BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 1(98), 2022, Pages 75–82
ISSN 1024–7696, E-ISSN 2587–4322

Equilibria in Pure Strategies for a Two-Player Zero-Sum Average Stochastic Positional Game

Dmitrii Lozovanu, Stefan Pickl

Abstract. The problem of the existence and determining equilibria in pure stationary strategies for a two-player zero-sum average stochastic positional game is considered. We show that for such a game there exists the value and players may achieve the value by applying pure stationary strategies of choosing the actions in their positions. Based on a constructive proof of these results we propose an algorithmic approach for determining the optimal pure stationary strategies of the players.

Mathematics subject classification: 90C15, 90A20, 91A50.
Keywords and phrases: Positional game, Two-player zero-sum stochastic game, Average payoff, Pure stationary equilibria.

1 Introduction

Average stochastic positional games have been introduced in [5, 6] where some preliminary results concerned with the existence of stationary Nash equilibria have been obtained. This class of games represents a generalization of deterministic positional games with mean payoffs studied by Ehrenfeucht and Mycielski [3], Gurvich et al [4], Zwick and Paterson [12] and Alpern [1]. In [3, 4] mainly two-player zero-sum mean payoff positional games has been studied for which the existence of the value and the optimal positional strategies are proven. Based on these results in [4] algorithms have been proposed for determining the value and the optimal positional strategies of the players in such games. Some possible applications of mean payoff games are describe in [2, 12]. Generalizations of mean payoff games to m-player games, have been considered in [1, 7, 8], however conditions for the existence of Nash equilibria in positional strategies have not been derived. The positional strategies for these dynamic games can be regarded as a pure stationary strategy and therefore Nash equilibria in pure strategies in the general case may not exist. This fact has been shown in [4], where an example of a non-zero-sum mean payoff game of two player for which Nash equilibria in pure stationary strategies does not exist has been constructed. The mean payoff games in mixed stationary strategies have been considered in [10] where the existence of Nash equilibria in mixed stationary strategies has been proved.

In this paper, we study the problem of the existence and determining of equilibria in pure stationary strategies for a two-player zero-sum average stochastic positional game. We show that for such a game there exist equilibria in pure stationary

DOI: https://doi.org/10.56415/basm.y2022.i1.p75

strategies. Based on a constructive proof of this result we propose an approach for determining the optimal pure stationary strategies of the players.

2 Some auxiliary results for an average Markov decision process

First we present the optimality conditions for the problem of determining the optimal stationary strategies in the average Markov decision process defined by the tuple $(X,\ \{A(x)\}_{x\in X},\ \{f(x,a)\}_{x\in X, a\in A(x)},\ p)$, where X is a finite set of states; $A(x)$ is a finite set of actions in $x\in X$; $f(x,a)$ is a reward step in $x \in X$ for $a\in A(x)$ and $p: X\times \prod\limits_{x\in X} A(x)\times X \ \rightarrow\ [0,1]$ is a probability transition function that satisfies the condition $\sum_{y\in X} p^a_{x,y}=1,\ \forall x\in X, a\in A(x)$. In [11] the following theorem is proven.

Theorem 1. *Let a Markov decision process* $(X,\ \{A(x)\}_{x\in X},\ \{f(x,a)\}_{x\in X,a\in A(x)},\ p)$ *be given. Then the system of equations*

$$\varepsilon_x+\omega_x=\max_{a\in A(x)}\Big\{f(x,a)+\sum_{y\in X}p^a_{x,y}\varepsilon_y\Big\},\quad \forall x\in X \tag{1}$$

has a solution under the set of solutions of the system of equations

$$\omega_x=\max_{a\in A(x)}\Big\{\sum_{y\in X}p^a_{x,y}\omega_y\Big\},\qquad \forall x\in X, \tag{2}$$

i.e., the system of equations (2) *has such a solution* $\omega^*_x,\ x\in X$, *for which there exists a solution* $\varepsilon^*_x,\ x\in X$, *of the system of equations*

$$\varepsilon_x+\omega^*_x=\max_{a\in A(x)}\Big\{f(x,a)+\sum_{y\in X}p^a_{x,y}\varepsilon_y\Big\},\quad \forall x\in X. \tag{3}$$

The values ω^*_x *for* $x\in X$ *represent the optimal average rewards for the Markov decision problem when the process starts in the corresponding states* $x\in X$ *and an optimal stationary strategy*

$$s^*: x\ \rightarrow\ a\in A(x)\ for\ x\in X$$

for the average Markov decision problem can be found by fixing $s^*(x)=a^*\in A(x)$ *such that*

$$a^*\in \arg\max_{a\in A(x)}\Big\{\sum_{y\in X}p^a_{x,y}\omega^*_y\Big\}$$

and

$$a^*\in \arg\max_{a\in A(x)}\Big\{f(x,a)+\sum_{y\in X}p^a_{x,y}\varepsilon^*_y\Big\}.$$

The strategy s^* *corresponds to an optimal pure stationary strategy for the average Markov decision problem with an arbitrary starting state* $x\in X$.

In the following we shall use this theorem for the proof of the existence of pure stationary equilibria in a two-player zero-sum average stochastic positional game.

3 Formulation of the zero-sum average stochastic positional game

A two-player zero-sum average stochastic game is determined by a tuple $(X = X_1 \cup X_2,\ \{A(x)\}_{x\in X},\ \{f(x,a)\}_{x\in X,\ a\in A(x)},\ p,\ x_0)$, where X is the set of states of the game, X_1 is the set of positions of first player, X_2 is the set of positions of second player, $A(x)$ is the set of actions in a state $x \in X$, $f(x,a)$ is the step reward in $x \in X$ for a fixed $a \in A(x)$, $p : X \times \cup_{x\in X} A(x) \times X \to [0,1]$ is a transition probability function that satisfies the condition $\sum_{y\in X} p^a_{x,y} = 1,\ \forall x \in X, a \in A(x)$ and x_0 is the starting state of the game.

The game starts at a given initial state x_0 where the player who is owner of this position fixes an action $a_0 \in A(x_0)$. So, if x_0 belongs to the set of positions of the first player then the action $a_0 \in A(x_0)$ in x_0 is chosen by the first player, otherwise the action $a_0 \in A(x_0)$ is chosen by the second one. After that the game passes randomly to a new position according to the probability distribution $\{p^{a_0}_{x_0,y}\}_{y\in X}$. At time moment $t = 1$ the players observe the position $x_1 \in X$. If x_1 belongs to the set of positions of the first player then the action $a_1 \in A(x_1)$ is chosen by the first player, otherwise the action is chosen by the second one and so on, indefinitely. In this process the first player chooses actions in his position set in order to maximize the average reward per transition $\lim_{t\to\infty} \inf \mathsf{E}\big(\frac{1}{t}\sum_{\tau=0}^{t} f(x_\tau, a_\tau)\big)$ while the second one chooses the actions in his position set in order to minimize the average reward per transition $\lim_{t\to\infty} \sup \mathsf{E}\big(\frac{1}{t}\sum_{\tau=0}^{t} f(x_\tau, a_\tau)\big)$. Here E is the expectation operator with respect to the probability measure in the Markov process induced by actions chosen by players in their position sets and fixed starting state x_0. Assuming that players choose actions in their state positions independently we show that for this game there exists a value ω_{x_0} such that the first player has strategy of choosing the actions in his position set that insures $\lim_{t\to\infty} \inf \mathsf{E}\big(\frac{1}{t}\sum_{\tau=0}^{t} f(x_\tau, a_\tau)\big) \ge \omega_{x_0}$ and the second player has strategy of choosing the actions in his position set that insures $\lim_{t\to\infty} \sup \mathsf{E}\big(\frac{1}{t}\sum_{\tau=0}^{t} f(x_\tau, a_\tau)\big) \le \omega_{x_0}$. Moreover, we show that players can achieve the value ω_{x_0} applying pure stationary strategies of selection of the actions in their position sets. We define the pure stationary strategies of the players as two maps

$$s^1 : x \to a \in A(x) \text{ for } x \in X_1; \quad s^2 : x \to a \in A(x) \text{ for } x \in X_2$$

and the sets of pure stationary strategies of the first player and of the second one we denote by $S^1 = \{s^1 |\ s^1 : x \to a \in A(x) \text{ for } x \in X_1\}$, $S^2 = \{s^2 |\ s^2 : x \to a \in A(x) \text{ for } x \in X_1\}$, respectively.

4 Pure Stationary Equilibria in the Game

Let s^1, s^2 be arbitrary pure stationary strategies of the players. Then the profile $s = (s^1, s^2)$ determines a Markov process induced by probability distributions $\{p_{x,y}^{s^i(x)}\}_{y\in X}$ in the states $x \in X_i,\ i = 1,2$ and a given starting state x_0. For this Markov process with step rewards $f(x, s^i(x))$, in the states $x \in X_i, i = 1,2$, we can determine the average reward per transition $\omega_{x_0}(s^1, s^2)$. The function $\omega_{x_0}(s^1, s^2)$ on $S = S^1 \times S^2$ defines an antagonistic game in normal form $\langle S^1, S^2, \omega_{x_0}(s^1, s^2)\rangle$ that in the extended form is determined by the tuple $(X = X_1 \cup X_2,\ \{A(x)\}_{x\in X},\ \{f(x,a)\}_{x\in X,\ a\in A(x)},\ p,\ x_0)$. Taking into account that the strategy sets S^1 and S^2 are finite sets we can regard $\langle S^1, S^2, \omega_{x_0}(s^1, s^2)\rangle$ as a matrix game and therefore for this game there exist the min-max strategies $\overline{s}^1,\ \overline{s}^2$ of the players and the max-min strategies $\overline{\overline{s}}^1,\ \overline{\overline{s}}^2$ of the players for which

$$\omega_{x_0}(\overline{s}^1, \overline{s}^2) = \min_{s^2\in S^2}\max_{s^1\in S^1} \omega_{x_0}(s^1, s^2); \qquad \omega_{x_0}(\overline{\overline{s}}^1, \overline{\overline{s}}^2) = \max_{s^1\in S^1}\min_{s^2\in S^2} \omega_{x_0}(s^1, s^2).$$

In this section we show that for the considered two-player zero-sum average stochastic positional game there exists a pure stationary strategy $s^{1^*} \in S^1$ of the first player and a pure stationary strategy $s^{2^*} \in S^2$ of the second player such that

$$\omega_x(s^{1^*}, s^{2^*}) = \max_{s^1\in S^1}\min_{s^2\in S^2} \omega_x(s^1, s^2) = \min_{s^2\in S^2}\max_{s^1\in S^1} \omega_x(s^1, s^2), \qquad \forall x \in X,$$

i.e we show that (s^{1^*}, s^{2^*}) is a pure stationary equilibrium of the game for an arbitrary starting position $x \in X$, in spite of the fact that the values of the games with different starting positions may be different.

In the following we will consider the game for which it is necessary to determine the optimal stationary strategies of the players for an arbitrary starting state $x \in X$ and we will denote such a game $(X = X_1 \cup X_2,\ \{A(x)\}_{x\in X},\ \{f(x,a)\}_{x\in X,\ a\in A(x)},\ p)$.

First we show that in a two-player zero-sum average stochastic positional game there exists a strategy $\overline{s}^1 \in S^1$ of the first player and a strategy $\overline{s}^2 \in S^2$ of the second player such that $(\overline{s}^1, \overline{s}^2)$ is a max-min strategy of the game for an arbitrary stating position $x \in X$, i. e.

$$\omega_x(\overline{s}^1, \overline{s}^2) = \min_{s^2\in S^2}\max_{s^1\in S^1} \omega_x(s^1, s^2), \qquad \forall x \in X.$$

To prove this we shall use the version of a two-player zero-sum average stochastic positional games in which the starting state is chosen randomly according to a given distribution $\{\theta_x\}$ on X. So, we consider the game in the case when the play starts in a state $x \in X$ with probability $\theta_x > 0$ where $\sum_{x\in X} \theta_x = 1$. We denote this game $(X = X_1 \cup X_2,\ \{A(x)\}_{x\in X},\ \{f(x,a)\}_{x\in X,\ a\in A(x)},\ p,\ \{\theta_x\}_{x\in X})$. This game looks more general, however it can easily be reduced to an auxiliary two-player zero-sum average stochastic positional game with a fixed starting position. Such an auxiliary game is determined by a new tuple obtained from $(X = X_1 \cup X_2,\ \{A(x)\}_{x\in X},\ \{f(x,a)\}_{x\in X,\ a\in A(x)},\ p)$ by adding to the set of positions of the first player a new state position z that has a unique action $a(z)$ for which the probability transitions $p_{z,x}^{a(z)} = \theta_x,\ \forall x \in X$ and the corresponding step reward $f(z, a(z)) = 0$. It is evident that for arbitrary strategies of the players in this game the first player will select in position z the unique action $a(z)$. If for the obtained game with a given starting position z we consider the normal form game in pure stationary strategies $\langle \hat{S}^1, \hat{S}^2, \omega_z(s^1, s^2)\rangle$ then for this game we can determine the min-max strategies of the players $\hat{s}^1, \hat{s}^2$ for which $\hat{\omega}_z(\hat{s}^1, \hat{s}^2) = \sum_{x\in X}\theta_x\omega_x(\overline{s}^1, \overline{s}^2)$. This means that the following lemmas hold.

Lemma 1. *For a two-player zero-sum average stochastic positional game characterized by a tuple* $(X = X_1 \cup X_2,\ \{A(x)\}_{x\in X},\ \{f(x,a)\}_{x\in X,\ a\in A(x)},\ p)$ *there exists a strategy* $\overline{s}^2 \in S^2$ *of the second player and a strategy* $\overline{s}^1 \in S^1$ *of the first player such that* $(\overline{s}^1, \overline{s}^2)$ *is a min-max strategy of the game for an arbitrary starting position* $x \in X$, *i. e.*

$$\omega_x(\overline{s}^1, \overline{s}^2) = \min_{s^2\in S^2}\max_{s^1\in S^1}\omega_x(s^1, s^2), \qquad \forall x \in X.$$

Lemma 2. *For a two-player zero-sum average stochastic positional game determined by a tuple* $(X = X_1 \cup X_2,\ \{A(x)\}_{x\in X},\ \{f(x,a)\}_{x\in X,\ a\in A(x)},\ p)$ *there exists a strategy* $\overline{\overline{s}}^1 \in S^1$ *of first player and a strategy* $\overline{\overline{s}}^2 \in S^2$ *of second player such that* $(\overline{\overline{s}}^1, \overline{\overline{s}}^2)$ *is a max-min strategy of the game for an arbitrary stating position* $x \in X$, *i. e.*

$$\omega_x(\overline{\overline{s}}^1, \overline{\overline{s}}^2) = \max_{s^1\in S^1}\min_{s^2\in S^2}\omega_x(s^1, s^2), \qquad \forall x \in X.$$

Using these lemmas we can prove the following theorem.

Theorem 2. *Let a two-player zero-sum average stochastic positional game determined by the tuple* $(X = X_1 \cup X_2,\ \{A(x)\}_{x\in X},\ \{f(x,a)\}_{x\in X,\ a\in A(x)},\ p)$ *be given. Then the system of equations*

$$\begin{cases} \varepsilon_x + \omega_x = \max\limits_{a\in A(x)}\Big\{f(x,a) + \sum\limits_{y\in X} p^a_{x,y}\varepsilon_y\Big\}, & \forall x \in X_1;\\ \varepsilon_x + \omega_x = \min\limits_{a\in A(x)}\Big\{f(x,a) + \sum\limits_{y\in X} p^a_{x,y}\varepsilon_y\Big\}, & \forall x \in X_2 \end{cases} \tag{4}$$

has a solution under the set of solutions of the system of equations

$$\begin{cases} \omega_x = \max\limits_{a\in A(x)}\Big\{\sum\limits_{y\in X} p^a_{x,y}\omega_y\Big\}, & \forall x \in X_1;\\ \omega_x = \min\limits_{a\in A(x)}\Big\{\sum\limits_{y\in X} p^a_{x,y}\omega_y\Big\}, & \forall x \in X_2, \end{cases} \tag{5}$$

i.e. the system of equations (5) has such a solution $\omega^*_x,\ x \in X$ *for which there exists a solution* $\varepsilon^*_x,\ x \in X$ *of the system of equations*

$$\begin{cases} \varepsilon_x + \omega^*_x = \max\limits_{a\in A(x)}\Big\{f(x,a) + \sum\limits_{y\in X} p^a_{x,y}\varepsilon_y\Big\}, & \forall x \in X_1;\\ \varepsilon_x + \omega^*_x = \min\limits_{a\in A(x)}\Big\{f(x,a) + \sum\limits_{y\in X} p^a_{x,y}\varepsilon_y\Big\}, & \forall x \in X_2. \end{cases}$$

The optimal pure stationary strategies s^{1^*}, s^{2^*} *of the players can be found by fixing arbitrary maps* $s^{1^*}(x) \in A(x)$ *for* $x \in X_1$ *and* $s^{2^*}(x) \in A(x)$ *for* $x \in X_2$ *such that*

$$s^{1*}(x) \in \left\{ Arg \max_{a\in A(x)} \left\{ \sum_{y\in X} p^a_{x,y}\omega^*_y \right\}\right\} \bigcap \left\{ Arg \max_{a\in A(x)} \left\{ f(x,a) + \sum_{y\in X} p^a_{x,y}\varepsilon^*_y \right\}\right\},\ x\in X_1,$$

$$s^{2*}(x) \in \left\{ Arg \min_{a\in A(x)} \left\{ \sum_{y\in X} p^a_{x,y}\omega^*_y \right\}\right\} \bigcap \left\{ Arg \min_{a\in A(x)} \left\{ f(x,a) + \sum_{y\in X} p^a_{x,y}\varepsilon^*_y \right\}\right\},\ x\in X_2$$

and $\omega_x(s^{1*}, s^{2*}) = \omega^*_x,\ \forall x \in X,$ *i.e.*

$$\omega_x(s^{1*}, s^{2*}) = \max_{s^1\in S^1} \min_{s^2\in S^2} \omega_x(s^1, s^2) = \min_{s^2\in S^2} \max_{s^1\in S^1} \omega_x(s^1, s^2), \qquad \forall x \in X.$$

Proof. According to Lemma 1 for the players in the considered game there exist the pure stationary strategies $\overline{s}^1 \in S^1,\ \overline{s}^2 \in S^2$ for which

$$\omega_x(\overline{s}^1, \overline{s}^2) = \min_{s^2\in S^2} \max_{s^1\in S^1} \omega_x(s^1, s^2), \quad \forall x \in X.$$

We show that

$$\omega_x(\overline{s}^1, \overline{s}^2) = \max_{s^1\in S^1} \min_{s^2\in S^2} \omega_x(s^1, s^2),, \quad \forall x \in X,$$

i.e. we show that $\overline{s}^1 = s^{1*},\ \overline{s}^2 = s^{2*}$.

Indeed, if we consider the Markov process induced by strategies $\overline{s}^1, \overline{s}^2$ then according to Theorem 1 for this process the system of linear equations

$$\begin{cases} \varepsilon_x + \omega_x = f(x,a) + \sum\limits_{y\in X} p^a_{x,y}\varepsilon_y, & \forall x \in X_1,\ a = \overline{s}^1(x); \\ \varepsilon_x + \omega_x = f(x,a) + \sum\limits_{y\in X} p^a_{x,y}\varepsilon_y, & \forall x \in X_2,\ a = \overline{s}^2(x); \\ \qquad \omega_x = \sum\limits_{y\in X} p^a_{x,y}\omega_y, & \forall x \in X_1,\ a = \overline{s}^1(x); \\ \qquad \omega_x = \sum\limits_{y\in X} p^a_{x,y}\omega_y, & \forall x \in X_2,\ a = \overline{s}^2(x) \end{cases} \tag{6}$$

has a basic solution $\varepsilon^*_x,\ \omega^*_x\,(x \in X)$. Now if we assume that in the game only the second payer fixes his strategy $\overline{s}^2 \in S^2$ then we obtain a Markov decision problem with respect to the first player and therefore according to Theorem 1 for this decision problem the system of linear equations

$$\begin{cases} \varepsilon_x + \omega_x \geq f(x,a) + \sum\limits_{y\in X} p^a_{x,y}\varepsilon_y, & \forall x \in X_1,\ a \in A(x); \\ \varepsilon_x + \omega_x = f(x,a) + \sum\limits_{y\in X} p^a_{x,y}\varepsilon_y, & \forall x \in X_2,\ a = \overline{s}^2(x); \\ \qquad \omega_x \geq \sum\limits_{y\in X} p^a_{x,y}\omega_y, & \forall x \in X_1,\ a \in A(x); \\ \qquad \omega_x = \sum\limits_{y\in X} p^a_{x,y}\omega_y, & \forall x \in X_2,\ a = \overline{s}^2(x) \end{cases}$$

has solutions. We can observe that $\epsilon^*_x,\ \omega^*_x\,(x \in X)$ represents a solution of this system and $\omega_x(\overline{s}^1, \overline{s}^2) = \omega^*_x,\ \forall x \in X$.

Taking into account that $\omega_x(\overline{s}^1,\overline{s}^2) = \min_{s^2\in S^2} F_x(\overline{s}^1, s^2)$ then for a fixed strategy $\overline{s}^1 \in S^1$ the following system has solutions

$$\begin{cases} \varepsilon_x + \omega_x = f(x,a) + \sum\limits_{y\in X} p^a_{x,y}\varepsilon_y, & \forall x \in X_1,\ a = \overline{s}^1(x); \\ \varepsilon_x + \omega_x \le f(x,a) + \sum\limits_{y\in X} p^a_{x,y}\varepsilon_y, & \forall x \in X_2,\ a \in A(x); \\ \omega_x = \sum\limits_{y\in X} p^a_{x,y}\omega_y, & \forall x \in X_1,\ a = \overline{s}^1(x); \\ \omega_x \le \sum\limits_{y\in X} p^a_{x,y}\omega_y, & \forall x \in X_2,\ a \in A(x). \end{cases}$$

and $\epsilon_x = \epsilon^*_x$, $\omega_x = \omega^*_x\ (x \in X)$ represents a solution of this system. This means that the following system

$$\begin{cases} \varepsilon_x + \omega_x \ge f(x,a) + \sum\limits_{y\in X} p^a_{x,y}\varepsilon_y, & \forall x \in X_1,\ a \in A(x); \\ \varepsilon_x + \omega_x \le f(x,a) + \sum\limits_{y\in X} p^a_{x,y}\varepsilon_y, & \forall x \in X_2,\ a \in A(x); \\ \omega_x \ge \sum\limits_{y\in X} p^a_{x,y}\omega_y, & \forall x \in X_1,\ a \in A(x); \\ \omega_x \le \sum\limits_{y\in X} p^a_{x,y}\omega_y, & \forall x \in X_2,\ a \in A(x) \end{cases}$$

has a solution which satisfies condition (6). Thus, we obtain that $\overline{s}^1 = s^{1^*},\ \overline{s}^2 = s^{2^*}$ and $\omega_x(s^{1^*}, s^{2^*}) = \omega^*_x, \forall x \in X$, i.e.

$$\omega_{\overline{x}}(s^{1^*}, s^{2^*}) = \max_{s^1\in S^1}\min_{s^2\in S^2} \omega_{\overline{x}}(s^1, s^2) = \min_{s^2\in S^2}\max_{s^1\in S^1} \omega_{\overline{x}}(s^1, s^2),\ \forall \overline{x} \in X.$$

So, the theorem holds. □

The formulation of Theorem 2 has been mentioned also in [9], however in [9] the full proof of this theorem is not presented. The obtained saddle point conditions for zero-sum stochastic games generalize the saddle point condition for deterministic average positional games from [3,4]. Based on Theorem 2 we may conclude that the optimal strategies of the players in the considered game can be found if we determine a solution of equations (4), (5). A solution of these equations can be determined using iterative algorithms like algorithms for determining the optimal solutions of an average Markov decision problem [11].

5 Conclusion

Two-player zero-sum games are an important class of average stochastic games that generalizes the deterministic positional games with mean payoffs from [3,4,7]. For such games there exists the value and the optimal pure stationary strategies of the players and these strategies can be found on the basis of Theorem 2.

Acknowledgement: This research was supported by the State Program of the Republic of Moldova, project 20.80009.5007.13 ”Deterministic and stochastic methods for solving optimization and control problems”.

References

[1] Alpern S. *Cycles in extensive form perfect information games.* J. Math. Anal. Appl., 159, 1-17, 1991.

[2] Codon A. *The complexity of stochastic games.* Inf. Comput, 96 (2), 203-224, 1992.

[3] Ehrenfeucht A., Mycielski J. *Positional strategies for mean payoff games.* Int. J. Game Theory, 8, 109-113,1979.

[4] Gurvich V., Karzaniv A., Khachyan L. *Cyclic games and an algorithm to find minimax mean cycles in directed graphs.* USSR Comput. Math. Math. Phys., 28, 85-91, 1988.

[5] Lozovanu D. *The game theoretical approach to Markov decision problems and determining Nash equilibria for stochastic positional games.* Int. J. Mathematical Modelling and Numerical Optimization. 2 (2), 162-174 (2011).

[6] Lozovanu, D. *Stationary Nash equilibria for average stochastic positional games.* In: Petrosyan et al (eds), Frontiers of dynamic games, Static and Dynamic Games Theory: Fondation and Applications, Birkhäuser, 139-163, 2018.

[7] Lozovanu D., Pickl S. *Nash equilibria conditions for cyclic games with p players.* Electron. Notes in Discrete Math, 25, 117-124, 2006.

[8] Lozovanu D., Pickl S. *Optimization and Multiobjective Control of Time-Discrete Systems.* Springer, 2009.

[9] Lozovanu D., Pickl S. *Determining the optimal strategies for zero-sum average stochastic positional games.* Electron. Notes Discrete Math., 55, 155-159, 2016.

[10] Lozovanu D., Pickl S. *Nash equilibria in Mixed Stationary Strategies for m-player mean payoff games on networks.* Contribution Game Theory Manag. 11, 103-112, 2018.

[11] Puterman M. *Markov Decision Processes: Discrete Dynamic Programming.* Wiley, 2005.

[12] Zwick U., Paterson M. *The complexity of mean payoff games on graphs.* Theoretical Computer Science, 158, 343-359, 1996.

Received February 16, 2022

Dmitrii Lozovanu
Institute of Mathematics and Computer Science,
5 Academiei str., Chişinău, MD−2028, Moldova
E-mail: *lozovanu@math.md*

Stefan Pickl
Institute for Theoretical Computer Science,
Mathematics and Operations Research,
Universität der Bundeswehr,München,
85577 Neubiberg-München, Germany,
E-mail: *stefan.pickl@unibw.de*