PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. IV, Fasc. 1 (1984), p. 13-32

ON ZERO-SUM STOCHASTIC GAMES WITH GENERAL STATE SPACE. I

BY

ANDRZEJ S. NOWAK (WROCŁAW)

Abstract. The present paper treats of discrete-time stationary models of stochastic games with an abstract measurable state space and separable metric action spaces. Under different assumptions on the state space, action spaces, the reward function, and the law of motion (assumptions (M_1), (M_2), (M_3)), a full solution of the finite horizon models is given. To ensure the existence of value in the infinite horizon models we impose some convergence conditions (conditions (D) and (P)) on the expected rewards, thus including the discounted case. The proofs of the existence of optimal (or ε-optimal) strategies for both players rely on iterative, finite horizon to infinite horizon methods.

Introduction. Stochastic games were introduced by Shapley [26] who assumed that the state space S and the action spaces X and Y of players I and II, respectively, are finite sets. He also assumed that each play would terminate in a finite number of stages with probability one and he considered stationary strategies only. Maitra and Parthasarathy studied in [17] and [18] two kinds of stochastic games with an infinite number of stages: discounted stochastic games where the reward function is bounded and there is discounting and positive stochastic games where the reward function is bounded and non-negative. In [17] and [18] Maitra and Parthasarathy were the first ones to consider the discounted stochastic games under the assumption that S, X, and Y are uncountable sets. More precisely, they assumed that S, X, and Y are compact metric spaces and the reward function and the law of motion satisfy some continuity requirements. More general results for discounted stochastic games with a standard Borel space of states and compact metric spaces of actions may be found in [3], [8], and [12]. Models of positive stochastic games with finite action spaces and a standard

Borel state space were first studied by Frid [5] and subsequently by Himmelberg et al. [8]. Stochastic games with unbounded reward function, metric state space and compact metric spaces of actions and under a convergence assumption which is more general than discounting were recently studied by Couwenbergh in [3]. He imposed similar continuity assumptions on the reward function and on the law of motion as Himmelberg et al. in [8], Idzik in [12], and Parthasarathy in [22].

The aim of this paper is to study stochastic games under a similar convergence assumption as in [3] (cf. Section 2, condition (D)) and positive stochastic games under a natural integrability condition (cf. Section 2, condition (P)). We consider three models ((M_1), (M_2), (M_3)) of stochastic games which differ themselves in the assumptions imposed on the state space, the action spaces, the reward function, and the law of motion. From the point of view of the assumptions involved these models are new. For each of the models (M_1), (M_2), (M_3) of the game with finite number of stages we obtain the existence of value and optimal Markov strategies for both players (Theorem 4.1). In Section 5 we show that under the assumption (D) the stochastic game has a value (Theorem 5.1) and both players have optimal stationary strategies (Theorem 5.4). The results obtained generalize and complete numerous earlier results on discounted stochastic games. More detailed discussion will be presented in the sequel. The study of positive stochastic games is postponed to part II of this paper [20].

The basic tools of this paper are topologies on spaces of probability measures, semi-continuous functions, multifunctions, measurable selections, and the uniformization of Borel and analytic sets (see Section 1).

1. Preliminaries. Throughout this section, we assume $(S, \mathscr{S})$ to be a measurable space and X to be a separable metric space endowed with the σ-algebra $\mathscr{B}(X)$ of its Borel subsets. Let R denote the set of real numbers. We write $M(S)$ for the set of all measurable functions $u\colon S \to R$, and $B(S)$ for the set of all functions in $M(S)$ which are bounded. By $C(X)$ we denote the set of all continuous functions in $B(X)$ and by $\tilde{C}(X)$ ($\underset{\sim}{C}(X)$) the set of all upper semi-continuous (lower semi-continuous) functions in $B(X)$.

We shall need the following

LEMMA 1.1 (cf. [25], Proposition 10.1). *Let $u_n \in \underset{\sim}{C}(X)$, $n \geqslant 1$, and let $u_n(x) \geqslant u_m(x)$ for $n \geqslant m$ and $x \in X$. Assume that the limit*

$$u(x) := \lim_n u_n(x)$$

exists for each $x \in X$. Then $u \in \underset{\sim}{C}(X)$ and, for each $x \in X$,

$$\lim_n \inf_{x \in X} u_n(x) = \inf_{x \in X} \lim_n u_n(x)$$

provided that X is compact.

Denote by $P(S)$ the set of all probability measures on $(S, \mathscr{S})$. Given any $\mathscr{F} \subset B(S)$ we may endow $P(S)$ with the $\mathscr{F}$-topology defined as the coarsest topology in which all mappings $\mu \to \int u d\mu$, $u \in \mathscr{F}$, are continuous. A net $\{\mu_\alpha\}$ will converge to μ_0 in the $\mathscr{F}$-topology if and only if $\int u d\mu_\alpha \to \int u d\mu_0$ for each $u \in \mathscr{F}$. In the case of $\mathscr{F} = C(X)$, the $\mathscr{F}$-topology on $P(X)$ is called the *w-topology* (*weak topology*). In the case of $\mathscr{F} = B(S)$, the $\mathscr{F}$-topology on $P(S)$ is called the *s-topology* (*strong topology*).

From the theorem of Dini and Theorem 2.6 of [6] we obtain the following

LEMMA 1.2 *For any $D \subset P(S)$, the following statements are equivalent:*
(i) *D is relatively compact in the s-topology;*
(ii) *For any sequence $\{u_n\}$ in $B(S)$ which decreases to 0,*

$$\int u_n d\mu \to 0 \text{ as } n \to \infty, \text{ uniformly in } \mu \in D.$$

For any separable metric space X, we shall assume that the set $P(X)$ is endowed with the w-topology. Then $P(X)$ is also a separable metric space (cf. [21], Theorem II 6.2).

LEMMA 1.3. *The σ-algebra $\mathscr{B}(P(X))$ of all Borel subsets of $P(X)$ coincides with the smallest σ-algebra on $P(X)$ such that, for each $E \in \mathscr{B}(X)$, the mapping $\mu \to \mu(E)$ is measurable.*

A proof of this fact can be found in [24].

A separable metric space X is called a *Polish space* if X is complete. X is called an *SB-space* (*standard Borel space*) if X is a Borel subset of a Polish space. X is called an *analytic space* if X is a continuous image of a Polish space. For a detailed discussion see [14].

We have the following facts:

(1.1) if X is Polish, then $P(X)$ is Polish (cf. [21], Theorem II 6.5);

(1.2) if X is compact, then $P(X)$ is compact (cf. [21], Theorem II 6.4);

(1.3) if X is an SB-space, then $P(X)$ is an SB-space (cf. [10], p. 91);

(1.4) if X is analytic, so is $P(X)$ (cf. [2], p. 933).

The following lemma can be obtained directly from the theorem of Baire and the bounded convergence theorem.

LEMMA 1.4. *If $u \in \tilde{C}(X)$, then the mapping $\mu \to \int u d\mu$ defined on $P(X)$ is upper semi-continuous.*

Using Lemma 1.4 and Fatou's lemma we obtain

LEMMA 1.5. *Let X and Y be separable metric spaces. Let $u\colon X \times Y \to R$ be a bounded Borel measurable function such that $u(\cdot, y) \in \tilde{C}(X)$ for each $y \in Y$ and $u(x, \cdot) \in \underset{\sim}{C}(Y)$ for each $x \in X$. Then the mapping $\hat{u}\colon P(X) \times P(Y) \to R$ defined by*

$$\hat{u}(\mu, \lambda) = \int_X \int_Y u(x, y)\, d\mu d\lambda$$

is such that

$\hat{u}(\mu, \cdot) \in \underset{\sim}{C}(P(Y))$ *for each* $\mu \in P(X)$ *and* $\hat{u}(\cdot, \lambda) \in \tilde{C}(P(X))$ *for each* $\lambda \in P(Y)$.

LEMMA 1.6. *Let X and Y be separable metric spaces. Let $(S, \mathscr{S})$ be a measurable space. Assume that u: $S \times X \times Y \to R$ is an $(\mathscr{S} \otimes \mathscr{B}(X) \otimes \mathscr{B}(Y))$-measurable function such that $u(s, \cdot, \cdot)$ is bounded for each $s \in S$. Then the mapping $\hat{u}$: $S \times P(X) \times P(Y) \to R$ defined by*

$$\hat{u}(s, \mu, \lambda) = \int_X \int_Y u(s, x, y)\, d\mu\, d\lambda$$

is an $(\mathscr{S} \otimes \mathscr{B}(P(X)) \otimes \mathscr{B}(P(Y)))$-measurable function.

Proof. If u is the characteristic function of the set $E_1 \times E_2 \times E_3$, where $E_1 \in \mathscr{S}$, $E_2 \in \mathscr{B}(X)$, $E_3 \in \mathscr{B}(Y)$, then the lemma follows immediately from Lemma 1.3. The rest of the proof is similar to that of Lemma 12.2 in [10].

Denote by $\mathscr{P}(X)$, $\mathscr{F}(X)$, and $\mathscr{K}(X)$ the sets of all non-empty, all non-empty closed, and all non-empty compact subsets, respectively, of a separable metric space X.

A function F: $S \to \mathscr{P}(X)$ is called a *multifunction*. A multifunction F is said to be *measurable* if

$$F^{-1}(U) = \{s \in S\colon F(s) \cap U \neq \emptyset\} \in \mathscr{S}$$

whenever $U \subset X$ is open.

A *selection* of F is a function f: $S \to X$ such that $f(s) \in F(s)$ for each $s \in S$. If F: $S \to \mathscr{P}(X)$ is a measurable multifunction with complete values, then by the theorem of Kuratowski and Ryll-Nardzewski F admits a measurable selection (cf. [15], Theorem 1). We denote by $\mathscr{S}_F$ the set of all measurable selections of a multifunction F.

The graph of F: $S \to \mathscr{P}(X)$ is defined by

$$\mathrm{Gr}\, F = \{(s, x) : x \in F(s)\}.$$

We say that F has a *measurable graph* if $\mathrm{Gr}\, F \in \mathscr{S} \otimes \mathscr{B}(X)$.

LEMMA 1.7 (cf. [9], Theorem 3). *Assume that X and Y are SB-spaces. Let F: $X \to \mathscr{K}(Y)$. Then the following statements are equivalent*:

(i) *F has a measurable graph*;

(ii) *F is measurable.*

LEMMA 1.8 (cf. [28], Theorem 4.2). *Let $(S, \mathscr{S})$ be a measurable space such that $\mathscr{S}$ is closed with respect to the Suslin operation. Let X be a metric analytic space and F: $S \to \mathscr{F}(X)$. Then the following statements are equivalent*:

(i) *F has a measurable graph*;

(ii) *F is measurable.*

LEMMA 1.9 (cf. [16], Theorem 5.5). *Let $(S, \mathscr{S})$ be a measurable space such that $\mathscr{S}$ is closed with respect to the Suslin operation. Let X be an analytic* *space and let $E \in \mathscr{S} \otimes \mathscr{B}(X)$. Then* $\mathrm{proj}_S E \in \mathscr{S}$, *where* proj_S *is the projection from $S \times X$ into S.*

LEMMA 1.10. *Let $(S, \mathscr{S})$ be a measurable space and X a separable metric space. Let $F: S \to \mathscr{K}(X)$ be a measurable multifunction. Assume that $u: S \times X \to R$ is an $(\mathscr{S} \otimes \mathscr{B}(X))$-measurable function such that $u(s, \cdot) \in C(F(s))$*(1) *for each $s \in S$ and put*

$$H(s) = \{x \in F(s): u(s, x) \leqslant w(s)\} \neq \emptyset \quad \textit{for each } s \in S,$$

where $w \in M(S)$. Then H is measurable.

Proof. Denote by cl the closure in X. Since F has closed (even compact) values, by Theorem 5.6 of [7] there is a sequence $\{f_m\}$ of measurable selections of F such that $F(s) = \mathrm{cl}\{f_m(s)\}$ for all $s \in S$.

Now, for each $n \geqslant 1$ we define a sequence of sets

$$E_{nm} = \{s: u(s, f_m(s)) < w(s) + 1/n\}.$$

Since u is an $(\mathscr{S} \otimes \mathscr{B}(X))$-measurable function, $w \in M(S)$, and $u(s, \cdot) \in C(F(s))$ for each $s \in S$, we have $E_{nm} \in \mathscr{S}$ for $n, m \geqslant 1$ and, moreover, we may assume without loss of generality that $E_{nm} \neq \emptyset$ for each $n, m \geqslant 1$.

Define $C_n(s) = \{f_k(s): s \in E_{nk}\}$ for each $n \geqslant 1$ and each $s \in S$. Let $U \subset X$ be an open set. Then

$$C_n^{-1}(U) = \{s: C_n(s) \cap U \neq \emptyset\} = \bigcup_{k=1}^{\infty} E_{nk} \cap \{s: f_k(s) \in U\} \in \mathscr{S}$$

and, consequently, C_n is a measurable multifunction for each $n \geqslant 1$. Now, we note that $C_n(s) \subset F(s)$ for each $s \in S$. This and Theorems 4.1 and 5.6 of [7] imply that $D(s) = \bigcap_{n=1}^{\infty} \mathrm{cl}\, C_n(s)$ is a measurable multifunction. It is easy to see that $H(s) = D(s)$ for each $s \in S$, which completes the proof.

For any multifunction $F: S \to \mathscr{K}(X)$ we define the multifunction $G: S \to \mathscr{P}(P(X))$ as $G(s) = P(F(s))$ for each $s \in S$.

We have the following

LEMMA 1.11. *Assume that $P(X)$ is endowed with the w-topology. Then G is compact-valued and if F is measurable, so is G.*

Proof. Since X is a separable metric space, so is $P(X)$ with the w-topology [21] and, by (1.2), G is compact-valued. Let F be a measurable multifunction. By Theorem 5.6 of [7] there is a sequence $\{f_n\}$ of measurable functions $f_n: S \to X$ such that $F(s) = \mathrm{cl}\,\{f_n(s)\}$ for each $s \in S$. It is clear that each f_n, $n \geqslant 1$, is measurable as a function $f_n: S \to P(X)$. Let Q be the set of all sequences $(\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots)$ of non-negative rational numbers such that

(1) $u(s, \cdot) \in C(F(s))$ means that the restriction of $u(s, \cdot)$ to $F(s)$ is continuous.

all but finitely many λ_n's are 0 and $\sum_n \lambda_n = 1$. Let $\{\lambda_n\} \in Q$. Then there is a finite sequence $\lambda_{n_1}, \lambda_{n_2}, \ldots, \lambda_{n_k}$ of rational numbers such that

$$\lambda_{n_m} > 0, \; m = 1, 2, \ldots, k, \quad \text{and} \quad \sum_{m=1}^{k} \lambda_{n_m} = 1.$$

By $\sum_n \lambda_n f_n(s)$ we denote a probability measure μ on $\mathscr{B}(X)$ with finite support $f_{n_1}(s), f_{n_2}(s), \ldots, f_{n_k}(s)$ such that $f_{n_m}(s) = \lambda_{n_m}$, $m = 1, 2, \ldots, k$. It is clear that for each $\{\lambda_n\} \in Q$ the function $\sum_n \lambda_n f_n$, as a function from S into $P(X)$, is measurable. Since the set Q is countable, so is $V = \{\sum_n \lambda_n f_n \colon \{\lambda_n\} \in Q\}$. Let

$$V(s) = \{\sum_n \lambda_n f_n(s) \colon \{\lambda_n\} \in Q\} \quad \text{for each } s \in S.$$

By Theorem 5.6 of [7], the closure cl $V(s)$ in the w-topology of $P(X)$ is a measurable multifunction. By Theorem II. 6.3 of [21] we know that $G(s) = \text{cl}\, V(s)$ for each $s \in S$, which completes the proof.

For any two measurable spaces $(S_1, \mathscr{S}_1)$ and $(S_2, \mathscr{S}_2)$, let $Q(S_2|S_1)$ be the set of all conditional probabilities on $(S_2, \mathscr{S}_2)$ given $(S_1, \mathscr{S}_1)$, i.e. the set of all functions q on $\mathscr{S}_2 \times S_1$ such that $q(E|\cdot) \in B(S_1)$ for each $E \in \mathscr{S}_2$ and $q(\cdot|s) \in P(S_2)$ for each $s \in S_1$ or, equivalently, the set of all functions $q \colon S_1 \to P(S_2)$ such that the mappings $s \to q(E|s)$ are measurable for each $E \in \mathscr{S}_2$. If Y is a metric space and $P(Y)$ is endowed with the w-topology, then every measurable function $q \colon S \to P(Y)$ is a conditional probability on Y given $(S, \mathscr{S})$. If Y is also separable, then $q \colon S \to P(Y)$ is a conditional probability on Y, given $(S, \mathscr{S})$, if and only if q is measurable (cf. [24], Lemma 6.1, and [4], Theorems 2.1 and 3.1).

Let $\{(S_i, \mathscr{S}_i)\}$, $i = 1, 2, \ldots, n$, be measurable spaces. We denote by $M(S_1 \times S_2 \ldots \times S_n)$ the set of all $(\mathscr{S}_1 \otimes \mathscr{S}_2 \otimes \ldots \otimes \mathscr{S}_n)$-measurable (product measurable) functions defined on the Cartesian product $S_1 \times S_2 \times \ldots \times S_n$. By $B(S_1 \times S_2 \times \ldots \times S_n)$ we denote the set of all bounded functions in $M(S_1 \times S_2 \times \ldots \times S_n)$.

2. Model formulation. The background for a *zero-sum stochastic game* may be provided by a *two-person stochastic control system*

$$\langle (S, \mathscr{S}), X, Y, A, B, r, q, \beta \rangle,$$

where

(2.1) $(S, \mathscr{S})$ is an abstract measurable space, S being a non-empty *set of states* of the system;

(2.2) X and Y are separable metric *spaces of actions* of players I and II, respectively;

(2.3) $A: S \to \mathscr{K}(X)$ and $B: S \to \mathscr{K}(Y)$ are measurable, $A(s)$ and $B(s)$ being the *sets of all admissible actions* to players I and II, respectively, when the system is at the state s;

(2.4) $r \in B(S \times X \times Y)$ is the *reward function*;

(2.5) $q \in Q(S|S \times X \times Y)$ is the *law of motion* of the system;

(2.6) $\beta \geqslant 0$ is the *discount factor*.

Periodically, players I and II observe the current state s of the system and choose actions $x \in A(s)$ and $y \in B(s)$, respectively; the choice of the actions is made with full knowledge of the history of the system, which has been evolved to the present. As a consequence of the actions chosen by the players, two things happen: player II pays player I $r(s, x, y)$ units and the system moves to a new state t according to the probability distribution $q(\cdot|s, x, y)$. The process is then repeated from the new state t. Player I maximizes his expected income and player II minimizes his expected loss as the game proceeds over the infinite future. (These notions will be defined more carefully below.)

Let us set $P_A(s) = P(A(s))$ and $P_B(s) = P(B(s))$ for each $s \in S$. Then by Lemma 1.11 P_A and P_B are measurable compact-valued multifunctions.

Put $H_1 = S$ and $H_n = S \times X \times Y \times H_{n-1}$ for $n \geqslant 2$.

A *strategy* for player I is a sequence $\pi = (\pi_1, \pi_2, \ldots)$, where $\pi_n \in Q(X|H_n)$ for each $n \geqslant 1$ and $\pi_n(A(s_n)|s_1, x_1, y_1, \ldots, s_n) = 1$ for each $(s_1, x_1, y_1, \ldots, s_n) \in H_n$, $n \geqslant 1$. We denote by Π the *set of all strategies* of player I. A strategy π is called *Markov* if $\pi_n \in Q(X|S)$ for each $n \geqslant 1$, i.e. if each π_n depends only on the actual state of the system. The *set of all Markov strategies* of player I will be denoted by Π_M. A strategy π is called *semi-stationary* if $\pi_n = f$ for each $n \geqslant 1$, where $f \in Q(X|S \times S)$. More precisely, a semi-stationary strategy is a strategy such that the action taken by player I at the n-th stage of the play depends only on the initial state and on the n-th state. We denote by Π_{ss} the *set of all semi-stationary strategies* of player I. Finally, a *stationary strategy* π of player I is a Markov strategy such that $\pi_n = f$ for each $n \geqslant 1$, where f is some measurable selection of the multifunction P_A. A stationary strategy determined by some $f \in \mathscr{S}_{P_A}$ will be denoted by $f^{(\infty)}$. We denote by Π_s the *set of all stationary strategies* of player I.

The sets Γ, Γ_M, Γ_{ss}, Γ_s of all strategies, all Markov strategies, all semi-stationary strategies, and all stationary strategies, respectively, of player II are defined similarly.

Remark. By Lemma 1.11 and by the selection theorem of Kuratowski and Ryll-Nardzewski ([15], Theorem 1), Π_s and Γ_s are non-empty.

Let us set

$$E_1 = S \quad \text{and} \quad E_n = \mathrm{Gr}\,(A \times B) \times E_{n-1} \quad \text{for } n \geqslant 2,$$

where $\mathrm{Gr}\,(A \times B) = \{(s, x, y)\colon x \in A(s) \text{ and } y \in B(s)\}$. Then E_n is the set of all *histories* of the model up to time n. The assumption (2.3) and Theorem 3.5 of [7] imply that, for each $n \geqslant 1$, E_n is a measurable subset of the product space H_n endowed with the product σ-algebra. The set E of all *plays* of the game is defined by setting

$$E = \bigcap_{n=1}^{\infty} (E_n \times X \times Y \times S \times X \times Y \times S \times \ldots).$$

It is clear that E is a measurable subset of the product space

$$H = S \times X \times Y \times S \times X \times Y \times \ldots$$

endowed with the product σ-algebra.

An application of the theorem of Ionescu Tulcea [19] yields that any pair $(\pi, \gamma) \in \Pi \times \Gamma$ together with the law of motion q and $s \in S$ uniquely defines a probability measure $m(\pi, \gamma, s)$ on the product σ-algebra of $K = X \times Y \times S \times X \times Y \times S \times \ldots$ Moreover, for any non-negative $u \in M(H)$, the expression $\int u\,dm(\pi, \gamma, \cdot)$ is an $\mathscr{S}$-measurable extended real-valued function. From the definition of π and γ it follows that if u is the characteristic function of the set E, then $\int u\,dm(\pi, \gamma, s) = 1$ for all $s \in S$.

Let $F_n = S \times X \times Y \times S \times X \times Y \times \ldots \times S \times X \times Y$ ($3n$ factors) and let $u \in B(F_n)$, $n \geqslant 1$. Then for each π, γ, and $s_1 \in S$ we have

$$\begin{aligned}\int u\,dm(\pi, \gamma, s_1) = \int\!\!\int\!\!\int \ldots \int\!\!\int\!\!\int u(s_1, x_1, y_1, \ldots, s_n, x_n, y_n)\,\pi_n(dx_n|s_1, x_1, y_1, \ldots, s_n) \times \\ \times \gamma_n(dy_n|s_1, x_1, y_1, \ldots, s_n)\,q(ds_n|s_{n-1}, x_{n-1}, y_{n-1}) \ldots \times \\ \times q(ds_2|s_1, x_1, y_1) \times \pi_1(dx_1|s_1)\,\gamma_1(dy_1|s_1).\end{aligned}$$

Now, for each $n \geqslant 1$ and each $h = (s_1, x_1, y_1, \ldots, s_n, x_n, y_n)$ we define the *n-stage reward function* $v_n \in B(F_n)$ as

$$v_n(h) = \sum_{k=1}^{n} \beta^{k-1} r(s_k, x_k, y_k),$$

and for any pair (π, γ) we define the *expected n-stage reward function* of (π, γ) and the *initial state* s_1 as

$$V_n(\pi, \gamma)(s_1) = \int v_n(h)\,dm(\pi, \gamma, s_1).$$

We shall often omit the variable s_1 in writing the expected reward functions. Let $r_n(\pi, \gamma)$ denote the *expected reward function at the n-th stage*, i.e.,

$$\begin{aligned}& r_n(\pi, \gamma)(s_1) \\ & = \int\!\!\int\!\!\int \ldots \int\!\!\int\!\!\int r(s_n, x_n, y_n)\,\pi_n(dx_n|s_1, x_1, y_1, \ldots, s_n)\,\gamma_n(dy_n|s_1, x_1, y_1, \ldots, s_n) \times \\ & \qquad \times q(ds_n|s_{n-1}, x_{n-1}, y_{n-1}) \ldots q(ds_2|s_1, x_1, y_1)\,\pi_1(dx_1|s_1)\,\gamma_1(dy_1|s_1).\end{aligned}$$

Then

(2.7) $$V_n(\pi, \gamma) = \sum_{k=1}^{n} \beta^{k-1} r_k(\pi, \gamma).$$

Put

$$\bar{v}_{nm}(h) = \sum_{k=n}^{m} \beta^{k-1} |r(s_k, x_k, y_k)|,$$

where $h = (s_1, x_1, y_1, \ldots, s_m, x_m, y_m)$, $m \geqslant n \geqslant 1$. Let us set

$$R_n^m(\pi, \gamma)(s_1) = \int \bar{v}_{nm}(h) dm(\pi, \gamma, s_1)$$

and

$$z_n(s_1) = \sup_m \sup_{\pi,\gamma} R_n^m(\pi, \gamma)(s_1).$$

The limit $\lim_n V_n(\pi, \gamma)$ may also be infinite or may be indefined. In this paper we shall study two cases in which the model is well defined:

(D) $z_n(s) \to 0$ pointwise as $n \to \infty$;

(P) $r \geqslant 0$, $\beta = 1$, and the structure of the model is such that

$$\sup_n \sup_{\pi,\gamma} V_n(\pi, \gamma)(s) < \infty \quad \text{for each } s \in S.$$

Now under (D) or (P) we define the *expected reward function over the infinite future* as follows:

$$V(\pi, \gamma) = \lim_n V_n(\pi, \gamma).$$

Remark. Let $r^+ = \max\{r, 0\}$ and $r^- = \min\{r, 0\}$. We define two functions $v^+ \colon H \to R \cup \{\infty\}$ and $v^- \colon H \to R \cup \{-\infty\}$ by

$$v^+(h) = \sum_{k=1} \beta^{k-1} r^+(s_k, x_k, y_k) \quad \text{and} \quad v^-(h) = \sum_{k=1} \beta^{k-1} r^-(s_k, x_k, y_k),$$

where $h = (s_1, x_1, y_1, \ldots) \in H$.

We claim that (D) yields

$$V(\pi, \gamma) = V^+(\pi, \gamma) + V^-(\pi, \gamma),$$

and (P) gives $V(\pi, \gamma) = V^+(\pi, \gamma)$, where

$$V^+(\pi, \gamma) = \int v^+(h) dm(\pi, \gamma) \quad \text{and} \quad V^-(\pi, \gamma) = \int v^-(h) dm(\pi, \gamma).$$

Remark. The condition (D) holds in the *discounted case*, i.e., when $\beta \in [0, 1)$, because r is bounded. For results in this direction see [1], [8], [9], [12], [23], etc. Stochastic games under the assumption similar to (D) were recently studied by Couwenbergh in [3]. He assumed the uniform convergence of the n-stage reward functions V_n on $\Pi \times \Gamma \times S$, but he allowed the reward function r to be unbounded. Stochastic games which satisfy the condition (P) are called *positive stochastic games*. For results in this direction see [5], [8], and [13]. The study of the condition (P) is contained in part II of this paper [20].

Let us set

$$\underline{v}^* = \sup_{\pi\in\Pi} \inf_{\gamma\in\Gamma} V(\pi, \gamma) \quad \text{and} \quad \bar{v}^* = \inf_{\gamma\in\Gamma} \sup_{\pi\in\Pi} V(\pi, \gamma).$$

We say that the stochastic game has a *value* if $\underline{v}^* = \bar{v}^*$, and we denote the value function by v^*.

A strategy π^* is ε-*optimal* for player I ($\varepsilon \geqslant 0$) if

$$\bar{v}^* \leqslant V(\pi^*, \gamma)+\varepsilon \quad \text{for each } \gamma\in\Gamma.$$

A strategy γ^* is ε-*optimal* for player II ($\varepsilon \geqslant 0$) if

$$\underline{v}^* \geqslant V(\pi, \gamma^*)-\varepsilon \quad \text{for each } \pi\in\Pi.$$

(The 0-optimal strategies are called *optimal*.)

3. Models with additional semi-continuity or continuity assumptions. To prove the main results of the paper we impose some additional conditions on the model presented in Section 2. We shall examine the stochastic games which satisfy one of the following assumptions:

(M_1) $r(s, \cdot, \cdot)$, $q(E|s, \cdot, \cdot)\in C(A(s)\times B(s))$ for each $s\in S$ and $E\in\mathcal{S}$.

(M_2) X and Y are analytic spaces, and $\mathcal{S}$ is a σ-algebra which is closed with respect to the Suslin operation. Moreover,

$$r(s, \cdot, y),\ q(E|s, \cdot, y)\in C(A(s)) \quad \text{and} \quad r(s, x, \cdot),\ q(E|s, x, \cdot)\in C(B(s))$$

for each $(s, x, y)\in \mathrm{Gr}\,(A\times B)$ and each $E\in\mathcal{S}$.

(M_3) S, X, and Y are SB-spaces and, moreover,

$$r(s, \cdot, y)\in \bar{C}(A(s)),\ r(s, x, \cdot)\in \underset{\sim}{C}(B(s))$$

and

$$q(E|s, \cdot, y)\in C(A(s)),\ q(E|s, x, \cdot)\in C(B(s))$$

for each $(s, x, y)\in \mathrm{Gr}\,(A\times B)$ and $E\in\mathcal{B}(S)$.

4. The finite horizon stochastic games. Let n be any positive integer. The stochastic game in which the players play up to time n is said to be a *finite horizon stochastic game*. Let π and γ be strategies of players I and II, respectively. Then the expected reward in that game is defined by (2.7). It is clear that $V_n(\pi, \gamma)$ involves only the first n terms of the strategies π and γ.

Before proving the main result of this section we give some auxiliary results. Let $u\in M(S\times X\times Y)$ be such that for each $(\mu, \lambda)\in P(X)\times P(Y)$ the expression

$$\hat{u}(s, \mu, \lambda) = \int_X \int_Y u(s, x, y)\, d\mu d\lambda,$$

is well defined and such that, for each $s \in S$,

$$\hat{u}^*(s) = \sup_{\mu \in P_A(s)} \inf_{\lambda \in P_B(s)} \hat{u}(s, \mu, \lambda) = \inf_{\lambda \in P_B(s)} \sup_{\mu \in P_A(s)} \hat{u}(s, \mu, \lambda).$$

Define

$$\mathcal{O}_1(s) = \{\mu \in P_A(s) \colon \hat{u}^*(s) \leqslant \inf_{\lambda \in P_B(s)} \hat{u}(s, \mu, \lambda)\}, \qquad s \in S,$$

and

$$\mathcal{O}_2(s) = \{\lambda \in P_B(s) \colon \sup_{\mu \in P_A(s)} \hat{u}(s, \mu, \lambda) \leqslant \hat{u}^*(s)\}, \qquad s \in S.$$

By Lemma 1.11, the multifunctions P_A and P_B are measurable and compact valued, and by Theorem 5.6 of [7] there exist sequences $\{f_n\} \subset \mathcal{S}_{P_A}$ and $\{g_m\} \subset \mathcal{S}_{P_B}$ such that

(4.1) $\quad P_A(s) = \mathrm{cl}\,\{f_n(s)\}$ and $P_B(s) = \mathrm{cl}\,\{g_m(s)\}$ for each $s \in S$.

The following lemmas are basic for this paper.

LEMMA 4.1. *Assume that $u(s, \cdot, \cdot) \in C(A(s) \times B(s))$ for each $s \in S$. Then $\hat{u}^* \in M(S)$ and there exist $f \in \mathcal{S}_{\mathcal{O}_1}$ and $g \in \mathcal{S}_{\mathcal{O}_2}$.*

LEMMA 4.2. *Let $\mathcal{S}$ be a σ-algebra which is closed with respect to the Suslin operation and let X and Y be metric analytic spaces. Assume that $u(s, \cdot, \cdot)$ is bounded for each $s \in S$ and*

$$u(s, \cdot, y) \in C(A(s)) \quad \text{and} \quad u(s, x, \cdot) \in C(B(s))$$

for each $(s, x, y) \in \mathrm{Gr}(A \times B)$. Then $\hat{u}^* \in M(S)$ and there exist $f \in \mathcal{S}_{\mathcal{O}_1}$ and $g \in \mathcal{S}_{\mathcal{O}_2}$.

LEMMA 4.3. *Let S, X, and Y be SB-spaces. Assume that $u(s, \cdot, \cdot)$ is bounded for each $s \in S$ and*

$$u(s, \cdot, y) \in \tilde{C}(A(s)) \quad \textit{and} \quad u(s, x, \cdot) \in \underset{\sim}{C}(B(s))$$

for each $(s, x, y) \in \mathrm{Gr}(A \times B)$. Then $\hat{u}^* \in M(S)$ and there exist $f \in \mathcal{S}_{\mathcal{O}_1}$ and $g \in \mathcal{S}_{\mathcal{O}_2}$.

Remark 4.1. The multifunctions $\mathcal{O}_1$ and $\mathcal{O}_2$ are well defined in all the cases considered in Lemmas 4.1 - 4.3 Moreover, $\mathcal{O}_1$ and $\mathcal{O}_2$ have non-empty and compact values. These sentences follow from the compactness of the sets $P(A(s))$ and $P(B(s))$, $s \in S$, Lemma 1.5, and the minimax theorem of Sion [27]. In order to prove that there exist $f \in \mathcal{S}_{\mathcal{O}_1}$ and $g \in \mathcal{S}_{\mathcal{O}_2}$ it is sufficient to show that $\mathcal{O}_1$ and $\mathcal{O}_2$ are measurable. Then the assertion follows from the selection theorem of Kuratowski and Ryll-Nardzewski [15].

Remark 4.2. Under the assumptions of Lemmas 4.1 - 4.3, respectively, we have

(4.2) $\quad \inf_{\lambda \in P_B(\cdot)} \hat{u}(\cdot, \cdot, \lambda) \in M(S \times P(X))$ and $\sup_{\mu \in P_A(\cdot)} \hat{u}(\cdot, \mu, \cdot) \in M(S \times P(Y))$.

If we take the assumptions of Lemma 4.3, then (4.2) can be shown by using Theorem 2 of [9] and Lemmas 1.5 and 1.6. (Recall that $P(X)$ and $P(Y)$ are SB-spaces in this case.) To verify (4.2) in the remaining two cases it is sufficient to use (4.1) and Lemma 1.6.

Proof of Lemma 4.1. We know that $P_A(s)$ and $P_B(s)$ are compact and $\hat{u}(s,\cdot,\cdot)\in C(P_A(s)\times P_B(s))$ for each $s\in S$. Hence we obtain

$$\text{(4.3)}\qquad \sup_{\mu\in P_A(s)}\hat{u}(s,\mu,\cdot)\in C(P_B(s)) \quad\text{and}\quad \inf_{\lambda\in P_B(s)}\hat{u}(s,\cdot,\lambda)\in C(P_A(s)),\qquad s\in S,$$

which together with (4.1) and (4.2) implies that

$$\text{(4.4)}\qquad \hat{u}^*(\cdot)=\inf_{\lambda\in P_B(\cdot)}\sup_{\mu\in P_A(\cdot)}\hat{u}(\cdot,\mu,\lambda)=\inf_m\sup_{\mu\in P_A(\cdot)}\hat{u}(\cdot,\mu,g_m(\cdot))\in M(S).$$

From (4.2) - (4.4) and Lemma 1.10 we infer that the multifunctions $\mathcal{O}_1$ and $\mathcal{O}_2$ are measurable, which completes the proof by virtue of Remark 4.1.

Proof of Lemma 4.2. First we prove that $\hat{u}^*\in M(S)$. Let c be any real number. By (1.4) and Lemmas 1.8 and 1.11 we have $\mathrm{Gr}\,P_A\in\mathcal{S}\otimes\mathcal{B}(P(X))$, so (4.2) implies

$$\begin{aligned}E&=\{(s,\mu)\in\mathrm{Gr}\,P_A:\ \inf_{\lambda\in P_B(s)}\hat{u}(s,\mu,\lambda)>c\}\\&=\{(s,\mu)\in S\times P(X):\ \inf_{\lambda\in P_B(s)}\hat{u}(s,\mu,\lambda)>c\}\cap\mathrm{Gr}\,P_A\in\mathcal{S}\otimes\mathcal{B}(P(X)).\end{aligned}$$

Since $P(X)$ endoved with the w-topology is a metric analytic space [2], by Lemma 1.9 we have

$$\{s\in S:\hat{u}^*(s)>c\}=\mathrm{proj}_S\,E\in\mathcal{S},$$

which implies $\hat{u}^*\in M(S)$.

It follows from the measurability of $\hat{u}^*$ and from (4.2) that the multifunctions $\mathcal{O}_1$ and $\mathcal{O}_2$ have measurable graphs. Consequently, by Lemma 1.8, $\mathcal{O}_1$ and $\mathcal{O}_2$ are measurable. (Recall that $\mathcal{O}_1$ and $\mathcal{O}_2$ are compact-valued.) Thus the lemma is proved.

Proof of Lemma 4.3. Using Lemma 1.5 it can be easily shown that

$$\text{(4.5)}\qquad \inf_{\lambda\in P_B(s)}\hat{u}(s,\cdot,\lambda)\in\tilde{C}(P_A(s)),\qquad s\in S.$$

Since S, $P(X)$, and $P(Y)$ are SB-spaces and $P_A(s)$ is compact for each $s\in S$, from (4.2), (4.5), and Theorem 2 of [9] we get $\hat{u}^*\in M(S)$. Moreover, applying Lemma 1.7 to our situation we infer that the multifunctions $\mathcal{O}_1$ and $\mathcal{O}_2$ are measurable, which completes the proof.

Let us set

$$M_q(S)=\{v\in M(S):\ \sup_{x\in A(s)}\sup_{y\in B(s)}\left|\int_S v(t)q(dt|s,x,y)\right|<\infty,\ s\in S\},$$

where q is the law of motion of the stochastic control system defined in Section 2. Clearly, $B(S) \subset M_q(S)$.

For any $v \in M_q(S)$ we set

$$L(s, x, y)(v) = r(s, x, y) + \beta \int_S v(t) q(dt|s, x, y)$$

and

$$L(s, \mu, \lambda)(v) = \int_X \int_Y L(s, x, y)(v)\, d\mu d\lambda,$$

where $(\mu, \lambda) \in P(X) \times P(Y)$. Further, for any pair $f \in \mathscr{S}_{P_A}$ and $g \in \mathscr{S}_{P_B}$ we put

$$T_{fg}(v)(s) = L(s, f(s), g(s))(v)$$

and

$$T(v)(s) = \sup_{\mu \in P_A(s)} \inf_{\lambda \in P_B(s)} L(s, \mu, \lambda)(v) = \inf_{\lambda \in P_B(s)} \sup_{\mu \in P_A(s)} L(s, \mu, \lambda)(v),$$

provided that the last equality holds.

Now we return to the finite horizon stochastic games described above. We shall prove the following

THEOREM 4.1. *For each of the models* (M_1), (M_2), (M_3) *the n-stage stochastic game ($n \geq 1$) has a value $v_n^* \in B(S)$ and both players have optimal Markov strategies. Moreover, for each $n \geq 2$, $T(v_{n-1}^*) = v_n^*$.*

Proof. The proof proceeds by induction. For $n = 1$ the theorem follows directly from Lemmas 4.1-4.3. Suppose the result holds for $n-1$ ($n \geq 2$). Let $\pi_{n-1} = (f_1, f_2, \ldots, f_{n-1})$ and $\gamma_{n-1} = (g_1, g_2, \ldots, g_{n-1})$ be a pair of optimal Markov strategies of players I and II, respectively, in the $(n-1)$-stage stochastic game. Then

$$v_{n-1}^* = V_{n-1}(\pi_{n-1}, \gamma_{n-1}). \tag{4.6}$$

Denote by $\bar{v}_n^*$ and $\underline{v}_n^*$ the upper and lower value functions, respectively, in the n-stage game. Let $\bar{\gamma} = (g, g_1, g_2, \ldots, g_{n-1})$, where $g \in \mathscr{S}_{P_B}$ and $(g_1, g_2, \ldots, g_{n-1}) = \gamma_{n-1}$. We note that

$$\bar{v}_n^* = \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} V_n(\pi, \gamma) \leq \sup_{\pi \in \Pi} V_n(\pi, \bar{\gamma}) \quad \text{for each } g \in \mathscr{S}_{P_B}.$$

Further, we observe that

$$\bar{v}_n^*(s) \leq \sup_{\mu \in P_A(s)} L(s, \mu, g(s))\big[\sup_{\pi \in \Pi} V_{n-1}(\pi, \gamma_{n-1})\big]$$

for each $g \in \mathscr{S}_{P_B}$ and each $s \in S$, i.e.,

$$\bar{v}_n^*(s) \leq \sup_{\mu \in P_A(s)} L(s, \mu, g(s))(v_{n-1}^*).$$

Hence

$$\bar{v}_n^*(s) \leq \inf_{\lambda \in P_B(s)} \sup_{\mu \in P_A(s)} L(s, \mu, \lambda)(v_{n-1}^*)$$

for each $s \in S$, and by the minimax theorem of Sion [27] we obtain

$$\bar{v}_n^* \leqslant T(v_{n-1}^*). \tag{4.7}$$

Similarly, we obtain

$$\underline{v}_n^* \geqslant T(v_{n-1}^*). \tag{4.8}$$

Combining (4.7) and (4.8) we get $\underline{v}_n^* = \bar{v}_n^* = T(v_{n-1}^*)$, which means that the n-stage stochastic game has a value v_n^* and $v_n^* = T(v_{n-1}^*)$. Further from (2.4) and Lemmas 4.1, 4.2, and 4.3 (applied to (M_1), (M_2), and (M_3), respectively) we conclude that $v_n^* \in B(S)$, and there exist $f_0 \in \mathcal{S}_{P_A}$ and $g_0 \in \mathcal{S}_{P_B}$ such that

$$T_{fg_0}(v_{n-1}^*) \leqslant v_n^* = T_{f_0 g_0}(v_{n-1}^*) \leqslant T_{f_0 g}(v_{n-1}^*) \tag{4.9}$$

for every $f \in \mathcal{S}_{P_A}$ and $g \in \mathcal{S}_{P_B}$.

Let $\pi_n = (f_0, f_1, f_2, \ldots, f_{n-1})$ and $\gamma_n = (g_0, g_1, g_2, \ldots, g_{n-1})$. Then from (4.6)-(4.9) it follows that π_n and γ_n are optimal for players I and II, respectively, in the n-stage stochastic game.

5. The infinite horizon stochastic games satisfying the condition (D).

THEOREM 5.1. *For each of the models* (M_1), (M_2), (M_3) *the value function* v^* *exists,* $v^* = \lim_n v_n^*$, *and* $v^* \in M(S)$.

Proof. It is easy to verify that, for each $s \in S$,

$$|\bar{v}^*(s) - \bar{v}_n^*(s)| \leqslant \sup_{\pi \in \Pi} \sup_{\gamma \in \Gamma} |V(\pi, \gamma)(s) - V_n(\pi, \gamma)(s)| \leqslant z_{n+1}(s)$$

and

$$|\underline{v}^*(s) - \underline{v}_n^*(s)| \leqslant \sup_{\pi \in \Pi} \sup_{\gamma \in \Gamma} |V(\pi, \gamma)(s) - V_n(\pi, \gamma)(s)| \leqslant z_{n+1}(s).$$

By Theorem 4.1 we have $\underline{v}_n^* = \bar{v}_n^* = v_n^* \in B(S)$ for each $n \geqslant 1$, so the condition (D) implies

$$v^* = \underline{v}^* = \bar{v}^* = \lim_n v_n^* \in M(S).$$

We denote by V^+ and V_n^+ (V^- and V_n^-), $n \geqslant 1$, the functions V and V_n, where the reward function r is replaced by r^+ (r^-), respectively. By z_n^+ and z_n^-, $n \geqslant 1$, we denote the function z_n, where $|r|$ is replaced by r^+ and $|r^-|$, respectively.

We have immediately the following

LEMMA 5.1. $z_n \to 0$ *pointwise as* $n \to \infty$ *if and only if* $z_n^+ \to 0$ *and* $z_n^- \to 0$ *pointwise as* $n \to \infty$.

We note that if r satisfies the assumption (M_1), then r^+ and r^- also satisfy this assumption. Similarly, the same holds for (M_2) and (M_3). Thus, Theorem 4.1 implies the following

LEMMA 5.2. *For each $n \geqslant 1$ and for each of the models* (M_1), (M_2), (M_3) *of the n-stage stochastic game with the reward function* r^+ (r^-) *the value function* v_n^{*+} (v_n^{*-}) *exists.*

Lemmas 5.1 and 5.2 and Theorem 5.1 imply

LEMMA 5.3. *For each of the models* (M_1), (M_2), (M_3) *of the stochastic games with the reward functions* r^+ *and* r^- *the value functions* v^{*+} *and* v^{*-} *exist. Moreover,*

$$v^{*+} = \lim_n v_n^{*+} \quad \text{and} \quad v^{*-} = \lim_n v_n^{*-}.$$

For each $u \in M(S)$ such that the right-hand sides below are well defined we set

$$T^1(u) = T(u), \qquad T^n(u) = T\big(T^{n-1}(u)\big),\ n \geqslant 2,$$

$$T_{fg}^1(u) = T_{fg}(u), \qquad T_{fg}^n(u) = T_{fg}\big(T_{fg}^{n-1}(u)\big),\ n \geqslant 2,$$

where $f \in \mathcal{S}_{P_A}$ and $g \in \mathcal{S}_{P_B}$. Further, we put

$$B_{fg}^1(u)(s) = \beta \iiint u(t) q(dt|s, x, y) f(dx|s) g(dy|s)$$

and

$$B_{fg}^n(u) = B_{fg}^1\big(B_{fg}^{n-1}(u)\big), \qquad n \geqslant 2.$$

LEMMA 5.4. *For each $f \in \mathcal{S}_{P_A}$ and $g \in \mathcal{S}_{P_B}$ and $n \geqslant 1$, $B_{fg}^n(v^*)$ is well defined and*

$$\lim_n B_{fg}^n(v^*) = 0.$$

Proof. By Theorem 5.1, the value function v^* exists and $v^* \in M(S)$. From Theorem 4.1 we infer that, for each $m \geqslant 1$, $v_m^{*+} = V_m^+(\pi_m, \gamma_m)$ for some Markov strategies π_m and γ_m of players I and II, respectively. Note that

$$B_{fg}^n(v_m^{*+}) = B_{fg}^n\big(V_m^+(\pi_m, \gamma_m)\big) \leqslant z_{n+1}^+ \tag{5.1}$$

for each $m \geqslant 1$ and $n \geqslant 1$. Similarly, we can show that

$$B_{fg}^n(|v_m^{*-}|) \leqslant z_{n+1}^- \tag{5.2}$$

for each $m \geqslant 1$ and $n \geqslant 1$. Observe that, for each $m \geqslant 1$,

$$0 \leqslant v_m^{*+} \leqslant v_{m+1}^{*+} \quad \text{and} \quad v_{m+1}^{*-} \leqslant v_m^{*-} \leqslant 0. \tag{5.3}$$

By (5.1)-(5.3) and the monotone convergence theorem, for each $n \geqslant 1$ we obtain

$$B_{fg}^n(v^{*+}) \leqslant z_{n+1}^+ \tag{5.4}$$

and

$$B_{fg}^n(|v^{*-}|) \leqslant z_{n+1}^-. \tag{5.5}$$

Since $v^{*-} \leqslant v^* \leqslant v^{*+}$, we have $|v^*| \leqslant v^{*+} + |v^{*-}|$, which together with (5.4) and (5.5) implies that, for each $n \geqslant 1$,

$$B_{fg}^n(|v^*|) \leqslant z_{n+1}^+ + z_{n+1}^-, \tag{5.6}$$

i.e., $B_{fg}^n(v^*)$ is well defined. By Lemma 5.1 we get

$$0 \leqslant \lim_n |B_{fg}^n(v^*)| \leqslant \lim_n B_{fg}^n(|v^*|) \leqslant \lim_n (z_{n+1}^+ + z_{n+1}^-) = 0.$$

COROLLARY 5.1. *For each $f \in \mathscr{S}_{P_A}$ and $g \in \mathscr{S}_{P_B}$ and $n \geqslant 1$, $T_{fg}^n(v^*)$ is well defined and*

$$\lim_n T_{fg}^n(v^*) = V(f^{(\infty)}, g^{(\infty)}).$$

Proof. Note that $T_{fg}^n(v^*) = V_n(f^{(\infty)}, g^{(\infty)}) + B_{fg}^n(v^*)$, $n \geqslant 1$, and this completes the proof.

THEOREM 5.2. $v^* \in M_q(S)$.

Proof. By (5.6) we have $B_{fg}^1(|v^*|) \leqslant z_2^+ + z_2^-$. Hence

$$\sup_{f \in \mathscr{S}_{P_A}} \sup_{g \in \mathscr{S}_{P_B}} B_{fg}^1(|v^*|) < \infty,$$

which implies immediately that $v^* \in M_q(S)$.

LEMMA 5.5. *For each $m > n$ and $s \in S$ we have*

$$\sup_{x \in A(s)} \sup_{y \in B(s)} |L(s, x, y)(v_m^*) - L(s, x, y)(v_n^*)| \leqslant z_{n+2}(s).$$

Proof. By Theorem 4.1 we have $v_m^* = V_m(\pi_m, \gamma_m)$ for some Markov strategies π_m and γ_m of players I and II, respectively. Similarly, $v_n^* = V_n(\pi_n, \gamma_n)$. We note that

$$v_m^* \leqslant v_n^* + R_{n+1}^m(\pi_m, \gamma_n) \tag{5.7}$$

and

$$v_m^* \geqslant v_n^* - R_{n+1}^m(\pi_n, \gamma_m). \tag{5.8}$$

Indeed,

$$\begin{aligned} v_m^* &= V_m(\pi_m, \gamma_m) \leqslant V_m(\pi_m, \gamma_n) \leqslant V_n(\pi_n, \gamma_n) + R_{n+1}^m(\pi_m, \gamma_n) \\ &= v_n^* + R_{n+1}^m(\pi_m, \gamma_n) \end{aligned}$$

and

$$\begin{aligned} v_m^* &= V_m(\pi_m, \gamma_m) \geqslant V_m(\pi_n, \gamma_m) \geqslant V_n(\pi_n, \gamma_n) - R_{n+1}^m(\pi_n, \gamma_m) \\ &= v_n^* - R_{n+1}^m(\pi_n, \gamma_m). \end{aligned}$$

By (5.7), for each $x \in A(s)$ and $y \in B(s)$ we obtain

$$L(s, x, y)(v_m^*) \leqslant L(s, x, y)(v_n^*) + z_{n+2}(s), \tag{5.9}$$

and by (5.8), for each $x \in A(s)$ and $y \in B(s)$ we get

$$(5.10) \qquad L(s, x, y)(v_m^*) \geqslant L(s, x, y)(v_n^*) - z_{n+2}(s).$$

Thus, (5.9) and (5.10) imply the required inequality.

From Lemma 5.5 and Theorems 4.1 and 5.1 we obtain immediately

LEMMA 5.6. (a) *If r and q satisfy the assumptions of the model* (M_1), *then the function* $L(\cdot, \cdot, \cdot)(v^*)$ *satisfies the assumptions of Lemma* 4.1.

(b) *If r and q satisfy the assumptions of the model* (M_2), *then the function* $L(\cdot, \cdot, \cdot)(v^*)$ *satisfies the assumptions of Lemma* 4.2.

(c) *If r and q satisfy the assumptions of the model* (M_3), *then the function* $L(\cdot, \cdot, \cdot)(v^*)$ *satisfies the assumptions of Lemma* 4.3.

THEOREM 5.3. *For each of the models* (M_1), (M_2), (M_3) *we have*

$$v^* = T(v^*) = T_{fg}(v^*) \quad \text{for some } f \in \mathscr{S}_{P_A} \text{ and } g \in \mathscr{S}_{P_B}.$$

Proof. By Theorem 5.2, $T(v^*)$ is well defined, i.e., $|T(v^*)| < \infty$. Using Theorem 5.1 and the bounded convergence theorem, for each $n \geqslant 1$ we obtain

$$\begin{aligned} |T(v^*)(s) - T(v_n^*)(s)| &\leqslant \sup_{x \in A(s)} \sup_{y \in B(s)} |L(s, x, y)(v^*) - L(s, x, y)(v_n^*)| \\ &= \sup_{x \in A(s)} \sup_{y \in B(s)} \lim_m |L(s, x, y)(v_m^*) - L(s, x, y)(v_n^*)| \\ &\leqslant \lim_m \sup_{x \in A(s)} \sup_{y \in B(s)} |L(s, x, y)(v_m^*) - L(s, x, y)(v_n^*)|. \end{aligned}$$

Hence and by Lemma 5.5 we have, for each $s \in S$,

$$|T(v^*)(s) - T(v_n^*)(s)| \leqslant z_{n+2}(s).$$

Consequently, the condition (D) implies $T(v^*) = \lim_n T(v_n^*)$. By Theorems 4.1 and 5.1 we get $T(v^*) = v^*$. Applying Lemma 5.6 and Lemmas 4.1, 4.2, and 4.3 to the models (M_1), (M_2), and (M_3), respectively, we obtain $f \in \mathscr{S}_{P_A}$ and $g \in \mathscr{S}_{P_B}$ such that $v^* = T(v^*) = T_{fg}(v^*)$.

Now we define, for any given $h \in \mathscr{S}_{P_B}$, an auxiliary one-person stochastic game

$$(+) \qquad \langle (S, \mathscr{S}), X, A, r', q', \beta \rangle,$$

where

$$r'(s, x) = \int_Y r(s, x, y)\, h(dy|s), \qquad q'(E|s, x) = \int_Y q(E|s, x, y)\, h(dy|s)$$

for each $s \in S$, $x \in A(s)$ and $E \in \mathscr{S}$.

Clearly, if the original two-person stochastic game satisfies the condition (D), so does the modified one-person stochastic game (+).

Let us set

$$v_{\mathrm{I}}^*(h) = \sup_{\pi \in \Pi} V(\pi, h^{(\infty)}).$$

LEMMA 5.7. *For each of the models* (M_1), (M_2), (M_3), *for any* $h \in \mathscr{S}_{P_B}$ *there is* $f \in \mathscr{S}_{P_A}$ *such that*

$$v_{\mathrm{I}}^*(h) = T_{fh}(v_{\mathrm{I}}^*(h)) = V(f^{(\infty)}, h^{(\infty)}).$$

Proof. Applying Theorem 5.3 to the game (+) we obtain

$$v_{\mathrm{I}}^*(h)(s) = \sup_{\mu \in P_A(s)} L(s, \mu, h(s))(v_{\mathrm{I}}^*(h)) = T_{fh}(v_{\mathrm{I}}^*(h))(s)$$

for each $s \in S$ and for some $f \in \mathscr{S}_{P_A}$. Hence $v_{\mathrm{I}}^*(h) = T_{fh}^n(v_{\mathrm{I}}^*(h))$ for each $n \geqslant 1$. Using Corollary 5.1 in the case (+) we get $v_{\mathrm{I}}^*(h) = V(f^{(\infty)}, h^{(\infty)})$, which completes the proof.

For any $f \in \mathscr{S}_{P_A}$ we put

$$v_{\mathrm{II}}^*(f) = \inf_{\gamma \in \Gamma} V(f^{(\infty)}, \gamma).$$

By similar arguments as above we can show the following

LEMMA 5.8. *For each of the models* (M_1), (M_2), (M_3), *for any* $f \in \mathscr{S}_{P_A}$ *there is* $g \in \mathscr{S}_{P_B}$ *such that*

$$v_{\mathrm{II}}^*(f) = T_{fg}(v_{\mathrm{II}}^*(f)) = V(f^{(\infty)}, g^{(\infty)}).$$

Now we are ready to state the main result of this section:

THEOREM 5.4. *For each of the models* (M_1), (M_2), (M_3) *both players have optimal stationary strategies.*

Proof. By Theorem 5.3, $v^* = T(v^*)$. Using Lemma 5.6 and Lemmas 4.1, 4.2, and 4.3 to the models (M_1), (M_2), and (M_3), respectively, we obtain $f \in \mathscr{S}_{P_A}$ and $g \in \mathscr{S}_{P_B}$ such that

$$T_{lg}(v^*) \leqslant v^* = T(v^*) = T_{fg}(v^*) \leqslant T_{fh}(v^*)$$

for each $l \in \mathscr{S}_{P_A}$ and $h \in \mathscr{S}_{P_B}$. Hence, for each $n \geqslant 1$, we have

$$T_{lg}^n(v^*) \leqslant v^* = T_{fg}^n(v^*) \leqslant T_{fh}^n(v^*).$$

By Corollary 5.1 we get

$$V(l^{(\infty)}, g^{(\infty)}) \leqslant v^* = V(f^{(\infty)}, g^{(\infty)}) \leqslant V(f^{(\infty)}, h^{(\infty)})$$

for each $l \in \mathscr{S}_{P_A}$ and $h \in \mathscr{S}_{P_B}$, which together with Lemmas 5.7 and 5.8 implies

$$\sup_{\pi \in \Pi} V(\pi, g^{(\infty)}) = V(f^{(\infty)}, g^{(\infty)}) = \inf_{\gamma \in \Gamma} V(f^{(\infty)}, \gamma),$$

i.e., $f^{(\infty)}$ and $g^{(\infty)}$ are optimal for players I and II, respectively.

Remark. Theorem 5.4 generalizes Theorem 3 of [8] and Theorem 3.23 of [11] (see also [12]), where S is an SB-space, $\beta \in [0, 1)$, and r and q satisfy the continuity assumptions of type (M_1).

Acknowledgment. I wish to express my gratitude to Professor Rastislav Telgársky for his encouragement and many valuable discussions.

References

[1] D. Blackwell, *Discounted dynamic programming*, Ann. Math. Statist. 36 (1965), p. 226-235.

[2] — D. Freedman and M. Orkin, *The optimal reward operator in dynamic programming*, Ann. Probability 2 (1974), p. 926-941.

[3] H. A. M. Couwenbergh, *Stochastic games with metric state space*, Internat. J. Game Theory 9 (1980), p. 25-36.

[4] L. Dubins and D. Freedman, *Measurable sets of measures*, Pacific J. Math. 14 (1965), p. 1211-1222.

[5] E. B. Frid, *On stochastic games*, Theor. Probability Appl. 18 (1973), p. 389-393.

[6] P. Gänssler, *Compactness and sequential compactness in spaces of measures*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 17 (1971), p. 124-146.

[7] C. J. Himmelberg, *Measurable relations*, Fund. Math. 87 (1975), p. 53-72.

[8] — T. Parthasarathy, T. E. S. Raghavan and F. S. Van Vleck, *Existence of p-equilibrium and optimal stationary strategies in stochastic games*, Proc. Amer. Math. Soc. 60 (1976), p. 245-251.

[9] C. J. Himmelberg, T. Parthasarathy and F. S. Van Vleck, *Optimal plans for dynamic programming problems*, Math. Oper. Res. 1 (1976), p. 390-394.

[10] K. Hinderer, *Foundations of non-stationary dynamic programming with discrete time parameter*, Lecture Notes in Operations Res. and Math. Systems 33, Springer, Berlin 1970.

[11] A. Idzik, *On discounted stochastic games* (in Polish), Ph. D. dissertation submitted to the Computation Centre of the Polish Academy of Sciences, Warsaw 1976.

[12] — *Remarks on discounted stochastic games. Information theory, statistical decision functions, random processes*, Trans. 8th Prague Conf., Prague 1978, Vol. C (1979), p. 165-174.

[13] D. B. Kamerud, *Repeated games and positive stochastic games*, Ph. D. dissertation submitted to the University of Minnesota, 1975.

[14] K. Kuratowski, *Topology*, Vol. I, II, Academic Press, New York 1966, 1968.

[15] — and C. Ryll-Nardzewski, *A general theorem on selectors*, Bull. Acad. Polon. Sci., Sér. Sci. Math. Astronom. Phys., 13 (1965), p. 379-403.

[16] S. J. Leese, *Measurable selections and the uniformization of Souslin sets*, Amer. J. Math. 100 (1978), p. 19-41.

[17] A. Maitra and T. Parthasarathy, *On stochastic games*, J. Optimization Theory Appl. 5 (1970), p. 289-300.

[18] — *On stochastic games. II*, ibidem 8 (1971), p. 154-160.

[19] J. Neveu, *Mathematical foundations of the calculus of probability*, Holden-Day, San Francisco 1965.

[20] A. S. Nowak, *On zero-sum stochastic games with general state space. II*, Probability and Mathematical Statistics 4 (1983) (to appear).

[21] K. R. Parthasarathy, *Probability measures on metric spaces*, Academic Press, New York 1967.

[22] T. Parthasarathy, *Discounted, positive and noncooperative stochastic games*, Internat. J. Game Theory 2 (1973), p. 25-37.
[23] — and M. Stern, *Markov games: a survey*, University of Illinois, Chicago, Illinois, 1976.
[24] U. Rieder, *Bayesian dynamic programming*, Advances in Appl. Probability 7 (1975), p. 330-348.
[25] M. Schäl, *Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 32 (1975), p. 179-196.
[26] L. S. Shapley, *Stochastic games*, Proc. Nat. Acad. Sci. U.S.A. 39 (1953), p. 1095-1100.
[27] M. Sion, *On general minimax theorems*, Pacific J. Math. 8 (1958), p. 171-176.
[28] D. H. Wagner, *Survey of measurable selection theorems*, SIAM J. Control Optimization 15 (1977), p. 859-903.

Institute of Mathematics
Technical University of Wrocław
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland

Received on 17. 12. 1980;
revised version on 8. 4. 1981