $x_1 = y_1 + (\vartheta - t)y_3$, $x_2 = y_2 + (\vartheta - t)y_4$, $x_3 = y_3$, $x_4 = y_4$, we obtain a system of equations, in which the first two equations have the form

$$x_i^\cdot = (\vartheta - t)(u_i + v_i), \quad i = 1, 2 \tag{2.3}$$

The criteria (2.2) in variables $x_1, x_2$ take the form

$$\sigma_i(x[\vartheta]) = (c^i, x[\vartheta]), \quad x = \left\| \begin{matrix} x_1 \\ x_2 \end{matrix} \right\|, \quad i = 1, 2 \tag{2.4}$$

Since the criteria (2.4) are determined only by the values of the coordinates $x_1, x_2$, and the right-hand side of system (2.3) is independent of the other coordinates, it suffices to analyse our differential game for the truncated system (2.3) with the criteria (2.4). The initial conditions for system (2.3) are $x_1[t_0] = x_{01} = \xi_{01} - (\vartheta - t_0)\xi_{01}^\cdot$, $x_2[t_0] = x_{02} = \xi_{02} - (\vartheta - t_0)\xi_{02}^\cdot$.

Define the initial conditions $t_0 = 0,\ \xi_{01} = \xi_{02} = \xi_{01}{}^{\cdot} = \xi_{02}{}^{\cdot} = 0$ and the following parameter values $\vartheta = \sqrt{2},\ c_1{}^1 = {}^1/_2,\ c_2{}^1 = \sqrt{3}/2,\ c_1{}^2 = -{}^5/_8,\ c_2{}^2 = 5\sqrt{3}/8, \lambda_1 = \lambda_2 = {}^1/_2,\ g_{12} = {}^3/_2,\ g_{21} = {}^6/_5$. We have $x_{01} = x_{02} = 0$.

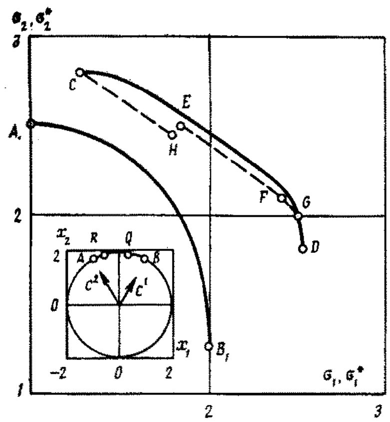


Fig.1

The circle in Fig.1 is the reachability set of system (2.3) at time $\vartheta = \sqrt{2}$ for these initial conditions. The solution of game 2 consists of the collections of strategies that generate the trajectories leading to points of the arc *AB*. In the reward plane $\sigma_1, \sigma_2$, the arc *AB* corresponds to the curve $A_1B_1$.

The curve *CD* describes the reward set $\sigma_1{}^*, \sigma_2{}^*$ (1.3) corresponding to the solutions of game 1; this curve is obtained in accordance with the theorem as the right upper envelope of the family of curves that describe unimprovable rewards for the collections of strategies solving game 2. Thus, the two-link polygonal line *EFG* specifies unimprovable rewards for the collection of strategies that lead to point *Q* in the plane $x_1, x_2$ and the segment *CH* gives the unimprovable rewards for the collection of strategies that lead to the point *R*.

## REFERENCES


1. KLEIMENOV A.F., Equilibrium coalition counterstrategies in many-person differential games, PMM, 46, 5, 1982.
2. KLEIMENOV A.F., Nash, Pareto, and Stackelberg solutions in non-zero-sum differential games, PMM, 51, 2, 1987.
3. KRASOVSKII N.N. and SUBBOTIN A.I., Positional Differential Games, Nauka, Moscow, 1974.
4. KRASOVSKII N.N., Control of Dynamic Systems, Nauka, Moscow, 1985.
5. VON NEUMANN J. and MORGENSTERN O., Theory of Games and Economic Behaviour, Nauka, Moscow, 1970.
6. KONONENKO A.F., On equilibrium positional strategies in non-zero-sum differential games, Dokl. Akad. Nauk SSSR, 231, 2, 1976.
7. PETROSYAN L.A. and DANILOV N.N., Cooperative Differential Games and their Applications, Izd. Tomsk, Univ., Tomsk, 1985.
8. PETROSYAN L.A. and CHISTYAKOV S.V., A sufficient condition of existence for strong $\varepsilon$-equilibria in a coalitionless differential game, in: Multistep, Hierarchical, Differential, and Cooperative Games, Kalinin. Univ., Kalinin, 1986.


Translated by Z.L.