VESTNIK UDMURTSKOGO UNIVERSITETA. MATEMATIKA. MEKHANIKA. KOMP'YUTERNYE NAUKI
MATHEMATICS 2017. Vol. 27. Issue 3

MSC2010: 91A23, 91A10, 91A05

© *Yu. V. Averboukh*

RANDOMIZED NASH EQUILIBRIUM FOR DIFFERENTIAL GAMES[1]

The paper is concerned with the randomized Nash equilibrium for a nonzero-sum deterministic differential game of two players. We assume that each player is informed about the control of the partner realized up to the current moment. Therefore, the game is formalized in the class of randomized non-anticipative strategies. The main result of the paper is the characterization of a set of Nash values considered as pairs of expected players' outcomes. The characterization involves the value functions of the auxiliary zero-sum games. As a corollary we get that the set of Nash values in the case when the players use randomized strategies is a convex hull of the set of Nash values in the class of deterministic strategies. Additionally, we present an example showing that the randomized strategies can enhance the outcome of the players.

Keywords: differential games, Nash equilibrium, randomized strategies.

DOI: 10.20537/vm170301

Introduction

The paper is concerned with a two players nonzero-sum differential game. We examine this problem under the framework of Nash equilibrium proposed for one-shot games in [13]. For the differential game the study of Nash equilibria started with papers [6, 9].

There exist two approaches to analyze Nash equilibria of differential game in the case when the players are well-informed (i. e. they know either the state of the system or the control of the partner realized up to current time). The first approach reduces the original problem to a system of Hamilton–Jacobi PDEs. Despite of some progress in the theory of systems of Hamilton–Jacobi PDEs [3, 4] the general theory of such systems is not developed. Moreover, there are intrinsic obstructions to such theory [4].

The second approach is based on the concept of punishment strategies. This concept goes back to the folks theorems in the theory of repeated games. The punishment strategy equilibrium is constructed in the following way. The players choose a line and arrange to move along this line; any unilateral deviation are punished by all other players. Note that the player using the punishment strategy is primary concerned not with her profit but with the maintenance of the arrangement. This is a main disadvantage of the punishment strategy approach. However, the punishment strategies are widely used to prove the existence of the Nash equilibrium.

The existence of the Nash equilibrium of differential game was proved using punishment techniques in [10–12, 16]. The set of Nash equilibria in the class of punishment strategies was characterized in [7, 10, 16]. In those papers feedback strategies or Friedman strategies were used i. e. it is assumed that the players are informed on the current position.

The equivalent approach works when the players are informed about the controls of each other realized up to the current time. In this case players use so called non-anticipative strategies. This formalization was first developed for zero-sum differential games [8, 14, 15]. For nonzero-sum differential games the Nash equilibria in the class of non-anticipative strategies were studied in [1, 2].

Note that all strategies mentioned above are deterministic. However, randomized strategies are widely used for the one-shot games. In the paper we develop an approach based on a randomization of pure strategies for the nonzero-sum differential game case. We will use non-anticipative strategies. Note that for the stochastic differential games the existence of the Nash equilibria was proved in the

[1]This work was supported by the Russian Science Foundation (project no. 17–11–01093).

class of stochastic non-anticipative strategies with a delay [5]. In that paper the set of Nash values was also characterized.

The approach introduced in this paper means that the randomized strategies for nonzero-sum two-player differential game are stochastic processes defined on the same probability space whereas the probability space is not fixed. This is the main difference between this paper and paper [5] where the probability space is fixed. In the paper we assume that the player are informed about the controls of each other realized up to the current time and use non-anticipative strategies. We obtain the characterization of Nash equilibria and prove that the set of Nash values is a convex hull of the set of Nash values in the case when only deterministic strategies are allowed. Additionally, we present an example showing that the usage of randomize strategies can enhance the outcome of the players.

§ 1. Randomized strategies

We consider the dynamical system governed by two players

$$\frac{d}{dt}x(t) = f(t, x(t), u(t), v(t)), \quad t \in [0, T], \quad x(t) \in \mathbb{R}^d, \quad u(t) \in U, \quad v(t) \in V.$$

Here $u(t)$ (respectively, $v(t)$) is the control of the first (respectively, second) player. We assume that the first (respectively, second) player wishes to maximize $\sigma_1(x(T))$ (respectively, $\sigma_2(x(T))$).

We impose the following conditions on the sets U and V and the functions f, σ_1 and σ_2:

1) the sets U and V are metric compact;

2) f, σ_1 and σ_2 are continuous;

3) f enjoys sublinear growth w.r.t. x;

4) f is locally Lipschitz continuous w.r.t. the space variable, i. e., for any compact $G \subset \mathbb{R}^d$ there exists a constant L such that, if $x', x'' \in G$, $t \in [0, T]$, $u \in U$, $v \in V$, then

$$\|f(t, x', u, v) - f(t, x'', u, v)\| \leq L\|x' - x''\|;$$

5) (Isaacs condition) for every $x, p \in \mathbb{R}^d$, $t \in [0, T]$,

$$\min_{u \in U} \max_{v \in V} \langle p, f(t, x, u, v)\rangle = \max_{v \in V} \min_{u \in U} \langle p, f(t, x, u, v)\rangle.$$

Below we consider relaxed controls. Let Ξ be a metric space. We assume that Ξ is endowed by the Borel σ-algebra. A measure ξ on $[t_0, t_1] \times \Xi$ is called consistent with the Lebesgue measure on $[t_0, t_1]$ if, for any $t', t'' \in [t_0, t_1]$, $t' < t''$,

$$\xi([t', t''] \times \Xi) = t'' - t'.$$

Below we say that the measures consistent with the Lebesgue measure ξ_1, ξ_2 on $[t_0, t_1] \times \Xi$ coincide on $[t_0, \tau]$ and write $\xi_1 \equiv_{[t_0, \tau]} \xi_2$ if, for any measurable $\Gamma \subset [t_0, \tau]$ and $\Lambda \subset \Xi$,

$$\xi_1(\Gamma \times \Lambda) = \xi_2(\Gamma \times \Lambda).$$

A measure on $[t_0, t_1] \times U$ (respectively, $[t_0, t_1] \times V$) consistent with the Lebesgue measure on $[t_0, t_1]$ is called a relaxed control of the first (respectively, second) player on time interval $[t_0, t_1]$. Further, we denote by $\mathcal{U}[t_0, t_1]$ (respectively, $\mathcal{V}[t_0, t_1]$) the set of relaxed controls of the first (respectively, second) player. A measure on $[t_0, t_1] \times U \times V$ consistent with the Lebesgue measure on $[t_0, t_1]$ is called a relaxed joint control of the players. The set of relaxed joint control of the players is denoted by $\mathcal{M}[t_0, t_1]$.

Given a relaxed joint strategy of the players ξ^c and an initial position (t_0, x_0), denote by $x(\cdot, t_0, x_0, \xi^c)$ a solution of the problem

$$x(t) = x_0 + \int_{[t_0,t]\times U\times V} f(\tau, x(\tau), u, v)\xi^c(d(\tau, u, v)).$$

Existence and uniqueness theorem for the solution of this problem can be found in [15].

Put

$$\mathcal{J}_i(t_0, x_0, \xi^c) \triangleq \sigma_i(x(T, t_0, x_0, \xi^c)).$$

We say that $\zeta^1 : \mathcal{V}[t_0, t_1] \to \mathcal{M}[t_0, t_1]$ is a non-anticipative strategy of the first player if

- for any $\xi^2 \in \mathcal{V}[t_0, t_1]$ and any measurable $\Gamma \subset [t_0, t_1]$, $\Lambda \subset V$,
$$\zeta^1[\xi^2](\Gamma \times U \times \Lambda) = \xi^2(\Gamma \times \Lambda);$$
- for any $\tau \in [t_0, t_1]$ and any $\xi^2_1, \xi^2_2 \in \mathcal{V}[t_0, t_1]$, the equality $\xi^2_1 \equiv_{[t_0,\tau]} \xi^2_2$ implies the equality $\zeta^1[\xi^2_1] \equiv_{[t_0,\tau]} \zeta^1[\xi^2_2]$.

The set of all non-anticipative strategies of the first player on $[t_0, t_1]$ is denoted by $\mathcal{Q}_1[t_0, t_1]$.

Analogously, a mapping $\zeta^2 : \mathcal{U}[t_0, t_1] \to \mathcal{M}[t_0, t_1]$ is a non-anticipative strategy of the second player if

- for any $\xi^1 \in \mathcal{U}[t_0, t_1]$ and any measurable $\Gamma \subset [t_0, t_1]$, $\Lambda \subset U$,
$$\zeta^2[\xi^1](\Gamma \times \Lambda \times V) = \xi^1(\Gamma \times \Lambda);$$
- for any $\tau \in [t_0, t_1]$ and any $\xi^1_1, \xi^1_2 \in \mathcal{U}[t_0, t_1]$, the equality $\xi^1_1 \equiv_{[t_0,\tau]} \xi^1_2$ implies the equality $\zeta^2[\xi^1_1] \equiv_{[t_0,\tau]} \zeta^2[\xi^1_2]$.

We denote the set of all non-anticipative strategies of the second player on $[t_0, t_1]$ by $\mathcal{Q}_2[t_0, t_1]$.

In [1] the following definition of the incentive Nash equilibrium is proposed. A triple $(\xi^{c,*}, \zeta^{1,*}, \zeta^{2,*})$ where $\xi^{c,*} \in \mathcal{M}[t_0, T]$, $\zeta^{i,*} \in \mathcal{Q}_i[t_0, T]$, is an incentive Nash equilibrium at (t_0, x_0) if

- there exist control $\xi^{1,*}$, $\xi^{2,*}$ such that $\xi^{c,*} = \zeta^{1,*}[\xi^{2,*}] = \zeta^{2,*}[\xi^{1,*}]$;
- for any $\xi^1 \in \mathcal{U}[t_0, T]$,
$$\mathcal{J}_1(t_0, x_0, \xi^{c,*}) \geq \mathcal{J}_1(t_0, x_0, \zeta^{2,*}[\xi^1]);$$
- for any $\xi^2 \in \mathcal{V}[t_0, T]$,
$$\mathcal{J}_2(t_0, x_0, \xi^{c,*}) \geq \mathcal{J}_2(t_0, x_0, \zeta^{1,*}[\xi^2]).$$

The equivalence of the incentive and feedback Nash equilibria is proved in [1].

In the paper we extend the class of strategies and use randomized strategies. We endow the sets $\mathcal{U}[t_0, t_1]$, $\mathcal{V}[t_0, t_1]$, and $\mathcal{M}[t_0, t_1]$ with the topology of narrow convergence. This topology is metrizable [17]. Thus, the Borel σ-algebras on this sets are well-defined.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A measurable function $\omega \mapsto \mu(\omega) \in \mathcal{U}[t_0, t_1]$ (respectively, $\omega \mapsto \nu(\omega) \in \mathcal{V}[t_0, t_1]$) is called a randomized relaxed control of the first (respectively, second) player. Additionally, a measurable function $\omega \mapsto \eta(\omega) \in \mathcal{M}[t_0, t_1]$ is called a randomized joint control of the players. If η is a randomized joint control on $[t_0, T]$, put

$$J_i(t_0, x_0, \eta) \triangleq E\mathcal{J}_i(t_0, x_0, \eta(\cdot)) = \int_\Omega \sigma_i(x(T, t_0, x_0, \eta(\omega)))P(d\omega).$$

A measurable mapping $\alpha\colon \Omega \times \mathcal{V}[t_0, t_1] \to \mathcal{M}[t_0, t_1]$ (respectively, $\beta\colon \Omega \times \mathcal{U}[t_0, t_1] \to \mathcal{M}[t_0, t_1]$) such that for any ω, $\alpha(\omega, \cdot) \in \mathcal{Q}_1[t_0, t_1]$ (respectively, $\beta(\omega, \cdot) \in \mathcal{Q}_2[t_0, t_1]$) is a randomized nonanticipative strategy of the second (respectively, first) player on $[t_0, t_1]$.

Definition 1. We say that the 6-tuple $(\Omega, \mathcal{F}, P, \eta^*, \alpha^*, \beta^*)$ is an incentive Nash equilibrium in the class of randomized strategies at the position (t_0, x_0) if

- $(\Omega, \mathcal{F}, P)$ is probability space;
- η^* is a randomized relaxed joint control of the players, whereas α^*, β^* are non-anticipative randomized strategies of the first and second players respectively;
- there exist a randomized control μ^* and ν^* such that $\eta^*(\cdot) = \alpha[\cdot, \nu^*(\cdot)]$, $\eta^*(\cdot) = \beta[\cdot, \mu^*(\cdot)]$;
- $J_1(t_0, x_0, \eta^*) \geq J_1(t_0, x_0, \beta^*[\cdot, \mu(\cdot)])$ for any randomized relaxed control of the first player $\mu(\cdot)$;
- $J_2(t_0, x_0, \eta^*) \geq J_2(t_0, x_0, \alpha^*[\cdot, \nu(\cdot)])$ for any randomized relaxed control of the second player $\nu(\cdot)$.

Definition 2. We call a pair $(c_1, c_2) \in \mathbb{R}^2$ a Nash value in the class of randomized strategies at the position (t_0, x_0) if there exists an incentive Nash equilibria at the position (t_0, x_0) $(\Omega, \mathcal{F}, P, \eta^*, \alpha^*, \beta^*)$ such that $c_i = J_i(t_0, x_0, \eta^*)$.

§ 2. Characterization of Nash equilibria

To characterize the set of Nash values in the class of randomized strategies we consider the auxiliary zero-sum differential games, where the i-th player wishes to maximize the outcome given by $\sigma_i(x(T))$ and the other player prevents her. Recall [15] that the value function of the zero-sum games where the first player maximizes $\sigma_1(x(T))$ is equal to

$$V_1(t_0, x_0) = \min_{\zeta^2 \in \mathcal{Q}_2[t_0,T]} \min_{\xi^1 \in \mathcal{V}[t_0,T]} \sigma_1(x(T, t_0, x_0, \zeta^2[\xi^1])), \tag{2.1}$$

whereas the value function of the zero-sum game where the second player wishes to maximize $\sigma_2(x(T))$ is given by

$$V_2(t_0, x_0) = \min_{\zeta^1 \in \mathcal{Q}_1[t_0,T]} \min_{\xi^2 \in \mathcal{V}[t_0,T]} \sigma_1(x(T, t_0, x_0, \zeta^1[\xi^2])).$$

Theorem 1. *A pair (c_1, c_2) is a Nash value in the class of randomized strategies at the position (t_0, x_0), if and only if there exists a probability $\mathbb{P}$ on $\mathcal{M}[t_0, T]$ such that*

$$c_i = \int_\Omega \mathcal{J}_i(t_0, x_0, \xi^c)\mathbb{P}(d\xi^c), \quad i = 1, 2, \tag{2.2}$$

and for any $t \in [t_0, T]$

$$\sigma_i(x(T, t_0, x_0, \xi^c)) \geq V_i(t, x(t, t_0, x_0, \xi^c)) \ \ \mathbb{P}\text{-a.s.}, \quad i = 1, 2. \tag{2.3}$$

The proof of the theorem involves the operation of concatenation of the relaxed controls. It is defined in the following way. If Ξ is a metric set, $t_0 < t_1$, $\tau \in [t_0, t_1]$, ξ' is a measure on $[t_0, t_1] \times \Xi$ consistent with the Lebesgue measure on $[t_0, t_1]$, whereas ξ'' is a measure on $[\tau, t_1]$ consistent with the Lebesgue measure on $[\tau, t_1]$, then denote by $\xi' \odot_\tau \xi''$ the probability ξ on $[t_0, t_1] \times \Xi$ defined by the rule: for $\Lambda \subset [t_0, t_1] \times \Xi$,

$$\xi(\Lambda) \triangleq \xi'(\Lambda \cap ([t_0, \tau] \times \Xi)) + \xi''(\Lambda \cap ([\tau, t_1] \times \Xi)).$$

Note that ξ is also consistent with the Lebesgue measure on $[t_0, t_1]$.

Additionally, define the projectors π_1, π_2 by the rules: if (t, u, v), then

$$\pi_1(t, u, v) \triangleq (t, u), \quad \pi_2(t, u, v) \triangleq (t, v).$$

Further, if $(\Omega',\mathcal{F}')$, $(\Omega'',\mathcal{F}'')$ are measurable spaces, $h:\Omega'\to\Omega''$ is measurable, m is a probability on $\mathcal{F}'$, then $h_{\#}m$ is a probability on $\mathcal{F}''$ such that, for any $\Upsilon\in\mathcal{F}''$,

$$(h_{\#}m)(\Upsilon)\triangleq m(h^{-1}(\Upsilon)).$$

Note that $\zeta^1:\mathcal{V}[t_0,t_1]\to\mathcal{M}[t_0,t_1]$ is a non-anticipative strategy of the first player if and only if $\pi_{2\#}(\zeta^1[\xi^2])=\xi^2$ for any $\xi^2\in\mathcal{V}[t_0,t_1]$. Similarly, $\zeta^2\colon\mathcal{U}[t_0,t_1]\to\mathcal{M}[t_0,t_1]$ is a non-anticipative strategy of the seconds player if and only if $\pi_{1\#}(\zeta^2[\xi^1])=\xi^1$ for any $\xi^1\in\mathcal{U}[t_0,t_1]$.

P r o o f o f T h e o r e m 1. First, let us assume that (c_1,c_2) is a Nash value. Thus, there exists an incentive Nash equilibrium in the class of randomized strategies $(\Omega,\mathcal{F},P,\eta^*,\alpha^*,\beta^*)$ such that $c_i=J_i(t_0,x_0,\eta^*)$. Let Γ_t^i be a set of $\omega\in\Omega$ such that inequality

$$\sigma_i(x(T,t_0,x_0,\eta^*(\omega)))<V_i(t,x(t,t_0,x_0,\eta^*(\omega))) \tag{2.4}$$

holds true.

Let us prove that $P(\Gamma_t^i)=0$. Without loss of generality, we can assume that $i=1$. Given $t\in[t_0,t_1]$, denote by $y(\omega)$ the state $x(t,t_0,x_0,\eta^*(\omega))$.

Let $\mu^*(\omega)$ denote $\pi_{1\#}\eta^*(\omega)$. Note that $\eta^*(\omega)=\beta^*[\omega,\mu^*(\omega)]$.

Let us define a randomized non-anticipative strategy on $[t,T]$ of the second player $\beta^\natural$ by the following rule: for $\omega\in\Omega$, $\xi^1\in\mathcal{U}[t,T]$,

$$\beta^\natural[\omega,\xi^1]\triangleq\beta^*[\omega,\mu^*(\omega)\odot_t\xi^1].$$

Further, (2.1) implies that

$$V_1(t,y(\omega))=\min_{\zeta^2\in\mathcal{Q}_2[t,T]}\max_{\xi^1\in\mathcal{U}[t,T]}\sigma_1(x(T,t,y(\omega),\zeta^2(\xi^1)))\le\max_{\xi^1\in\mathcal{U}[t,T]}\sigma_1(x(T,t,y(\omega),\beta^\natural[\omega,\xi^1])). \tag{2.5}$$

Let $\gamma(\omega)\in\mathcal{U}[t,T]$ maximize the right-hand side of (2.5), i. e.,

$$\max_{\xi^1\in\mathcal{U}[t,T]}\sigma_1(x(T,t,y(\omega),\beta^\natural(\omega,\xi^1)))=\sigma_1(x(T,t,y(\omega),\beta^\natural(\omega,\gamma(\omega)))). \tag{2.6}$$

Without loss of generality one may assume that the mapping $\omega\mapsto\gamma(\omega)$ is measurable.

Define a randomized control $\mu^\natural$ of the first player on $[t_0,T]$ by the following rule:

$$\mu^\natural(\omega)\triangleq\begin{cases}\mu^*(\omega)\odot_t\gamma(\omega), & \omega\in\Gamma_t^1\\ \mu^*(\omega), & \omega\notin\Gamma_t^1.\end{cases}$$

Note that, if $\tau>t$, $\omega\notin\Gamma_t^1$,

$$x(\tau,t_0,x_0,\eta^*(\omega))=x(\tau,t_0,x_0,\beta^*[\omega,\mu^\natural(\omega)]). \tag{2.7}$$

Moreover, for $\tau>t$, $\omega\in\Gamma_t^1$,

$$x(\tau,t,y(\omega),\beta^\natural[\omega,\gamma(\omega)])=x(\tau,t_0,x_0,\beta^*[\omega,\mu^\natural(\omega)]). \tag{2.8}$$

If one assume that $P(\Gamma_t^1)>0$, then from (2.4) we get

$$\begin{aligned}c_1=J_1(t_0,x_0,\eta^*)=\int_\Omega\mathcal{J}_1(t_0,x_0,\eta^*(\omega))P(d\omega)=\\ =\int_{\Gamma_t^1}\sigma_1(x(T,t_0,x_0,\eta^*(\omega)))P(d\omega)+\int_{\Omega\setminus\Gamma_t^1}\sigma_1(x(T,t_0,x_0,\eta^*(\omega)))P(d\omega)<\\ <\int_{\Gamma_t^1}V_1(t,y(\omega))P(d\omega)+\int_{\Omega\setminus\Gamma_t^1}\sigma_1(x(T,t_0,x_0,\eta^*(\omega)))P(d\omega).\end{aligned}$$

Further, under assumption $P(\Gamma_t^1) > 0$ (2.6), (2.5), (2.7), (2.8) imply that

$$c_1 < \int_{\Gamma_t^1} \max_{\xi^1 \in \mathcal{U}[t,T]} \sigma_1(x(T,t,y(\omega),\beta^\natural[\omega,\xi^1]))P(d\omega) + \int_{\Omega\setminus\Gamma_t^1} \sigma_1(x(T,t_0,x_0,\beta^*[\omega,\mu^*(\omega)]))P(d\omega) =$$
$$= \int_{\Gamma_t^1} \sigma_1(x(T,t_0,x_0,\beta^*[\omega,\mu^\natural(\omega)]))P(d\omega) + \int_{\Omega\setminus\Gamma_t^1} \sigma_1(x(T,t_0,x_0,\beta^*[\omega,\mu^\natural(\omega)]))P(d\omega) =$$
$$= \int_{\Omega} \sigma_1(x(T,t_0,x_0,\beta^*[\omega,\mu^\natural(\omega)]))P(d\omega).$$

This contradicts with the assumption that $(\Omega,\mathcal{F},P,\eta^*,\alpha^*,\beta^*)$ is an incentive Nash equilibrium in the class of randomized strategies. The equality $P(\Gamma_t^2) = 0$ is proved in the same way.

To complete the proof let us define the probability $\mathbb{P}$ on $\mathcal{M}[t_0,T]$ by the rule: for $\Upsilon \subset \mathcal{M}[t_0,T]$,

$$\mathbb{P}(\Upsilon) = P\{\omega : \eta^*(\omega) \in \Upsilon\}.$$

Since inequality (2.4) can violates only on P-negligible set, we have that (2.2) and (2.3) hold true.

Now assume that (2.2) and (2.3) are valid for some probability $\mathbb{P}$ on $\mathcal{M}[t_0,T]$. We shall prove that (c_1,c_2) is a Nash value in the class of randomized strategies. Using continuity of V_i and σ one can deduce that

$$\sigma_i(x(T,t_0,x_0,\xi^c)) \geq V_i(t,x(t,t_0,x_0,\xi^c)), \quad i = 1,2$$

for $\mathbb{P}$-a.e. ξ^c and λ-a.e. $t \in [t_0,T]$ (here λ stands for the Lebesgue measure).

Recall that we endow $\mathcal{M}[t_0,T]$ with the topology of narrow convergence which is metrizable. Let $\mathcal{B}$ stand for Borel σ-algebra on $\mathcal{M}[t_0,T]$. Put $\Omega \triangleq \mathcal{M}[t_0,T]$, $\mathcal{F} \triangleq \mathcal{B}(\mathcal{M}[t_0,T])$, $P \triangleq \mathbb{P}$. For $\omega \in \Omega = \mathcal{M}[t_0,T]$, put $\eta^*(\omega) \triangleq \omega$, $\mu^*(\omega) \triangleq \pi_{1\#}\omega$, $\nu^*(\omega) \triangleq \pi_{2\#}\omega$.

Further, if $\omega \in \mathcal{M}[t_0,T]$, $\xi_2 \in \mathcal{V}[t_0,T]$, then let $\tau^1[\omega,\xi^2]$ be the greatest time τ such that

$$\pi_{2\#}\omega \equiv_{[t_0,\tau]} \xi_2.$$

For $\theta \in [t_0,T]$, $z \in \mathbb{R}^d$, let $\zeta^{1,*}_{\theta,z}$ minimize the right-hand side of

$$V_2(\theta,z) = \min_{\zeta^1 \in \mathcal{Q}_1[\theta,T]} \max_{\xi^2 \in \mathcal{V}[\theta,T]} \sigma_2(x(T,\theta,z,\zeta^1[\xi^2])).$$

For $\omega \in \Omega = \mathcal{M}[t_0,T]$, $\xi^2 \in \mathcal{U}[t_0,T]$, $\Lambda \subset [t_0,T] \times U \times V$, put

$$(\alpha^*[\omega,\xi^2])(\Lambda) \triangleq \omega(\Lambda \cap ([t_0,\tau^1(\omega,\xi^2)] \times U \times V)) +$$
$$+ (\zeta^{1,*}_{\tau^1(\omega,\xi^2),x(\tau^1(\omega,\xi^2),t_0,x_0,\omega)}[\xi^2])(\Lambda \cap ([\tau^1(\omega,\xi^2),T] \times U \times V)).$$

Analogously, if $\omega \in \mathcal{M}[t_0,T]$, $\xi^1 \in \mathcal{U}[t_0,T]$, then let $\tau^2[\omega,\xi^1]$ be the greatest time τ such that

$$\pi_{1\#}\omega \equiv_{[t_0,\tau]} \xi^1.$$

As above, given $\theta \in [t_0,T]$, $z \in \mathbb{R}^d$, choose a non-anticipative strategy of the second player $\zeta^{2,*}_{\theta,z}$ by the rule

$$V_1(\theta,z) = \max_{\xi^1 \in \mathcal{U}[\theta,T]} \sigma_1(x(T,\theta,z,\zeta^{2,*}_{\theta,z}[\xi^1])).$$

Finally, for $\omega \in \Omega = \mathcal{M}[t_0,T]$, $\xi^1 \in \mathcal{V}[t_0,T]$, $\Lambda \subset [t_0,T] \times U \times V$, put

$$(\beta^*[\omega,\xi^1])(\Lambda) \triangleq \omega(\Lambda \cap ([t_0,\tau^2(\omega,\xi^1)] \times U \times V)) +$$
$$+ (\zeta^{2,*}_{\tau^2(\omega,\xi^1),x(\tau^2(\omega,\xi^1),t_0,x_0,\omega)}[\xi^1])(\Lambda \cap ([\tau^2(\omega,\xi^1),T] \times U \times V)).$$

We have that $\omega = \alpha^*[\omega,\nu^*(\omega)] = \beta^*[\omega,\mu^*(\omega)]$. Further, for any $\omega \mapsto \mu(\omega) \in \mathcal{U}[t_0,t_1]$, $\omega \mapsto \nu(\omega) \in \mathcal{V}[t_0,t_1]$, and for $\mathbb{P}$-a.e. ω,

$$\sigma_1(x(T,t_0,x_0,\beta^*[\omega,\mu(\omega)])) = V_1(\tau^2(\omega,\mu(\omega)),x(\tau^2(\omega,\mu(\omega))),t_0,x_0,\omega) \leq \sigma_1(x(T,t_0,x_0,\omega)),$$

$$\sigma_2(x(T,t_0,x_0,\alpha^*[\omega,\nu(\omega)])) = V_2(\tau^1(\omega,\nu(\omega)), x(\tau^1(\omega,\nu(\omega))), t_0, x_0, \omega) \leq \sigma_2(x(T,t_0,x_0,\omega)).$$

Thus, $(\Omega,\mathcal{F},P,\eta^*,\alpha^*,\beta^*)$ is an incentive Nash equilibrium at (t_0,x_0) in the class of randomized strategies. Therefore, (c_1,c_2) is a Nash value. □

Below we denote the set of Nash values at (t_0,x_0) in the class of randomized strategies by $\mathcal{N}^r(t_0,x_0)$. One can also consider the set of Nash values at (t_0,x_0) in the class of deterministic strategies. Denote this set by $\mathcal{N}^d(t_0,x_0)$. Recall [1] that the pair (c_1,c_2) belongs to $\mathcal{N}^r(t_0,x_0)$, if and only if there exists a control $\xi^c \in \mathcal{M}[t_0,T]$ such that

$$c_i = \sigma_i(x(T,t_0,x_0,\xi^c)) \tag{2.9}$$

and, for any $t \in [t_0,T]$,

$$c_i \geq V_i(t,x(t,t_0,x_0,\xi^c)). \tag{2.10}$$

Corollary 1. $\mathcal{N}^r(t_0,x_0) = \operatorname{co}\mathcal{N}^d(t_0,x_0)$.

Proof of Corollary 1. The proof directly follows from Theorem 1 and (2.9), (2.10). □

§ 3. Example

Let us highlight the difference between Nash values in the class of the deterministic and randomized strategies by the following simple example. Let

$$\dot{x}_1 = u, \quad \dot{x}_2 = v, \quad |u|,|v| \leq 1, \quad T = 1,$$

$$\sigma_i(x_1,x_2) = |x_{3-i}|(1 - |\sin(2\mathrm{arctg}(|x_1/x_2|))|).$$

We restrict our attention to the case when initial time is 0 and initial position is $x_{1,0} = x_{2,0} = 0$. Since the dynamics of the game is symmetric the sets of Nash values at $(0,0,0)$ are also symmetric.

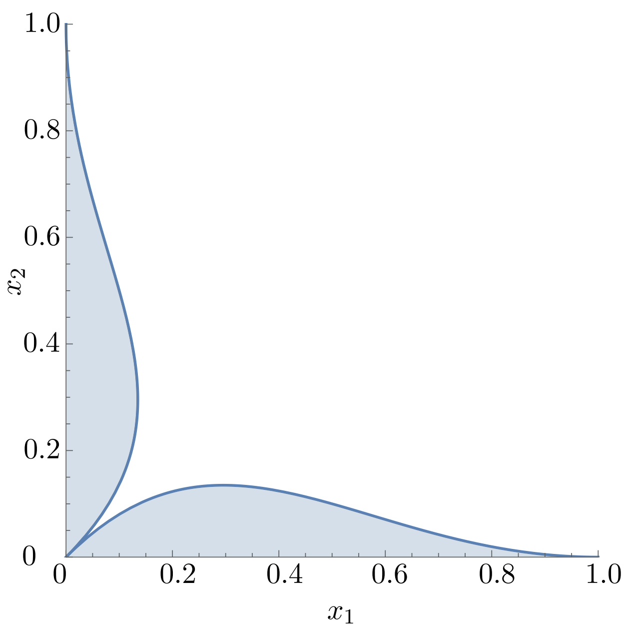

Fig 1. The set containing the set of Nash values in the class of deterministic strategies

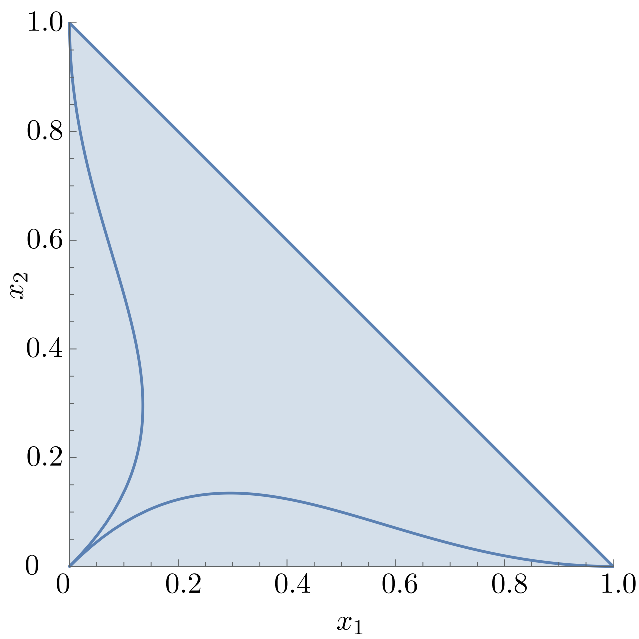

Fig 2. The set of Nash values in the class of randomized strategies

Further, the set $\mathcal{N}^d(0,0,0)$ is contained in the set

$$\{(\sigma_1(x_1,x_2),\sigma_2(x_1,x_2))\colon x_1,x_2 \in [-1,1]\} =$$
$$= \{(x_1,x_2)\colon x_1 \in [0, x_2(1-\sin(2\mathrm{arctg}(x_2)))], x_2 \in [0,1]\} \cup$$
$$\cup \{(x_1,x_2)\colon x_1 \in [0,1],\ x_2 \in [0, x_1(1-\sin(2\mathrm{arctg}(x_1)))]\}.$$

This set is shown in Fig. 1.

Now let us prove that

$$\{0\}\times[0,1],[0,1]\times\{0\}\subset\mathcal{N}^d(0,0,0).$$

We have that, for $x_{1,0}\in\mathbb{R}$,

$$V_1(t_0,x_{1,0},0)=0,\quad V_2(t_0,x_{1,0},0)\leq\min\{|x_{1,0}-(1-t_0)|,0\}. \tag{3.1}$$

Analogously,

$$V_1(t_0,0,x_{2,0})\leq\min\{|x_{2,0}-(1-t_0)|,0\},\quad V_2(t_0,0,x_{2,0})=0. \tag{3.2}$$

Since

$$\sigma_1(r,0)=\sigma_2(0,r)=0,\quad \sigma_1(0,r)=\sigma_2(r,0)=|r| \tag{3.3}$$

we have by (3.1), (3.2) that, for any $r\in[0,1]$,

$$\sigma_i(r,0)\geq V_i(t,rt,0),\quad \sigma_i(0,r)\geq V_i(t,0,rt).$$

Hence, any controls $u=r$, $v=0$ and $u=0$, $v=r$ provide the deterministic Nash equilibrium in the class of punishment strategies. Using (3.3) we get that the sets $[0,1]\times\{0\}$, $\{0\}\times[0,1]$ lies in $\mathcal{N}^d(0,0)$. Consequently, by Corollary 1 we get that $\mathcal{N}^r(0,0)=\{(x_1,x_2):x_1,x_2\in[0,1],x_1+x_2\leq 1\}$. The set $\mathcal{N}^r(0,0,0)$ is presented in Fig. 2.

It is instructive to compare the optimal symmetric Nash values. If the players use deterministic strategies the best symmetric Nash value is $(0,0)$, whereas players using randomized strategies can reach the outcomes equal to $1/2$.

REFERENCES

1. Averboukh Yu. Nash equilibrium for differential games and nonanticipative strategies, *IFAC Proceedings Volumes*, 2011, vol. 44, issue 1, pp. 9340–9342. DOI: 10.3182/20110828-6-IT-1002.00967
2. Averboukh Yu. Nash equilibrium in differential games and the quasi-strategy formalism, *Automation and Remote Control*, 2014, vol. 75, issue 8, pp. 1491–1502. DOI: 10.1134/S000511791408013X
3. Bressan A., Shen W. Semi-cooperative strategies for differential games, *Internat. J. Game Theory*, 2004, vol. 32, issue 4, pp. 561–593. DOI: 10.1007/s001820400180
4. Bressan A., Shen W. Small BV solutions of hyperbolic noncooperative differential games, *SIAM J. Control Optim.*, 2004, vol. 43, issue 1, pp. 194–215. DOI: 10.1137/S0363012903425581
5. Buckdahn R., Cardaliaguet P., Rainer C. Nash equilibrium payoffs for nonzero-sum stochastic differential games, *SIAM J. Control Optim.*, 2004, vol. 43, issue 2, pp. 624–642. DOI: 10.1137/S0363012902411556
6. Case J.H. Towards a theory of many players differential games, *SIAM Journal on Control*, 1969, vol. 7, issue 2, pp. 179–197. DOI: 10.1137/0307013
7. Chistyakov S.V. On noncooperative differential games, *Dokl. Akad. Nauk SSSR*, 1981, vol. 259, pp. 1052–1055 (in Russian).
8. Elliot R.J., Kalton N. The existence of value in differential games, in *Memoir Am. Math. Soc.*, Providence: AMS, 1972, vol. 126.
9. Friedman A. *Differential games*, New York: Wiley, 1971.
10. Kleimenov A.F. *Neantagonisticheskie pozitsionnye differentsial'nye igry* (Non-antagonistic positional differential games), Ekaterinburg: Nauka, 1993.
11. Kononenko A.F. On equilibrium positional strategies in nonantagonistic differential games, *Dokl. Akad. Nauk SSSR*, 1976, vol. 231, pp. 285–288 (in Russian).
12. Kononenko A.F., Chistyakov Yu.E. On equilibrium positional strategies in n-person differential games, *Soviet Math. Dokl.*, 1988, vol. 37, issue 2, pp. 514–517.
13. Nash J. Equilibrium points in n-person games, *Proceedings of the National Academy of Sciences of the United States of America*, 1950, vol. 36, no. 1, pp. 48–49. DOI: 10.1073/pnas.36.1.48
14. Roxin E. Axiomatic approach in differential games, *J. Optimiz. Theory Appl.*, 1969, vol. 3, no. 3, pp. 153–163. DOI: 10.1007/BF00929440
15. Subbotin A.I., Chentsov A.G. *Optimizatsiya garantii v zadachakh upravleniya* (Optimization of the guarantee in control problems), Moscow: Nauka, 1981.

16. Tolwinski B., Haurie A., Leitman G. Cooperate equilibria in differential games, *J. Math. Anal. Appl.*, 1986, vol. 119, pp. 182–202. DOI: 10.1016/0022-247X(86)90152-6
17. Warga J. *Optimal control of differential and functional equations*, New York: Academic Press, 1972.

Received 02.08.2017

Averboukh Yurii Vladimirovich, Candidate of Physics and Mathematics, Senior Researcher, Department of Control Systems, N.N. Krasovskii Institute of Mathematics and Mechanics, Ural Branch of the Russian Academy of Sciences, ul. S. Kovalevskoi, 16, Yekaterinburg, 620990, Russia.
E-mail: ayv@imm.uran.ru

Ю. В. Авербух

Рандомизированное равновесие по Нэшу в дифференциальных играх

Цитата: Вестник Удмуртского университета. Математика. Механика. Компьютерные науки. 2017. Т. 27. Вып. 3. С. 299–308.

Ключевые слова: дифференциальные игры, равновесие по Нэшу, рандомизированные стратегии.

УДК 517.977.8

DOI: 10.20537/vm170301

Работа посвящена исследованию равновесия по Нэшу в неантагонистической детерминированной дифференциальной игре двух лиц в классе рандомизированных стратегий. Предполагается, что игроки информированы об управлении своего партнера, реализовавшегося к текущему времени. Поэтому игра формализуется в классе рандомизированных квазистратегий. В работе получена характеризация множества выигрышей (пар ожидаемых выигрышей игроков) в ситуациях равновесия по Нэшу с использованием вспомогательных антагонистических игр. Показано, что множество выигрышей в ситуациях рандомизированного равновесия по Нэшу является выпуклой оболочкой множества выигрышей в классе детерминированных стратегий. Приведен пример, показывающий дополнительные возможности, которые возникают при переходе к рандомизированным стратегиям.

СПИСОК ЛИТЕРАТУРЫ

1. Averboukh Yu. Nash equilibrium for differential games and nonanticipative strategies // IFAC Proceedings Volumes. 2011. Vol. 44. Issue 1. P. 9340–9342. DOI: 10.3182/20110828-6-IT-1002.00967
2. Авербух Ю.В. Равновесие по Нэшу в дифференциальных играх и формализм квазистратегий // Математическая теория игр и ее приложения. 2012. Т. 4. № 3. С. 3–20.
3. Bressan A., Shen W. Semi-cooperative strategies for differential games // Internat. J. Game Theory. 2004. Vol. 32. Issue 4. P. 561–593. DOI: 10.1007/s001820400180
4. Bressan A., Shen W. Small BV solutions of hyperbolic noncooperative differential games // SIAM J. Control Optim. 2004. Vol. 43. Issue 1. P. 194–215. DOI: 10.1137/S0363012903425581
5. Buckdahn R., Cardaliaguet P., Rainer C. Nash equilibrium payoffs for nonzero-sum stochastic differential games // SIAM J. Control Optim. 2004. Vol. 43. Issue 2. P. 624–642. DOI: 10.1137/S0363012902411556
6. Case J.H. Towards a theory of many players differential games // SIAM Journal on Control. 1969. Vol. 7. Issue 2. P. 179–197. DOI: 10.1137/0307013
7. Чистяков С.В. О бескоалиционных дифференциальных играх // Доклады АН СССР. 1981. Т. 259. № 5. С. 1052–1055.
8. Elliot R.J., Kalton N. The existence of value for differential games // Memoir Am. Math. Soc. Providence: AMS, 1972. Vol. 126.
9. Friedman A. Differential games. New York: Wiley, 1971.
10. Клейменов А.Ф. Неантагонистические позиционные дифференциальные игры. Екатеринбург: Наука, 1993.
11. Кононенко А.Ф. О равновесных позиционных стратегиях в неантагонистических дифференциальных играх // Доклады АН СССР. 1976. Т. 231. № 2. С. 285–288.

12. Кононенко А.Ф., Чистяков Ю.Е. О равновесных позиционных стратегиях в дифференциальных играх многих лиц // Доклады АН СССР. 1988. Т. 299. № 5. С. 1053–1056.
13. Nash J. Equilibrium points in n-person games // Proceedings of the National Academy of Sciences of the United States of America. 1950. Vol. 36. No. 1. P. 48–49. DOI: 10.1073/pnas.36.1.48
14. Roxin E. Axiomatic approach in differential games // J. Optimiz. Theory Appl. 1969. Vol. 3. No. 3. P. 153–163. DOI: 10.1007/BF00929440
15. Субботин А.И., Ченцов А.Г. Оптимизация гарантии в задачах управления. М.: Наука, 1981.
16. Tolwinski B., Haurie A., Leitman G. Cooperate equilibria in differential games // J. Math. Anal. Appl. 1986. Vol. 119. P. 182–202. DOI: 10.1016/0022-247X(86)90152-6
17. Warga J. Optimal control of differential and functional equations. New York: Academic Press, 1972.

Поступила в редакцию 02.08.2017

Авербух Юрий Владимирович, к. ф.-м. н., старший научный сотрудник, отдел управляемых систем, Институт математики и механики им. Н. Н. Красовского УрО РАН, 620990, Россия, г. Екатеринбург, ул. С. Ковалевской, 16.
E-mail: ayv@imm.uran.ru