Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Zero-Sum Ergodic Stochastic Games

Anna Jaśkiewicz and Andrzej S. Nowak

Abstract— **Zero-sum stochastic games with Borel state spaces satisfying a generalized geometric ergodicity condition are considered. The main objective of this paper is to establish a minimax theorem for a class of ergodic stochastic games with the Feller transition probabilities.**

I. INTRODUCTION

This paper deals with zero-sum Borel state space stochastic games under the average payoff criterion and are based on [14]. We make certain stochastic stability assumptions on the transition structure of the game which imply the so-called V-uniform geometric ergodicity of the Markov chains governed by stationary strategies of the players. Such conditions are inspired by the works of Kartashov [15], Meyn and Tweedie [18], [19], and some recent results concerning Markov control processes, see [7], [8] and the references cited therein. Other related papers on stochastic games are [9], [11], [12], [17], [23], [27]. A common feature of the aforementioned works is the assumption that the transition probabilities are strongly continuous in actions of the players. Such a restriction is often not satisfied in applications. Weakly continuous transition probabilities are in some situations (e.g. dynamic programming problems or stochastic games) much more natural, see [6] for a relevant example. The aim of this paper is to establish a new minimax theorem for a class of ergodic zero-sum stochastic games with the Borel state space and Feller transition (i.e. weakly continuous) probability function. To our best knowledge no results of this type are known in the existing literature. Nevertheless, one can refer to [13] for a related result in the context of dynamic programming (one person games). Futhermore, we also study the so-called optimality or Shapley equation for the games under consideration. We show that stationary strategies obtained from this equation are strong average optimal. The proof makes use of the vanishing discount factor approach. In order to overcome some difficulties (compared with earlier papers) we need to use Michael's theorem on continuous selections [20] and a version of Serfozo's extension of Fatou's lemma for varying probability measures.

The ergodicity assumptions which are made in our paper are crucial and may be regarded as quite restrictive. However, they allow us to prove the existence of optimal (or even strong optimal) strategies for a very large class of stochastic games. It is worth mentioning that many optimization models and dynamic games in engineering and economics are ergodic, [3], [5], [16], [18].

A. Jaśkiewicz is with the Institute of Mathematics, Wrocław University of Technology, 50-370 Wrocław, Poland ajaskiew@im.pwr.wroc.pl

A.S. Nowak is with the Faculty of Mathematics, Computer Science and Econometrics, Zielona Góra University, 65-246 Zielona Góra, Poland a.nowak@wmie.uz.zgora.pl

II. MODEL AND ASSUMPTIONS

Let Y be a Borel space, i.e., a non-empty Borel subset of a complete separable metric space. By $\mathcal{B}(Y)$, we denote the σ-algebra of all Borel subsets of Y. Let $P(Y)$ be the space of all probability measures on $\mathcal{B}(Y)$, endowed with the weak topology (see p. 124 in [2]). This topology, for any Borel space Y, can be characterized in terms of convergent sequences (Proposition 7.21 in [2]). Namely, a sequence $\{p_n\}$ converges to some $p \in P(Y)$ in the weak topology if and only if

$$\int_Y u(y)p_n(dy) \to \int_Y u(y)p(dy)$$

for every $u \in C(Y)$. Here $C(Y)$ denotes the space of all bounded continuous functions on Y, endowed with the supremum metric.

For the reader's convenience, we recall some basic results which we shall be using later on. If Y is compact, then $P(Y)$ is compact as well (see Proposition 7.22 in [2]). If Y is Borel, then $P(Y)$ is Borel too (see Corollary 7.25.1 in [2]).

Let X and Y be Borel spaces. By a Borel measurable transition probability from X to Y we mean a function $\phi : \mathcal{B}(Y) \times X \mapsto [0, 1]$ such that, for each $B \in \mathcal{B}(Y)$, $\phi(B|\cdot)$ is a Borel measurable function on X, and $\phi(\cdot|x) \in P(Y)$ for each $x \in X$. It is well-known that every Borel measurable mapping $g : X \mapsto P(Y)$ induces a transition probability by setting $\phi(\cdot|x) := g(x)(\cdot)$ (see Proposition 7.25 in [2]). If g is continuous (with respect to the weak topology on $P(Y)$), then the corresponding transition probability is called *weakly continuous* or *Feller*.

Let $M : X \mapsto \mathcal{B}(Y)$ be a set-valued mapping. For any $D \subset Y$, define

$$M^{-1}(D) := \{x \in X : M(x) \cap D \neq \emptyset\}.$$

If $M^{-1}(D)$ is a closed [an open] subset in X for each closed [open] subset D of Y, then M is said to be upper [lower] semicontinuous. A mapping $M : X \mapsto \mathcal{B}(Y)$ is called continuous if it is both lower and upper semicontinuous. For a further discussion of semicontinuous set-valued mappings consult [25].

Let $\Phi : X \mapsto \mathcal{B}(P(Y))$ be the set-valued mapping defined by $\Phi(x) := P(M(x))$, $x \in X$.

Lemma 1.1: (Theorem 3 in [10]). If M is non-empty compact valued and continuous, then Φ is also continuous and compact-valued.

0-7803-9568-9/05/$20.00 ©2005 IEEE

The next lemma concerns continuous selections for lower semicontinuous set-valued mappings and is essentially due to Michael [20].

Lemma 1.2. Assume that M is non-empty compact-valued and continuous. Let $x_0 \in X$ and $\nu_0 \in \Phi(x_0)$. Then Φ admits a continuous selection whose graph contains the point (x_0, ν_0), that is, there exists a continuous mapping $\phi : X \mapsto P(Y)$ such that $\phi(x) \in \Phi(x)$ for each $x \in X$ and $\phi(x_0) = \nu_0$.

Note that $\Psi(x) := \Phi(x)$ for $x \neq x_0$ and $\Psi(x_0) := \{\nu_0\}$ is lower semicontinuous and compact convex valued. The proof of Lemma 1.2 relies on showing that Ψ has a continuous selection and follows along the same lines as that of Theorem (1.5)* in [25], which in turn is a modification of Michael's proof (see [20]) given for lower semicontinuous set-valued mappings with closed convex values in a Banach space.

A *zero-sum stochastic game* is described by the following objects:

(i) S is the *set of states* for the game and is assumed to be a Borel space.

(ii) A and B are the *action spaces* for players 1 and 2, respectively, and are also assumed to be Borel spaces.

(iii) $A(s) \subset A$ and $B(s) \subset B$ are non-empty compact *sets of actions available* to players 1 and 2, respectively, in state $s \in S$. It is assumed that the set-valued mappings $s \mapsto A(s)$ and $s \mapsto B(s)$ are continuous. Let

$$K = \{(s, a, b) : s \in S,\ a \in A(s) \text{ and } b \in B(s)\}$$

and

$$\bar{K} = \{(s, \nu, \rho) : s \in S,\ \nu \in P(A(s)) \text{ and } \rho \in P(B(s))\}.$$

It is obvious that K is a closed subset of $S \times A \times B$. By Lemma 1.1, $\bar{K}$ is a closed subset of the Borel space $S \times P(A) \times P(B)$.

(iv) q is a Borel measurable transition probability from K to S, called the *law of motion among states*. If s is a state at some stage of the game and the players select actions $a \in A(s)$ and $b \in B(s)$, then $q(\cdot|s, a, b)$ is the probability distribution of the next state of the game.

(v) $r : K \mapsto R$ is a Borel measurable *(daily) reward function* for player 1 (*cost function* for player 2).

Strategies for player 1 are denoted by $\pi = \{\pi_n\}$ whereas for player 2 by $\gamma = \{\gamma_n\}$ and are defined in an usual way. By Π $[\Gamma]$ we denote the *class of all strategies for player* 1 [*player* 2]. Let F $[G]$ be the set of all stationary strategies for player 1 [player 2].

Let $H = K \times K \times K \times \cdots$ be the space of all infinite histories of the game endowed with the Borel σ-algebra. For any $\pi \in \Pi$ and $\gamma \in \Gamma$, every initial state $s_0 = s \in S$ a probability measure $P_s^{\pi\gamma}$ and a stochastic process $\{s_m, a_m, b_m\}$ are defined on H in a canonical way, where the random variables s_m, a_m and b_m describe the state and the action chosen by players 1 and 2, respectively, on the m-th stage of the game (see Ionescu-Tuclcea's Theorem in Chapter 7 in [2] for a formal construction). Thus, for each initial state $s_0 = s \in S$, any strategies $\pi \in \Pi$, $\gamma \in \Gamma$ and any finite horizon n, the total *expected n-stage reward* to player 1 is

$$J_n(s, \pi, \gamma) = E_s^{\pi\gamma} \left[\sum_{m=0}^{n-1} r(s_m, a_m, b_m) \right],$$

where $E_s^{\pi\gamma}$ means the expectation operator with respect to the probability measure $P_s^{\pi\gamma}$ with $s = s_0$. (The assumptions imposed on r and q below guarantee that all expectations considered in the sequel are well-defined.) If β is a fixed real number in $(0, 1)$, called the *discount factor*, then the *expected discounted reward* to player 1 is

$$J_\beta(s, \pi, \gamma) = E_s^{\pi\gamma} \left[\sum_{m=0}^{\infty} \beta^m r(s_m, a_m, b_m) \right].$$

The *expected average reward per unit time* to player 1 is defined as

$$J(s, \pi, \gamma) = \liminf_{n \to \infty} \frac{J_n(s, \pi, \gamma)}{n}.$$

For any initial state $s \in S$, define

$$L(s) := \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J(s, \pi, \gamma)$$

and

$$U(s) := \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J(s, \pi, \gamma).$$

Then L (U) is called the *lower (upper) value*, respectively, in the average payoff stochastic game. It is always true that $L(s) \leq U(s)$ for $s \in S$. If $L(s) = U(s)$ for all $s \in S$, then this common function is called the *value* of the stochastic game and is denoted by ξ.

Suppose that the average reward stochastic game has a value ξ. A strategy $\pi^* \in \Pi$ is called *optimal for player* 1 in the average payoff stochastic game iff

$$\inf_{\gamma \in \Gamma} J(s, \pi^*, \gamma) = \xi(s)$$

for all $s \in S$, and a strategy $\gamma^* \in \Gamma$ is called *optimal for player* 2 in the average payoff stochastic game iff

$$\sup_{\pi \in \Pi} J(s, \pi, \gamma^*) = \xi(s)$$

for all $s \in S$. Of course, the value and optimal strategies are defined similarly for the β-discounted and n-stage stochastic games.

We impose the following continuity assumptions.

C1: $r : K \mapsto R$ is continuous.

C2: $q : K \mapsto P(S)$ is weakly continuous.

C3: There exists a continuous function $V : S \mapsto [1, \infty)$ such that $|r(s, a, b)| \leq V(s)$ for every $(s, a, b) \in K$.

C4: The function

$$(s, a, b) \mapsto \int_S V(y) q(dy|s, a, b)$$

is continuous on K.

The following assumption has been recently made in the theory of Markov control processes and stochastic games [7], [8], [13], [27]. Some special cases were considered in [9], [11], [12].

C5: (1) There exist a Borel function $\delta : K \mapsto [0,1]$ and a probability measure φ on S such that

$$q(B|s,a,b) \geq \delta(s,a,b)\varphi(B)$$

for any Borel set $B \subset S$ and $s \in S$.

(2) $\int_S \inf_{a\in A(s)} \inf_{b\in B(s)} \delta(s,a,b)\varphi(ds) > 0$.

(3) $\varphi(V) := \int_S V(s)\varphi(ds) < \infty$.

(4) For some $\lambda \in (0,1)$ and every $(s,a,b) \in K$, it holds

$$\int_S V(y)q(dy|s,a,b) \leq \lambda V(s) + \delta(s,a,b)\varphi(V).$$

By Proposition 7.50 in [2] the function $s \mapsto \inf_{a\in A(s)} \inf_{b\in B(s)} \delta(s,a,b)$ is universally measurable. Therefore, the integral in **C5**(2) is well-defined.

Let $C_V(S)$ denote the subset of all continuous functions on S for which the so-called V-norm

$$\|u\|_V := \sup_{s\in S} \frac{|u(s)|}{V(s)}$$

is finite. The space of Borel measurable functions u on S for which $\|u\|_V < \infty$ is denoted by $L_V^\infty(S)$. Let μ be a finite signed measure on $\mathcal{B}(S)$. The V-norm of μ is defined by

$$\|\mu\|_V := \sup_{\|u\|_V \leq 1} \left| \int_S u(s)\mu(ds) \right| = \int_S V(s)|\mu|(ds),$$

where $|\mu| = \mu^+ + \mu^-$ denotes the total variation of μ, and μ^+, μ^- stand for the positive and negative parts of μ, respectively. From **C5**(3)-(4), it follows that the intergral $\int_S |u(y)|q(dy|s,a,b)$ is finite for all $(s,a,b) \in K$ and each $u \in L_V^\infty(S)$.

Let $s \in S$, $\nu \in P(A(s))$ and $\rho \in P(B(s))$. We define

$$r(s,\nu,\rho) = \int_{A(s)} \int_{B(s)} r(s,a,b)\rho(db)\nu(da)$$

and for any Borel set $D \subset S$, we put

$$q(D|s,\nu,\rho) = \int_{A(s)} \int_{B(s)} q(D|s,a,b)\rho(db)\nu(da).$$

Thus, for any $f \in F$ and $g \in G$, $r(s,f(s),g(s))$ and $q(D|s,f(s),g(s))$ ($s \in S$, $D \in \mathcal{B}(S)$) have clear meaning.

An important consequence of **C5** is that for any $f \in F$ and $g \in G$, the state process $\{s_n\}$ is a positive recurrent aperiodic Markov chain with the unique invariant probability measure (also called stationary distribution), denoted by π_{fg}. In addition, $\{s_n\}$ is V-uniformly ergodic, that is, there exist $\theta > 0$ and $\alpha \in (0,1)$ such that

$$\|q^n(\cdot|s,f(s),g(s)) - \pi_{fg}(\cdot)\|_V \leq V(s)\theta\alpha^n. \qquad (1)$$

From (1), we conclude that

$$J(f,g) := J(s,f,g) = \int_S r(s,f(s),g(s))\pi_{fg}(ds), \qquad (2)$$

for every $f \in F$ and $g \in G$, that is, the expected average payoff is independent of the initial state.

Let $u \in L_V^\infty(S)$. Define

$$u'(s,a,b) = r(s,a,b) + \beta \int_S u(y)q(dy|s,a,b)$$

and

$$\bar{u}(s,\nu,\rho) = \int_{A(s)} \int_{B(s)} u'(s,a,b)\rho(db)\nu(da),$$

where $(s,a,b) \in K$ and $(s,\nu,\rho) \in \bar{K}$.

Lemma 2.1: Let **C1** through **C4** hold. If u is lower [upper] semicontinuous, then u' [$\bar{u}$] is lower [upper] semicontinuous on K [$\bar{K}$].

To simplify our notation, we shall use the following (lower) value operators. For each function $u \in L_V^\infty(S)$ and $\beta \in (0,1]$, we put

$$\begin{aligned}(T_\beta u)(s) &= \sup_{\nu\in P(A(s))} \inf_{\rho\in P(B(s))} [r(s,\nu,\rho) + \\ &\quad \beta \int_S u(y)q(dy|s,\nu,\rho)], \end{aligned} \qquad (3)$$

where $s \in S$ and we set $Tu = T_\beta u$ when $\beta = 1$. We close this section with some auxiliary results on minimax selections which are closely related to Theorem 5.1 in [22]. We remind that r is continuous on K and q is Feller. Moreover, by Lemma 1.1 the set-valued mappings $s \mapsto P(A(s))$ and $s \mapsto P(B(s))$ are continuous. If u is upper semicontinuous, then by Lemma 2.1 and Berge's theorems (see pp. 115-116 in [1]) $T_\beta u$ is upper semicontinuous and by a minimax selection theorem [21], there exists some $f \in F$ such that

$$\begin{aligned}(T_\beta u)(s) &= \max_{\nu\in P(A(s))} \inf_{\rho\in P(B(s))} [r(s,\nu,\rho) + \\ &\quad \beta \int_S u(y)q(dy|s,\nu,\rho)] \\ &= \inf_{\rho\in P(B(s))} [r(s,f(s),\rho) + \\ &\quad \beta \int_S u(y)q(dy|s,f(s),\rho)], \end{aligned} \qquad (4)$$

for each $s \in S$. Similarly, if u is lower semicontinuous, then by Lemma 2.1 and Berge's theorems, $T_\beta u$ is lower semicontinuous and by Fan's minimax theorem [4] and a measurable selection theorem [21], there exists some $g \in G$ such that

$$\begin{aligned}(T_\beta u)(s) &= \min_{\rho\in P(B(s))} \sup_{\nu\in P(A(s))} [r(s,\nu,\rho) + \\ &\quad \beta \int_S u(y)q(dy|s,\nu,\rho)] \\ &= \sup_{\nu\in P(A(s))} [r(s,\nu,g(s)) + \\ &\quad \beta \int_S u(y)q(dy|s,\nu,g(s))], \end{aligned} \qquad (5)$$

for each $s \in S$.

Let $u \in C_V(S)$. Then, by Fan's minimax theorem [4], (4) and (5), $T_\beta u \in C_V(S)$.

III. FATOU'S LEMMA FOR VARYING PROBABILITY MEASURES

Let $\{w_n\}$ be a sequence of functions in $L_V^\infty(S)$. As in [26], we consider the following "generalized liminf (limsup)":

$$w_*(s) := \inf\{\liminf_{n\to\infty} w_n(s_n) : s_n \to s\} \tag{6}$$

and

$$w^*(s) := \sup\{\limsup_{n\to\infty} w_n(s_n) : s_n \to s\}. \tag{7}$$

Lemma 3.1: The function w_* (w^*) is lower (upper) semi-continuous.

Lemma 3.2: Let $\{\mu_n\} \subset P(S)$ be converging weakly to some $\mu_0 \in P(S)$. If $\{v_n\}$ is a sequence of nonnegative Borel measurable functions on S and v_* is defined as in (6), then

$$\int_S v_*(s)\mu_0(ds) \le \liminf_{n\to\infty} \int_S v_n(s)\mu_n(ds), \tag{8}$$

and if the functions $\{v_n\}$ are nonpositive, then

$$\int_S v^*(s)\mu_0(ds) \ge \limsup_{n\to\infty} \int_S v_n(s)\mu_n(ds) \tag{9}$$

with v^* defined as in (7).

Proof: Inequality (8) easily follows Lemma 3.2 in [26].Obviously, (9) can be easily concluded from (8) by taking into account the sequence $\{-v_n\}$. □

Lemma 3.3: Assume that $\{\mu_n\}$ converges weakly to some $\mu_0 \in P(S)$ and $\{w_n\}$ is a sequence of functions in $C_V(S)$ such that $\|w_n\|_V \le b$ for all n and some constant $b > 0$. If V is a continuous function and $\int_S V(s)\mu_m(ds) < \infty$ for every $m \ge 0$ and

$$\int_S V(s)\mu_m(ds) \to \int_S V(s)\mu_0(ds) \tag{10}$$

as $m \to \infty$, then

$$\int_S w_*(s)\mu_0(ds) \le \liminf_{n\to\infty} \int_S w_n(s)\mu_n(ds), \tag{11}$$

and

$$\int_S w^*(s)\mu_0(ds) \ge \limsup_{n\to\infty} \int_S w_n(s)\mu_n(ds). \tag{12}$$

Proof: Define $v_n(s) := w_n(s) + bV(s)$ and note that $v_n \ge 0$. For any $s \in S$ and arbitrary sequence $s_n \to s$ as $n \to \infty$, we have

$$\liminf_{n\to\infty} v_n(s_n) = bV(s) + \liminf_{n\to\infty} w_n(s_n).$$

Hence $v_*(s) = bV(s) + w_*(s)$, $s \in S$, and consequently

$$\int_S v_*(s)\mu_0(ds) = b\int_S V(s)\mu_0(ds) + \int_S w_*(s)\mu_0(ds).$$

Applying (8) to the sequence $\{v_n\}$ and (10), we easily get

$$\begin{aligned}
&\liminf_{n\to\infty} \int_S w_n(s)\mu_n(ds) + b\int_S V(s)\mu_0(ds) \\
&= \liminf_{n\to\infty} \int_S v_n(s)\mu_n(ds) \\
&\ge \int_S v_*(s)\mu_0(ds) = \int_S w_*(s)\mu_0(ds) + \\
&\quad b\int_S V(s)\mu_0(ds)
\end{aligned}$$

which immediately gives (11).

Similarly, (12) can be concluded from (9) by taking $v_n(s) := w_n(s) - bV(s) \le 0$. □

IV. MAIN RESULTS

Theorem 4.1: Assume (i)–(v), **C1-C4**, and **C5**(4). (a) The finite horizon discounted stochastic game has a value and both players have optimal Markov strategies. Moreover, if ξ_k is the value function for the k-stage game, then $\xi_k \in C_V(S)$ and

$$\xi_{n+1}(s) = (T_\beta \xi_n)(s) \quad \text{for each} \quad s \in S,\ n \ge 1.$$

(b) The discounted stochastic game has a value ξ_β and both players have optimal stationary strategies $f_\beta \in F$ and $g_\beta \in G$. Moreover, $\xi_\beta \in C_V(S)$ and

$$\begin{aligned}
\xi_\beta(s) &= (T_\beta\xi_\beta)(s) = r(s, f_\beta(s), g_\beta(s)) + \\
&\quad \beta\int_S \xi_\beta(y)q(dy|s, f_\beta(s), g_\beta(s)) \\
&= \max_{\nu\in P(A(s))}[r(s,\nu,g_\beta(s)) + \\
&\quad \beta\int_S \xi_\beta(y)q(dy|s,\nu,g_\beta(s))] \\
&= \min_{\rho\in P(B(s))}[r(s,f_\beta(s),\rho) + \\
&\quad \beta\int_S \xi_\beta(y)q(dy|s,f_\beta(s),\rho)]
\end{aligned} \tag{13}$$

for every $s \in S$.

Proof: For the proof the reader is referred to [14]. □

Theorem 4.2: Assume (i)–(v) and **C1-C5**. Then the average payoff stochastic game has a value ξ which is independent of the initial state and both players have optimal stationary strategies. Moreover, $\xi = \lim_{\beta\to 1}(1-\beta)\xi_\beta(s^*)$ for any state $s^* \in S$.

Proof: By Theorem 4.1 the value ξ_β of the β-discounted game and stationary strategies $f_\beta \in F$, $g_\beta \in G$ exist for every $\beta \in (0,1)$. Fix a state $s^* \in S$ and consider a sequence $\{\beta_n\}$ of discount factors converging to one. Define

$$w_n(s) = \xi_{\beta_n}(s) - \xi_{\beta_n}(s^*), \quad \xi_n = (1-\beta_n)\xi_{\beta_n}(s^*).$$

Then, from (13), it follows that

$$\begin{aligned}
\xi_n + w_n(s) &= \min_{\rho\in P(B(s))}\max_{\nu\in P(A(s))}[r(s,\nu,\rho) + \\
&\quad \beta_n\int_S w_n(t)q(dt|s,\nu,\rho)] \\
&= \max_{\nu\in P(A(s))}[r(s,\nu,g_{\beta_n}(s)) + \\
&\quad \beta_n\int_S w_n(t)q(dt|s,\nu,g_{\beta_n}(s))] \quad s\in S.
\end{aligned} \tag{14}$$

By an argument given on page 135 in [8], the sequence $\{\xi_n\}$ is bounded and there is no loss of generality to assume that ξ_n converges to some real number ξ^* as $n \to \infty$. Moreover, by Lemma 10.4.2 in [8], it follows that $\|w_n\|_V \le b$ for some constant b. Now, we fix a state $s_0 \in S$ and consider an arbitrary sequence of states $\{s_n\}$ such that $s_n \to s_0$ as $n \to \infty$. Then from (14), putting $g_n := g_{\beta_n}$, we obtain

$$\xi_n + w_n(s_n) = \max_{\nu \in P(A(s_n))} [r(s_n, \nu, g_n(s_n)) + \beta_n \int_S w_n(y) q(dy|s_n, \nu, g_n(s_n))]. \quad (15)$$

Let $\{n_k\}$ be a subsequence of positive integers for which

$$\liminf_{n\to\infty} w_n(s_n) = \lim_{k\to\infty} w_{n_k}(s_{n_k}).$$

Obviously, $\lim_{k\to\infty} \xi_{n_k} = \xi^*$ and

$$\xi^* + \liminf_{n\to\infty} w_n(s_n) = \liminf_{n\to\infty} [\xi_n + w_n(s_n)] = \lim_{k\to\infty} [\xi_{n_k} + w_{n_k}(s_{n_k})].$$

Consequently, from (15), we obtain

$$\xi^* + \liminf_{n\to\infty} w_n(s_n) = \lim_{k\to\infty} \max_{\nu \in P(A(s_{n_k}))} [r(s_{n_k}, \nu, g_{n_k}(s_{n_k})) + \beta_{n_k} \int_S w_{n_k}(y) q(dy|s_{n_k}, \nu, g_{n_k}(s_{n_k}))].$$

Let F_C be the set of all continuous stationary strategies for player 1. By Lemmas 1.1 and 1.2, F_C is non-empty. Choose any $f \in F_C$. Then, we have

$$\xi^* + \liminf_{n\to\infty} w_n(s_n) \geq \liminf_{k\to\infty} [r(s_{n_k}, f(s_{n_k}), g_{n_k}(s_{n_k})) + \beta_{n_k} \int_S w_{n_k}(y) q(dy|s_{n_k}, f(s_{n_k}), g_{n_k}(s_{n_k}))]. \quad (16)$$

Note that $Z := \{s_0\} \cup \{s_n\}$ is compact in S. We know that the set-valued mapping $s \mapsto P(B(s))$ is continuous and compact-valued. These facts together with Berge's theorem (see [1]) imply that $\bigcup_{z\in Z} P(B(z))$ is compact in $P(B)$. Therefore, $\{g_{n_k}(s_{n_k})\}$ has a subsequence converging to some $\rho_0 \in P(B)$. Without loss of generality, let $g_{n_k}(s_{n_k}) \to \rho_0$, as $k \to \infty$. By the continuity of $s \mapsto P(B(s))$, $\rho_0 \in P(B(s_0))$. On the other hand, $f(s_{n_k}) \to f(s_0) \in P(A(s_0))$, because $f \in F_C$. Clearly, $q(\cdot|s_{n_k}, f(s_{n_k}), g_{n_k}(s_{n_k})) \to q(\cdot|s_0, f(s_0), \rho_0)$ weakly as $k \to \infty$. By (16) and Lemma 3.3, we infer that

$$\begin{aligned}\xi^* + \liminf_{n\to\infty} w_n(s_n) &\geq \liminf_{k\to\infty} r(s_{n_k}, f(s_{n_k}), g_{n_k}(s_{n_k})) + \liminf_{k\to\infty} \int_S w_{n_k}(y) q(dy|s_{n_k}, f(s_{n_k}), g_{n_k}(s_{n_k})) \\ &\geq r(s_0, f(s_0), \rho_0) + \liminf_{k\to\infty} \int_S w_{n_k}(y) q(dy|s_{n_k}, f(s_{n_k}), g_{n_k}(s_{n_k})) \\ &\geq r(s_0, f(s_0), \rho_0) + \int_S \bar{w}_*(y) q(dy|s_0, f(s_0), \rho_0),\end{aligned} \quad (17)$$

where $\bar{w}_*$ is the generalized liminf of the sequence $\bar{w}_k = w_{n_k}$. Let $w_*(s_0)$ be the generalized liminf of $\{w_n\}$ defined in (6). Then $w_* \leq \bar{w}_*$ and applying this fact to (17), we get

$$\xi^* + \liminf_{n\to\infty} w_n(s_n) \geq r(s_0, f(s_0), \rho_0) + \int_S w_*(y) q(dy|s_0, f(s_0), \rho_0).$$

Furthermore, by the fact that a continuous selector $f \in F_C$ can be chosen in such a way that an arbitrary value from $P(A(s))$ is assigned to the point s_0 (see Lemma 1.2), we infer that

$$\begin{aligned}\xi^* + \liminf_{n\to\infty} w_n(s_n) &\geq \sup_{f\in F_C} [r(s_0, f(s_0), \rho_0) + \int_S w_*(y) q(dy|s_0, f(s_0), \rho_0)] \\ &= \sup_{\nu \in P(A(s_0))} [r(s_0, \nu, \rho) + \int_S w_*(y) q(dy|s_0, \nu, \rho)].\end{aligned}$$

Since w_* is lower semicontinuous (see Lemma 3.1), we can write

$$\xi^* + \liminf_{n\to\infty} w_n(s_n) \geq \min_{\rho \in P(B(s_0))} \sup_{\nu \in P(A(s_0))} [r(s_0, \nu, \rho) + \int_S w_*(y) q(dy|s_0, \nu, \rho)].$$

By the definition of w_*, see (6), we have

$$\xi^* + w_*(s_0) \geq \min_{\rho \in P(B(s_0))} \sup_{\nu \in P(A(s_0))} [r(s_0, \nu, \rho) + \int_S w_*(y) q(dy|s_0, \nu, \rho)] \quad (18)$$

Since s_0 was chosen arbitrarily, then (18) holds with s_0 replaced by any $s \in S$.

We already know that the function w_* is lower semicontinuous. By (18) and a minimax measurable selection theorem (see (5)), there exists some $g^* \in G$ such that

$$\xi^* + w_*(s) \geq r(s, \nu, g^*(s)) + \int_S w_*(y) q(dy|s, \nu, g^*(s)) \quad (19)$$

for every $s \in S$ and $\nu \in P(A(s))$. Iterating (19), one can show in a standard manner (see [8]) that

$$\xi^* \geq \sup_{\pi\in\Pi} J(s, \pi, g^*) \geq U(s) \quad (20)$$

for each $s \in S$.

Now, let $f_{\beta_n} = f_n$ be a stationary optimal strategy to player 1 in the β_n discounted stochastic game. We need to prove that $\xi^* \leq L(s)$. From (14) and Fan's minimax theorem [4], it can be easily seen that

$$\xi_n + w_n(s_n) = \min_{\rho \in P(B(s_n))} [r(s_n, f_n(s_n), \rho) + \beta_n \int_S w_n(y) q(dy|s_n, f_n(s_n), \rho)].$$

Assume that $s_n \to s_0$ and consider the generalized limsup $w^*(s_0)$ (see (7)). Proceeding along similar lines, we take again a subsequence $\{n_k\}$ of positive integers such that

$$\limsup_{n\to\infty} w_n(s_n) = \lim_{k\to\infty} w_{n_k}(s_{n_k}).$$

Then using Lemma 3.3 and (4), we obtain some $f^* \in F$ such that

$$\xi^* + w^*(s) \leq r(s, f^*(s), \rho) + \int_S w^*(y) q(dy|s, f^*(s), \rho) \quad (21)$$

for every $s \in S$ and $\rho \in P(B(s))$. Again standard dynamic programming arguments, based on (21), show that

$$\xi^* \leq \inf_{\gamma \in \Gamma} J(s, f^*, \gamma) \leq L(s), \quad (22)$$

for each $s \in S$. By (20) and (22), we have

$$\begin{aligned} \xi^* &= \sup_{\pi \in \Pi} J(s, \pi, g^*) = \inf_{\gamma \in \Gamma} J(s, f^*, \gamma) \\ &= J(s, f^*, g^*) = J(f^*, g^*), \end{aligned} \quad (23)$$

that is, the game has a value $\xi(s) = \xi^*$, and f^*, g^* are stationary optimal strategies for players 1 and 2, respectively. □

V. FINAL REMARKS

Theorem 4.2 is a *first* result on ergodic stochastic games with Feller transition probabilities satisfying fairly general assumptions. All related papers [11], [12], [17], [23], [27] are based on the strong continuity assumption on q saying that the mapping $(s, a, b) \mapsto q(D|s, a, b)$ is continuous in (a, b) for every Borel subset D of S. The payoffs and transitions in the mentioned papers need not be continuous with respect to the state variable. The situation studied in this paper is somewhat more delicate. Note that to overcome some technical difficulties we have to use Michael's theorem on continuous selections [20], which plays no part in the proofs given in the aforementioned. Feller transition probabilities are in some applications more natural than strongly continuous ones.

REFERENCES

[1] E. Berge, *Topological Spaces*, MacMillan, New York; 1963.
[2] D.P. Bertsekas and S.E. Shreve, *Stochastic Optimal Control: The Discrete Time Case*, Academic Press, New York; 1978.
[3] L.O. Curtat, Markov equilibria of stochastic games with complementarities, *Games Econ. Behavior*, vol. 17, 1996, pp 177-199.
[4] K. Fan, Minimax theorems, *Proc. Nat. Acad. Sci. U.S.A.*, vol. 39, 1953, pp 42-47.
[5] D. Duffie, J. Geanakoplos, A. Mas-Colell and A. McLennan, Stationary Markov equilibria, *Econometrica*, vol. 62, 1994, pp 745-782.
[6] E.A. Feinberg and M.E. Lewis, Optimality of four-treshold policies in inventory systems with customer returns and borrowing/storage options, *Probab. in Engineering and Informational Science*, 2004.
[7] J.I. González-Trejo, O. Hernández-Lerma and L.F. Hoyos-Reyes, Minimax control of discrete-time stochastic systems, *SIAM J. Control Optim.*, vol. 41, 2003, pp 1626-1659.
[8] O. Hernández-Lerma and J.B. Lasserre, *Further Topics on Discrete-Time Markov Control Processes*, Springer-Verlag, New York; 1999.
[9] O. Hernández-Lerma and J.B. Lasserre, Zero-sum stochastic games in Borel spaces: average payoff criteria, *SIAM J. Control Optim.*, vol. 39, 2001, pp 1520-1539.
[10] C.J. Himmelberg and F.S. Van Vleck, Multifunctions with values in a space of probability measures, *J. Math. Anal. Appl.*, vol. 50, 1975, pp 108-112.
[11] A. Jaśkiewicz, Zero-sum semi-Markov games, *SIAM J. Control Optim.*, vol. 41, 2002, pp 723-739.
[12] A. Jaśkiewicz and A.S. Nowak, On the optimality equation for zero-sum ergodic stochastic games, *Math. Methods Oper. Res.*, vol. 54, 2001, pp 291-301.
[13] A. Jaśkiewicz and A.S. Nowak, On the average cost optimality equation for Markov control processes with Feller transition probabilities, 2004, submitted.
[14] A. Jaśkiewicz and A.S. Nowak, Zero-sum ergodic stochastic games with Feller transition probabilities, 2004, submitted.
[15] N.V. Kartashov, *Strong Stable Markov Chains*, VSP, Utrecht, The Netherlands; 1996.
[16] H.J. Kushner, Numerical approximations for stochastic differential games: the ergodic case, *SIAM J. Control Optim.*, vol. 42, 2004, pp 1911-1933.
[17] H.-U. Küenle and R. Schurath, The optimality equation and ε-optimal strategies in Markov games with average reward criterion, *Math. Methods Oper. Res.*, vol. 56, 2003, pp 439-449.
[18] S.P. Meyn and R.L. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag, New York; 1993.
[19] S.P. Meyn and R.L. Tweedie, Computable bounds for geometric convergence rates of Markov chains, *Ann. Appl. Probab.*, vol. 4, 1994, pp 981-1011.
[20] E. Michael, Continuous selections I, *Ann. Math.* vol. 63, 1956, pp 361-382.
[21] A.S. Nowak, Measurable selection theorems for minimax stochastic optimization problems, *SIAM J. Control Optim.*, vol. 23, 1985, pp 466-476.
[22] A.S. Nowak, Semicontinuous nonstationary stochastic games, *J. Math. Anal. Appl.*, vol. 117, 1986, pp 84-99.
[23] A.S. Nowak, Zero-sum average payoff stochastic games with general state space, *Games and Econ. Behavior*, vol. 7, 1994, pp 221-232.
[24] K.R. Parthasarathy, *Probability Measures on Metric Spaces*, Academic Press, New York; 1967.
[25] D. Repovš, and P.V. Semenov, *Continuous Selections of Multivalued Mappings*, Kluwer Acad. Publishers, Dordrecht; 1998.
[26] R. Serfozo, Convergence of Lebesgue integrals with varying measures, *Sankhyã, Ser. A*, vol. 44, 1982, pp 380-402.
[27] O. Vega-Amaya, Zero-sum average semi-Markov games: fixed-point solutions of the Shapley equation, *SIAM J. Control Optim.*, vol. 42, 2003, pp 1876-1894.