

NASH, PARETO, AND STACKELBERG SOLUTIONS IN NON-ANTAGONISTIC TWO-PERSON GAMES*

A.F. KLEIMENOV

Non-antagonistic differential games are formalized below on the basis of the formal theory of positional antagonistic differential games /1, 2/. There is a large literature (for part of it, see survey /3/) concerned with the types of solution indicated in the title. The Nash solution is best known in coalitionless games; Pareto optimality is a basic concept in cooperative games without collateral pay-offs; and finally, the Stackelberg solution /4, 5/ is typical for hierarchical games. All three types of solution are considered below in a unified approach. Our concept of a Pareto solution differs from the usual one in allowing for the individual scope of each player. A single structure of strategies for all types of solution is discovered for two-person games. Relations between the sets of solutions of different types are established. It is shown that the set of solutions of each type is characterized by the solutions of appropriate non-standard (optimal) control problems. The results are illustrated by the example of the plane motion of a material particle subject to the total action of control forces generated by the different players. The paper is related to /6-8/ and continues the studies of /9, 10/. (See also: A.F. Kleimenov. On the theory of hierarchical two-person differential games, Preprint Inst. matematika i Mekhanika, UNTs AN SSSR, Sverdlovsk, 1985).

Let the dynamics of the controlled system be described by the equation

$$\dot{x} = f(t, x, u, v), u \in P, v \in Q, x[t_0] = x_0 \quad (1.1)$$

where the function $f: G \times P \times Q \rightarrow R^n$ is continuous with respect to the set of its arguments, satisfies a Lipschitz condition with respect to x , and a condition ensuring that the solutions of (1.1) can be continued into a given interval $[t_0, \theta]$. Here, G is a compactum in the space of the variables t, x ; P and Q are compacta in the appropriate finite-dimensional spaces.

The first and second players have at their disposal the choice of controls u and v respectively, and aim at minimizing their performance factors, which have the form

$$I_i = \sigma_i(x[\theta]), \quad i = 1, 2 \quad (1.2)$$

where the functions $\sigma_i: R^n \rightarrow R$ are continuous. Both players have available perfect information on the current position of the game $(t, x[t])$, and hence can use positional strategies /1, 2/ when forming their controls. We assume here for simplicity that it is sufficient for both players to confine themselves to the class of pure positional strategies, in the sense that an extension of this class (say, to mixed or counter-strategies) for a player does not lead to improvement of his factor. The more general case when such an extension is worth while can be considered similarly (see e.g., the author's paper mentioned at the end of the introduction).

The formalization below of positional strategies and their generated motions in non-antagonistic differential games is based on the formal theory for antagonistic differential games /1, 2/.

The first player's pure positional strategy (or simply strategy) is identified with the pair $U = \{u(t, x, \epsilon), \beta_1(\epsilon)\}$, where $u(\cdot, \cdot, \cdot)$ is a function of the position $(t, x) \in G$ and $\epsilon > 0$ with values in P . The function $\beta_1: (0, \infty) \rightarrow (0, \infty)$ is continuous and monotonic, and satisfies the condition $\beta_1(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$. The introduction of the accuracy parameter ϵ as the argument of the strategy is justified in the theory of antagonistic differential games /2/. By using it, we can ensure that the optimal strategies are universal with respect to the initial position. Essential use of this fact is made below. We also show that the introduction of the parameter ϵ plays an entirely new role when discovering the structure of solutions in

non-antagonistic games. The addition of the function $\beta_1(\cdot)$ is due to the specific nature of non-antagonistic games. For fixed ε the quantity $\beta_1(\varepsilon)$ is the constraint of the step of the discrete scheme which the first player uses when constructing the Euler step-lines.

The control law $Z(U, \varepsilon_1, \Delta_1)$, corresponding to strategy U , is defined by three components: the function $u(\cdot, \cdot, \cdot)$, appearing in strategy U ; the value of the parameter ε_1 ; and the division $\Delta_1 \{t_i^{(1)}\}$ of the interval $[t_0, \theta]$, $i = 1, \dots, k_1 + 1$, $t_1^{(1)} = t_0$, $t_{i+1}^{(1)} > t_i^{(1)}$, $t_{k_1+1}^{(1)} = \theta$. Here we must have the condition on the division step

$$\delta(\Delta_1) = \max_i (t_{i+1}^{(1)} - t_i^{(1)}) \leq \beta_1(\varepsilon_1)$$

The second player's strategy $V = \{v(t, x, \varepsilon), \beta_2(\varepsilon)\}$ is similarly defined, along with the corresponding control law $Z(V, \varepsilon_2, \Delta_2)$, where $\delta(\Delta_2) \leq \beta_2(\varepsilon_2)$.

Let strategies U and V be chosen. Let $Z(U, \varepsilon_1, \Delta_1)$ and $Z(V, \varepsilon_2, \Delta_2)$ be the control laws corresponding to the accuracy parameters ε_1 and ε_2 and the divisions $\Delta_1 \{t_i^{(1)}\}$, $\Delta_2 \{t_j^{(2)}\}$ of $[t_0, \theta]$, chosen by the players. The Euler's step-line generated by control laws $Z(U, \varepsilon_1, \Delta_1)$ and $Z(V, \varepsilon_2, \Delta_2)$ from the initial position (t_0, x_0) is the piecewise differentiable function

$$x_{\Delta}^{\varepsilon}[t] = x[t, t_0, x_0, Z(U, \varepsilon_1, \Delta_1), Z(V, \varepsilon_2, \Delta_2)] \quad (1.3)$$

which is the stepped solution of the differential equation

$$\begin{aligned} x_{\Delta}^{\varepsilon}[t] &= f(t, x_{\Delta}^{\varepsilon}[t], u_{\Delta_1}^{\varepsilon_1}[t], v_{\Delta_2}^{\varepsilon_2}[t]), \quad x_{\Delta}^{\varepsilon}[t_0] = x_0 \\ u_{\Delta_1}^{\varepsilon_1}[t] &= u(t_i^{(1)}, x_{\Delta}^{\varepsilon}[t_i^{(1)}], \varepsilon_1), \quad t_i^{(1)} \leq t < t_{i+1}^{(1)} \\ v_{\Delta_2}^{\varepsilon_2}[t] &= v(t_j^{(2)}, x_{\Delta}^{\varepsilon}[t_j^{(2)}], \varepsilon_2), \quad t_j^{(2)} \leq t < t_{j+1}^{(2)} \end{aligned} \quad (1.4)$$

The continuous function $x[t] = x[t, t_0, x_0, U, V]$, which is the uniform limit in $[t_0, \theta]$ of the sequence of Euler step-lines $\{x[t, t_0^k, x_0^k, Z(U, \varepsilon_1^k, \Delta_1^k), Z(V, \varepsilon_2^k, \Delta_2^k)]\}$ as $k \rightarrow \infty$, $\varepsilon_i^k \rightarrow 0$, $t_0^k \rightarrow t_0$, $x_0^k \rightarrow x_0$, $\delta(\Delta_i^k) \leq \beta_i(\varepsilon_i^k)$, will be called the motion generated by strategies U and V from the initial position (t_0, x_0) . At least one motion $x[t]$ exists. Whereas the pair of control laws $Z(U, \varepsilon_1, \Delta_1)$ and $Z(V, \varepsilon_2, \Delta_2)$ defines a unique Euler step-line, the pair of strategies U and V generates in general a set (pencil) of motions, which we henceforth denote by $X(t_0, x_0, U, V)$. It is compact in $C[t_0, \theta]$.

This formalization can be interpreted as follows. In essence, two interconnected models, the descriptive and constructive, of the game are obtained. Either model is built up from the equations of dynamics, formalization of the player's actions, and the motions generated by these actions. In both models, the dynamics are described by Eq.(1.1). A player's action in the descriptive model is formalized in the form of strategies U and V , and the motions generated, in the form of the pencil $X(t_0, x_0, U, V)$. In theoretical arguments it is more convenient to work with the descriptive model, whereas the constructive model is best used for calculation purposes. To construct the Euler step-lines, it is not, in general, necessary to calculate the motions in the descriptive model.

We can define differently the solution in a non-antagonistic differential game. Here we study the three familiar types of solution (see surve /3/): Nash, Pareto, and Stackelberg. However, we shall quote the relevant definitions, first in order to refine the details connected with non-uniqueness of the motions in the descriptive model, and second, so that the Pareto solution will differ from the classical solution.

Definition 1. The pair of strategies (U^N, V^N) forms a Nash equilibrium solution (N -solution) in our differential game if, given any strategies U and V , we have the inequalities

$$\begin{aligned} \min \sigma_1(x[\theta, t_0, x_0, U, V^N]) &\geq \max \sigma_1(x[\theta, t_0, x_0, U^N, V]) \\ \min \sigma_2(x[\theta, t_0, x_0, U^N, V]) &\geq \max \sigma_2(x[\theta, t_0, x_0, U, V^N]) \end{aligned} \quad (1.5)$$

The operations min and max are taken here with respect to all motions of the relevant pencils. Obviously, the value of each player's factor is the same in all the pencil of motions generated by the N -solution. The set of N -solutions is denoted by N .

Definition 2. The N -solution (U^P, V^P) forms a modified Pareto solution (P^* -solution) if, given any $(U, V) \in N$, we have one of the alternatives:

$$\begin{aligned} a) \sigma_i(x[\theta, t_0, x_0, U, V]) &= \sigma_i(x[\theta, t_0, x_0, U^P, V^P]), \quad i = 1, 2 \\ b) \exists j \in \overline{1, 2}: \sigma_j(x[\theta, t_0, x_0, U, V]) &> \sigma_j(x[\theta, t_0, x_0, U^P, V^P]) \end{aligned} \quad (1.6)$$

Our concept of P^* -solution differs from the classical Pareto solution, introduced for multicriterion problems, in that the P^* -solution is sought in the set of N -solutions. This is because, in a differential game, as distinct from a multicriterion problem, the control resources are distributed between the players, and hence the individual scope of each player must be taken into account. Let P^* denote the set of P^* -solutions. By definition, $P^* \subset N$.

The set of classical Pareto (P -solutions) is denoted by P .

Before defining the Stackelberg solution, we make the following assumptions, governing the sequence in which the players choose the strategies.

1^o. The first player, called the leader, explains his strategy $U^* = \{u^*(t, x, \varepsilon), \beta_1^*(\varepsilon)\}$ to the second player before the start of the game.

2^o. The second player, knowing the first player's strategy U^* , chooses a sensible strategy V^* from the condition

$$\max \sigma_2(x|\vartheta, t_0, x_0, U^*, V) \rightarrow \min_V \tag{1.7}$$

where max is taken over the pencil $X(t_0, x_0, U^*, V)$.

When assumptions 1^o, 2^o hold, the differential game is called hierarchical or a Stackelberg game with first player as leader. The set of second player's sensible strategies is denoted by $K_2(U^*)$.

The first player's task is to find the strategy U , which minimizes the factor $\sigma_1(x|\vartheta)$ (1.2) when the second player chooses a sensible strategy. (For more details, including the different cases when the second player chooses a strategy from the set $K_2(U)$, see /10/). Let the strategy U^{S1} be the solution of this problem. Then, the pair (U^{S1}, V^{S1}) , where $V^{S1} \in K_2(U^{S1})$, is called the Stackelberg (S_1 -) solution in the hierarchical game with first player as leader. Denote by S_1 the set of S_1 -solutions.

The hierarchical differential game with second player as leader can be similarly formalized. We denote by S_2 the set of S_2 -solutions in this game. The trajectories $x[t, t_0 \leq t \leq \vartheta]$, generated by the N -, P^* -, S_1 -, and S_2 -solutions will be called N -, P^* -, S_1 - and S_2 -trajectories respectively.

Under our assumptions, the sets N, P^*, S_1 , and S_2 are not empty. From /10/ and the present author's cited paper, we have:

Theorem 1. $S_i \cap P^* \neq \emptyset, S_i \subset N, i = 1, 2$.

Notice that the set P^* cannot in general be replaced by P in the statement of Theorem 1.

We shall now show how to find each of sets N, P^*, S_1, S_2 , by solving the appropriate non-standard control problem. In order to state these problems, we consider the auxiliary antagonistic differential games Γ_1 and Γ_2 . In both games the dynamics are described by Eq. (1.1). In Γ_1 the first player minimizes the factor $\sigma_1(x|\vartheta)$ (1.2), and the second player acts against him. In Γ_2 the second player minimizes the factor $\sigma_2(x|\vartheta)$ (1.2), and the first player acts against him. By the theory of positional antagonistic differential games /1, 2/, both our games have continuous value functions $\gamma_1(t, x)$ and $\gamma_2(t, x)$ and universal saddle points

$$\{u^{(i)}(t, x, \varepsilon), v^{(i)}(t, x, \varepsilon)\} \tag{1.8}$$

in the game Γ_i ($i = 1, 2$).

Problem 1 (non-standard control problem). Let the dynamics of the controlled system be described by Eq.(1.1). It is required to find the admissible measurable controls $u(t), v(t), t_0 \leq t \leq \vartheta$, such that the trajectory $x(t), t_0 \leq t \leq \vartheta$, generated by them satisfies the inequalities

$$\gamma_i(t, x(t)) \geq \gamma_i(\vartheta, x(\vartheta)) = \sigma_i(x(\vartheta)), t_0 \leq t \leq \vartheta, i = 1, 2 \tag{1.9}$$

Problem 2 (non-standard optimal control problem). For fixed $\alpha \in [0, 1]$, it is required to find the admissible measurable controls $u(t), v(t), t_0 \leq t \leq \vartheta$, which minimize the factor $\alpha\sigma_1(x(\vartheta)) + (1-\alpha)\sigma_2(x(\vartheta))$ under conditions (1.9).

Problem 3 is problem 2 with $\alpha = 1$; condition (1.9) with $i = 1$ is omitted.

Problem 4 is problem 2 with $\alpha = 0$; condition (1.9) with $i = 2$ is omitted.

Let $u^*(t), v^*(t) (t_0 \leq t \leq \vartheta)$ be admissible measurable controls, and $x^*(t) (t_0 \leq t \leq \vartheta)$ the generated trajectories. Using Luzin's theorem, we can indicate functions $u_*(t, \varepsilon), v_*(t, \varepsilon)$, piecewise continuous with respect to the first argument, such that, for the motion $x_*(t, \varepsilon)$ generated by them, we have $\|x_*(t, \varepsilon) - x^*(t)\| \leq \varepsilon$ for all $t \in [t_0, \vartheta]$, $\varepsilon > 0$. We consider the strategies $U^\circ = \{u^\circ(t, x, \varepsilon), \beta_1^\circ(\varepsilon)\}, V^\circ = \{v^\circ(t, x, \varepsilon), \beta_2^\circ(\varepsilon)\}$, where, with $t_0 \leq t \leq \vartheta, \varepsilon > 0$,

$$\begin{aligned} u^\circ &= u_*(t, \varepsilon), v^\circ = v_*(t, \varepsilon), \|x - x_*(t, \varepsilon)\| \leq \varepsilon \\ u^\circ &= u^{(2)}(t, x, \varepsilon), v^\circ = v^{(1)}(t, x, \varepsilon), \|x - x_*(t, \varepsilon)\| > \varepsilon \end{aligned} \tag{1.10}$$

and the functions $\beta_1^\circ(\cdot), \beta_2^\circ(\cdot)$ are chosen so that the following inequality is satisfied for the Euler step-lines:

$$\begin{aligned} \|x[t, t_0, x_0, Z(U^\circ, \varepsilon_1, \Delta_1), Z(V^\circ, \varepsilon_2, \Delta_2)] - x_*(t, \varepsilon)\| < \varepsilon \\ t_0 \leq t \leq \vartheta, \varepsilon_1 \leq \varepsilon, \varepsilon_2 \leq \varepsilon, \delta(\Delta_i) \leq \beta_i^\circ(\varepsilon), i = 1, 2 \end{aligned} \tag{1.11}$$

The functions $u^{(2)}(\cdot, \cdot, \cdot)$ and $v^{(1)}(\cdot, \cdot, \cdot)$ are defined in (1.8).

Notice that the pair of strategies (U^0, V^0) generates a unique motion identical with $x^*(\cdot)$.

Theorem 2. Let the measurable controls $u(t), v(t), t_0 \leq t \leq \theta$, solve Problem 1 (or Problem 2, 3, 4). Then, the pair of strategies (U^0, V^0) (1.10), (1.11) is the N -solution (or P^* -, S_1 -, S_2 -solution). Conversely, all the n -solutions (or P^* -, S_1 -, S_2 -solutions) are exhausted up to equivalence by the pairs of strategies (U^0, V^0) (1.10), (1.11), where $u^*(\cdot), v^*(\cdot)$ is the solution of Problem 1 (or of Problem 2, 3, 4).

As regards the N -, S_1 -, and S_2 -solutions, the author proved the theorem in the paper cited. The proof for the P^* -solution is similar.

Corollary. The set of trajectories $x^*(t)$ ($t_0 \leq t \leq \theta$) of system (1.1), generated by controls $u^*(t), v^*(t)$ ($t_0 \leq t \leq \theta$), solving Problem 1 (or Problem 2, 3, 4), is the same as the set of N -trajectories (or of P^* -, S_1 -, S_2 -trajectories).

Notes. 1^o. Strategies $u^{(i)}(\cdot, \cdot, \cdot), v^{(i)}(\cdot, \cdot, \cdot)$ in structure (1.10) can be interpreted as universal strategies of penalizing the opponent if he refuses at any instant $t \in [t_0, \theta]$ to follow trajectory $x^*(\cdot)$. An approach involving penalization strategies is proposed in /5/ for static games and is developed in /6/ for dynamic games.

2^o. When constructively tracking trajectory $x^*(\cdot)$ information exchange is required in the Euler step-lines between the players concerning the accuracy parameter ε_1 and ε_2 values.

3^o. We can consider similarly the case when the player's factors (1.2) contain integral as well as terminal terms (see the author's paper cited above).

2. Take an example. The equation

$$\dot{\xi} = u + v, \quad \xi, u, v \in R^2, \quad \|u\| \leq 1, \quad \|v\| \leq 1 \quad (2.1)$$

describes the plane motion of a material particle under the joint action of forces u and v , available to the first and second players respectively. We are given the initial conditions $\xi[t_0] = \xi_0, \dot{\xi}[t_0] = \dot{\xi}_0$ and the instant of termination θ . The i -th player's aim is to move point $\xi[\theta]$ as close as possible to the target point $a^{(i)}$, i.e.,

$$\sigma_i(\xi[\theta]) = \|\xi[\theta] - a^{(i)}\| \rightarrow \min, \quad i = 1, 2 \quad (2.2)$$

Putting $y_1 = \xi_1, y_2 = \xi_2, y_3 = \dot{\xi}_1, y_4 = \dot{\xi}_2$ in Eq. (2.1) and making the change of variables $x_1 = y_1 + (\theta - t)y_3, x_2 = y_2 + (\theta - t)y_4, x_3 = y_3, x_4 = y_4$, we obtain a system whose first two equations are

$$\dot{x}_i = (\theta - t)(u_i + v_i), \quad i = 1, 2 \quad (2.3)$$

In variables x_1, x_2 , the factors (2.2) become

$$\sigma_i(x[\theta]) = \|x[\theta] - a^{(i)}\|, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (2.4)$$

Since the factors (2.4) are defined by the values of the coordinates x_1 and x_2 only, while the right-hand side of system (2.3) is independent of the other coordinates, we can conclude that it suffices to study the differential game only for the truncated system (2.3) with factors (2.4). The initial conditions for system (2.3) will then be

$$x_i[t_0] = x_{0i} = \xi_{0i} - (\theta - t_0)\dot{\xi}_{0i}, \quad i = 1, 2$$

Clearly, the value functions of antagonistic games Γ_1 and Γ_2 are

$$\gamma_i(t, x) = \|x - a^{(i)}\|, \quad i = 1, 2$$

and the universal optimal strategies (1.8) are

$$u^{(i)}(t, x, \varepsilon) = -v^{(i)}(t, x, \varepsilon) = (-1)^i \frac{x - a^{(i)}}{\|x - a^{(i)}\|}, \quad i = 1, 2$$

We specify the initial conditions $t_0 = 0, \xi_{01} = 2.2, \dot{\xi}_{01} = -0.8, \xi_{02} = 1.3, \dot{\xi}_{02} = -0.2$ and the parameter values $\theta = 2, a_1^{(1)} = -1, a_2^{(1)} = 5, a_1^{(2)} = 5, a_2^{(2)} = 4$. We then have $x_{01} = 0.6, x_{02} = 0.9$.

The auxiliary Problem 1, whose solutions appear in structure (1.10) of the N -solutions, is formalized as follows: to find measurable vector functions $u(t), v(t), 0 \leq t \leq 2$, which satisfy the conditions

$$\|x(\theta) - a^{(i)}\| \leq \|x(t) - a^{(i)}\|, \quad 0 \leq t \leq 2, \quad i = 1, 2$$

where $x(\cdot)$ is the trajectory of the system

$$\dot{x}(t) = (\theta - t)[u(t) + v(t)], \quad \|u(t)\| \leq 1, \quad \|v(t)\| \leq 1, \quad x(0) = x_0 \quad (2.5)$$

Problems 2-4 are stated in a similar way. Problems 1-4 were solved. Without going into details, we note e.g., that, in essence, all the solutions of Problem 1 can be obtained in the class of continuous controls $u(t), v(t), 0 \leq t \leq 2$, generating phase trajectories of system (2.5) of just three types: segments of a straight line, arcs of a circle, and "pasteing" of a segment with an arc of a circle. By using Theorem 2, the sets N, P^*, S_1 and S_2 were found.

