

Existence of Nash equilibria in nonzero-sum ergodic stochastic games in Borel spaces *

Rafael Benítez-Medina

Abstract

In this paper we study nonzero-sum stochastic games with Borel state and action spaces, and the average payoff criterion. Under suitable assumptions we show the existence of Nash equilibria in stationary strategies. Our hypotheses include ergodicity conditions and an ARAT (additive reward, additive transition) structure.

2000 Mathematics Subject Classification: 91A15, 91A10.

Keywords and phrases: nonzero-sum stochastic games, ARAT games, Nash equilibria, expected average payoff.

1 Introduction

This paper concerns nonzero-sum stochastic games with Borel state and action spaces, and the average payoff criterion with possibly unbounded payoffs. This class of games has many applications, for instance, in queueing and economic theory (see [1], [2], [12], [27]).

The problem we are interested in is the existence of Nash equilibria in stationary strategies. To do this we impose ergodicity conditions already used by several authors for markov games and control problems (e.g. [1], [6], [9], [10], [14], [15], [19], [22]) together with a so-called ARAT (additive reward, additive transition law) structure. Similar results have been obtained by Ghosh and Bagchi [5] and Küenle [14] for games with bounded payoffs. Other related works include [18], which deals with

*Research partially supported by a CONACyT scholarship. This paper is a part of the author's M. Sc. Thesis presented at the Department of Mathematics of CINVESTAV-IPN.

Borel state space and bounded payoffs, and [27], in which the state space is countable.

For stochastic games with a *discounted* payoff criterion there is a larger literature. For instance, for zero-sum problems in countable spaces, see [1], [17], [26]; for uncountable spaces, see [11], [13], [21], [25]. On the other hand, for the nonzero-sum case in countable spaces, see [27], and for uncountable spaces, see [11], [23], [24].

The remainder of the paper is organized as follows. Section 2 introduces standard material on stochastic games and strategies, and the optimality criteria. The core of the paper is contained in section 3: after introducing some assumptions, we present our main result, Theorem 3.10, on the existence of Nash equilibria. Finally, after some technical preliminaries in section 4, the proof of Theorem 3.10 is presented in section 5.

2 The game model

For notational ease, we shall consider a stochastic game with only two players. For $N > 2$ players, the situation is completely analogous. We begin with the following remark on terminology and notation.

2.1 Remark.

- (a) A Borel subset X of a complete and separable metric space is called a *Borel space*, and its Borel σ -algebra is denoted by $\mathcal{B}(X)$. We only deal with Borel spaces, and so measurable always means “Borel measurable”. Given a Borel space X , we denote by $\mathbb{P}(X)$ the family of probability measures on X , endowed with the weak topology $\sigma(\mathbb{P}(X), C_b(X))$, where $C_b(X)$ stands for the space of continuous bounded functions on X . In this case, $\mathbb{P}(X)$ is a Borel space. Moreover, if X is compact, then so is $\mathbb{P}(X)$.
- (b) Let X and Y be Borel spaces. A measurable function $\phi : Y \rightarrow \mathbb{P}(X)$ is called a *transition probability* from Y to X , and we denote by $\mathbb{P}(X|Y)$ the family of all those transition probabilities. If ϕ is in $\mathbb{P}(X|Y)$, then we write its values either as $\phi(y)(B)$ or as $\phi(B|y)$, for all $y \in Y$ and $B \in \mathcal{B}(X)$. Finally, if $X = Y$ then ϕ is said to be a *Markov transition probability* on X .

The stochastic game model. We shall consider the two-person nonzero-sum stochastic game model

$$(1) \quad GM := (X, A, B, \mathbb{K}_A, \mathbb{K}_B, Q, r_1, r_2),$$

where X is the state space, and A and B are the action spaces for players 1 and 2, respectively. These spaces are all assumed to be Borel spaces. The sets $\mathbb{K}_A \in \mathcal{B}(X \times A)$ and $\mathbb{K}_B \in \mathcal{B}(X \times B)$ are the constraint sets. That is, for each $x \in X$, the x -section in \mathbb{K}_A , namely

$$A(x) := \{a \in A \mid (x, a) \in \mathbb{K}_A\},$$

represents the set of admissible actions for player 1 in the state x . Similarly, the x -section in \mathbb{K}_B , i.e.

$$B(x) := \{b \in B \mid (x, b) \in \mathbb{K}_B\},$$

stands for the family of admissible actions for player 2 in the state x . Let

$$\mathbb{K} := \{(x, a, b) \mid x \in X, a \in A(x), b \in B(x)\},$$

which is a Borel subset of $X \times A \times B$. Then $Q \in \mathbb{P}(X \mid \mathbb{K})$ is the game's transition law, and, finally, $r_i : \mathbb{K} \rightarrow \mathbb{R}$ is a measurable function representing the reward function for player $i = 1, 2$.

The game is played as follows. At each stage $t = 0, 1, \dots$, the players 1 and 2 observe the current state $x \in X$ of the system, and independently choose actions $a \in A(x)$ and $b \in B(x)$, respectively. As a consequence of this, the following happens: (1) player i receives an immediate reward $r_i(x, a, b)$, $i = 1, 2$; and (2) the system moves to a new state with distribution $Q(\cdot \mid x, a, b)$. The goal of each player is to maximize, in the sense of Definition 2.2, below, his long-run expected average reward (or payoff) per unit time.

2.2 Strategies

Let $\mathcal{H}_0 := X$ and $\mathcal{H}_t := \mathbb{K} \times \mathcal{H}_{t-1}$ for $t = 1, 2, \dots$. For each t , an element $h_t = (x_0, a_0, b_0, \dots, x_{t-1}, a_{t-1}, b_{t-1}, x_t)$ of \mathcal{H}_t represents a “history” of the game up to time t . A *strategy for player 1* is then defined as a sequence $\pi^1 = \{\pi_t^1, t = 0, 1, \dots\}$ of transition probabilities π_t^1 in $\mathbb{P}(A \mid \mathcal{H}_t)$ such that

$$\pi_t^1(A(x_t) \mid h_t) = 1 \quad \forall h_t \in \mathcal{H}_t, t = 0, 1, \dots$$

We denote by Π_1 the family of all strategies for player 1.

Now define $\mathbb{A}(x) := \mathbb{P}(A(x))$ for each state $x \in X$, and let S_1 be the class of all transition probabilities $\phi \in \mathbb{P}(A|X)$ such that $\phi(x)$ is in $\mathbb{A}(x)$ for all $x \in X$. Then a strategy $\pi^1 = \{\pi_t^1\} \in \Pi_1$ is called *stationary* if there exists $\phi \in S_1$ such that

$$\pi_t^1(\cdot|h_t) = \phi(x_t)(\cdot) \quad \forall h_t \in \mathcal{H}_t, t = 0, 1, \dots$$

We will identify S_1 with the family of stationary strategies for player 1.

The sets of strategies Π_2 and S_2 for player 2 are defined similarly, writing $B(x)$ and $\mathbb{B}(x) := \mathbb{P}(B(x))$ in lieu of $A(x)$ and $\mathbb{A}(x)$, respectively.

Let (Ω, \mathcal{F}) be the canonical measurable space that consists of the sample space $\Omega := (X \times A \times B)^\infty$ and its product σ -algebra \mathcal{F} . Then for each pair of strategies $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ and each initial state $x \in X$ there exists a probability measure $P_x^{\pi^1, \pi^2}$ and a stochastic process $\{(x_t, a_t, b_t), t = 0, 1, \dots\}$ defined on (Ω, \mathcal{F}) in a canonical way, where x_t, a_t and b_t represent the state and the actions of players 1 and 2, respectively, at each stage $t = 0, 1, \dots$. The expectation operator with respect to $P_x^{\pi^1, \pi^2}$ is denoted by $E_x^{\pi^1, \pi^2}$.

2.3 Average payoff criteria

For each $n = 1, 2, \dots$ and $i = 1, 2$, let

$$J_n^i(\pi^1, \pi^2, x) := E_x^{\pi^1, \pi^2} \left[\sum_{t=0}^{n-1} r_i(x_t, a_t, b_t) \right]$$

be the n -stage expected total payoff (or reward) of player i when the players use the strategies $\pi^1 \in \Pi_1$ and $\pi^2 \in \Pi_2$, given the initial state $x_0 = x$.

The corresponding long-run expected average payoff (EAP) per unit time is then defined as

$$(2) \quad J^i(\pi^1, \pi^2, x) := \liminf_{n \rightarrow \infty} J_n^i(\pi^1, \pi^2, x)/n.$$

The EAP is also known as the ergodic payoff (or ergodic reward) criterion.

2.2 Definition. A pair of strategies (π^{1*}, π^{2*}) is called a *Nash equilibrium* (for the EAP criterion) if

$$J^1(\pi^{1*}, \pi^{2*}, x) \geq J^1(\pi^1, \pi^{2*}, x) \text{ for all } \pi^1 \in \Pi_1, \quad x \in X,$$

and

$$J^2(\pi^{1*}, \pi^{2*}, x) \geq J^2(\pi^{1*}, \pi^2, x) \text{ for all } \pi^2 \in \Pi_2, \quad x \in X.$$

Our aim is to establish, under certain assumptions, the existence of a Nash equilibrium (ϕ^*, ψ^*) in $S_1 \times S_2$.

We introduce the following notation. For any given function $f : \mathbb{K} \rightarrow \mathbb{R}$ and probability measures $\phi \in \mathbb{A}(x)$ and $\psi \in \mathbb{B}(x)$, we write

$$f(x, \phi, \psi) := \int_{A(x)} \int_{B(x)} f(x, a, b) \psi(db) \phi(da)$$

whenever the integrals are well defined. In particular, for r_i and Q as in (1),

$$r_i(x, \phi, \psi) := \int_{A(x)} \int_{B(x)} r_i(x, a, b) \psi(db) \phi(da)$$

and

$$Q(\cdot | x, \phi, \psi) := \int_{A(x)} \int_{B(x)} Q(\cdot | x, a, b) \psi(db) \phi(da).$$

3 Main result

We first introduce our assumptions, and then present our main result.

3.1 Assumption. (a) For each state $x \in X$, the sets $A(x)$ and $B(x)$ of admissible actions are compact.

(b) For each (x, a, b) in \mathbb{K} , $r_1(x, \cdot, b)$ is upper semicontinuous (u.s.c) on $A(x)$, and $r_2(x, a, \cdot)$ is u.s.c on $B(x)$.

(c) For each (x, a, b) in \mathbb{K} and each bounded measurable function v on X , the functions

$$\int_X v(y) Q(dy | x, \cdot, b) \text{ and } \int_X v(y) Q(dy | x, a, \cdot)$$

are continuous on $A(x)$ and $B(x)$, respectively.

(d) There exists a constant \bar{r} and a measurable function $w(\cdot) \geq 1$ on X such that

$$(3) \quad |r_i(x, a, b)| \leq \bar{r} w(x) \quad \forall (x, a, b) \in \mathbb{K}, \quad i = 1, 2,$$

and, in addition, part (c) holds when v is replaced with w .

The next two assumptions are used to guarantee that the state process $\{X_t\}$ is ergodic in a suitable sense.

3.2 Assumption. There exists a probability measure $\nu \in \mathbb{P}(X)$, a positive number $\alpha < 1$, and a measurable function $\beta : \mathbb{K} \rightarrow [0, 1]$ for which the following holds for all $(x, a, b) \in \mathbb{K}$ and $D \in \mathcal{B}(X)$:

- (a) $Q(D|x, a, b) \geq \beta(x, a, b)\nu(D)$;
- (b) $\int_X w(y)Q(dy|x, a, b) \leq \alpha w(x) + \beta(x, a, b)\|\nu\|_w$, where $w(\cdot) \geq 1$ is the function in Assumption 3.1(d), and $\|\nu\|_w := \int w d\nu$.
- (c) $\inf \int_X \beta(x, \phi(x), \psi(x))\nu(dx) > 0$, where the infimum is over all the pairs (ϕ, ψ) in $S_1 \times S_2$.

3.3 Assumption. There exists a σ -finite measure λ on X with respect to which, for each pair $(\phi, \psi) \in S_1 \times S_2$, the Markov transition probability $Q(\cdot|x, \phi(x), \psi(x))$ is λ -irreducible.

We next introduce some notation and then we mention some important consequences of the above assumptions.

3.4 Definition. $\mathbb{B}_w(X)$ denotes the linear space of real-valued measurable functions u on X with a finite w -norm, which is defined as

$$(4) \quad \|u\|_w := \sup_{x \in X} |u(x)|/w(x),$$

and $\mathbb{M}_w(X)$ stands for the normed linear space of finite signed measures μ on X such that

$$(5) \quad \|\mu\|_w := \int_X w d|\mu| < \infty.$$

Note that the integral $\int u d\mu$ is finite for each $u \in \mathbb{B}_w(X)$ and μ in $\mathbb{M}_w(X)$, because, by (4) and (5),

$$|\int u d\mu| \leq \|u\|_w \int w d|\mu| = \|u\|_w \|\mu\|_w < \infty.$$

3.5 Remark. Suppose that Assumptions 3.2 and 3.3 are satisfied. Then:

- (a) For each pair $(\phi, \psi) \in S_1 \times S_2$, the state (Markov) process $\{X_t\}$ is positive Harris recurrent; hence, in particular, the Markov transition probability $Q(\cdot|x, \phi(x), \psi(x))$ admits a unique invariant probability measure in $\mathbb{M}_w(X)$ which will be denoted by $q(\phi, \psi)$; thus

$$q(\phi, \psi)(D) = \int_X Q(D|x, \phi(x), \psi(x))q(\phi, \psi)(dx) \quad \forall D \in \mathcal{B}(X).$$

(b) $\{X_t\}$ is w -geometrically ergodic, that is, there exist positive constants $\theta < 1$ and M such that

$$(6) \quad \left| \int_X u(y) Q^n(dy|x, \phi(x), \psi(x)) - \int_X u(y) q(\phi, \psi)(dy) \right| \leq w(x) \|u\|_w M \theta^n$$

for every $u \in \mathbb{B}_w(X)$, $x \in X$, and $n = 0, 1, \dots$, where Q^n denotes the n -step Markov transition probability. This result follows from Lemmas 3.3 and 3.4 in [6] where it was assumed the positive Harris recurrence in part (a). However, as shown in Lemma 4.1 of [15], the latter recurrence is a consequence of Assumptions 3.2 and 3.3.

3.6 Assumption. There exists a probability measure γ in $\mathbb{M}_w(X)$ (i.e. $\int w d\gamma < \infty$) and a strictly positive density function $g(x, a, b, \cdot)$ such that

$$Q(D|x, a, b) = \int_D g(x, a, b, y) \gamma(dy)$$

for all $D \in \mathcal{B}(\mathcal{X})$ and $(x, a, b) \in \mathbb{K}$.

Note that Assumption 3.6 implies 3.3 with $\lambda = \gamma$.

3.7 Assumption. The transition density $g(x, a, b, y)$ is such that

$$(7) \quad \lim_{n \rightarrow \infty} \int_X |g(x, a^n, b^n, y) - g(x, a, b, y)| w(y) \gamma(dy) = 0 \quad \forall x \in X$$

if $a^n \rightarrow a$ in $A(x)$ and $b^n \rightarrow b$ in $B(x)$, where $w(\cdot)$ is the function in Assumption 3.1(d).

The next two assumptions require that the game model (1) has a so-called *ARAT* (additive reward, additive transition law) structure.

3.8 Assumption. There exist substochastic kernels $Q_1 \in \mathbb{P}(X|\mathbb{K}_A)$ and $Q_2 \in \mathbb{P}(X|\mathbb{K}_B)$ such that

$$Q(\cdot|x, a, b) = Q_1(\cdot|x, a) + Q_2(\cdot|x, b)$$

for all $x \in X$, $a \in A(x)$, $b \in B(x)$. Further, $Q_1(D|x, \cdot)$ and $Q_2(D|x, \cdot)$ are continuous on $A(x)$ and $B(x)$, respectively, for each $D \in \mathcal{B}(X)$.

3.9 Assumption. For $i = 1, 2$ there exist measurable functions

$$r_{i1} : \mathbb{K}_A \rightarrow \mathbb{R}, \quad r_{i2} : \mathbb{K}_B \rightarrow \mathbb{R},$$

such that

(a) $r_i(x, a, b) = r_{i1}(x, a) + r_{i2}(x, b)$ for all $x \in X$, $a \in A$, $b \in B$.

Moreover, for each $x \in X$,

(b) the functions $r_{i1}(x, \cdot)$ and $r_{i2}(x, \cdot)$ are continuous on $A(x)$ and $B(x)$, respectively, and

(c) $\max_{a \in A(x)} |r_{i1}(x, a)| \leq w(x)$, and $\max_{b \in B(x)} |r_{i2}(x, b)| \leq w(x)$.

Observe that (c) and the condition $\gamma \in \mathbb{M}_w(X)$ in Assumption 3.6 yield that

$$\int_X \max_{a \in A(x)} |r_{i1}(x, a)| \gamma(dx) < \infty, \quad \int_X \max_{b \in B(x)} |r_{i2}(x, b)| \gamma(dx) < \infty.$$

3.10 Theorem. *Under Assumptions 3.1, 3.2 and 3.6-3.9, there is a pair $(\phi^*, \psi^*) \in S_1 \times S_2$ that is a Nash equilibrium.*

The remainder of this work is devoted to prove Theorem 3.10.

4 Preliminaries

Suppose that one of the players, say player 2, selects a *fixed* stationary strategy ψ in S_2 . Then the game model GM in (1) reduces to a *Markov control model*

$$(8) \quad MCM_1(\psi) = (X, A, \mathbb{K}_A, Q_\psi, r_{1,\psi})$$

where X , A and \mathbb{K}_A are as in (1), and the transition law Q_ψ in $\mathbb{P}(X | \mathbb{K}_A)$ and the reward function $r_{1,\psi} : \mathbb{K}_A \rightarrow \mathbb{R}$ are given by

$$Q_\psi(\cdot | x, a) := Q(\cdot | x, a, \psi(x)) \quad \text{and} \quad r_{1,\psi}(x, a) := r_1(x, a, \psi(x)),$$

respectively. Then from Corollary 5.12 in [10], for instance, we get the following.

4.1 Lemma. *Suppose that Assumptions 3.1, 3.2 and 3.3 are satisfied. Then for each fixed $\psi \in S_2$, there exists a stationary strategy $\phi^* \in S_1$ that is expected average reward (EAR) optimal for the Markov control model in (8), i.e.,*

$$(9) \quad J^1(\phi^*, \psi, x) = \max_{\pi^1 \in \Pi_1} J^1(\pi^1, \psi, x) =: \rho_1^*(\psi) \quad \forall x \in X.$$

Moreover, there exists a function $h_{\phi^*, \psi}^1 \in \mathbb{B}_w(X)$ such that $(\rho_1^*(\psi), h_{\phi^*, \psi}^1)$ is the unique solution in $\mathbb{R} \times \mathbb{B}_w(X)$ of the equation

$$(10) \quad \rho_1^*(\psi) + h_{\phi^*, \psi}^1(x) = r_1(x, \phi^*(x), \psi(x)) + \int_X h_{\phi^*, \psi}^1(y) Q(dy|x, \phi^*(x), \psi(x))$$

$$(11) \quad = \max_{\phi \in \mathbb{A}(x)} [r_1(x, \phi, \psi(x)) + \int_X h_{\phi^*, \psi}^1(y) Q(dy|x, \phi, \psi(x))]$$

for all $x \in X$, and such that $\int_X h_{\phi^*, \psi}^1(y) q(\phi^*, \psi)(dy) = 0$, with $q(\phi^*, \psi)$ as in the Remark 3.5(a).

In other words, (9) states that $\phi^* \in S_1$ is an *optimal response* of player 1, given that player 2 uses the fixed stationary strategy $\psi \in S_2$. Similarly, we can obtain an optimal response $\psi^* \in S_2$ of player 2 if player 1 uses a fixed strategy $\phi \in S_1$.

We now wish to express the optimal average reward $\rho_1^*(\psi)$ in (9), in a more convenient form. We will use the following fact, which is borrowed from Proposition 10.2.3 in [9].

4.2 Lemma. *Suppose that Assumptions 3.1, 3.2 and 3.3 are satisfied, and let $(\phi, \psi) \in S_1 \times S_2$ be an arbitrary pair of stationary strategies. Then for $i = 1, 2$ we have:*

(a) *The EAP in (2) satisfies that*

$$(12) \quad J^i(\phi, \psi, x) = \lim_{n \rightarrow \infty} J_n^i(\phi, \psi, x)/n = \rho_i(\phi, \psi),$$

where

$$(13) \quad \rho_i(\phi, \psi) := \int_X r_i(x, \phi(x), \psi(x)) q(\phi, \psi)(dx)$$

with $q(\phi, \psi)$ as in Remark 3.5(a).

(b) *The function $h_{\phi, \psi}^i$ defined on X as*

$$h_{\phi, \psi}^i(x) := \sum_{t=0}^{\infty} E_x^{\phi, \psi} [r_i(x_t, \phi(x_t), \psi(x_t)) - \rho_i(\phi, \psi)]$$

belongs to $\mathbb{B}_w(X)$, and, moreover, its w -norm is independent of (ϕ, ψ) :

$$(14) \quad \|h_{\phi, \psi}^i\|_w \leq \bar{r}M/(1 - \theta),$$

where \bar{r} is the constant in (3), and M and θ are as in (6).

(c) The pair $(\rho_i(\phi, \psi), h_{\phi, \psi}^i)$ is the unique solution in $\mathbb{R} \times \mathbb{B}_w(X)$ of the so-called Poisson equation

$$(15) \quad \rho_i(\phi, \psi) + h_{\phi, \psi}^i(x) = r_i(x, \phi(x), \psi(x)) + \int_X h_{\phi, \psi}^i(y) Q(dy | x, \phi(x), \psi(x))$$

that satisfies the condition

$$\int_X h_{\phi, \psi}^i(y) q(\phi, \psi)(dx) = 0.$$

5 Proof of Theorem 3.10

From (12) and Corollary 5.12(a) in [10], we can write $\rho_1^*(\psi)$ in (9) as

$$(16) \quad \rho_1^*(\psi) = \rho_1(\phi^*, \psi) = \max_{\phi \in S_1} \rho_1(\phi, \psi).$$

Similarly, for each $\phi \in S_1$ there exists $\psi^* \in S_2$ such that

$$(17) \quad \rho_2^*(\phi) = \rho_2(\phi, \psi^*) = \max_{\psi \in S_2} \rho_2(\phi, \psi)$$

We next use (16) and (17) to introduce a multifunction τ from $S_1 \times S_2$ to $2^{S_1 \times S_2}$ as follows: for each pair (ϕ, ψ) in $S_1 \times S_2$

$$(18) \quad \tau(\phi, \psi) := \{(\phi^*, \psi^*) | \rho_1(\phi^*, \psi) = \rho_1^*(\psi), \rho_2(\phi, \psi^*) = \rho_2^*(\phi)\}.$$

To complete the proof of Theorem 3.10 we shall proceed in two steps, which is in fact a standard procedure (see Ghosh and Bagchi [5], Himmelberg et. al. [11], Parthasarathy [23], for instance).

Step 1. Introduce a topology on S_i ($i = 1, 2$) with respect to which S_i is compact and metrizable.

Step 2. Show that the multifunction τ is *upper semicontinuous* (u.s.c.), that is, if (i) $(\phi_n, \psi_n) \rightarrow (\phi_\infty, \psi_\infty)$ in $S_1 \times S_2$, and (ii) $(\phi_n^*, \psi_n^*) \in \tau(\phi_n, \psi_n)$ is such that $(\phi_n^*, \psi_n^*) \rightarrow (\phi_\infty^*, \psi_\infty^*)$, then $(\phi_\infty^*, \psi_\infty^*)$ is in $\tau(\phi_\infty, \psi_\infty)$.

From these two steps and Fan's fixed point theorem (Theorem 1 in [4]), it will follow that the multifunction τ has a fixed point (ϕ^*, ψ^*) in $S_1 \times S_2$, that is

$$(19) \quad (\phi^*, \psi^*) \in \tau(\phi^*, \psi^*).$$

Finally, from (16) – (18) and (19) we shall conclude that (ϕ^*, ψ^*) is a Nash equilibrium.

In step 1 we shall use the topology introduced by Warga (see Theorem IV.3.1 in [28]): Let F_1 be the Banach space of measurable functions $f : \mathbb{K}_A \rightarrow \mathbb{R}$ such that $f(x, a)$ is continuous in $a \in A(x)$ for each $x \in X$ and

$$\|f\| := \int_X \max_{a \in A(x)} |f(x, a)| \gamma(dx) < \infty,$$

with γ as in Assumption 3.6. We shall identify two stationary strategies ϕ and ϕ' in S_1 if $\phi = \phi'$ γ -a.e. (almost everywhere), and, on the other hand, $\phi \in S_1$ can be identified with the linear functional $\Delta_\phi \in F_1^*$ given by

$$\Delta_\phi(f) := \int_X \int_A f(x, a) \phi(da|x) \gamma(dx).$$

Thus S_1 can be identified with a subset of F_1^* , and endowing S_1 with the weak* topology it can be shown that S_1 is compact and metrizable [28]. The set S_2 is topologized analogously.

To proceed with step 2, suppose that

$$(20) \quad (\phi_n, \psi_n) \rightarrow (\phi_\infty, \psi_\infty) \text{ in } S_1 \times S_2,$$

and that

$$(21) \quad (\phi_n^*, \psi_n^*) \in \tau(\phi_n, \psi_n) \quad \forall n$$

is such that

$$(22) \quad (\phi_n^*, \psi_n^*) \rightarrow (\phi_\infty^*, \psi_\infty^*).$$

By (21) and the definition (18) of τ , together with (10) (or (15)), for all $x \in X$ we have

$$(23) \quad \begin{aligned} \rho_1^*(\psi_n) + h_{\phi_n^*, \psi_n}^1(x) &= r_1(x, \phi_n^*(x), \psi_n(x)) \\ &\quad + \int_X h_{\phi_n^*, \psi_n}^1(y) Q(dy|x, \phi_n^*(x), \psi_n(x)) \end{aligned}$$

and, similarly,

$$(24) \quad \begin{aligned} \rho_2^*(\phi_n) + h_{\phi_n, \psi_n^*}^2(x) &= r_2(x, \phi_n(x), \psi_n^*(x)) \\ &\quad + \int_X h_{\phi_n, \psi_n^*}^2(y) Q(dy|x, \phi_n(x), \psi_n^*(x)). \end{aligned}$$

Now observe that, by Assumptions 3.8 and 3.9, for each $D \in \mathcal{B}(X)$, the functions $Q_1(D|x, a)$ and $r_{i1}(x, a)$ are in F_1 , whereas $Q_2(D|x, b)$ and $r_{i2}(x, b)$ are in F_2 . Therefore, by (20) and (22),

$$(25) \quad \int_X r_1(x, \phi_n^*(x), \psi_n(x)) \gamma(dx) \rightarrow \int_X r_1(x, \phi_\infty^*(x), \psi_\infty(x)) \gamma(dx),$$

and similarly for $i = 2$. Moreover, for any $D \in \mathcal{B}(X)$,

$$(26) \quad \int_X Q(D|x, \phi_n^*(x), \psi_n(x))\gamma(dx) \rightarrow \int_X Q(D|x, \phi_\infty^*(x), \psi_\infty(x))\gamma(dx),$$

and similarly for $(\phi_n, \psi_n^*) \rightarrow (\phi_\infty, \psi_\infty^*)$.

5.1 Lemma. *There is a subsequence $\{m\}$ of $\{n\}$ and numbers $\hat{\rho}_1$ and $\hat{\rho}_2$ such that*

$$(27) \quad \rho_1^*(\psi_m) = \rho_1(\phi_m^*, \psi_m) \rightarrow \hat{\rho}_1$$

and

$$(28) \quad \rho_2^*(\phi_m) = \rho_2(\phi_m, \psi_m^*) \rightarrow \hat{\rho}_2.$$

Proof: Let $\rho_i(\phi, \psi)$ be as in (13). We next show that, for $i = 1, 2$,

$$(29) \quad |\rho_i(\phi, \psi)| \leq \bar{r} \|\nu\|_w / (1 - \alpha) \quad \forall (\phi, \psi) \in S_1 \times S_2,$$

with \bar{r} as in (3), and ν and α as in Assumption 3.2. Clearly, (29) implies (27) and (28).

To prove (29), note that Assumption 3.2(b) yields

$$(30) \quad \int_X w(y)Q(dy|x, a, b) \leq \alpha w(x) + \|\nu\|_w$$

because $\beta(x, a, b) \leq 1$. Now let (ϕ, ψ) be an arbitrary pair in $S_1 \times S_2$. Integrating both sides of (30) with respect to $\phi(da|x)$ and $\psi(db|x)$, and then integrating with respect to the invariant probability measure $q(\phi, \psi)$ yields

$$\int_X w(y)q(\phi, \psi)(dy) \leq \alpha \int_X w(y)q(\phi, \psi)(dy) + \|\nu\|_w,$$

and, therefore,

$$\int_X w(y)q(\phi, \psi)(dy) \leq \|\nu\|_w / (1 - \alpha).$$

The latter inequality, together with (3) and (13), gives

$$\begin{aligned} |\rho_i(\psi, \phi)| &\leq \int_X |r_i(x, \phi(x), \psi(x))|q(\phi, \psi)(dx) \\ &\leq \bar{r} \int_X w(y)q(\phi, \psi)(dy) \\ &\leq \bar{r} \|\nu\|_w / (1 - \alpha), \end{aligned}$$

i.e., (29) holds. This in turn gives that the sequences $\{\rho_1(\phi_n^*, \psi_n)\}$ and $\{\rho_2(\phi_n, \psi_n^*)\}$ are uniformly bounded, and so the lemma follows. \square

For notational convenience, we shall write the subsequence $\{m\} \subset \{n\}$ in (27) and (28) as the original sequence, $\{n\}$. Moreover, let

$$(31) \quad u_n(\cdot) := h_{\phi_n^*, \psi_n}^1(\cdot), \text{ and } \tilde{u}_n(\cdot) := u_n(\cdot)/w(\cdot).$$

By (14), the constant $m_0 := \bar{r}M/(1 - \theta)$ satisfies that

$$|\tilde{u}_n(x)| \leq m_0 \quad \forall x, n.$$

Let U be the space of all γ -equivalence classes of real-valued measurable functions u on X such that $|u(x)| \leq m_0$ γ -a.e. By the Alaoglu (or Banach-Alaoglu) Theorem (see page 424 in [3], for instance), U is a compact and metrizable subset of $L^\infty(\gamma) \equiv L^\infty(X, \mathcal{B}(X), \gamma)$ equipped with the relative weak* topology $\sigma(L^\infty(\gamma), L^1(\gamma))$. Therefore, we can assume that $\{\tilde{u}_n\}$ converges in the weak* topology to some function \tilde{u}_* in $L^\infty(\gamma)$. Let $u_*(x) := \tilde{u}_*(x)w(x)$ for all $x \in X$. Then, as in the proof of Theorem 4 in [19], using Assumption 3.7, one can show that as $n \rightarrow \infty$.

$$(32) \quad \max_{a \in A(x)} \max_{b \in B(x)} \left| \int_X (u_n(y) - u_*(y)) Q(dy|x, a, b) \right| \rightarrow 0 \quad \forall x \in X$$

with $u_n(\cdot)$ as in (31). In turn, (32) and Assumption 3.8 yield that

$$(33) \quad \max_{\phi \in \mathbf{A}(x)} \max_{\psi \in \mathbf{B}(x)} \left| \int_X (u_n(y) - u_*(y)) Q(dy|x, \phi, \psi) \right| \rightarrow 0 \quad \forall x \in X.$$

We also have the following.

5.2 Lemma. *If $(\phi_n, \psi_n) \rightarrow (\phi, \psi)$ in $S_1 \times S_2$, then, as $n \rightarrow \infty$,*

$$(34) \quad \int_X \int_X u(y) Q(dy|x, \phi_n(x), \psi_n(x)) \gamma(dx) \\ \rightarrow \int_X \int_X u(y) Q(dy|x, \phi(x), \psi(x)) \gamma(dx)$$

for any function $u \in \mathbb{B}_w(X)$.

Proof: Choose an arbitrary function $u \in \mathbb{B}_w(X)$. By definition of the weak convergence of $\phi_n \rightarrow \phi$ and $\psi_n \rightarrow \psi$ in S_1 and S_2 , respectively, and Assumption 3.8, to prove the lemma it suffices to show that the functions

$$(35) \quad (x, a) \rightarrow \int_X u(y) Q_1(dy|x, a)$$

and

$$(36) \quad (x, b) \rightarrow \int_X u(y)Q_2(dy|x, b)$$

are in F_1 and F_2 , respectively. With this in mind, first note that $\int_X u(y)Q_i(dy|x, \cdot)$ is continuous in $a \in A(x)$ and $b \in B(x)$, for $i = 1$ and $i = 2$, respectively, (see Lemma 8.3.7(a) in [9]). Moreover, by (4) and Assumption 3.2(b) (using that $\beta(x, a, b) \leq 1$),

$$\max_{a \in A(x)} \left| \int_X u(y)Q_1(dy|x, a) \right| \leq \|u\|_w(\alpha w(x) + \|\nu\|_w) \quad \forall x \in X.$$

Hence, as $\int w d\gamma < \infty$ (by Assumption 3.6), the function in (35) is in F_1 . Similarly, the function in (36) is in F_2 . \square

By Lemmas 5.1 and 5.2, together with (25), (26) and (33), letting $n \rightarrow \infty$ in (23) we obtain γ -a.e.

$$(37) \quad \begin{aligned} \hat{\rho}_1 + u_*(x) &= r_1(x, \phi_\infty^*(x), \psi_\infty(x)) \\ &\quad + \int_X u_*(y)Q(dy|x, \phi_\infty^*(x), \psi_\infty(x)) \\ &= \max_{\phi \in \mathbb{A}(x)} [r_1(x, \phi, \psi_\infty(x)) \\ &\quad + \int_X u_*(y)Q(dy|x, \phi, \psi_\infty(x))], \end{aligned}$$

where the second equality comes from (10)- (11) replacing (ϕ^*, ψ) with (ϕ_n^*, ψ_n) .

Finally, arguing as in the last part of the proof of Theorem 5.8 in [10], let $D \in \mathcal{B}(X)$ be the set with $\gamma(D) = 1$ on which (37) holds, and let $h_* : X \rightarrow \mathbb{R}$ be such that $h_*(x) := u_*(x)$ for $x \in D$, and

$$h_*(x) := \max_{\phi \in \mathbb{A}(x)} [r_1(x, \phi, \psi_\infty(x)) + \int_X u_*(y)Q(dy|x, \phi, \psi_\infty(x))] - \hat{\rho}_1$$

for all x in the complement D^c of D . As $\gamma(D^c) = 0$, by Lemma 6.3 in [10], we have $Q(D^c|x, a, b) = 0$ for all $(x, a, b) \in \mathbb{K}$. Therefore, (37) holds for all $x \in X$ when $u_*(\cdot)$ is replaced with $h_*(\cdot)$. This implies (by Lemma 4.1) that

$$(38) \quad \hat{\rho}_1 = \rho_1^*(\psi_\infty) = \rho_1(\phi_\infty^*, \psi_\infty).$$

An analogous argument using (21), (22) and (24) with obvious changes, shows that

$$(39) \quad \hat{\rho}_2 = \rho_2^*(\phi_\infty) = \rho_2(\phi_\infty, \psi_\infty^*).$$

In other words, (38) and (39) state that, under (20)- (22), the pair $(\phi_\infty^*, \psi_\infty^*)$ is in $\tau(\phi_\infty, \psi_\infty)$, and so the set-valued map defined by (18) is u.s.c. Thus, as was already noted, it follows from Fan's fixed point theorem that τ has a fixed point (as in (19), say), which completes the proof of Theorem 3.10. \square

Acknowledgement

The author wishes to thank Dr. Onésimo Hernández-Lerma for his very valuable comments and suggestions.

Rafael Benítez-Medina
Departamento de Matemáticas,
 CINVESTAV-IPN,
 A. Postal 14-740,
 07000, México D.F., MEXICO,
 rbenitez@math.cinvestav.mx.

References

- [1] Altman E., Hordijk A. and Spieksma F. M., *Contraction conditions for average and α -discount optimality in countable state Markov games with unbounded rewards*, Math. Oper. Res., **22** (1997), 588-618.
- [2] Curtat L. O., *Markov equilibria of stochastic games with complementarities*, Games and Economic Behavior, **17** (1996), 177-199.
- [3] Dunford N. and Schwartz J. T., *Linear Operators*, Vol. 1, Interscience Publishers, New York, 1958.
- [4] Fan K., *Fixed point and minimax theorems in locally convex topological spaces*, Proc. Nat. Acad. Sci., U. S. A., **38** (1952), 539-560.
- [5] Ghosh M. K. and Bagchi A., *Stochastic games with average payoff criterion*, Appl. Math. Optim., **38** (1998), 283-301.
- [6] Gordienko E. and Hernández-Lerma O., *Average cost Markov control processes with weighted norms: existence of canonical policies*, Appl. Math. (Warsaw), **23** (1995), 199-218.
- [7] Hernández-Lerma O., *Adaptive Markov Control Process*, Springer-Verlag, New York, 1989.

- [8] Hernández-Lerma O. and Lasserre J. B., *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, Springer-Verlag, New York, 1996.
- [9] Hernández-Lerma O. and Lasserre J. B., *Further Topics on Discrete-Time Markov Control Processes*, Springer-Verlag, New York, 1999.
- [10] Hernández-Lerma O. and Lasserre J. B., *Zero-sum stochastic games in Borel spaces: average payoff criteria*, SIAM J. Control Optim., **39** (2001), 1520-1539.
- [11] Himmelberg C., Parthasarathy T., Raghavan T. E. S. and Van Vleck F., *Existence of ρ -equilibrium and optimal stationary strategies in stochastic games*, Proc. Amer. Math. Soc., **60** (1976), 245-251.
- [12] Karatzas I., Shubik M. and Sudderth W. D., *Construction of stationary Markov equilibria in a strategic market game*, Math. Oper. Res., **19** (1994), 975-1006.
- [13] Kumar P. R. and Shiau T. H., *Existence of value and randomized strategies in zero-sum discrete time stochastic dynamic games*, SIAM J. Control Optim., **19** (1981), 561-585.
- [14] Küenle H. U., *Equilibrium strategies in stochastic games with additive cost and transition structure and Borel state and action spaces*, International Game Theory Review, **1** (1999), 131-147.
- [15] Luque-Vázquez F. and Hernández-Lerma O., *Semi-Markov control models with average costs*, Appl. Math. (Warsaw), **26** (1999), 315-331.
- [16] Maitra A. and Parthasarathy T., *On stochastic games*, J. Optim. Theory Appl., **5** (1970), 289-300.
- [17] Mohan S. R., Neogy S. K., Parthasarathy T. and Sinha S., *Vertical linear complementarity and discounted zero-sum stochastic games with ARAT structure*, Math. Program., Ser. A, **86** (1999), 637-648.
- [18] Nowak A. S., *Stationary equilibria for nonzero-sum average payoff ergodic stochastic games with general space state*, Annals of the International Society of Dynamic Games, **1** (1993), 232-246.

- [19] Nowak A. S., *Optimal strategies in a class of zero-sum ergodic stochastic games*, Math. Meth. Oper. Res., **50** (1999), 399-419.
- [20] Nowak, A. S., *Zero-sum average payoff stochastic games with general state space*, Games and Economic Behavior , **7** (1994), 221-232.
- [21] Nowak, A. S., *On zero-sum stochastic games with general state space*, I. Prob. Math. Stat., **4** (1984), 13-32.
- [22] Nowak A. S. and Altman E., *ϵ -Nash equilibria for stochastic games with uncountable state space and unbounded costs*, (1998). Preprint.
- [23] Parthasarathy T., *Existence of equilibrium stationary strategies in discounted stochastic games*, Shankhya. The Indian Journal of Statistics, **44** (1982), Series A, 114-127.
- [24] Parthasarathy T. and Sinha S., *Existence of stationary equilibrium strategies in nonzero-sum discounted stochastic games with uncountable state space, and state independent transitions*, International Journal of Game Theory, **18** (1989), 189-194.
- [25] Ramirez-Reyes F., *Existence of optimal strategies for zero-sum stochastic games with discounted payoff*, Morfismos, **5** (2001), 63-83.
- [26] Sennot L. I., *Zero-sum stochastic games with unbounded cost: discounted and average cost cases*, Math. Meth. Oper. Res., **39** (1994), 209-225.
- [27] Sennot L. I., *Nonzero-sum stochastic games with unbounded cost: discounted and average cost cases*, Math. Meth. Oper. Res., **40** (1994), 145-162.
- [28] Warga J., *Optimal Control of Differential and Funtional Equations*, Academic Press, New York, 1972.