

A. S. NOWAK (Wrocław)

ON APPROXIMATIONS OF NONZERO-SUM
UNIFORMLY CONTINUOUS
ERGODIC STOCHASTIC GAMES

Abstract. We consider a class of uniformly ergodic nonzero-sum stochastic games with the expected average payoff criterion, a separable metric state space and compact metric action spaces. We assume that the payoff and transition probability functions are uniformly continuous. Our aim is to prove the existence of stationary ε -equilibria for that class of ergodic stochastic games. This theorem extends to a much wider class of stochastic games a result proven recently by Bielecki [2].

1. Introduction. The existence of stationary equilibria for discounted or limiting average payoff stochastic games with uncountable state spaces is still an open problem. Positive answers to this question are known only for some special classes of stochastic games where the transition probabilities satisfy certain very specific conditions. Some results concerning correlated equilibria are also available. For a good survey of the existing literature the reader is referred to [11, 12].

The existence of stationary ε -equilibria for discounted stochastic games was first proved by Whitt [13] who considered a separable metric state space model in which the daily payoff and transition probability functions are uniformly continuous. His proof is based on an approximation technique of the original model by “smaller” models with countably many states for which stationary equilibria are known to exist. Another ε -equilibrium theorem for a class of discounted stochastic games with an abstract measurable state space and the transition probability having a density function is included in

1991 *Mathematics Subject Classification*: Primary 90D10, 90D20; Secondary 90D05, 93E05.

Key words and phrases: nonzero-sum Markov game, Nash equilibrium, long run average reward criterion, general state space.

This research was supported by KBN grant 2PO3A 05810.

our paper [9]. The proof in [9] is based on a completely different approximation method compared with Whitt's technique [13]. (That is because the state space need not be metric.) Some extensions of the result in [9] are given in our recent paper with Altman [11] where unbounded (in the state variable) daily payoff functions are allowable and the expected average payoff criterion is also considered.

Bielecki [2] extended Whitt's theorem [13] to a class of limiting average payoff stochastic games satisfying a very strong ergodicity condition. Stochastic games studied in [2] are in some sense equivalent to discounted stochastic games [2, 3]. In this paper, we consider uniformly continuous stochastic games as studied by Whitt [13] and Bielecki [2] but under a much weaker ergodicity assumption. Using an extension of Ueno's inequality given in [10], we prove that any stochastic game with uniformly ergodic transition structure can be approximated (in some sense) by games studied by Bielecki [2]. This result combined with those of Bielecki [2] and Whitt [13] implies a new existence theorem for stationary ε -equilibria for nonzero-sum stochastic games with a separable metric state space.

2. The model and results. Let Y be a Borel space, i.e., a non-empty Borel subset of a complete separable metric space. Let $\mathcal{B}(Y)$ be the family of all Borel subsets of Y .

An N -person nonzero-sum stochastic game is defined by the following objects:

- S is a Borel space of states for the game,
- X_k is a non-empty compact metric space of actions for player k . We put $X = X_1 \times \dots \times X_N$,
- $r_k : S \times X \rightarrow \mathbb{R}$ is a bounded Borel measurable payoff function for player k .
- q is a Borel measurable transition probability from $S \times X$ to S , called the law of motion among states. If s is a state at some stage of the game and the players select an $x \in X$, then $q(\cdot | s, x)$ is the probability distribution of the next state of the game.

In this paper, we adopt the following conditions:

C1: The functions r_k are uniformly continuous on $S \times X$.

C2: The function $q(B|s, x)$ is uniformly continuous on $S \times X$, uniformly in $B \in \mathcal{B}(S)$.

The game is played in discrete time with past history as common knowledge for all the players. Suppose that at the beginning of each period n of the game the players are informed of the outcome of the preceding period and the current state s_n . Then the information available to them at time n is

a vector $h^n = (s_1, x^1, \dots, s_{n-1}, x^{n-1}, s_n)$, where $s_i \in S$, $x^i = (x_1^i, \dots, x_N^i) \in X$. We denote the set of such vectors by H^n and assume that it is endowed with the Borel (or equivalently product) σ -algebra.

A *strategy* for player k is a sequence $\pi_k = (\pi_k^1, \pi_k^2, \dots)$, where every π_k^n is a Borel measurable transition probability from H^n into X_k . A *stationary strategy* for player k is a strategy $\pi_k = (\pi_k^1, \pi_k^2, \dots)$ such that each π_k^n depends on the current state s_n only. In other words, a strategy π_k for player k is called stationary if there exists a transition probability f_k from S into X_k such that for every period n of the game and each history $h^n \in H^n$, we have $\pi_k^n(\cdot | h^n) = f_k(\cdot | s_n)$. We put $\Pi_k(F_k)$ to denote the set of all strategies (stationary strategies) for player k .

Let $H = S \times X \times S \times \dots$ be the space of all infinite histories of the game, endowed with the product σ -algebra. For any multi-strategy $\pi = (\pi_1, \dots, \pi_N)$ of the players and every initial state $s_1 = s \in S$, a probability measure P_s^π and a stochastic process $\{\sigma_n, \alpha_n\}$ are defined on H in a canonical way, where the random variables σ_n and α_n describe the state and the actions chosen by the players, respectively, on the n th stage of the game (cf. Chapter 7 of [1]). Thus, for each multi-strategy $\pi = (\pi_1, \dots, \pi_N)$, any finite horizon T , and every initial state $s \in S$, the *expected T -stage payoff* to player k is

$$J_k^T(s, \pi) = E_s^\pi \left(\sum_{n=1}^T r_k(\sigma_n, \alpha_n) \right).$$

Here E_s^π is the expectation operator with respect to the probability measure P_s^π .

The expected *average payoff per unit time* for player k is defined as

$$J_k(s, \pi) = \limsup_{T \rightarrow \infty} \frac{1}{T} J_k^T(s, \pi).$$

Let $\pi^* = (\pi_1^*, \dots, \pi_N^*)$ be a fixed multi-strategy for the players. For any strategy π_k of player k , we write (π_{-k}^*, π_k) to denote the multi-strategy obtained from π^* by replacing π_k^* with π_k .

Let $\varepsilon \geq 0$. A multi-strategy $\pi^* = (\pi_1^*, \dots, \pi_N^*)$ is called an ε -*equilibrium* for the average payoff stochastic game if and only if

$$\varepsilon + J_k(s, \pi^*) \geq J_k(s, (\pi_{-k}^*, \pi_k))$$

for every player k and $\pi_k \in \Pi_k$. A 0-equilibrium is called a *Nash equilibrium*.

Before we formulate further assumptions and state our results, we adopt some useful notation. Let F be the set of all stationary multi-strategies of the players. For any $f = (f_1, \dots, f_N) \in F$,

$$r_k(s, f) = \int_{X_1} \dots \int_{X_N} r_k(s, x_1, \dots, x_N) f_1(dx_1 | s) \dots f_N(dx_N | s)$$

and

$$q(B | s, f) = \int_{X_1} \dots \int_{X_N} q(B | s, x_1, \dots, x_N) f_1(dx_1 | s) \dots f_N(dx_N | s)$$

for any $B \in \mathcal{B}(S)$.

The basic assumption made in this paper is:

C3 (*Uniform geometric ergodicity*): There exist scalars $\alpha \in (0, 1)$ and $\gamma > 0$ for which the following holds: for any $f \in F$, there exists a probability measure p_f on $\mathcal{B}(S)$ such that

$$\|q^n(\cdot | s, f) - p_f(\cdot)\| \leq \gamma \alpha^n \quad \text{for each } n \geq 1.$$

Here $q^n(\cdot | s, f)$ is the n -step transition probability of the Markov chain induced by q and $f \in F$, $p_f(\cdot)$ is the unique invariant probability distribution of this Markov chain, and $\|\cdot\|$ is the total variation norm in the space of finite signed measures on $\mathcal{B}(S)$.

It is well known that **C3** is equivalent to the following condition:

C4: There exist a positive integer m and a positive number $c < 1$ such that

$$\|q^m(\cdot | s, f) - q^m(\cdot | z, f)\| \leq 2c \quad \text{for all } s, z \in S \text{ and } f \in F.$$

Clearly **C3** implies **C4**. **C4** implies **C3** by Ueno's inequality [10]. For details consult pages 275 and 276 of [4]. Condition **C3** (or **C4**) is rather difficult to check. Much easier to verify are assumptions implying **C3** which can be formulated on the basis of a recent result by Meyn and Tweedie (see Theorem 2.3 of [7]).

C5 (*Drift inequality*): There exist a bounded Borel measurable function $w : S \rightarrow [1, \infty)$ and a set $C \in \mathcal{B}(S)$ such that for some $\lambda \in (0, 1)$ and $\eta > 0$, we have

$$\int_S w(t) q(dt | s, x) \leq \lambda w(s) + \eta 1_C(s)$$

for each $(s, x) \in S \times X$. Here 1_C is the characteristic function of the set C .

C6: There exist $b \in (0, 1)$ and a probability measure p concentrated on the Borel set C such that

$$q(D | s, x) \geq bp(D)$$

for each $D \in \mathcal{B}(C)$, $x \in X$ and $s \in C$.

For a further discussion of uniform geometric ergodicity of Markov chains consult Theorem 16.0.2 in [6].

In a recent paper Bielecki [2] considered stochastic games under the following strong assumption:

M (*Minorization property*): There exist a constant $b \in (0, 1)$ and a probability measure p on $\mathcal{B}(S)$ such that

$$q(D | s, x) \geq bp(D)$$

for every $D \in \mathcal{B}(S)$, $x \in X$ and $s \in S$.

It is easy to prove that **M** implies **C3** (see page 185 in Neveu [8]). Note that **M** is much stronger than **C6** and if it is satisfied then **C5** holds trivially with $\eta = \sup_{s \in S} w(s)$.

Let \mathcal{G} (resp. \mathcal{G}_M) be the class of nonzero-sum stochastic games satisfying condition **C3** (resp. **M**). Every game $G \in \mathcal{G}_M$ is said to satisfy a *strong ergodicity condition*. Using some ideas from Chapter 7 of [3] Bielecki showed that the problem of finding a stationary ε -equilibrium in any game $G \in \mathcal{G}_M$ which satisfies **C1** and **C2** can be reduced to finding a stationary $\varepsilon\beta$ -equilibrium in some β -discounted stochastic game with a new transition probability and $\beta = 1 - b < 1$ (see Theorem 2.1 of [2]). In other words, stationary ε -equilibria can be constructed in any game $G \in \mathcal{G}_M$ by using the approximation technique developed for discounted stochastic games by Whitt [13]. In this paper, we prove that any game $G \in \mathcal{G}$ can be in some sense approximated by games from the class \mathcal{G}_M .

Before we state our results we point out that for any multi-strategy $f \in F$ the expected average payoff $J_k(s, f)$ to player k is under condition **C3** independent of the initial state $s \in S$ and therefore it will be denoted by $J_k(f)$. Clearly, **C3** implies that

$$J_k(f) = \int_S r_k(s, f) p_f(ds).$$

Our approximation result is formulated as follows:

THEOREM 1. *For any $\varepsilon > 0$ and any game $G \in \mathcal{G}$ there exists a game $G^\varepsilon \in \mathcal{G}_M$ such that*

$$\max_{1 \leq k \leq N} \sup_{f \in F} |J_k(f) - J_k^\varepsilon(f)| \leq \varepsilon/2$$

where $J_k^\varepsilon(f)$ is the expected average payoff to player k in G^ε , $f \in F$.

PROOF. Assume that $G \in \mathcal{G}$. Then **C4** holds with some positive integer m and $c \in (0, 1)$. Choose any probability measure p on $\mathcal{B}(S)$. For any $\delta \in (0, 1)$ define

$$q_\delta(\cdot | s, f) = (1 - \delta)q(\cdot | s, f) + \delta p(\cdot)$$

where $s \in S$ and $f \in F$. Clearly, q_δ satisfies condition **M** for any $\delta \in (0, 1)$. It is easy to show that

$$(1) \quad q_\delta^m(\cdot | s, f) = (1 - \delta)^m q_\delta^m(\cdot | s, f) + \sum_{n=0}^{m-1} \delta(1 - \delta)^n p q_f^n(\cdot)$$

where $pq_f^0 := p$ and

$$pq_f^n(\cdot) = \int_S q(\cdot | s, f) p(ds)$$

for $n = 1, \dots, m - 1$. Note that

$$(1 - \delta)^m + \sum_{n=0}^{m-1} \delta(1 - \delta)^n = 1.$$

Using this and (1), we obtain

$$\begin{aligned} \|q^m(\cdot | s, f) - q_\delta^m(\cdot | s, f)\| &\leq \sum_{n=0}^{m-1} \delta(1 - \delta)^n \|pq_f^n(\cdot) - q^m(\cdot | s, f)\| \\ &\leq 2 \sum_{n=0}^{m-1} \delta(1 - \delta)^n = 2[1 - (1 - \delta)^m] \end{aligned}$$

for every $s \in S$ and $f \in F$. Put $\xi = 2[1 - (1 - \delta)^m]$. By the Theorem of [10], we have

$$(2) \quad \|q^{mn}(\cdot | s, f) - q_\delta^{mn}(\cdot | s, f)\| \leq \xi(1 + c + \dots + c^{n-1}) + 2c^n$$

for each $s \in S, f \in F$ and $n \geq 1$. Hence

$$(3) \quad \|p_f - p_f^\delta\| \leq \frac{\xi}{1 - c}$$

for each $f \in F$, where p_f^δ is the invariant probability distribution for the Markov chain induced by q_δ and f .

Put

$$K = \max_{1 \leq k \leq N} \sup_{s \in S, x \in X} |r_k(s, x)|.$$

Using (3) we obtain

$$\begin{aligned} |J_k(f) - J_k^\delta(f)| &\leq \left| \int_S r_k(s, f) p_f(ds) - \int_S r_k(s, f) p_f^\delta(ds) \right| \\ &\leq K \|p_f - p_f^\delta\| \leq K \frac{\xi}{1 - c} \end{aligned}$$

for all $f \in F$. Let $\varepsilon > 0$ be fixed. Choose any $\delta_\varepsilon \in (0, 1)$ with $K\xi/(1 - c) = K[1 - (1 - \delta_\varepsilon)^m]/(1 - c) \leq \varepsilon/2$. The game $G^\varepsilon \in \mathcal{G}_M$ we are looking for is simply the stochastic game with the transition probability q_{δ_ε} . ■

THEOREM 2. *Any nonzero-sum stochastic game satisfying conditions **C1–C3** has a stationary ε -equilibrium for any $\varepsilon > 0$.*

PROOF. Let $\varepsilon > 0$. Consider a game that satisfies **C1–C3**. By Corollary 2.1 of [2] and Theorem 1 there exists some $f^* \in F$ such that

$$(4) \quad \varepsilon + J_k(s, f^*) \geq J_k(s, (f_{-k}^*, f_k))$$

for every player k and any stationary strategy $f_k \in F_k$. From standard results in dynamic programming [5], it follows that

$$(5) \quad \sup_{f_k \in F_k} J_k(s, (f_{-k}^*, f_k)) = \sup_{\pi_k \in \Pi_k} J_k(s, (f_{-k}^*, \pi_k))$$

for every player k . Clearly, (4) and (5) complete the proof. ■

3. Concluding remarks. In this paper, we obtain stationary ε -equilibria for uniformly continuous ergodic nonzero-sum stochastic games with a separable metric state space using a simple approximation of any such game by stochastic games (from the class \mathcal{G}_M) studied in [2] with strongly ergodic transition structure. Our result is based on an extension of Ueno's inequality given in [10]. We point out that Whitt's approximation technique [13] (which is in fact applied in [2]) makes use of the separability of the state space. The approach taken in [9] and further extended in [11] is based on a different idea. In [11] we make weaker ergodicity assumptions than **C1**, allow for unbounded daily payoff functions but assume that the transition probability has a density function. In [11], we approximate the original game by games with countably many states having analogous ergodic properties. In the present case, we approximate uniformly ergodic stochastic games by games having the same state space but satisfying a much stronger ergodicity condition.

References

- [1] D. P. Bertsekas and S. E. Shreve, *Stochastic Optimal Control: The Discrete Time Case*, Academic Press, New York, 1979.
- [2] T. R. Bielecki, *Approximations of dynamic Nash games with general state and action spaces and ergodic costs for the players*, Appl. Math. (Warsaw) 24 (1996), 195–202.
- [3] E. B. Dynkin and A. A. Yushkevich, *Controlled Markov Processes*, Springer, New York, 1979.
- [4] J. P. Georgin, *Contrôle de chaînes de Markov sur des espaces arbitraires*, Ann. Inst. H. Poincaré Sér. B 14 (1978), 255–277.
- [5] O. Hernández-Lerma and J. B. Lasserre, *Discrete Time Markov Control Processes: Basic Optimality Criteria*, Springer, New York, 1996.
- [6] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer, New York, 1993.
- [7] —, —, *Computable bounds for geometric convergence rates of Markov chains*, Ann. Appl. Probab. 4 (1994), 981–1011.
- [8] J. Neveu, *Mathematical Foundations of the Calculus of Probability*, Holden-Day, San Francisco, 1965.
- [9] A. S. Nowak, *Existence of equilibrium stationary strategies in discounted noncooperative stochastic games with uncountable state space*, J. Optim. Theory Appl. 45 (1985), 591–602.

- [10] A. S. Nowak, *A generalization of Ueno's inequality for n -step transition probabilities*, Appl. Math. (Warsaw) 25 (1998), 295-299.
- [11] A. S. Nowak and E. Altman, *ε -Nash equilibria for stochastic games with uncountable state space and unbounded costs*, technical report, Inst. Math., Wrocław Univ. of Technology, 1998 (submitted).
- [12] A. S. Nowak and K. Szajowski, *Nonzero-sum stochastic games*, Ann. Dynamic Games 1999 (to appear).
- [13] W. Whitt, *Representation and approximation of noncooperative sequential games*, SIAM J. Control Optim. 18 (1980), 33-48.

Andrzej S. Nowak
Institute of Mathematics
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland
E-mail: nowak@im.pwr.wroc.pl

*Received on 23.12.1998;
revised version on 13.1.1999*