The Annals of Probability
1985, Vol. 13, No. 1, 269–287

UNIVERSALLY MEASURABLE STRATEGIES IN ZERO-SUM STOCHASTIC GAMES

BY ANDRZEJ S. NOWAK

Technical University of Wrocław

This paper deals with zero-sum discrete-time stationary models of stochastic games with Borel state and action spaces. A mathematical framework introduced here for such games refers to the minimax theorem of Ky Fan involving certain asymmetric assumptions on the primitive data. This approach ensures the existence and the universal measurability of the value functions and the existence for either or both players of optimal or ε-optimal universally measurable strategies in the finite horizon games as well as in certain infinite horizon games. The fundamental result of this paper is a minimax selection theorem extending a selection theorem of Brown and Purves. As applications of this basic result, we obtain some new theorems on absorbing, discounted, and positive stochastic games.

1. Introduction. Sequential competitive decision processes, such as stochastic games, represent the behavior of several competing decision makers interacting over time under uncertainty and therefore they play an increasing role as useful models for phenomena in the social sciences. We only mention here the papers of Deshmukh and Winston [9], Kirman and Sobel [22], and the book of Friedman [14], in which oligopoly situations are related to stochastic games.

The theory of stochastic games started with the fundamental paper of Shapley [43], in which two-person zero-sum stochastic games with finite state and action spaces were considered. The results of Shapley were extended in various directions; for a good survey see [32] and [36]. In recent years many researchers have been interested in formulating general mathematical frameworks for zero-sum stochastic games with uncountable state space which are broad enough to include the many applications of the methods and well-behaved enough to permit analysis. First results in this direction were given by Maitra and Parthasarathy (see [27] and [28]). For some extensions of their results we refer to [8], [13], [18], [21], [23], [25], [29], [30], [34] and [38]. Most of these papers present only so-called Borel space frameworks for stochastic games where the state space is an uncountable Borel set and the primitive data and strategies are Borel measurable. Stochastic games with more general state space are studied in [8], [29] and [30].

It is well-known that the value function of a stochastic game need not exist. As noted by Rieder, in contrast with the dynamic programming, the value function of a Borel space stochastic game which satisfies neither semi-continuity nor compactness conditions need not be universally measurable (see [38, Example 4.1] and [48, Theorem 7.1] or [47, Theorem 1]). The main purpose of this paper

Received January 1983; revised September 1983.
AMS 1980 *subject classifications.* Primary 90D15, 28A05; secondary 60K99, 93C55.
Key words and phrases. Zero-sum discrete-time stochastic games, optimal stationary strategies, universally measurable strategies, minimax selection theorem.

is to present a general Borel space framework for two-person zero-sum stochastic games that ensures the existence and the universal (not necessarily Borel) measurability of the value functions of the games as well as the existence for either or both players of optimal or ε-optimal universally measurable strategies. All the previous treatments, however, have mainly been directed toward formulating some additional assumptions on the Borel space stochastic game model that would guarantee the existence of a value function, its Borel measurability, and the existence of ε-optimal or optimal Borel measurable strategies for both players. The model we introduce here assumes certain asymmetric conditions on the primitive data (see Section 5, the assumption (RA)) inspired by the minimax theorem of Ky Fan [12, Theorem 2]. The compactness of action spaces and the semicontinuity or continuity of both the payoff per stage and the transition law, which were usually assumed in the literature for both players, are here imposed on one player only, say the minimizer. Thus, the other player need not have compact action spaces and the pay-off per stage and the transition law need not satisfy any semicontinuity requirements with respect to his actions. Therefore our model involving also a convergence condition (see Section 5, the assumption (FA)) generalizes the well-known discounted and positive stochastic dynamic programming ones studied by Blackwell [4], [5] and Strauch [48]. The fundamental result of this paper is Theorem 5.1 which is an extension of a selection theorem of Brown and Purves [7, Theorem 2]. The remaining results may be regarded as game theoretic extensions of finite horizon, discounted and negative dynamic programming results closely related to those of Shreve and Bertsekas from [46] and [47].

The organization of this paper is as follows. Section 2 gives some terminology and basic facts concerning Borel and analytic sets. Section 3 reviews some auxiliary measure theoretic facts. The stochastic game model is described in Section 4 with a minimum of assumptions. Section 4 also contains definitions of the dynamic programming operators which are useful in our investigation. The main results are provided in Section 5. Some auxiliary selection results are given in Section 6. Finally, Section 7 presents the proofs of the main results.

2. Borel and analytic sets. Semi-analytic functions. Throughout this paper a separable metric space X is called a *Borel space* or a *Borel set* if X is a Borel subset of some *Polish space*, i.e. complete separable metric space, and is endowed with the σ-algebra $\mathscr{B}(X)$ of all its Borel subsets.

We shall need the following facts.

(F 2.1) Let X and Y be Borel spaces and E be a Borel subset of $X \times Y$ such that the set $E(x) = \{y \in Y\colon (x, y) \in E\}$ is nonempty and compact for each $x \in X$. Then by [19, Theorem 3] and [17, Theorem 5.6], there is a sequence $\{f_n\}$ of Borel measurable functions on X into Y such that

$$E(x) = \mathrm{cl}\{f_n(x)\} \quad \text{for each } x \in X,$$

where cl denotes the closure in Y.

(F 2.2) If X and Y are Borel spaces, then the product space $X \times Y$ endowed with the product topology is also a Borel space and $\mathscr{B}(X \times Y)$ equals the product σ-algebra $\mathscr{B}(X) \otimes \mathscr{B}(Y)$ on $X \times Y$ [33, Chapter 1, Theorem 1.10].

Let N^N be the set of sequences of positive integers, endowed with the product topology. So N^N is a Polish space. Let X be a separable metric space. Then X is called an *analytic space* or an *analytic set* provided there is a continuous function f on N^N whose range $f(N^N)$ is X.

In this section, we list some properties of analytic sets that we shall be using.

(F 2.3) Every Borel set is analytic [24, Section 38 VI].

(F 2.4) The countable union, intersection and product of analytic sets is analytic [33, Chapter 1, Theorems 3.1 and 3.2].

(F 2.5) Let E be an analytic subset of an analytic space X. Then E is *universally measurable*, that is, if p is any probability measure on the Borel subsets of X, then E is in the completion of the Borel σ-algebra with respect to p [40].

The complement of an analytic set relative to a Borel space is called *complementary analytic.* We have the following fact.

(F 2.6) According to Gödel [16], it is consistent with the usual axioms of set theory to assume there is a complementary analytic subset of the unit square whose projection on the horizontal axis is not universally measurable.

For any Borel space X, we denote by $\mathscr{U}(X)$ the σ-algebra of all universally measurable subsets of X. It is known that $\mathscr{U}(X)$ is closed with respect to the Suslin operation (cf. [40, page 50]). Let $\mathscr{S}(X)$ be the smallest σ-algebra containing the Borel subsets of X and closed with respect to the Suslin operation. The σ-algebra $\mathscr{S}(X)$ was studied by Selivanovskij [42] and is discussed in Appendix B of [3] and in [45]. It is known that $\mathscr{S}(X)$ is contained in $\mathscr{U}(X)$.

Let X and Y be Borel spaces. Following Bertsekas and Shreve [3], we say that a function $f: X \to Y$ is *limit measurable* if $f^{-1}(B) \in \mathscr{S}(X)$ for every $B \in \mathscr{B}(Y)$. We say f is *universally measurable* if $f^{-1}(B) \in \mathscr{U}(X)$ for every $B \in \mathscr{B}(Y)$. Clearly, if f is limit measurable, then it is universally measurable.

By Theorem 5.5 of Leese [26] we have:

(F 2.7) Let X and Y be Borel spaces, and $C \in \mathscr{S}(X) \otimes \mathscr{B}(Y)$. Then the projection $\text{proj}_X C$ of C on X belongs to $\mathscr{S}(X)$, and, moreover, there is a limit measurable function $f: X \to Y$ such that $(x, f(x)) \in C$ for every $x \in \text{proj}_X C$.

If X is an analytic space and f is an extended real-valued function on X, then we say f is *upper semi-analytic* (u.s.a.) if the set $\{x \in X: f(x) > c\}$ (equivalently, $\{x \in X: f(x) \geq c\}$) is analytic for each real number c. By (F 2.3) every Borel measurable function is u.s.a., and by (F 2.5) every u.s.a. function is universally measurable. Let X be a Borel space and f be an extended real-valued u.s.a. function on X. Since analytic subsets of X are in $\mathscr{S}(X)$ [45], so f is limit measurable.

3. Auxiliary measure theoretic facts. Throughout this section, let X be a separable metric space, endowed with the σ-algebra $\mathscr{B}(X)$ of all its Borel subsets. We write $B(X)$ for the set of all bounded Borel measurable real-valued functions on X, and $C(X)$ for the set of all continuous functions in $B(X)$. Let $P(X)$ be the set of all probability measures on $\mathscr{B}(X)$. Given any $\mathscr{F} \subset B(X)$, we may endow $P(X)$ with the $\mathscr{F}$-topology defined as the coarsest topology on $P(X)$ in which all mappings $\mu \to \int u(x)\mu(dx)$, $u \in \mathscr{F}$, are continuous. In the case $\mathscr{F} = C(X)$ the $\mathscr{F}$-topology on $P(X)$ is called the *weak topology*, and in the case $\mathscr{F} = B(X)$ the $\mathscr{F}$-topology on $P(X)$ is called the *strong topology*.

From the theorem of Dini [39, Proposition 9.2.11] and Theorem 2.6 of Gänssler [15] we get

(F 3.1) For any $D \subset P(X)$, the following statements are equivalent:

(a) D is relatively compact in the strong topology.
(b) For any sequence $\{u_n\}$ in $B(X)$ which decreases to 0,

$$\int u_n(x)p(dx) \downarrow 0 \quad \text{as} \quad n \to \infty, \quad \text{uniformly in } p \in D.$$

Let $P(X)$ be endowed with the weak topology. We shall need the following facts.

(F 3.2) By embedding X in a countable product of unit intervals and using the fact that the unit ball in the space of uniformly continuous functions on a totally bounded metric space (with the supremum norm $\|\cdot\|$) is separable we get: there is a sequence $\{u_n\}$ of real-valued continuous functions on X with $\|u_n\| \leq 1$, $n \in N$, such that the metric ρ defined on $P(X)$ by

$$(3.1) \quad \rho(\mu, \lambda) = \sum_{n=1}^{\infty} 2^{-n} \left| \int u_n(x)\mu(dx) - \int u_n(x)\lambda(dx) \right|, \quad \mu, \lambda \in P(X),$$

is equivalent to the weak topology on $P(X)$ [33, page 47].

(F 3.3) If X is a Borel space, then $P(X)$ is a Borel space too [20, page 91].

(F 3.4) If X is compact, so is $P(X)$ [33, Theorem II 6.4].

(F 3.5) The σ-algebra $\mathscr{B}(P(X))$ of all Borel subsets of $P(X)$ coincides with the smallest σ-algebra on $P(X)$ for which the mapping $p \to p(E)$ is measurable for each $E \in \mathscr{B}(X)$ (cf. [37]).

(F 3.6) Let u be a bounded below real-valued lower semicontinuous function on X. Then $p \to \int u(x)p(dx)$ is an extended real-valued lower semicontinuous function on $P(X)$. This fact follows from the theorem of Baire [1, page 390] and the monotone convergence theorem.

(F 3.7) Let X and Y be analytic spaces and u be a bounded below extended real-valued u.s.a. function on $X \times Y$. Then from Corollary 31 of [6], it follows that $(x, p) \to \int u(x, y)p(dy)$ is an extended real-valued u.s.a. function on $X \times P(Y)$.

If X and Y are Borel spaces, $t(\cdot \mid x)$ is a probability measure on $\mathscr{B}(Y)$ for each $x \in X$, and the function $t(B \mid \cdot)$ from X into $[0, 1]$ is Borel (universally) measurable for each $B \in \mathscr{B}(Y)$, we say that t is a Borel (universally) measurable *transition probability* from X into Y. It can be shown that t is a Borel (universally) measurable transition probability from X into Y if and only if the mapping $x \to t(\cdot \mid x)$ from X into $P(Y)$ is Borel (universally) measurable (cf. [37, Lemma 6.1] and [10, Theorems 2.1 and 3.1]).

By a modification of Lemma 29 of [6] (cf. [44]) we can obtain the following fact.

(F 3.8) If f is a real-valued universally measurable (u.s.a., Borel measurable) function on $X \times Y$ which is bounded below, and $t\colon X \to P(Y)$ is universally measurable (Borel measurable, Borel measurable), then $x \to \int f(x, y)t(dy \mid x)$ is an extended real-valued universally measurable (u.s.a., Borel measurable) function on X. (A partial discussion of (F 3.8) is contained in appendix of [2]).

Finally, we give the following fact.

(F 3.9) Let f be a bounded real-valued universally measurable function on a Borel space Y, and t be a Borel measurable transition probability from a Borel space X into Y such that $t(B \mid \cdot)$ is continuous on X for each $B \in \mathscr{B}(Y)$. Then the function $x \to \int f(y)t(dy \mid x)$ is continuous on X.

For a *proof*, let $x_n \to x_0$ as $n \to \infty$. For each $m \geq 0$, there is a Borel measurable function f_m on Y and there is a Borel subset B_m of Y such that $f(y) = f_m(y)$ for all $y \in B_m$ and $t(B_m \mid x_m) = 1$. Let $B = \bigcup_{m=0}^{\infty} B_m$. Then $t(B \mid x_m) = 1$ for each $m \geq 0$, and since f is bounded we have

$$\int_Y f(y)t(dy \mid x_n) = \int_B f(y)t(dy \mid x_n) \to \int_B f(y)t(dy \mid x_0) = \int_Y f(y)t(dy \mid x_0)$$

as $n \to \infty$, which terminates the proof.

4. The stochastic game model. The stochastic game model we consider is defined by a two-person discrete-time stochastic control system $\langle S, A, B, F, G, q, r, \beta \rangle$ of the following meaning:

(i) S stands for the *state space* and is assumed to be a nonempty Borel space.

(ii) A and B are the *action spaces* for players I and II, respectively. A and B are assumed to be nonempty Borel spaces.

(iii) F and G are Borel subsets of $S \times A$ and $S \times B$, respectively. For any $s \in S$, the nonempty s-section $F(s) = \{a \in A\colon (s, a) \in F\}$ of F is called the *set of all admissible actions* for player I when the system is at the state s. Analogously for G. We assume that

$$H = \{(s, a, b)\colon s \in S,\ a \in F(s),\ b \in G(s)\} \tag{4.1}$$

is a Borel subset of $S \times A \times B$.

(iv) $q\colon H \to P(S)$ is a Borel measurable transition probability from H into S called the *transition law* of the system. Here, $q(\cdot \mid s, a, b)$ is the probability distribution of the state next visited by the system if the system is at the state s and the actions a and b are taken by the players I and II, respectively.

(v) $r: H \to R_+$ is a nonnegative Borel measurable function called the *immediate pay-off function* or the *pay-off function per stage.*

(vi) $\beta \in [0, 1]$ and is called the *discount factor.*

We write $H_1 = S$, $H_{n+1} = H \times H_n$ for $n \in N$, and $H_\infty = \times_{n=1}^{\infty} \tilde{H}_n$, where $\tilde{H}_n = H$ for each $n \in N$.

As usual, a *strategy* for player I is defined as a sequence $\pi = \{\pi_n\}$ of universally measurable transition probabilities $\pi_n: H_n \to P(A)$ such that $\pi_n(\cdot \mid s_1, a_1, b_1, \cdots, s_n)$ assigns probability one to $F(s_n)$ for each $(s_1, a_1, b_1, \cdots, s_n) \in H_n$, $n \in N$. We write Π for the set of all strategies for player I. We denote by D_A the set of all *decision functions* of player I, i.e., $f \in D_A$ if and only if f is a universally measurable function $f: S \to P(A)$ such that $f(s) \in P(F(s))$ for every $s \in S$. A *Markov strategy* for player I is a sequence $\{f_n\}$ where $f_n \in D_A$ for each $n \in N$. We write Π_M for the set of all Markov strategies for player I. A *stationary strategy* for player I is a Markov strategy $\{f_n\}$ where $f_n = f$ is independent of n. We identify D_A with the set of all stationary strategies for player I. Similarly, the sets Γ, D_B and Γ_M of all strategies, all decision functions (stationary strategies), and all Markov strategies, respectively, are defined for player II.

Let ζ_n, α_n, and β_n, $n \in N$, denote the projections from H_∞ onto the nth state space, nth action space for player I, and the nth action space for player II, respectively. Then the random variables ζ_n, α_n, and β_n describe the state of the system at stage n, and the actions chosen by the players I and II, respectively, at stage n.

The transition law q and a pair π, γ of strategies of players I and II define a conditional probability $m_{\pi\gamma}: S \to P(H_\infty)$. There, $m_{\pi\gamma}$ associates with each $s_1 \in S$ a probability measure on H_∞ which is concentrated on $\{s_1\} \times F(s_1) \times G(s_1) \times H \times H \times H \times \ldots$ and is defined through s_1, π, γ, and q according to the theorem of Ionescu Tulcea (see, e.g., [20, page 80]. We write $E_{\pi\gamma}[\cdot \mid \zeta_1 = s]$ for the respective conditional expectation.

Define

$$I_n(\pi, \gamma)(s) = E_{\pi\gamma}[\textstyle\sum_{k=1}^{n} \beta^{k-1} r(\zeta_k, \alpha_k, \beta_k) \mid \zeta_1 = s] \tag{4.2}$$

where $n \in \bar{N} = N \cup \{\infty\}$, $\pi \in \Pi$, and $\gamma \in \Gamma$.

Under our assumptions I_n is a nonnegative extended real-valued function on $\Pi \times \Gamma \times S$ which is universally measurable in s. The function $I_n(\pi, \gamma)(\cdot)$ is the *total expected n-stage pay-off function* of the initial state when players I and II use strategies π and γ, respectively. For $n = \infty$ we shall write $I(\pi, \gamma)$ instead of $I_n(\pi, \gamma)$.

Let

$$\underline{v}_n = \sup_{\pi\in\Pi}\inf_{\gamma\in\Gamma} I_n(\pi, \gamma) \quad \text{and} \quad \bar{v}_n = \inf_{\gamma\in\Gamma}\sup_{\pi\in\Pi} I_n(\pi, \gamma), \quad n \in \bar{N}.$$

The function $\underline{v}_n(\bar{v}_n)$ is called the *lower* (*upper*) *value function* of the n-stage stochastic game. We say that the n-stage stochastic game has a *value function* v_n if $v_n = \underline{v}_n = \bar{v}_n$. Assume that $\bar{v}_n < \infty$ where $n \in \bar{N}$ is fixed. Let $\varepsilon \geq 0$ be given.

A strategy $\pi^* \in \Pi$ is called *ε-optimal* for player I at the n-stage stochastic game if

$$\bar{v}_n \leq \inf_{\gamma\in\Gamma} I_n(\pi^*, \gamma) + \varepsilon.$$

A strategy $\gamma^* \in \Gamma$ is called ε-optimal for player II at the n-stage stochastic game if

$$\underline{v}_n \geq \sup_{\pi \in \Gamma} I_n(\pi, \gamma^*) - \varepsilon.$$

The 0-optimal strategies are called *optimal*.

We shall write $\underline{v}$, $\bar{v}$, and v instead of $\underline{v}_n$, $\bar{v}_n$, and v_n, respectively, when $n = \infty$.

Putting $P_F(s) = P(F(s))$ and $P_G(s) = P(G(s))$ we define the sets:

$$K_{\mathrm{I}} = \{(s, \mu): s \in S, \mu \in P_F(s)\},$$

$$K_{\mathrm{II}} = \{(s, \lambda): s \in S, \lambda \in P_G(s)\}, \tag{4.3}$$

$$K = \{(s, \mu, \lambda): s \in S, (s, \mu) \in K_{\mathrm{I}}, (s, \lambda) \in K_{\mathrm{II}}\}.$$

Let $M_+(S)$ be the set of all nonnegative universally measurable functions on S. Many of our results will be stated in terms of the following *dynamic programming operators* defined on $M_+(S)$. If $u \in M_+(S)$ and $(s, \mu, \lambda) \in K$, then we define

$$L(s,\mu,\lambda)(u) = \int\int \left[r(s,a,b) + \beta \int u(t) q(dt \mid s,a,b) \right] \mu(da)\lambda(db). \tag{4.4}$$

If $f \in D_A$, $g \in D_B$, and $u \in M_+(S)$, then we put

$$(L_{fg}u)(s) = L(s, f(s), g(s))(u), \quad s \in S, \tag{4.5}$$

$$(U_g u)(s) = \sup_{f \in D_A} (L_{fg}u)(s), \quad s \in S. \tag{4.6}$$

Finally, we define

$$(Uu)(s) = \inf_{g \in D_B} (U_g u)(s), \quad s \in S. \tag{4.7}$$

We shall often omit the variable s in writing the operators (4.5)–(4.7).

5. Main results. The main results of this paper concern the existence and the universal (limit) measurability of the value functions v_n, $n \in \bar{N}$, of the n-stage stochastic games, and existence of ε-optimal strategies for both players.

Let $n \in \bar{N}$ be fixed. Suppose the value function v_n of an n-stage stochastic game exists. Then an important question arises as to whether the value function v_n is universally measurable (cf. [35, Chapter 11] and [36, Problem 6.8]). A negative answer to the above question was recently given by Rieder [38, Example 4.1]. In particular, Rieder, using (F 2.5), showed that it is consistent with the usual axioms of set theory to assume there is an n-stage stochastic game satisfying (i) – (vi) that possesses the value function v_n, but v_n is not universally measurable. Because of the Rieder's example we impose some regularity conditions on the model described in Section 4.

We shall study the stochastic games under the following additional assumption:

(RA) $G(s)$ is compact for each $s \in S$, $r(s, a, \cdot)$ is lower semicontinuous, and $q(E \mid s, a, \cdot)$ is continuous on $G(s)$ for each $(s, a) \in F$ and $E \in \mathscr{B}(S)$.

Because of our concept of ε-optimality, we shall assume that the following statement holds:

(FA) If we consider an n-stage stochastic game ($n \in \bar{N}$), then there is a stationary strategy $g \in D_B$ for player II such that

$$\sup_{\pi \in \Pi} I_n(\pi, g)(s) < \infty \quad \text{for each} \quad s \in S.$$

Before we describe our main results, let us briefly comment on the regularity conditions (RA) and compare them to those appearing in the literature. We require here the strong continuity of the transition law q in the minimizer's actions. However, we do not impose any continuity conditions on q and r with respect to the maximizer's actions and the state variable. Often in the literature q is assumed to be strongly continuous in the action variables of both players. Similarly, the semicontinuity or even continuity condition concerning the payoff function r has been made for both players. For details see [8], [23], [29], [30], [38] and their references. It should be noted that there is an alternative approach to stochastic games with an uncountable state space, where the set H satisfies some additional regularity conditions, the payoff function r is continuous on H, and q: $H \to P(S)$ is continuous on H in the weak topology of $P(S)$. For results in that approach we refer to [8], [23], [38] and the references therein.

Now, let us state our main results. Let u be a nonnegative u.s.a. function on S. First, we consider an auxiliary one stage game with *terminal reward* u. If the players choose $f \in D_A$ and $g \in D_B$, respectively, then the total expected pay-off is given by $L_{fg}u$.

The following result is basic for this paper.

THEOREM 5.1. *Let u be a nonnegative upper semianalytic function on S such that $(Uu)(s) < \infty$ for each $s \in S$. Suppose* (RA) *holds. Then the function $(Uu)(\cdot)$ is the value function of the one stage game with terminal reward u. Moreover, $(Uu)(\cdot)$ is upper semianalytic, player* II *has an optimal limit measurable strategy, and for any $\varepsilon > 0$ player* I *has an ε-optimal limit measurable strategy.*

REMARK 5.1. The conclusions of Theorem 5.1 hold if instead of assuming that r and u are nonnegative, we assume only that r and u are bounded below.

Define

$$O_u = \{s \in S\colon (Uu)(s) = \inf_{g \in P_G(s)} L(s, \mu_s, g(s))(u) \quad \text{for some } \mu_s \in P_F(s)\},$$

where u is any nonnegative u.s.a. function on S.

REMARK 5.2. If $O_u = S$, then (RA) implies that player I has an optimal limit measurable strategy in the one stage game (cf. Lemma 6.1 and the proof of Theorem 5.1).

The next theorem deals with the *finite horizon stochastic games* where $n < \infty$.

THEOREM 5.2. *Assume* (FA) *and* (RA). *Then for any $n \in N$, the n-stage* *stochastic game has a value function* v_n, *the function* v_n *is upper semianalytic, and*

$$v_n = Uv_{n-1} \quad \text{for} \quad n \in N,$$

where $v_0 = 0$ *is the function which vanishes identically.*

Moreover, player II *has an optimal Markov strategy, and for any* $\varepsilon > 0$, *player* I *has an* ε-*optimal Markov strategy.*

REMARK 5.3. From Theorem 5.2 we infer that $v_n = U^n v_0$ for each $n \in N$. Here $U^1 = U$ and $U^n = UU^{n-1}$ for $n \geq 2$.

REMARK 5.4. Let $O_u = S$ for $u = v_0, v_1, \cdots, v_{n-1}$. Then player I has an optimal Markov strategy in the n-stage game, which follows from Remark 5.2 and the proof of Theorem 5.2. One can say that the above condition is implicit and it cannot be checked a priori. Therefore, we would like to indicate some explicit conditions that guarantee the existence of optimal Markov strategies for player I in finite horizon games. Here they are: besides the semicontinuity and compactness assumptions concerning player II we assume that $F(s)$ is compact for each $s \in S$, r is bounded, $r(s, \cdot, b)$ is upper semicontinuous, and $q(E \mid s, \cdot, b)$ is continuous on $F(s)$ for each $(s, b) \in G$ and $E \in \mathscr{B}(S)$. It should also be noted here that under such additional assumptions the value function of any finite horizon game is Borel measurable and optimal Markov strategies for both players in that game may be chosen to be Borel measurable (cf. [29, Theorem 4.1]).

The next theorems concern the *infinite horizon stochastic games* where $n = \infty$. In this case we write I, v instead of I_∞, v_∞, respectively, and so on. First, we consider so-called *discounted stochastic games* where the payoff function r is bounded and $\beta < 1$. In this case the condition (FA) is satisfied trivially.

THEOREM 5.3. *Assume* (RA). *Then the discounted stochastic game has a value function* v, *the function* v *is bounded and upper semianalytic, and* v *is the unique solution of the equation*:

$$v = Uv.$$

Moreover, player II *has an optimal stationary strategy, and for any* $\varepsilon > 0$ *player* I *has an* ε-*optimal stationary strategy.*

REMARK 5.5 If $\mathscr{O}_v = S$, then by the proof of Theorem 5.3 and Remark 5.2 player I has an optimal stationary strategy.

Let $\{\beta_n\}$ be any sequence of real numbers $\beta_n \in [0, 1]$ such that $\beta_n \uparrow \beta$ as $n \to \infty$. We associate with each $n \in N$ an auxiliary discounted stochastic game DSG_n where the immediate payoff function is given by $r_n = \min\{r, n\}$, and the discount factor is equal to β_n. From Theorem 5.3, it follows that each game DSG_n has a bounded upper semianalytic value function w_n.

The next theorems concern so-called *positive stochastic games* where $\beta = 1$, and so-called *absorbing stochastic games* where $\beta < 1$, but the immediate payoff function r is unbounded.

THEOREM 5.4. *Assume* (FA) *and* (RA). *Then the infinite horizon positive or absorbing stochastic game has a value function* v, *the function* v *is upper semianalytic, and*

$$v = Uv = \lim_n w_n = \lim_n v_n,$$

where $w_n(v_n)$ *denotes the value function of the game* DSG_n *(respectively, the n-stage stochastic game).*

Moreover, player II *has an optimal stationary strategy* $g \in D_B$ *satisfying*

$$v = Uv = U_g v,$$

and there is a sequence $\{f_n\} \subset D_A$ *of stationary strategies for player* I *such that*

$$v = \lim_n \inf_{\gamma \in \Gamma} I(f_n, \gamma).$$

REMARK 5.6. Theorem 5.4 implies that for any $\varepsilon > 0$ there is a measurable partition $\{S_k\}$ of the state space S and a subsequence $\{f_{n_k}\}$ of $\{f_n\} \subset D_A$ such that

$$v(s) < \inf_{\gamma \in \Gamma} I(f_{n_k}, \gamma)(s) + \varepsilon \quad \text{for every} \quad s \in S_k, \quad k \in N.$$

A strategy $\pi = \{\pi_n\}$ where $\pi_n(\cdot \mid s_1, a_1, b_1, \cdots, s_n) = f_{n_k}(s_n)$ whenever $s_1 \in S_k$, $n \in N$, is called *semistationary*. Thus, Theorem 5.4 implies that for any $\varepsilon > 0$ player I has an ε-optimal semistationary strategy.

REMARK 5.7. Van der Wal has given an example of the positive stochastic game with countable state space in which player I has no ε-optimal stationary (even Markov) strategies although he has finite action spaces [49, Example 2.26]. The optimal strategies for player I (stationary or otherwise) need not exist in the positive stochastic game even if the state and action spaces are finite [23, Example 1].

REMARK 5.8. All the results of this paper regarding the finite horizon, discounted, and absorbing stochastic games remain true if instead of assuming in the definitions of these games that r is nonnegative, we assume only that r is bounded below.

THEOREM 5.5. *Let* (FA) *and* (RA) *be satisfied. Assume further that the value function of the positive or absorbing stochastic game is bounded and the set*

$$\{q(\cdot \mid s, a, b) : (s, a, b) \in H\} \tag{5.1}$$

is relatively compact in the strong topology in $P(S)$. *Then the sequence* $\{w_n\}$ *of value functions of the games* DSG_n, $n \in N$, *converges to* v *uniformly in* $s \in S$. *Thus, for any* $\varepsilon > 0$ *player* I *has an* ε*-optimal stationary strategy.*

For a detailed discussion of the relative compactness in the space of all probability measures on arbitrary measurable space endowed with the strong topology, we refer to [11, IV 8.9, IV 9.1, 9.2] and [15, Theorem 2.6]. In this paper we would like to point out two simple examples interesting from the point of view of stochastic games in which the set (5.1) is relatively compact in the strong topology of $P(S)$.

EXAMPLE 5.1. Assume that H is a compact subset of $S \times A \times B$, and $P(S)$ is endowed with the strong topology. Then the set (5.1) is compact in $P(S)$ for each continuous q: $H \to P(S)$.

EXAMPLE 5.2. Assume that the set (5.1) is dominated by some probability measure m on $\mathscr{B}(S)$, and the set of Radon-Nikodym derivatives of the measures $q(\cdot \mid s, a, b)((s, a, b) \in H)$ with respect to m is bounded in some Lebesgue space $L^p(S, \mathscr{B}(S), m), p > 1$. Then the set (5.1) is relatively compact in $P(S)$ endowed with the strong topology by [11, IV 8.4]. (This fact can be also verified by means of (F 3.1), the Hölder inequality and the monotone convergence theorem.)

THEOREM 5.6. *The condition* (FA) *holds for each* $n \in \bar{N}$ *if and only if there are* $u \in M_+(S)$ *and* $g \in D_B$ *such that*

$$U_g u \leq u. \tag{5.2}$$

REMARK 5.9. The function u satisfying (5.2) is called *excessive* for U_g. If (5.2) holds for $n \in \bar{N}$, then by the proof of Theorem 5.6 we have

$$\bar{v}_n \leq \sup_{\pi \in \Pi} I_n(\pi, g) \leq u \quad \text{for} \quad n \in \bar{N}.$$

REMARK 5.10. Theorem 5.1 is a game theoretic extension of Theorem 2 of Brown and Purves [7]. Similar minimax selection theorems with the asymmetric assumptions corresponding with the assumptions of the Fan minimax theorem [12, Theorem 2] can be found in [31]. For results related to Theorem 5.2 we refer to [8, Lemma 3.5], [29, Theorem 4.1] and [38, Theorem 6.1]. Theorems 5.3 and 5.4 have predecessors in [29], [30], [38] and in their references. A closely related result to Theorem 5.5 is Theorem 7.2 from [30]. For predecessors of Theorem 5.6 consult [5, Theorem 2] and [23, Theorem 1 (P)].

6. Measurable selections of extrema. Let X and Y be Borel spaces, and $E \subset X \times Y$ be such that $E(x) = \{y \in Y: (x, y) \in E\} \neq 0$ for each $x \in X$.

Let u: $E \to R$ be such that

$$u^*(x) = \sup_{y \in E(x)} u(x, y) < \infty \quad \text{for each} \quad x \in X.$$

Define $\mathscr{O} = \{x \in X: u^*(x) = u(x, y_x) \text{ for some } y_x \in E(x)\}$.

A function f: $X \to Y$ is called an *ε-maximizer* of u if $(x, f(x)) \in E$ for each $x \in X$ and

$$u^*(x) = u(x, f(x)) \quad \text{for} \quad x \in \mathscr{O},$$

and

$$u^*(x) < u(x, f(x)) + \varepsilon \quad \text{for} \quad x \in X - \mathscr{D}.$$

If $\mathscr{D} = X$ then an ε-maximizer of u is called a *maximizer* of u.

We shall need the following results:

LEMMA 6.1 ([46, page 968 and Remark on page 971]). *Assume that E is an analytic set, u is an upper semianalytic function on E. Then u^* is upper semianalytic, $\mathscr{D} \in \mathscr{S}(X)$ (is limit measurable), and for any $\varepsilon > 0$ there is a limit measurable ε-maximizer of u.*

LEMMA 6.2. *Assume that $E \in \mathscr{S}(X) \otimes \mathscr{B}(Y)$, and u is an $\mathscr{S}(X) \otimes \mathscr{B}(Y)$-measurable function. Then u^* is limit measurable, $\mathscr{D} \in \mathscr{S}(X)$, and for any $\varepsilon > 0$ there is a limit measurable ε-maximizer of u.*

PROOF. Note that for each real number c

$$C = \{x \in X: u^*(x) > c\} = \text{proj}_X\{(x, y) \in E: u(x, y) > c\}.$$

By (F 2.7) the set C belongs to $\mathscr{S}(X)$. This obviously proves the measurability of u^*.

Define

$$D_0 = \{(x, y) \in E: u^*(x) = u(x, y)\},$$

and, for any given $\varepsilon > 0$,

$$D = \{(x, y) \in E: u^*(x) < u(x, y) + \varepsilon\} - D_0.$$

It is clear that D_0 and D belong to $\mathscr{S}(X) \otimes \mathscr{B}(Y)$, and $\mathscr{D} = \text{proj}_X D_0$. Now the lemma follows from (F 2.7).

7. Proofs of the main results. We start with some auxiliary lemmas. By means of Proposition 10.1 of Schäl [41] we can easily prove the following:

LEMMA 7.1. *Let X and Y be separable metric spaces and u_n: $X \times Y \to R$, $n \in N$. Assume that $u_n \leq u_{n+1}$, and $u_n(x, \cdot)$ is lower semicontinuous on Y for each $x \in X$, $n \in N$. Then*

(a) $$\lim_n \sup_{x \in X} \inf_{y \in Y} u_n(x, y) = \sup_{x \in X} \inf_{y \in Y} \lim_n u_n(x, y),$$

and

(b) $$\lim_n \inf_{y \in Y} \sup_{x \in X} u_n(x, y) = \inf_{y \in Y} \sup_{x \in X} \lim_n u_n(x, y),$$

provided that Y is a compact metric space.

Let $\tilde{M}_+(S)$ be the set of all nonnegative upper semianalytic functions on S, and $\tilde{B}_+(S)$ be the set of all bounded functions in $\tilde{M}_+(S)$.

LEMMA 7.2. *Assume* (4.1) *and* (RA). *Then*

(a) *The sets* K_{I}, K_{II}, *and* K *defined by* (4.3) *are Borel sets, and* $P_G(s)$ *is compact for each* $s \in S$.

(b) *For any* $u \in \tilde{M}_+(S)$ *the (extended real-valued) function* $L(\cdot, \cdot, \cdot)(u)$ *defined on* K *by* (4.4) *is upper semianalytic.*

(c) *If* r *is bounded and* $r(s, a, \cdot)$ *is continuous on* $G(s)$, $(s, a) \in F$, *and* $u \in \tilde{B}_+(S)$, *then the function* $L(s, \mu, \cdot)(u)$ *is continuous on* $P_G(s)$, $(s, \mu) \in K_{\mathrm{I}}$.

(d) *For any* $u \in \tilde{M}_+(S)$, *the function* $L(s, \mu, \cdot)(u)$ *is lower semicontinuous on* $P_G(s)$, $(s, \mu) \in K_{\mathrm{I}}$.

PROOF. The part (a) follows from (F 2.2) and (F 3.3) – (F 3.5). To prove (b) it is sufficient to use (F 2.2), (F 2.3), (F 3.7), and (F 3.8). The part (c) follows immediately from (F 3.9). For proving (d), let $u_n = \min\{u, n\}$, $n \in N$. Then by (F 3.6) and (F 3.9) each function $L(s, \mu, \cdot)(u_n)$ is lower semicontinuous on $P_G(s)$, $n \in N$, and by the monotone convergence theorem

$$L(s, \mu, \cdot)(u_n) \uparrow L(s, \mu, \cdot)(u), \quad (s, \mu) \in K_{\mathrm{I}}.$$

This obviously implies (d).

The following lemma is apparent from (4.2) and (4.5).

LEMMA 7.3. *Let* $\pi = \{f_n\} \in \Pi_M$ *and* $\gamma = \{g_n\} \in \Gamma_M$. *For any* $u \in M_+(S)$ *we define a sequence*

$$L_{f_1 g_1} L_{f_2 g_2} \cdots L_{f_n g_n} u, \quad n \in N. \tag{7.1}$$

The sequence (7.1) *is convergent to* $I(\pi, \gamma)$ *(uniformly in* $s \in S$*) if* $u = 0$ *(if* r *and* u *are bounded and* $\beta < 1$*).*

LEMMA 7.4. *Let* $f \in D_A$ *and* $g \in D_B$. *Then for any* $n \in \bar{N}$, $s \in S$, $\pi \in \Pi$, *and* $\gamma \in \Gamma$ *there are* $\pi' \in \Pi_M$ *and* $\gamma' \in \Gamma_M$ *such that*

$$I_n(f, \gamma)(s) = I_n(f, \gamma')(s) \quad \text{and} \quad I_n(\pi, g)(s) = I_n(\pi', g)(s).$$

PROOF. The proof follows similar lines as that of Proposition 1 of Shreve and Bertsekas from [47].

REMARK 7.1. We do not assume (FA) in Lemma 7.4.

Now we are ready to prove the theorems.

PROOF OF THEOREM 5.1. The fact that Uu is the value function of the one stage game with terminal reward u follows from the compactness of sets $P_G(s)$, $s \in S$, Lemma 7.2(d), and the Fan minimax theorem [12, Theorem 2].

Define

$$\Phi(s, \mu) = \inf_{\lambda \in P_G(s)} L(s, \mu, \lambda)(u), \quad (s, \mu) \in K_{\mathrm{I}}.$$

Note that

$$(Uu)(s) = \sup_{\mu \in P_F(s)} \Phi(s, \mu), \quad s \in S.$$

To prove that Uu is u.s.a., and player I has an ε-optimal limit measurable strategy for each $\varepsilon > 0$, it is sufficient to show that Φ is u.s.a. on the Borel space K_{I} and apply Lemma 6.1. In order to show that Φ is u.s.a. on K_{I}, we construct an auxiliary sequence $\{\Phi_n\}$ of u.s.a. functions on K_{I} that converges to Φ. Thus, Φ becomes a u.s.a. function on K_{I}.

Let

$$\varphi_n = \min\{\psi_n, n\}, \quad n \in N, \tag{7.2}$$

where

$$\psi_n(s, a, b) = \inf_{y \in G(s)}[r(s, a, y) + nd(b, y)], \quad (s, a, b) \in F \times B, \quad n \in N,$$

and d is the metric in B.

By [19, Theorem 2], ψ_n is a Borel measurable function on $F \times B$, and so is φ_n, $n \in N$. It is easy to check that $\varphi_n(s, a, \cdot)$ is continuous on B for each $(s, a) \in F$, $n \in N$. By the proof of the theorem of Baire [1, page 390], $\psi_n \uparrow r$ on H. Hence $\varphi_n \uparrow r$ on H.

Let $L_n(\cdot, \cdot, \cdot)(u_n)$ be defined by (4.4) where the function r is replaced by φ_n and u is replaced by $u_n = \min\{u, n\}$. Clearly, the facts listed in Lemma 7.2 for $L(\cdot, \cdot, \cdot)(u)$ carry over to $L_n(\cdot, \cdot, \cdot)(u_n)$.

Define

$$\Phi_n(s, \mu) = \inf_{\lambda \in P_G(s)} L_n(s, \mu, \lambda)(u_n), \quad (s, \mu) \in K_{\mathrm{I}}, \ n \in N.$$

Because $\varphi_n \uparrow r$ on H, and $u_n \uparrow u$ on S, from the monotone convergence theorem we get

$$L_n(\cdot, \cdot, \cdot)(u_n) \uparrow L(\cdot, \cdot, \cdot)(u) \quad \text{on } K. \tag{7.3}$$

This fact, the compactness of $P_G(s)$, $s \in S$, Lemma 7.2(c), and Proposition 10.1 of [41] imply that $\Phi_n \uparrow \Phi$ on K_{I}. Now, it remains to show that Φ_n is u.s.a. on K_{I} for each $n \in N$.

By Lemma 7.2(a) and (F 2.1) there is a sequence $\{g_k\}$ of Borel measurable mappings $g_k: S \to P(B)$ such that

$$P_G(s) = \mathrm{cl}\{g_k(s)\}, \quad \text{for each} \quad s \in S, \tag{7.4}$$

where cl denotes the closure in the weak topology on $P(B)$. This together with Lemma 7.2(c) implies

$$\Phi_n(s, \mu) = \inf_k L_n(s, \mu, g_k(s))(u_n), \quad (s, \mu) \in K_I, \quad n \in N. \tag{7.5}$$

From the Borel measurability of g_k and (F 3.8), we infer that the function

$$(s, a) \to L_n(s, \mu_a, g_k(s))(u_n), \quad (s, a) \in F,$$

where $\mu_a(\{a\}) = 1$, is u.s.a. on F for each $k, n \in N$. Using this fact and (F 3.7) we can easily show that $L_n(\cdot, \cdot, g_k(\cdot))(u_n)$ is u.s.a. on K_{I}, for each $k, n \in N$, which together with (7.5) and (F 2.4) (used for the intersection) implies that so is Φ_n, $n \in N$.

For proving that player II has an optimal limit measurable strategy we define the function

$$\Psi(s, \lambda) = \sup_{\mu \in P_F(s)} L(s, \mu, \lambda)(u), \quad (s, \lambda) \in K_{\mathrm{II}}.$$

Note that

$$(Uu)(s) = \inf_{\lambda \in P_G(s)} \Psi(s, \lambda) = \Psi(s, \lambda_s) \tag{7.6}$$

for each $s \in S$ and some $\lambda_s \in P_G(s)$. The last equality follows from the compactness of the sets $P_G(s)$, $s \in S$, and Lemma 7.2(d). We shall show that Ψ is an $\mathscr{S}(S) \otimes \mathscr{B}(P(B))$-measurable function on K_{II} (being Borel space). Then the existence of the required strategy for player II follows immediately from (7.6) and Lemma 6.2.

In order to prove the measurability of Ψ we use the following sequences of functions:

$$\Psi_n(s, \lambda) = \sup_{\mu \in P_F(s)} L_n(s, \mu, \lambda)(u_n), \quad (s, \lambda) \in K_{\mathrm{II}},\ n \in N,$$

and

$$\Delta_{nm}(s, \lambda) = \inf_{\eta \in P_G(s)} [\Psi_n(s, \eta) + m\rho(\eta, \lambda)], \quad n, m \in N,$$

where $(s, \lambda) \in S \times P(B)$, and ρ is the metric on $P(B)$ defined according to (3.1). Let $n, m \in N$ be arbitrary. Note that $\Delta_{nm}(s, \cdot)$ is continuous on $P(B)$ for each $s \in S$. We shall prove that $\Delta_{nm}(\cdot, \lambda)$ is u.s.a. on S for each $\lambda \in P(B)$.

Denote

$$w_{nm}(s, \mu, \eta, \lambda) = L_n(s, \mu, \eta)(u_n) + m\rho(\eta, \lambda),$$

where $(s, \mu, \eta) \in K$ and $\lambda \in P(B)$. From the properties of $L_n(s, \cdot, \cdot)(u_n)$ and (3.1) we infer that $w_{nm}(s, \cdot, \eta, \lambda)$ is linear on a convex set $P_F(s)$ and $w_{nm}(s, \mu, \cdot, \lambda)$ is convex and continuous on $P_G(s)$ being compact convex space. Applying the Fan minimax theorem [12, Theorem 2] to the function $w_{nm}(s, \cdot, \cdot, \lambda)$ we get:

$$\Delta_{nm}(s, \lambda) = \sup_{\mu \in P_F(s)} \inf_{\eta \in P_G(s)} w_{nm}(s, \mu, \eta, \lambda), \quad (s, \lambda) \in S \times P(B). \tag{7.7}$$

It is clear that $w_{nm}(\cdot, \cdot, \cdot, \lambda)$ is u.s.a. on K, and since $w_{nm}(s, \mu, \cdot, \lambda)$ is continuous on $P_G(s)$, $(s, \mu) \in K_{\mathrm{I}}$, so using the sequence $\{g_k\}$ satisfying (7.4) we can show that the function

$$(s, \mu) \longrightarrow \inf_{\eta \in P_G(s)} w_{nm}(s, \mu, \eta, \lambda) \quad \text{is u.s.a. on } K_{\mathrm{I}}.$$

This fact together with (7.7) and Lemma 6.1 implies that $\Delta_{nm}(\cdot, \lambda)$ is u.s.a. on S. Thus, we have shown that $\Delta_{nm}(\cdot, \lambda)$ is $\mathscr{S}(S)$-measurable on S for each $\lambda \in P(B)$, and $\Delta_{nm}(s, \cdot)$ is continuous on $P(B)$ for each $s \in S$. By [17, Theorem 6.1], the function Δ_{nm} is $\mathscr{S}(S) \otimes \mathscr{B}(P(B))$-measurable on $S \times P(B)$ because $P(B)$ endowed with the weak topology is a separable metric space.

Now observe that $\Psi_n(s, \cdot)$ is lower semicontinuous on $P_G(s)$ for each $s \in S$,

$n \in N$. By the proof of the theorem of Baire [1, page 390] we obtain $\Delta_{nm} \uparrow \Psi_n$ as $m \to \infty$. Hence, it follows that Ψ_n is $\mathscr{S}(S) \otimes \mathscr{B}(P(B))$-measurable on K_{II} for each $n \in N$, and from (7.3) we can easily derive that $\Psi_n \uparrow \Psi$ as $n \to \infty$. Thus, Ψ is also an $\mathscr{S}(S) \otimes \mathscr{B}(P(B))$-measurable function on K_{II}, which terminates the proof.

REMARK 7.2. The Borel measurability of the functions $\{g_k\}$ is very important in the proof of Theorem 5.1 because of the fact that the composition of two analytically measurable functions need not be analytically measurable (cf. [3, page 187] or [6, Example 24]).

PROOF OF THEOREM 5.2. The proof is based on Theorem 5.1 and proceeds along similar lines as that of Theorem 4.1 in [29], or Lemma 3.5 in [8].

PROOF OF THEOREM 5.3. Assume that $\tilde{B}_+(S)$ is endowed with the supremum metric. Then $\tilde{B}_+(S)$ is a complete metric space. With the help of Theorem 5.1 we can easily show that U is a contraction mapping from $\tilde{B}_+(S)$ into $\tilde{B}_+(S)$. By the Banach fixed point theorem there is unique function v^* in $\tilde{B}_+(S)$ such that $v^* = Uv^*$. We shall prove that v^* is the value function of the discounted stochastic game.

Let $\varepsilon > 0$ be arbitrary. By Theorem 5.1 there are $f^* \in D_A$ and $g^* \in D_B$ such that

$$L_{fg^*}v^* \le v^* = Uv^* \le L_{f^*g}v^* + \varepsilon(1-\beta), \tag{7.8}$$

for every $f \in D_A$ and $g \in D_B$.

Let $\pi' = \{f_n\} \in \Pi_M$ and $\gamma' = \{g_n\} \in \Gamma_M$ be any Markov strategies for players I and II. By means of (7.8) we can show that, for each n,

$$\begin{aligned} L_{f_1g^*}L_{f_2g^*} \cdots L_{f_ng^*}v^* &\le v^* \\ &= Uv^* \le L_{f^*g_1}L_{f^*g_2}L_{f^*g_n}v^* + \varepsilon(1-\beta)(1+\beta+\cdots+\beta^{n-1}). \end{aligned}$$

This and Lemma 7.3 imply

$$I(\pi', g^*) \le v^* = Uv^* \le I(f^*, \gamma') + \varepsilon. \tag{7.9}$$

Since π' and γ' are arbitrary, so from (7.9) and Lemma 7.4 we infer

$$\bar{v} \le \sup_{\pi\in\Pi} I(\pi, g^*) \le v^* = Uv^* \le \inf_{\gamma\in\Gamma} I(f^*, \gamma) + \varepsilon \le \underline{v} + \varepsilon.$$

This implies that the value function v of the game exists and $v = v^*$. At the same time we have shown that player II has an optimal stationary strategy g^* and player I has an ε-optimal stationary strategy f^*, which completes the proof.

REMARK 7.3. If there are Borel measurable functions $f^* \in D_A$ and $g^* \in D_B$ which satisfy (7.8), then the rest part of the proof of Theorem 5.3 after (7.8) can be reduced to a direct application of dynamic programming results of Shreve and Bertsekas [47, Theorems 1 and 2(D)]. However, f^* and g^* need not be Borel measurable [6, Example 47], so we use slight different arguments.

PROOF OF THEOREM 5.4. Let L_n and U_n denote the operators defined by (4.4) and (4.7), respectively, where β is replaced by β_n and r is replaced by $r_n = \min\{r, n\}$.

By Theorem 5.3, we have $w_n = U_n w_n$ where w_n is the value function of the game DSG$_n$. Clearly, $w_n \leq w_{n+1} \leq \underline{v}$ for each $n \in N$. Hence

$$0 \leq w = \lim_n w_n \leq \underline{v}, \tag{7.10}$$

and by (FA) and Theorem 5.3 we have $w \in \tilde{M}_+(S)$.

From the monotone convergence theorem, Lemma 7.2(a), (d) (applied to L_n), and Lemma 7.1 we infer that

$$w = \lim_n w_n = \lim_n U_n w_n = U \lim_n w_n = Uw.$$

By Theorem 5.1, there is $g \in D_B$ such that

(7.11) $w = Uw = U_g w$, where U_g is the operator defined by (4.6).

Let $U_g^1 = U_g$ and $U_g^n = U_g U_g^{n-1}$ for $n \geq 2$. Then from (7.11) we get

$$w = U_g^n w \quad \text{for each} \quad n \in N.$$

Note that

$$w = U_g^n w \geq I_n(\pi', g)$$

for each $n \in N$ and each $\pi' \in \Pi_M$. Hence

$$w \geq \lim_n I_n(\pi', g) = I(\pi', g)$$

for each $\pi' \in \Pi_M$. This and Lemma 7.4 imply that

$$w \geq \sup_{\pi' \in \Pi_M} I(\pi', g) = \sup_{\pi \in \Pi} I(\pi, g) \geq \bar{v}. \tag{7.12}$$

Combining (7.10), (7.11), and (7.12) we get

$$w = Uw = U_g w = \underline{v} = \bar{v} = v \tag{7.13}$$

for some $g \in D_B$. This and (7.12) imply also that g is an optimal stationary strategy for player II.

The proof of $v = \lim_n v_n$ follows similar lines as that of (7.13), but we use Theorem 5.2 instead of Theorem 5.3.

Now, let $\varepsilon_n = 1/n$, $n \in N$. By Theorem 5.3, for each $n \in N$, there is an ε_n-optimal stationary strategy f_n for player I in the discounted stochastic game DSG$_n$. It is easy to check that $\{f_n\}$ is the required sequence of stationary strategies for player I.

PROOF OF THEOREM 5.5. Suppose that $|v(s)| \leq c$ for all $s \in S$ and some $c > 0$. Then for each $s \in S$ we have

$$\begin{aligned} |v(s) - w_n(s)| &\leq c(\beta - \beta_n) \\ &\quad + \sup_{(s,a,b)\in H} \int \left[v(t) - w_n(t) \right] q(dt \mid s, a, b). \end{aligned} \tag{7.14}$$

By Theorem 5.3, $(v - w_n) \downarrow 0$, so the theorem follows from (7.14) and (F 3.1).

PROOF OF THEOREM 5.6. If (FA) holds for each $n \in \bar{N}$, then (5.2) follows from Theorem 5.4. Suppose that (5.2) holds for some $u \in M_+(S)$ and $g \in D_B$. Then

$$u \geq L_{fg}u \quad \text{for each} \quad f \in D_A. \tag{7.15}$$

Let $\pi' = \{f_n\} \in \Pi_M$ be arbitrary Markov strategy for player I. From (7.15) we obtain

$$u \geq L_{f_1 g} L_{f_2 g} \cdots L_{f_n g} u \geq I_n(\pi', g) \quad \text{for each} \quad n \in N.$$

Hence

$$u \geq \lim_n I_n(\pi', g) = I(\pi', g), \quad \pi' \in \Pi_M.$$

This and Lemma 7.4 imply that (FA) holds, which completes the proof.

REFERENCES

[1] ASH, R. B. (1972). *Real Analysis and Probability.* Academic, New York.

[2] BARBAROSHIE, A. (1977). On the theory of controlled Markov processes. *Theory Probab. Appl.* **22** 53–69.

[3] BERTSEKAS, D. P. and SHREVE, S. E. (1978). *Stochastic Optimal Control: The Discrete Time Case.* Academic, New York.

[4] BLACKWELL, D. (1965). Discounted dynamic programming. *Ann. Math. Statist.* **36** 226–235.

[5] BLACKWELL, D. (1967). Positive dynamic programming. *Proc. Fifth Berkeley Symp. Math. Statist. Probab.* **2** 415–418. Univ. of California Press.

[6] BLACKWELL, D., FREEDMAN, D. and ORKIN, M. (1974). The optimal reward operator in dynamic programming. *Ann. Probab.* **2** 926–941.

[7] BROWN, L. D. and PURVES, R. (1973). Measurable selections of extrema. *Ann. Statist.* **1** 902–912.

[8] COUWENBERG, H. A. M. (1980). Stochastic games with metric state space. *Internat. J. Game Theory* **9** 25–36.

[9] DESHMUKH, S. D. and WINSTON, W. (1978). A zero-sum stochastic game model of duopoly. *Internat. J. Game Theory* **7** 19–30.

[10] DUBINS, L. E. and FREEDMAN, D. (1965). Measurable sets of measures. *Pacific J. Math.* **14** 1211–1222.

[11] DUNFORD, N. and SCHWARTZ, J. T. (1958). *Linear Operators, Part I: General Theory.* Interscience, New York.

[12] FAN, K. (1953). Minimax theorems. *Proc. Nat. Acad. Sci. U.S.A.* **39** 42–47.

[13] FRID, E. B. (1973). On stochastic games. *Theory Probab. Appl.* **18** 389–393.

[14] FRIEDMAN, J. W. (1977). *Oligopoly and the Theory of Games.* North Holland, Amsterdam.

[15] GÄNSSLER, P. (1971). Compactness and sequential compactness in spaces of measures. *Z. Wahrsch. verw. Gebiete* **17** 124–146.

[16] GÖDEL, K. (1938). The consistency of the axiom of choice and of the generalized continuum hypothesis. *Proc. Nat. Acad. Sci. U.S.A.* **24** 556–557.

[17] HIMMELBERG, C. J. (1975). Measurable relations. *Fund. Math.* **87** 53–72.

[18] HIMMELBERG, C. J., PARTHASARATHY, T. RAGHAVAN, T. E. S. and VAN VLECK, F. S. (1976). Existence of p-equilibrium and optimal stationary strategies in stochastic games. *Proc. Amer. Math. Soc.* **60** 245–251.

[19] HIMMELBERG, C. J., PARTHASARATHY, T. and VAN VLECK, F. S. (1976). Optimal plans for dynamic programming problems. *Math. Oper. Res.* **1** 390–394.

[20] HINDERER, K. (1970). Foundations of non-stationary dynamic programming with discrete-time parameter. *Lecture Notes in Operations Research and Mathematical Systems* (M. Beckman and H. P. Künzi, eds.) **33**. Springer, New York.

[21] IDZIK, A. (1979). Remarks on discounted stochastic games. *Trans. Eighth Prague Conf. Information Theory, Statist. Decision Functions, Random Processes. Vol. C*, Academia, Prague.

[22] KIRMAN, A. P. and SOBEL, M. J. (1974). Dynamic oligopoly with inventories. *Econometrica* **42** 279–287.

[23] KUMAR, P. R. and SHIAU, T. H. (1981). Existence of value and randomized strategies in zero-sum discrete-time stochastic dynamic games. *SIAM J. Control Optim.* **19** 617–634.

[24] KURATOWSKI, K. (1966). *Topology I.* Academic, New York.

[25] KÜENLE, H. U. (1982). On ε-optimal strategies in discounted Markov games. *Banach Center Publications, Ser. Math. Control Theory.* To appear.

[26] LEESE, S. J. (1978). Measurable selections and the uniformization of Souslin sets. *Amer. J. Math.* **100** 19–41.

[27] MAITRA, A. and PARTHASARATHY, T. (1970). On stochastic games. *J. Optim. Theory Appl.* **5** 289–300.

[28] MAITRA, A. and PARTHASARATHY, T. (1971). On stochastic games II. *J. Optim. Theory Appl.* **8** 154–160.

[29] NOWAK, A. S. (1984). On zero-sum stochastic games with general state space I. *Probab. Math. Statist.* **4** 13–32.

[30] NOWAK, A. S. (1984). On zero-sum stochastic games with general state space II. *Probab. Math. Statist.* **4** 143–152.

[31] NOWAK, A. S. (1984). Minimax selection theorems. *J. Math. Anal. Appl.* **103** 106–116.

[32] NUNEN VAN, J. and WESSELS, J. (1979). Successive approximations for Markov decision processes and Markov games with unbounded rewards. *Math. Oper. Statist., Ser. Optim.* **10** 431–455.

[33] PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces.* Academic, New York.

[34] PARTHASARATHY, T. (1973). Discounted, positive and non-cooperative stochastic games. *Internat. J. Game Theory* **2** 25–37.

[35] PARTHASARATHY, T. and RAGHAVAN, T. E. S. (1971). *Some Topics in Two Person Games.* American Elsevier, New York.

[36] PARTHASARATHY, T. and STERN, M. (1977). Markov games—a survey. In *Differential Games and Control Theory II.* (E. Roxin, P. Liu and R. Sternberg, eds.). Dekker, New York.

[37] RIEDER, U. (1975). Bayesian dynamic programming. *Adv. Appl. Probab.* **7** 330–348.

[38] RIEDER, U. (1978). On semi-continuous dynamic games. Preprint Univ. Karlsruhe.

[39] ROYDEN, H. L. (1968). *Real Analysis.* MacMillan, London.

[40] SAKS, S. (1937). *Theory of the Integral.* Stechert, New York.

[41] SCHÄL, M. (1975). Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal. *Z. Wahrsch. verw. Gebiete* **32** 179–196.

[42] SELIVANOVSKIJ, E. (1928). Ob odnom klasse effektivnyh mnozestv (mnozestva C). *Mat. Sb.* **35** 379–413.

[43] SHAPLEY, L. S. (1953). Stochastic games. *Proc. Nat. Acad. Sci. U.S.A.* **39** 1095–1100.

[44] SHREVE, S. E. (1977). Dynamic programming in complete separable spaces. Ph.D. thesis, Dept. of Math., Univ. of Illinois, Urbana, IL.

[45] SHREVE, S. E. (1978). Probability measures and the C-sets of Selivanovskij. *Pacific J. Math.* **79** 189–196.

[46] SHREVE, S. E. and BERTSEKAS, D. P. (1978). Alternative theoretical frameworks for finite horizon discrete-time stochastic optimal control. *SIAM J. Control Optim.* **16** 953–978.

[47] SHREVE, S. E. and BERTSEKAS, D. P. (1979). Universally measurable policies in dynamic programming. *Math. Oper. Res.* **4** 15–30.

[48] STRAUCH, R. E. (1966). Negative dynamic programming. *Ann. Math. Statist.* **37** 871–890.

[49] WAL VAN DER, J. (1981). Stochastic dynamic programming. *Math. Centre Tract* **139**. Mathematisch Centrum Amsterdam.

INSTITUTE OF MATHEMATICS
TECHNICAL UNIVERSITY OF WROCŁAW
WYBRZEŻE WYSPIAŃSKIEGO 27
50-370 WROCŁAW, POLAND