stichting
mathematisch
centrum

AFDELING MATHEMATISCHE BESLISKUNDE BW 86/77 DECEMBER
(DEPARTMENT OF OPERATIONS RESEARCH)

S.H. TIJS & O.J. VRIEZE

PERTURBATION THEORY FOR GAMES IN NORMAL FORM AND STOCHASTIC GAMES

Preprint

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O).

AMS(MOS) subject classification scheme (1970): 90D10, 90D15, 47A55

Perturbation theory for games in normal form and stochastic games *

by

S.H. Tijs** and O.J. Vrieze

ABSTRACT

In this paper the effect on values and optimal strategies of perturbations of game parameters (as payoff function, transition probability function and discount factor) is studied for the class of zero-sum games in normal form and for the class of stationary discounted two-person zero-sum stochastic games.

A main result is that, under certain conditions, the value depends in a (pointwise Lipschitz) continuous way on these parameters and that the sets of (ε-)optimal strategies for both players are upper semicontinuous multifunctions of the game parameters.

Extensions to general sum games and non-stationary stochastic games are also given.

KEYWORDS & PHRASES: *Game in normal form, stochastic game, non-stationary stochastic game, perturbation of "payoffs, transition probability function and discount factor".*

** Department of Mathematics, Catholic University, Nijmegen, The Netherlands.

* This report will be submitted for publication elsewhere.

1. INTRODUCTION AND SUMMARY

In this paper the main question in various settings is: What is the influence of perturbation of the "game parameters" on the "solutions" of the game? This question is not only of theoretical importance but also of practical utility, because "favourable" answers to this question will give greater confidence in the use of game models in applications. Roughly speaking, "favourable" means here that small changes in the game parameters induce only small changes in the values, and "good" strategies in the original game are not "bad" in a slightly perturbed game.

For papers in the same spirit - but in a different context - we refer to KRABS [4], SCHWEITZER [8], TIJS [11] and WHITT [14]. Here we focus our attention upon two classes of games, namely games in normal form (sections 2 and 5) and stochastic games (sections 3 and 4).

In section 2 subclasses of two-person zer-sum games in normal form with fixed strategy spaces are considered. The value appears to depend in a (Lipschitz) continuous manner on the payoff function (theorem 2.1) and ε-optimal strategies of the original game are $(\varepsilon+2\delta)$-optimal in a δ-perturbed game (theorem 2.2). Under additional topological conditions on strategy spaces and the payoff function, the space of (ε-)optimal strategies appears to depend in an upper semicontinuous manner (in the multi-function sense) on the payoff function (theorem 2.3). Furthermore, in this section special attention is paid to a subclass of games with unique optimal strategies for both players (theorems 2.7 and 2.8).

In section 5 for general sum games in normal form sufficient conditions are given to guarantee that the ε-equilibrium point set depends in an upper semicontinuous way on the payoff functions (theorem 5.4).

In section 3, for a family of discounted two-person zero-sum stochastic games with fixed (countable) state space and (compact metric) action spaces, the effect of perturbations of the reward function and the transition probability function on the value and the set of stationary (ε-)optimal strategies is studied. "Favourable" answers are also obtained here in theorems 3.4, 3.5 and 3.7.

In section 4 some of the results of section 3 are extended to two-person zero-sum stochastic games where discount factor, reward and transition probability functions are all time-dependent and where Markov strategies take over the role of stationary strategies.

2. PERTURBATIONS OF TWO-PERSON ZERO-SUM GAMES IN NORMAL FORM

A *two-person zero-sum game in normal form* is an ordered triplet $<\Pi_1,\Pi_2,p>$, in which Π_1 and Π_2 are non-empty sets and $p: \Pi_1 \times \Pi_2 \to \mathbb{R}$ is a real-valued function. Π_1 and Π_2 are called the *strategy spaces* of player 1 and player 2, respectively, and p the *payoff function* of player 1. Such a game is played as follows. Player 1 and player 2 choose, independently of one another, a strategy $\pi_1 \in \Pi_1$ and a strategy $\pi_2 \in \Pi_2$, respectively; then player 1 receives a payoff $p(\pi_1,\pi_2)$ from player 2. For such a game $<\Pi_1,\Pi_2,p>$ let

$$\underline{V}(p) := \sup_{\pi_1\in\Pi_1} \inf_{\pi_2\in\Pi_2} p(\pi_1,\pi_2)$$

and

$$\overline{V}(p) := \inf_{\pi_2\in\Pi_2} \sup_{\pi_1\in\Pi_1} p(\pi_1,\pi_2)$$

If $\underline{V}(p) = \overline{V}(p)$, then we say that the *game* is *strictly determined,* and then $\mathrm{val}(p) := \underline{V}(p)$ is called the *value* of the game $<\Pi_1,\Pi_2,p>$. Denote by $BV(\Pi_1,\Pi_2)$ the set of those bounded functions $p: \Pi_1 \times \Pi_2 \to \mathbb{R}$ for which the game $<\Pi_1,\Pi_2,p>$ is strictly determined. For $p \in BV(\Pi_1,\Pi_2)$ and $\varepsilon \geq 0$ let

$$O_1^{\varepsilon}(p) := \{\pi_1 \in \Pi_1 \mid \inf_{\pi_2\in\Pi_2} p(\pi_1,\pi_2) \geq \mathrm{val}(p) - \varepsilon\},$$

$$O_2^{\varepsilon}(p) := \{\pi_2 \in \Pi_2 \mid \sup_{\pi_1\in\Pi_1} p(\pi_1,\pi_2) \leq \mathrm{val}(p) + \varepsilon\}.$$

For each $\varepsilon > 0$ the set $O_i^{\varepsilon}(p) \neq \emptyset$; the elements of $O_i^{\varepsilon}(p)$ are called *ε-optimal strategies* for player $i (i \in \{1,2\})$. The elements of the (possibly empty) set $O_i(p) := O_i^0(p)$ are called *optimal strategies* for player i. Note that

$$(2.1) \qquad O_1(p) = \bigcap_{\varepsilon>0} O_1^{\varepsilon}(p), \quad O_2(p) = \bigcap_{\varepsilon>0} O_2^{\varepsilon}(p)$$

and that val is a monotone function on $BV(\Pi_1,\Pi_2)$ (i.e. $p \le q$ implies $val(p) \le val(q)$). We provide $BV(\Pi_1,\Pi_2)$ with the metric $d: BV(\Pi_1,\Pi_2) \times BV(\Pi_1,\Pi_2) \to \mathbb{R}$ where $d(p,q) := \|p-q\|$ for each $p,q \in BV(\Pi_1,\Pi_2)$.
[Throughout this paper, for a bounded function f on a set S the number $\sup_{x\in S}|f(x)|$ is denoted by $\|f\|$].

We show now that the value function $val: BV(\Pi_1,\Pi_2) \to \mathbb{R}$ is Lipschitz continuous with constant 1.

<u>THEOREM 2.1</u> *For each* $p,q \in BV(\Pi_1,\Pi_2)$ *we have*

$$|val(p) - val(q)| \le d(p,q).$$

<u>PROOF</u>: Let $i: \Pi_1 \times \Pi_2 \to \mathbb{R}$ be the function with $i(\pi_1,\pi_2) = 1$ for each $(\pi_1,\pi_2) \in \Pi_1 \times \Pi_2$. Take $p,q \in BV(\Pi_1,\Pi_2)$. Then $q + d(p,q)i$ and $q - d(p,q)i$ are elements of $BV(\Pi_1,\Pi_2)$ and

$$val(q \pm d(p,q)i) = val(q) \pm d(p,q).$$

Since $q - d(p,q)i \le p \le q + d(p,q)i$, the monotony of the value function implies

$$val(q) - d(p,q) \le val(p) \le val(q) + d(p,q)$$

and so the theorem is proved. □

<u>THEOREM 2.2</u> *Let* $\varepsilon \ge 0$, $\delta > 0$. *Let* $p,q \in BV(\Pi_1,\Pi_2)$ *such that* $d(p,q) \le \delta$. *Then* $O_i^{\varepsilon}(p) \subset O_i^{\varepsilon+2\delta}(q)$ *for each* $i \in \{1,2\}$.

<u>PROOF</u>: We only show the inclusion for $i = 1$. Let $\tilde{\pi}_1 \in O_1^{\varepsilon}(p)$. The following three inequalities hold:

$$q(\tilde{\pi}_1,\pi_2) \ge p(\tilde{\pi}_1,\pi_2) - \delta \text{ for each } \pi_2 \in \Pi_2,$$

$$p(\tilde{\pi}_1,\pi_2) \ge val(p) - \varepsilon \text{ for each } \pi_2 \in \Pi_2,$$

$$val(p) \ge val(q) - \delta.$$

The last inequality follows from theorem 2.1. Combining these three inequalities, we have for each $\pi_2 \in \Pi_2$:

$$q(\tilde{\pi}_1,\pi_2) \geq \mathrm{val}(q) - \varepsilon - 2\delta.$$

Hence $\tilde{\pi}_1 \in O_1^{\varepsilon+2\delta}(q)$. □

So far we have not needed topological assumptions on strategy spaces and the payoff function. Now we look at two-person zero-sum games for which the strategy spaces Π_1 and Π_2 are topological Hausdorff spaces. We shall call a function $p\colon \Pi_1 \times \Pi_2 \to \mathbb{R}$ semicontinuous if for each $\pi_2 \in \Pi_2$ the function $\pi_1 \mapsto p(\pi_1,\pi_2)$ is an upper semicontinuous function on Π_1, and if for each $\pi_1 \in \Pi_1$ the function $\pi_2 \mapsto p(\pi_1,\pi_2)$ is lower semicontinuous on Π_2. Let $SBV(\Pi_1,\Pi_2) := \{p \in BV(\Pi_1,\Pi_2) \mid p \text{ is semicontinuous}\}$. Then for each $\varepsilon > 0$, the set $O_1^{\varepsilon}(p)$ is a closed subset of Π_1 if $p \in SBV(\Pi_1,\Pi_2)$ because $O_1^{\varepsilon}(p) = f^{-1}([\mathrm{val}(p)-\varepsilon,\infty))$, where f is the upper semi-continuous function on Π_1, defined by $f(\pi_1) := \inf_{\pi_2\in\Pi_2} p(\pi_1,\pi_2)$. If, moreover, Π_1 is compact then it follows from (2.1) that $O_1(p) \neq \emptyset$. Analogously, $O_2^{\varepsilon}(p)$ is closed if $p \in SBV(\Pi_1,\Pi_2)$.

Following BERGE [1] pp. 114,115, we call a multifunction f from X into Y, where X and Y are topological spaces, an *upper semicontinuous multifunction* if for each open set $U \subset Y$ the set $\{x \in X \mid f(x) \subset U\}$ is an open subset of X.

<u>THEOREM 2.3</u> *Let* Π_1 *and* Π_2 *be Hausdorff spaces and let* $\varepsilon \geq 0$. If Π_i *is compact* ($i \in \{1,2\}$) *then* $O_i^{\varepsilon}\colon SBV(\Pi_1,\Pi_2) \to \Pi_i$ *is an upper semicontinuous multifunction.*

<u>PROOF</u>: Let Π_1 be a compact space. In view of the corollary on page 118 of BERGE [1], O_1^{ε} is upper semicontinuous if (and only if)

$$G := \{(p,\pi_1) \in SBV(\Pi_1,\Pi_2) \times \Pi_1 \mid \pi_1 \in O_1^{\varepsilon}(p)\}$$

is a closed subset of $SBV(\Pi_1,\Pi_2) \times \Pi_1$. We prove that the complement of G is open. Let $(p,\tilde{\pi}_1) \in (SBV(\Pi_1,\Pi_2) \times \Pi_1)-G$. Then we can take a $\delta > 0$ such that $\inf_{\pi_2\in\Pi_2} p(\tilde{\pi}_1,\pi_2) < \mathrm{val}(p) - \varepsilon - \delta$. Put

$$U := \{q \in SBV(\Pi_1,\Pi_2);\ d(p,q) < \tfrac{1}{2}\delta\}$$

and

$$V := \{\pi_1 \in \Pi_1;\ \inf_{\pi_2\in\Pi_2} p(\pi_1,\pi_2) < val(p) - \varepsilon - \delta\}.$$

Then U is open, and V is also open, because $\pi_1 \mapsto \inf_{\pi_2\in\Pi_2} p(\pi_1,\pi_2)$ is an upper semicontinuous function on Π_1. Hence U x V is an open neighbourhood of $(p,\tilde{\pi}_1)$. Since for each $(q,\pi_1) \in U \times V$ we have

$$\inf_{\pi_2\in\Pi_2} q(\pi_1,\pi_2) \le \inf_{\pi_2\in\Pi_2} p(\pi_1,\pi_2) + \tfrac{1}{2}\delta < val(p) - \varepsilon - \tfrac{1}{2}\delta < val(q) - \varepsilon,$$

we may conclude that $(U\times V) \cap G = \emptyset$. This implies that $(SBV(\Pi_1,\Pi_2)\times \Pi_1) - G$ is an open set. Hence G is closed, and so 0_1^ε is an upper semicontinuous function. □

<u>EXAMPLE 2.4</u>. Now we want to show that the multifunction 0_1^ε ($\varepsilon\ge0$) is not necessarily a lower semicontinuous multifunction. Take $\Pi_1 := [-1,1]$ and $\Pi_2 := \{0\}$. Then Π_1 and Π_2 are compact sets w.r.t. the usual topology. We assert that for each $\varepsilon \ge 0$ the multifunction 0_1^ε is not lower semicontinuous. To prove this assertion, we will show that the lower inverse (cf. [1],p.25 and p. 115)

$$(0_1^\varepsilon)^{-1}((-1,0)) := \{p \in SBV(\Pi_1,\Pi_2) \mid 0_1^\varepsilon(p) \cap (-1,0) \neq \phi\}$$

of the open interval $(-1,0) \subset \Pi_1$ is not an open subset of $SBV(\Pi_1,\Pi_2)$. For that purpose we introduce the sequence of functions $p,p_1,p_2,\ldots$ in $SBV(\Pi_1,\Pi_2)$, where for each $n \in \mathbb{N}$

$$p_n(\pi_1,0) := \begin{cases} 0 & \text{if } \pi_1 \in [-1,0) \\ (\varepsilon+n^{-1})\pi_1 & \text{if } \pi_1 \in [0,1] \end{cases}$$

and

$$p(\pi_1,0) := \begin{cases} 0 & \text{if } \pi_1 \in [-1,0) \\ \varepsilon\, \pi_1 & \text{if } \pi_1 \in [0,1] \end{cases} .$$

Then $\lim_{n\to\infty} d(p_n,p) = 0$. It is easy to see that $val(p) = \varepsilon$, $0_1^\varepsilon(p) = [-1,1]$; $val(p_n) = \varepsilon + n^{-1}$ and $0_1^\varepsilon(p_n) = [(1+n\varepsilon)^{-1},1]$ for each $n \in \mathbb{N}$. Hence

$$p \in (0_1^\varepsilon)^{-1}((-1,0)) \text{ and } p_n \notin (0_1^\varepsilon)^{-1}((-1,0)) \text{ for each } n \in \mathbb{N}.$$

This implies that $(0_1^\varepsilon)^{-1}((-1,0))$ is not open.

From theorem 2.3 we infer the following

COROLLARY 2.5. *Let* Π_1 *and* Π_2 *be Hausdorff spaces and suppose that* Π_1 *is compact. Let* $p_1,p_2,p_3,\ldots$ *be a sequence in* $SBV(\Pi_1,\Pi_2)$ *and* $\pi_1,\pi_2,\pi_3,\ldots$ *a sequence in* Π_1 *such that* $\pi_n \in 0_1(p_n)$ *for each* $n \in \mathbb{N}$. *Let* $p \in SBV(\Pi_1,\Pi_2)$ *such that* $0_1(p)$ *consists of exactly one element, say* π, *and suppose that* $\lim_{n\to\infty} d(p,p_n) = 0$. *Then* $\lim_{n\to\infty} \pi_n$ *exists and is equal to* π.

Let $\pi_1^* \in \Pi_1$, $\pi_2^* \in \Pi_2$. It is well-known that $(\pi_1^*,\pi_2^*) \in 0_1(p) \times 0_2(p)$ if and only if

$$(2.2) \qquad p(\pi_1,\pi_2^*) \le p(\pi_1^*,\pi_2^*) \le p(\pi_1^*,\pi_2) \quad \text{for all } \pi_1 \in \Pi_1,\ \pi_2 \in \Pi_2.$$

In view of (2.2), the elements of $0_1(p) \times 0_2(p)$ are called *saddle-points* of the game $\langle\Pi_1,\Pi_2,p\rangle$. We now want to study the subset $US(\Pi_1,\Pi_2)$ of $SBV(\Pi_1,\Pi_2)$ consisting of those functions p for which $\langle\Pi_1,\Pi_2,p\rangle$ has a unique saddle-point. Note that for each $(\pi_1^*,\pi_2^*) \in \Pi_1 \times \Pi_2$ the bounded semicontinuous function $s_{(\pi_1^*,\pi_2^*)} : \Pi_1 \times \Pi_2 \to \mathbb{R}$ defined by

$$(2.3) \qquad s_{(\pi_1^*,\pi_2^*)}(\pi_1,\pi_2) := \begin{cases} -1 & \text{if } \pi_2 = \pi_2^* \text{ and } \pi_1 \neq \pi_1^* \\ 1 & \text{if } \pi_1 = \pi_1^* \text{ and } \pi_2 \neq \pi_2^* \\ 0 & \text{elsewhere} \end{cases}$$

is an element of $US(\Pi_1,\Pi_2)$ with unique saddle-point (π_1^*,π_2^*). For the proof of the next theorem we need a lemma.

LEMMA 2.6. *Let* $p \in SBV(\Pi_1,\Pi_2)$ *for which* $(\pi_1^*,\pi_2^*) \in 0_1(p) \times 0_2(p)$. *Then* $p + \varepsilon\, s_{(\pi_1^*,\pi_2^*)} \in US(\Pi_1,\Pi_2)$ *for each* $\varepsilon > 0$ *and* $0_i(p+\varepsilon s_{(\pi_1^*,\pi_2^*)}) = \{\pi_i^*\}$ *for* $i \in \{1,2\}$.

PROOF: Since (π_1^*,π_2^*) is a saddle-point in the games $<\Pi_1,\Pi_2,p>$ and $<\Pi_1,\Pi_2,\ \varepsilon\, s_{(\pi_1^*,\pi_2^*)}>$, it is obvious that for each $(\pi_1,\pi_2) \in \Pi_1 \times \Pi_2$:

$$h(\pi_1,\pi_2^*) < h(\pi_1^*,\pi_2^*) < h(\pi_1^*,\pi_2), \quad \text{where } h := p + \varepsilon\, s_{(\pi_1^*,\pi_2^*)}.$$

Hence, in view of (2.2), $(\pi_1^*,\pi_2^*) \in 0_1(h) \times 0_2(h)$. Now suppose that (π_1',π_2') is also a saddle-point of the game $<\Pi_1,\Pi_2,h>$. If $\pi_1' \neq \pi_1^*$, then on one hand we have

$$h(\pi_1^*,\pi_2^*) \le h(\pi_1^*,\pi_2') \le h(\pi_1',\pi_2')$$

and on the other hand

$$h(\pi_1',\pi_2') \le h(\pi_1',\pi_2^*) < h(\pi_1^*,\pi_2^*)$$

which is impossible. Hence $\pi_1' = \pi_1^*$. In a similar manner we may conclude that $\pi_2' = \pi_2^*$. Thus $<\Pi_1,\Pi_2,h>$ has (π_1^*,π_2^*) as unique saddle-point. □

THEOREM 2.7. *Let* Π_1,Π_2 *be compact Hausdorff spaces. Then*

(a) the restriction of $0_i : SBV(\Pi_1,\Pi_2) \to \mathbb{R}$ *to the subset* $US(\Pi_1,\Pi_2)$ *is a continuous map,*

(b) $US(\Pi_1,\Pi_2)$ *is a dense subset of* $SBV(\Pi_1,\Pi_2)$.

PROOF: (a) follows from the fact that a single-valued map, which is upper semicontinuous in the multi-valued sense, is continuous.
(b) Let $p \in SBV(\Pi_1,\Pi_2)$ and $\varepsilon > 0$. We have to prove that there is a $q \in US(\Pi_1,\Pi_2)$ such that $d(p,q) < \varepsilon$. Take $(\pi_1^*,\pi_2^*) \in 0_1(p) \times 0_2(p)$ ($0_1(p) \times 0_2(p) \neq \emptyset$ because Π_1 and Π_2 are compact). Now let $q := p + \frac{1}{2}\varepsilon\, s_{(\pi_1^*,\pi_2^*)}$, where $s_{(\pi_1^*,\pi_2^*)}$ is the function defined in (2.3). Then $q \in US(\Pi_1,\Pi_2)$ by lemma (2.6) and $d(p,q) \le \frac{1}{2}\varepsilon < \varepsilon$. □

THEOREM 2.8. *Let* Π_1, Π_2 *be compact metric spaces with metrics* d_1, d_2. *Then* $US(\Pi_1, \Pi_2)$ *is connected if and only if both strategy spaces* Π_1 *and* Π_2 *are connected.*

PROOF: (a) First suppose that, say, Π_1 is not connected. Let Π_{11} and Π_{12} be two disjunct non-empty open subsets of Π_1 with $\Pi_1 = \Pi_{11} \cup \Pi_{12}$. Let $US_i := \{p \in US \mid O_1(p) \subset \Pi_{1i}\}$ for each $i \in \{1,2\}$. It is obvious that $US = US_1 \cup US_2$ and that $US_1 \cap US_2 = \emptyset$. Further, US_1 and US_2 are open in US, because O_1 and O_2 are upper semicontinuous multifunctions. If we can show that $US_1 \neq \phi$, and $US_2 \neq \phi$, then we have proved that $US(\Pi_1, \Pi_2)$ is not connected if Π_1 is not connected. Now $\Pi_{11} \neq \phi$, $\Pi_{12} \neq \phi$. Take $\pi' \in \Pi_{11}$, $\pi'' \in \Pi_{12}$ and $\pi \in \Pi_2$. Then it is obvious that $s_{(\pi',\pi)} \in US_1$ and $s_{(\pi'',\pi)} \in US_2$, where $s_{(\pi',\pi)}$, $s_{(\pi'',\pi)}$ are defined in an analogous manner to $s_{(\pi_1^*,\pi_2^*)}$ in (2.3). So $US_i \neq \phi$ for $i \in \{1,2\}$; and we have proved the implication to the right in the theorem.
(b) Now we suppose that Π_1 and Π_2 are connected sets. Let U_1 and U_2 be disjunct open subsets of the metric space $US(\Pi_1, \Pi_2)$ such that $US(\Pi_1, \Pi_2) = U_1 \cup U_2$. If we can show that $U_1 = \phi$ or $U_2 = \phi$, then $US(\Pi_1, \Pi_2)$ is connected.
(b.1) For each $(\pi_1^*, \pi_2^*) \in \Pi_1 \times \Pi_2$, let $z_{(\pi_1^*,\pi_2^*)} : \Pi_1 \times \Pi_2 \to \mathbb{R}$ be the function defined by

$$z_{(\pi_1^*,\pi_2^*)}(\pi_1, \pi_2) = d_2(\pi_2, \pi_2^*) - d_1(\pi_1, \pi_1^*) \quad \text{for each } (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2.$$

Then

$$z_{(\pi_1^*,\pi_2^*)} \in US(\Pi_1, \Pi_2) \quad \text{and } O_i(z_{(\pi_1^*,\pi_2^*)}) = \{\pi_i^*\} \quad \text{for } i \in \{1,2\}.$$

Let $F: \Pi_1 \times \Pi_2 \to US(\Pi_1, \Pi_2)$ be the map defined by $F(\pi_1^*, \pi_2^*) = z_{(\pi_1^*,\pi_2^*)}$. Then it is straightforward to show that $\| F(\pi_1^*, \pi_2^*) - F(\pi_1^{**}, \pi_2^{**}) \| \leq$ $\leq d_1(\pi_1^*, \pi_1^{**}) + d_2(\pi_2^*, \pi_2^{**})$ for all (π_1^*, π_2^*), $(\pi_1^{**}, \pi_2^{**}) \in \Pi_1 \times \Pi_2$. Hence F is a continuous map from the connected set $\Pi_1 \times \Pi_2$ into $US(\Pi_1, \Pi_2)$. This implies that either $F(\Pi_1 \times \Pi_2) \subset U_1$ or $F(\Pi_1 \times \Pi_2) \subset U_2$. Without loss of generality we suppose that $F(\Pi_1 \times \Pi_2) \subset U_1$ i.e.

$$(2.4) \qquad z_{(\pi_1^*,\pi_2^*)} \in U_1 \quad \text{for each } (\pi_1^*, \pi_2^*) \in \Pi_1 \times \Pi_2.$$

(b.2) Now take an arbitrary $p \in US(\Pi_1,\Pi_2)$. Let $0_1(p) \times 0_2(p) = \{(\pi_1,\pi_2)\}$. For each $t \in [0,1]$ let $G(t) := t\,p + (1-t)z_{(\pi_1,\pi_2)}$. Then it is easy to show (cf. the proof of lemma (2.6)) that $G(t) \in US(\Pi_1,\Pi_2)$ for each $t \in [0,1]$. Furthermore,

$$\|G(s) - G(t)\| \le |s-t| \; (\|p\|+\|z_{(\pi_1,\pi_2)}\|) \quad \text{for each } s,\, t \in [0,1].$$

Hence $G: [0,1] \to US(\Pi_1,\Pi_2)$ is continuous. Since $[0,1]$ is connected and $G(0) = z_{(\pi_1,\pi_2)} \in U_1$ by (2.4), we may conclude that $G(1) = p \in U_1$, as well. So we have proved that $US(\Pi_1,\Pi_2) \subset U_1$. Thus $U_2 = \phi$; this completes the proof of the theorem. □

Note that the metric property of Π_1 and Π_2 in theorem 2.8 is only used in the proof of the implication to the left of that theorem.

<u>REMARKS</u>.

<u>2.9.1</u>. The set $US(\Pi_1,\Pi_2)$ is not necessarily an open subset of $SBV(\Pi_1,\Pi_2)$ as the following example shows. Take $\Pi_1 := [0,1]$, $\Pi_2 := \{0\}$. Then $p \in US(\Pi_1,\Pi_2)$ if $p(\pi_1,0) := \pi_1$ for each $\pi_1 \in \Pi_1$. For each $\varepsilon > 0$, the ε-neighbourhood of p contains the function $q \in SBV(\Pi_1,\Pi_2)$, defined by $q(\pi_1,0) := \min\{\pi_1, 1-\frac{1}{2}\varepsilon\}$, but $q \notin US(\Pi_1,\Pi_2)$. Hence $US(\Pi_1,\Pi_2)$ is not open.

<u>2.9.2</u>. BOHNENBLUST, KARLIN & SHAPLEY proved in [3] that the set U_{mn} of those $m \times n$-matrix games ($m,n \in \mathbb{N}$), for which the mixed extension has a unique saddle-point, is an open and dense subset of the set of all $m \times n$-matrix games (provided with the usual topology). With some labour one can prov that U_{mn} is not connected for all $(m,n) \neq (1,1)$. We will not do this here but remark that in case $(m,n) = (1,2)$ we have:

$$U_{12} = \{[a,b] \mid a \neq b\} = \{[a,b] \mid a > b \cup \{[a,b] \mid a < b\}.$$

Hence U_{12} is the union of two disjunct open subsets. Thus U_{12} is not connected.

<u>2.9.3</u>. For semi-infinite matrix games the influence of perturbations of the payoffs on value and (ε-)optimal strategies was studied in TIJS [10],

pp. 65-70.

2.9.4. Let Π_1 and Π_2 be compact metric spaces. Let $CV(\Pi_1,\Pi_2)$ be the family of continuous functions $p\colon \Pi_1 \times \Pi_2 \to \mathbb{R}$ such that the value of the game $\langle\Pi_1,\Pi_2,p\rangle$ exists. Let $UC(\Pi_1,\Pi_2) := \{p \in CV(\Pi_1,\Pi_2) \mid \langle\Pi_1,\Pi_2,p\rangle$ has a unique saddle-point$\}$. Then $UC(\Pi_1,\Pi_2)$ is a dense subset of $CV(\Pi_1,\Pi_2)$; also $UC(\Pi_1,\Pi_2)$ is connected iff Π_1 and Π_2 are connected. The proofs of these facts can be given by a minor modification of the proofs of 2.7 and 2.8. [E.g. the role of the function $s_{(\pi_1^*,\pi_2^*)}$ in the proof of 2.7 can be taken over by the function $z_{(\pi_1^*,\pi_2^*)}$ defined in the proof of 2.8.]

3. PERTURBATIONS IN STOCHASTIC GAMES

In this section we extend some results of section 2 to stochastic games. Stochastic games (or Markov games) were introduced in 1953 by SHAPLEY [9]. For a recent survey of the theory of stochastic games we refer to PARTHASARATHY & STERN [7]. In this section we restrict our attention to *discounted two-person zero-sum stochastic games*, characterized by an ordered six-tuple $\langle S,A_1,A_2,r,q,\beta\rangle$, where

(3.1) S is a non-empty countable set, called the *state space*,

(3.2) A_1 and A_2 are non-empty compact metric spaces, called the *action spaces* of player 1 and player 2, respectively,

(3.3) $r\colon S \times A_1 \times A_2 \to \mathbb{R}$ is a bounded function, called the *reward function*, for which for each $s \in S$ the map $(a_1,a_2) \mapsto r(s,a_1,a_2)$ is a measurable function on $A_1 \times A_2$ (the measurability is taken with respect to the product σ-algebra of $\mathcal{A}_1$ and $\mathcal{A}_2$, where $\mathcal{A}_i$ is the σ-algebra generated by the Borel sets of $A_i (i=1,2)$),

(3.4) $q\colon S \times A_1 \times A_2 \to P$ is a function from $S \times A_1 \times A_2$ into the family P of probability measures on S, such that for all $s, s' \in S$ the map $(a_1,a_2) \mapsto q(s' \mid s,a_1,a_2) := q(s,a_1,a_2)\{s'\}$ is a measurable function on $A_1 \times A_2$. q is called the *transition probability function*,

(3.5) β is a real number in $[0,1)$, called the *discount factor*.

Such a stochastic game corresponds with a dynamic system with state space S, where the dynamic behaviour as well as the rewards are influenced by the players at discrete points in time, say $t = 0,1,2,\ldots$, in the following way. At each time $t \in \{0,1,2,\ldots\}$ the players observe the current state of the system. They, then, have to select, independently of one another, an action. If at time t the system is in state s and if player 1 selects action $a_1 \in A_1$ and player 2 action $a_2 \in A_2$, then two things happen:

(1) player 1 obtains an immediate reward $r(s,a_1,a_2)$ from player 2.

(2) the system moves with probability $q(s' \mid s,a_1,a_2)$ to the state $s' \in S$, which is observed at time $t + 1$.

Furthermore, one supposes that a reward r to player 1 (or 2) at time t has worth $\beta^t r$ at time 0 ($\beta^t r$ is called the *discounted reward*) and that player 1 (player 2) wants to maximize (minimize) the total discounted expected reward.

<u>DEFINITION 3.1</u>. Let $\langle S,A_1,A_2,r,q,\beta\rangle$ be a stochastic game. Let P_i be the set of probability measures on $\langle A_i,\mathcal{A}_i\rangle$ $(i=1,2)$. Then each map $\pi_i : S \to P_i$ is called a *stationary strategy* for player i. The set of stationary strategies is denoted by Π_i.

Playing a stationary strategy $\pi_i \in \Pi_i$ means for player i that, each time $t \in \{0,1,2,\ldots\}$ that the system is in state $s \in S$, he chooses his action according to the probability measure $\pi_i(s)$.

Let us suppose that the players 1 and 2 decide to play $\pi_1 \in \Pi_1$ and $\pi_2 \in \Pi_2$. Suppose further that the initial state (the state at $t = 0$) of the system is $s \in S$. Then the expected reward of player 1 at time $t \in \{0,1,2,\ldots\}$ exists and is denoted by $f^t_{srq}(\pi_1,\pi_2)$; the *total discounted expected reward* $\sum_{t=0}^{\infty} \beta^t f^t_{srq}(\pi_1,\pi_2)$ is denoted by $f_{srq}(\pi_1,\pi_2)$. Note that $\|f_{srq}\| \le \sum_{t=0}^{\infty} \beta^t \|r\| = (1-\beta)^{-1} \|r\|$. Furthermore, it can be seen that the function $s \mapsto f_{srq}(\pi_1,\pi_2)$ satisfies the relation:

$$(3.6) \qquad f_{srq}(\pi_1,\pi_2) = \tilde{r}(s,\pi_1(s),\pi_2(s)) + \\ + \beta \sum_{s'\in S} \tilde{q}(s'|s,\pi_1(s),\pi_2(s))\, f_{s'rq}(\pi_1,\pi_2)$$

for all $s \in S$, where

$$\tilde{r}(s,\pi_1(s),\pi_2(s)) := \int_{A_2}\int_{A_1} r(s,a_1,a_2)d\pi_1(s)(a_1)d\pi_2(s)(a_2)$$

and

$$\tilde{q}(s'|s,\pi_1(s),\pi_2(s)) := \int_{A_2}\int_{A_1} q(s'|s,a_1,a_2)\, d\pi_1(s)(a_1)\, d\pi_2(s)(a_2).$$

($\tilde{r}(s,\pi_1(s),\pi_2(s))$ is the expected reward at time 0 and if at time 1 the state is $s' \in S$ (chance $\tilde{q}(s'|s,\pi_1(s),\pi_2(s))$, then s' can be seen as a new starting state, so the discounted expected reward from time 1 on is $\beta \sum_{s' \in S} \tilde{q}(s'|s,\pi_1(s),\pi_2(s)) f_{s'rq}(\pi_1,\pi_2)$.)

DEFINITION 3.2. Let $\langle S,A_1,A_2,r,q,\beta\rangle$ be a stochastic game and $\varepsilon \geq 0$. A pair of stationary strategies $(\pi_1^\varepsilon,\pi_2^\varepsilon) \in \Pi_1 \times \Pi_2$, such that

$$-\varepsilon + f_{srq}(\pi_1,\pi_2^\varepsilon) \leq f_{srq}(\pi_1^\varepsilon,\pi_2^\varepsilon) \leq f_{srq}(\pi_1^\varepsilon,\pi_2) + \varepsilon$$

for all $s \in S$ and all $(\pi_1,\pi_2) \in \Pi_1 \times \Pi_2$ is called an *ε-saddle-point* if $\varepsilon > 0$, and a *saddle-point* if $\varepsilon = 0$. If, for each $\varepsilon > 0$, there are ε-saddle-points, then we say that the *stochastic game* is *strictly determined*. In that case, for each $s \in S$, the two-person game in normal form $\langle\Pi_1,\Pi_2,f_{srq}\rangle$ is strictly determined and the function $V_{rq} : S \to \mathbb{R}$, where $V_{rq}(s)$ is the value of $\langle\Pi_1,\Pi_2,f_{srq}\rangle$, is called the *value* of the stochastic game. By an *ε-optimal (optimal) strategy* $\pi_i \in \Pi_i$ for player i *in the stochastic game* we mean a strategy such that $\pi_i(s)$ is ε-optimal (optimal) in $\langle\Pi_1,\Pi_2,f_{srq}\rangle$ for all $s \in S$.

For the remainder of this section S, A_1, A_2 and β are fixed. Let DV be the family of pairs of functions (r,q) satisfying (3.3) and (3.4) such that for each bounded function $Y : S \to \mathbb{R}$ and all $s \in S$ the *(dummy-) game* in normal form

$$\langle P_1,\ P_2,\ \tilde{r}(s,.,.) + \beta \sum_{s' \in S} \tilde{q}(s' \mid s,.,.)\, Y(s')\rangle \text{ has a value.}$$

THEOREM 3.3. *Let* $(r,q) \in DV$ *and* $\varepsilon \geq 0$. *Then* $<S,A_1,A_2,r,q,\beta>$ *is strictly determined. The value of the stochastic game is the unique solution of the following functional equation in* $Y : S \to \mathbb{R}$:

$$Y(s) = \mathrm{val}(\tilde{r}(s,.,.) + \beta \sum_{s' \in S} \tilde{q}(s' \mid s,.,.)\, Y(s')) \text{ for all } s \in S.$$

Furthermore, if for each $s \in S$ *an* ε*-optimal strategy* $\pi_i^{\varepsilon}(s)$ *is given for player* i *in the game in normal form*

$$<P_1,P_2,\tilde{r}(s,.,.) + \beta \sum_{s' \in S} \tilde{q}(s' | s,.,.)\, V_{rq}(s')>,$$

then the map $s \mapsto \pi_i^{\varepsilon}(s)$ *is a* $(1-\beta)^{-1}\varepsilon$*-optimal strategy for the stochastic game.*

PROOF: Let B(S) be the family of bounded realvalued functions on S. Let $T : B(S) \to B(S)$ be the map defined by

$$(TY)(s) := \mathrm{val}(\tilde{r}(s,.,.) + \beta \sum_{s' \in S} \tilde{q}(s' | s,.,.)\, Y(s'))$$

for all $Y \in B(S)$ and $s \in S$.

Then, using theorem 2.1, we have

$$\| TY_1 - TY_2 \| \leq \beta \, \| Y_1 - Y_2 \| \quad \text{for each } Y_1,\ Y_2 \in B(S).$$

Hence T is a contraction with factor $\beta \in [0,1)$, so that by the Banach-Picard fixed point theorem T has a unique fixed point, say V. So V satisfies: $V(s) = \mathrm{val}(\tilde{r}(s,.,.) + \beta \sum_{s' \in S} \tilde{q}(s' | s,.,.)\, V(s'))$ for all $s \in S$. We now show that V is the value of the game $<S,A_1,A_2,r,q,\beta>$. For $\varepsilon > 0$ let $\pi_1^{\varepsilon}(s) \in P_1$ be ε-optimal in $<P_1,P_2,\tilde{r}(s,.,.) + \beta \sum_{s' \in S} \tilde{q}(s' | s,.,.) V(s')>$. For $(\pi_1,\pi_2) \in \Pi_1 \times \Pi_2$ let $Q_{\pi_1\pi_2} : B(S) \to B(S)$ be the map defined by

$$(Q_{\pi_1\pi_2} Y)(s) := \sum_{s' \in S} \tilde{q}(s' | s,\pi_1,\pi_2)\, Y(s')$$

for all $Y \in B(S)$ and $s \in S$.

Consider the strategy $s \mapsto \pi_1^\varepsilon(s)$, denoted by π_1^ε. Then, for each $\pi_2 \in \Pi_2$, we have:

$$(3.7) \qquad \bar{r}(\pi_1^\varepsilon,\pi_2) + \beta\, Q_{\pi_1^\varepsilon\pi_2} V + \varepsilon_S \geq_S V,$$

where $\bar{r}(\pi_1^\varepsilon,\pi_2)$ is the function $s \mapsto \tilde{r}(s,\pi_1^\varepsilon(s),\pi_2(s))$ on S, ε_S is the function $s \mapsto \varepsilon$ on S and $\geq_S$ is the relation on B(S), defined by $Y_1 \geq_S Y_2 \iff$ $\iff Y_1(s) \geq Y_2(s)$ for all $s \in S$. When we repeatedly substitute for V in the left hand side of (3.7) the entire left hand side, then we see, that for each $t \in \mathbb{N}$ the following inequality holds:

$$(3.8) \qquad \sum_{\tau=0}^{t-1} \beta^\tau\, Q^\tau_{\pi_1^\varepsilon\pi_2}\, \bar{r}(\pi_1^\varepsilon,\pi_2) + \beta^t\, Q^t_{\pi_1^\varepsilon\pi_2} V + \sum_{\tau=0}^{t-1} \beta^\tau\, Q^\tau_{\pi_1^\varepsilon\pi_2}(\varepsilon_S) \geq_S V,$$

where $Q^0_{\pi_1^\varepsilon\pi_2}$ is the identity map and $Q^t_{\pi_1^\varepsilon\pi_2} := Q\, Q^{t-1}_{\pi_1^\varepsilon\pi_2}$ $(t \in \mathbb{N})$. Note that $\sum_{\tau=0}^{t-1} \beta^\tau Q^\tau_{\pi_1^\varepsilon\pi_2}\, \bar{r}(\pi_1^\varepsilon,\pi_2)$ equals the total expected discounted reward until time t, if the players use the strategies π_1^ε and π_2. So letting $t \to \infty$ in (3.8) we get

$$(3.9) \qquad \bar{f}_{rq}(\pi_1^\varepsilon,\pi_2) \geq_S V - (1-\beta)^{-1}\, \varepsilon_S \qquad \text{for each } \pi_2 \in \Pi_2$$

where $\bar{f}_{rq}(\pi_1^\varepsilon,\pi_2)$ is the function $s \mapsto f_{srq}(\pi_1^\varepsilon,\pi_2)$ on S. As $\varepsilon > 0$ was arbitrary, (3.9) yields:

$$\sup_{\pi_1} \inf_{\pi_2} f_{srq}(\pi_1,\pi_2) \geq V(s) \qquad \text{for all } s \in S.$$

Similarly, we can show that $\inf_{\pi_2} \sup_{\pi_1} f_{srq}(\pi_1,\pi_2) \leq V(s)$ and, as always, $\sup_{\pi_1} \inf_{\pi_2} f_{srq}(\pi_1,\pi_2) \leq \inf_{\pi_2} \sup_{\pi_1} f_{srq}(\pi_1,\pi_2)$. Consequently, we arrive at the desired result:

$$(3.10) \qquad \sup_{\pi_1} \inf_{\pi_2} f_{srq}(\pi_1,\pi_2) = \inf_{\pi_2} \sup_{\pi_1} f_{srq}(\pi_1,\pi_2) = V(s) \qquad \text{for all } s \in S.$$

So V is the value of the stochastic game and then (3.9) yields, for $\varepsilon > 0$, the second assertion in the theorem. Now it is easy to show that this assertion also holds for $\varepsilon = 0$; this is left to the reader. □

Now we provide DV with the metric d defined by

$$d((r,q),(r'q')) := \max\{\|r-r'\|,\ \rho(q,q')\},$$

where $\rho(q,q') := \sup_{s',s,a_1,a_2} |q(s' \mid s,a_1,a_2)-q'(s' \mid s,a_1,a_2)|$.

<u>THEOREM 3.4</u>. *The map* $(r,q) \mapsto V_{rq}$ *from* DV *into* B(S) *is a continuous map (even pointwise Lipschitz continuous).*

<u>PROOF</u>: Let $(r,q),(r',q') \in DV$. First note that, in view of theorem 2.1, we have

$$(3.11) \qquad |V_{rq}(s)-V_{r'q'}(s)| = |\mathrm{val}(f_{srq})-\mathrm{val}(f_{sr'q'})| \le \|f_{srq}-f_{sr'q'}\|.$$

Take $(\pi_1,\pi_2) \in \Pi_1 \times \Pi_2$ and put $x(s) := f_{srq}(\pi_1,\pi_2)$ and $x'(s) := f_{sr'q'}(\pi_1,\pi_2)$ for each $s \in S$. Then it follows from (3.6) that for each $s \in S$

$$|x(s)-x'(s)| \le \|r-r'\| + \beta\ \|x-x'\| + \beta\ \rho\ (q,q')\ \|x\|,$$

so

$$\|x-x'\| \le \|r-r'\| + \beta\ \|x-x'\| + \beta\ \rho\ (q,q')\ \|x\|.$$

Recall that $\|x\| \le (1-\beta)^{-1}\ \|r\|$ and put

$$(3.12) \qquad C_r := (1-\beta)^{-1}\ (1+\beta(1-\beta)^{-1}\|r\|).$$

Then

$$(3.13) \qquad \|f_{srq}-f_{sr'q'}\| \le C_r\ d((r,q),(r',q')) \quad \text{for each } s \in S.$$

Combining (3.11) and (3.13), we obtain:

$$\|V_{rq}-V_{r'q'}\| \le C_r\ d((r,q),(r',q'))$$

and this implies that V is (pointwise Lipschitz) continuous in (r,q). □

Let $\varepsilon > 0$ and $(r,q) \in DV$. Denote the set of ε-optimal strategies for player i of the game $<P_1,P_2,\tilde{r}(s,.,.) + \beta \sum_{s'\in S} \tilde{q}(s'|s,.,.)V_{rq}(s')>$ by $O_i^{\varepsilon}(s,r,q)$ and the set of optimal strategies by $O_i(s,r,q)$, $(i=1,2)$. Then $X_{s\in S}\ O_i^{\varepsilon}(s,r,q)$ can be seen as a subset of the set of $(1-\beta)^{-1}$ ε-optimal strategies of the stochastic game $<S,A_1,A_2,r,q,\beta>$ and $X_{s\in S}\ O_i(s,r,q)$ can be identified with the set of optimal strategies (cf. theorem 3.3). The influence of perturbations of (r,q) on this subset $X_{s\in S}\ O_i^{\varepsilon}(s,r,q)$ of the set of $(1-\beta)^{-1}$ ε-optimal strategies can be studied by looking at $O_i^{\varepsilon}(s,r,q)$ for each $s \in S$.

THEOREM 3.5. *Let* $\varepsilon \geq 0$ *and* (r,q), $(r',q') \in DV$, *such that* $d((r,q),(r',q')) \leq \delta$. *Then for each* $s \in S$ *we have*

$$O_i^{\varepsilon}(s,r,q) \subset O_i^{\varepsilon+2C_r\delta}(s,r',q')$$

with C_r *as defined in (3.12).*

PROOF: This theorem is a direct consequence of (3.13) and theorem 2.2. □

Let CDV be the subset of DV consisting of the elements (r,q), such that for each $s,s' \in S$ the realvalued functions on $A_1 \times A_2 : (a_1,a_2) \mapsto r(s,a_1,a_2)$ and $(a_1,a_2) \mapsto q(s'|s,a_1,a_2)$ are continuous. Now endow P_i with the weak topology. Then P_i is compact (cf. PARTHASARATHY [5],th.6.4, p.45), and so $\Pi_i = P_i^S$, provided with the product topology, is also compact.

THEOREM 3.6. *Let* $(r,q) \in CDV$. *Then*

(1) for each $s \in S$ *the function* $f_{srq} : \Pi_1 \times \Pi_2 \to \mathbb{R}$ *is continuous.*

(2) $O_i(s,r,q) \neq \emptyset$ *for each* $s \in S$ *and* $i \in \{1,2\}$.

(3) There is a one-to-one correspondence between the set of optimal stationary strategies for player i *in the stochastic game and the set* $X_{s\in S}\ O_i(s,r,q)$, $i \in \{1,2\}$.

PROOF: The statements in this theorem are special cases of more general statements in VRIEZE [13] (especially lemma 2.1 and theorem 2.1). □

THEOREM 3.7. *For each* $s \in S$ *and* $i \in \{1,2\}$ *we have*

$$0_i(s,.,.): CDV \to P_i$$

is an upper semicontinuous multifunction and therefore also
$X_{s \in S} 0_i(s,.,.): CDV \to \Pi_i$.

PROOF: This statement is a direct consequence of theorem 2.3. □

REMARKS.

3.8.1. In this section we have studied stochastic games, for which in each state $s \in S$ player i has available to him the same set A_i of pure actions. However this restriction is not serious with respect to models in which in each state player i can only use a state-dependent subset of A_i, as is pointed out in PARTHASARATHY [6] (remark 3.2, p.30). So our results also apply to that case.

3.8.2. So far we have restricted ourselves to the classes of stationary strategies for both players. A reason to do so is given in the following theorem. The proof of this theorem runs along the same lines as the proof of the analogous statements for Markov-decision problems (BLACKWELL [2], th. 6, p. 232) and will be omitted here.

THEOREM. *If a stochastic game* $\langle S,A_1,A_2,r,q,\beta \rangle$ *is strictly determined within the classes of all behavioural strategies, then it is strictly determined within the classes of stationary strategies and the value in both cases is the same. Furthermore, if a player has an optimal behavioural strategy, then he has an optimal stationary strategy, which is optimal within the classes of all behavioural strategies.*

3.8.3. Now, let A_1 and A_2 be finite sets consisting of m and n elements, respectively ($m,n \in \mathbb{N}$). Once again let S be a countable set. Let B(S,m,n) consist of the pairs (r,q) with r as in (3.3) and q as in (3.4). As SHAPLEY [9] proved, for each pair $(r,q) \in B(S,m,n)$ the stochastic game $\langle S,A_1,A_2,r,q,\beta \rangle$ has a value and both players have stationary optimal strategies. Now for each $s \in S$ the dummy game $\langle P_1,P_2,\tilde{r}(s,.,.) +$

$+ \beta \sum_{s' \in S} \tilde{q}(s'|s,.,.)\ V_{rq}(s')>$ with value $V_{rq}(s)$ can be seen as a mixed extension of an m x n-matrix game. Let U(S,m,n) be the subset of pairs (r,q) $\in$ B(S,m,n) for which the game $<S,A_1,A_2,r,q,\beta>$ has a unique pair of optimal strategies. Now, if (r,q) $\in$ B(S,m,n), then we can see from remark 2.9.2 that, for each $\varepsilon > 0$, there exists a pair $(r_u,q) \in$ B(S,m,n) with $\|r-r_u\| < \varepsilon$ and such that for each $s \in S$ the game in normal form

$$<P_1,P_2,\tilde{r}_u(s,.,.) + \beta \sum_{s' \in S} \tilde{q}(s'|s,.,.)\ V_{rq}(s')>$$

possesses a unique pair of optimal strategies and, furthermore, we may suppose that the game has value $V_{rq}(s)$. But this means (cf. theorem 3.3) that the stochastic game $<S,A_1,A_2,r_u,q,\beta>$ has value V_{rq} and possesses a unique pair of optimal strategies. So $(r_u,q) \in$ U(S,m,n). The following theorem is now immediate.

THEOREM. *The set* U(S,m,n) *is an open and dense subset of* B(S,m,n).

3.8.4. A criterion other than the discounted reward criterion, which is also often considered, is the average reward per unit of time criterion. Note that, in deducing the theorems of this section, the main argument we use is, that small perturbations of the game parameters (r,q) cause for each pair of stationary strategies small deviations of the expected discounted reward. In general this is not the case, when we look at the average reward per unit of time, because small perturbations for q may cause a change in the chainstructures, belonging to the divers pairs of strategies. When we only admit perturbations of r, then small deviations cause small deviations in the average reward per unit of time for each pair of stationary strategies. In SCHWEITZER [8] one can find, that small perturbations of q, which cause no change in the chain structure for each pair of strategies, yield small deviations of the average reward per unit of time (a sufficient condition). So by choosing an appropriate family of pairs (r,q) (it is not yet known if every game with finite S,A_1 and A_2 has a value with respect to the average reward criterion) adaptions of the theorems 3.4, 3.5 and 3.7 hold true for stochastic games under the average reward per unit of time criterion.

In the next section we direct our attention to non-stationary stochastic games and shall see, that analogous theorems to those in this section can be stated.

4. NON-STATIONARY STOCHASTIC GAMES

So far we only have considered perturbations of the game parameters r and q and, furthermore, we have assumed them time-independent. In this section we look at stochastic games in which r, q and β are each time-dependent and we shall study the influence of perturbations of them. In the following the set $\{0,1,2,\ldots\}$ of non-negative integers is denoted by $\mathbb{N}_0$.

A *non-stationary two-person zero-sum stochastic game* is characterized by a quadruplet $<S,A_1,A_2,<r_t,q_t,\beta_t>>$, where S,A_1 and A_2 are as in (3.1) and (3.2) and $<r_t,q_t,\beta_t>$ denotes the infinite sequence of triples

$$(r_0,q_0,\beta_0),\ (r_1,q_1,\beta_1),\ (r_2,q_2,\beta_2),\ldots$$

with the property that r_t,q_t and β_t statisfy (3.3), (3.4) and (3.5), respectively, for each $t \in \mathbb{N}_0$. Now r_t,q_t and β_t, respectively, are called *reward function, transition probability function* and *discount factor at time* $t \in \mathbb{N}_0$.

DEFINITION 4.1. Let $<S,A_1,A_2,<r_t,q_t,\beta_t>>$ be a non-stationary two-person zero-sum stochastic game. Let P_i be the set of probability measures on $<A_i,\mathcal{A}_i>$. Then each map $\pi_i^M: \mathbb{N}_0 \times S \to P_i$ is called a *Markov-strategy* (or *memoryless strategy*) *for player* i. The set of Markov-strategies of player i is denoted by Π_i^M.

Fix $\tilde{\beta} \in [0,1)$ and $M \in [0,\infty)$. Let

$$F_{\tilde{\beta}M} := \{(r,q,\beta) \mid (r,q) \in DV,\ \beta \in [0,\tilde{\beta}],\ \|r\| \le M\}.$$

Let $<S,A_1,A_2,<r_t,q_t,\beta_t>>$ be a game such that $<r_t,q_t,\beta_t> \in (F_{\tilde{\beta}M})^{\mathbb{N}_0}$ i.e. $(r_t,q_t,\beta_t) \in F_{\tilde{\beta}M}$ for each $t \in \mathbb{N}_0$. Let us suppose that the players 1 and 2 decide to play $\pi_1^M \in \Pi_1^M$ and $\pi_2^M \in \Pi_2^M$, respectively. Then, for each initial state $s \in S$, the expected reward of player 1 at time $t \in \mathbb{N}_0$ exists and is denoted by $f^t_{s<r_t q_t>}(\pi_1^M,\pi_2^M)$. The total expected discounted reward $\sum_{t=0}^{\infty} \Pi_{\tau=0}^{t-1} \beta_\tau f^t_{s<r_t q_t>}(\pi_1^M,\pi_2^M)$ is denoted by $f_{s<r_t,q_t,\beta_t>}(\pi_1^M,\pi_2^M)$, where $\Pi_{\tau=0}^{-1} \beta_\tau := 1$. We note that

$$\| f_{s<r_t,q_t,\beta_t>} \| \le \sum_{t=0}^{\infty} \prod_{\tau=0}^{t-1} \beta_\tau \|r_t\| \le (1-\tilde{\beta})^{-1} M.$$

The notions of *value* and (ε-) *optimal strategies* are defined in a similar way as in section 3, but now the roles of Π_1 and Π_2 in section 3 are taken over by Π_1^M and Π_2^M. In the following, for a sequence $<r_t,q_t,\beta_t>$ and a $\tau \in \mathbb{N}_0$ the sequence

$$(r_\tau,q_\tau,\beta_\tau),\ (r_{\tau+1},q_{\tau+1},\beta_{\tau+1}),\ldots$$

is denoted by $<r_t^{-\tau},q_t^{-\tau},\beta_t^{-\tau}>$. Further, for a game $<S,A_1,A_2,<r_t,q_t,\beta_t>>$ with value, this value is denoted by $s \mapsto V(s,<r_t,q_t,\beta_t>)$ or by $V(.,<r_t,q_t,\beta_t>)$.

THEOREM 4.2. *Let* $<S,A_1,A_2,<r_t,q_t,\beta_t>>$ *be a non-stationary stochastic two-person zero-sum game, such that* $<r_t,q_t,\beta_t> \in (F_{\tilde{\beta}M})^{\mathbb{N}_0}$. *Then the game is strictly determined. Let* $\varepsilon \ge 0$ *and let* $\pi_i^{M_\varepsilon}$ *be a Markov strategy for player* i*, such that (for each* $\tau \in \mathbb{N}_0$ *and* $s \in S$*)* $\pi_i^{M_\varepsilon}(\tau,s)$ *is an* ε*-optimal strategy in the following game in normal form:*

$$<P_1,P_2,\tilde{r}_\tau(s,.,.) + \beta_\tau \sum_{s'\in S} \tilde{q}_\tau(s'|\, s,..)V(s',<r_t^{-\tau-1},q_t^{-\tau-1},\beta_t^{-\tau-1}>)>.$$

Then $\pi_i^{M_\varepsilon}$ *is a* $(1-\tilde{\beta})^{-1}\varepsilon$*-optimal strategy in the stochastic game.*

PROOF. Let B be the family of bounded realvalued functions on $S \times (F_{\tilde{\beta}M})^{\mathbb{N}_0}$. Then B is a complete metric space, if we provide B with the metric derived from the sup-norm. Let $\tilde{T}: B \to B$ be the map such that

$$(\tilde{T}f)(s,<r_t,q_t,\beta_t>) := \mathrm{val}(\tilde{r}_0(s,.,.) +$$

$$+ \beta_0 \sum_{s'\in S} \tilde{q}_0(s' \mid s,.,.)f(s',<r_t^{-1},q_t^{-1},\beta_t^{-1}>))$$

for each $f \in B$ and each $s \in S$ and $\langle r_t, q_t, \beta_t \rangle \in (F_{\tilde{\beta}M})^{\mathbb{N}_0}$. Then $\|\tilde{T}f_1 - \tilde{T}f_2\| \leq \tilde{\beta}\|f_1 - f_2\|$ for each $f_1, f_2 \in B$. So $\tilde{T}: B \to B$ is a contraction map with factor $\tilde{\beta} < 1$. This implies that $\tilde{T}$ has a unique fixed point. The proof of the theorem can now be concluded in a similar way as the proof of theorem 3.3, whereby the role of the stationary strategies there is taken over by the Markov strategies. □

As an inequality similar to (3.13) can be compounded, an easy extension of the theorems 2.1 and 3.4 leads to

THEOREM 4.3. *The map* $\langle r_t, q_t, \beta_t \rangle \mapsto V(., \langle r_t, q_t, \beta_t \rangle)$ *from* $(F_{\tilde{\beta}M})^{\mathbb{N}_0}$ *into* $B(S)$, *where* $B(S)$ *is the metric space of bounded realvalued functions on* S, *is a continuous map.*

Also the theorems 3.5 and 3.7 can be extended to the case of non-stationary stochastic games. We only indicate the extension of theorem 3.7. Let $CF_{\tilde{\beta}M} := \{(r,q,\beta) \in CDV \times [0,\tilde{\beta}] \mid \|r\| \leq M\}$. Let $\langle S, A_1, A_2, \langle r_t, q_t, \beta_t \rangle\rangle$ be a game with $\langle r_t, q_t, \beta_t \rangle \in (CF_{\tilde{\beta}M})^{\mathbb{N}_0}$. Then for the game in normal form

$$\langle P_1, P_2, \tilde{r}_0(s,.,.) + \beta_0 \sum_{s' \in S} \tilde{q}_0(s'|s,..)V(s', \langle r_t^{-1}, q_t^{-1}, \beta_t^{-1} \rangle)\rangle$$

the set $O_i(s, \langle r_t, q_t, \beta_t \rangle)$ of optimal strategies for player i is a non-empty set. Furthermore, it can be shown that there is a one-to-one correspondence between the set $X_{\tau \in \mathbb{N}_0} X_{s \in S} O_i(s, \langle r_t^{-\tau}, q_t^{-\tau}, \beta_t^{-\tau} \rangle)$ and the set of optimal Markov strategies for player i (cf. theorem 3.6).

THEOREM 4.4. *For each* $s \in S$ *the multivalued map* $O_i(s, \langle .,.,. \rangle): (CF_{\tilde{\beta}M})^{\mathbb{N}_0} \to P_i$ *is an upper semicontinuous multifunction.*

In the next section, where we return to games in normal form, we shall concern ourselves with non-zero-sum games.

5. NON-COOPERATIVE GAMES IN NORMAL FORM

So far we only have looked at zero-sum games. In this section we study perturbations of general sum two-person games in normal form. We emphasize that all results in this section can easily be extended to N-person games in normal form with $N \geq 2$; only for notational reasons we restrict our attention to two-person games.

DEFINITION. 5.1. A (*general sum*) *two-person game in normal form* is an ordered quadruplet $<\Pi_1,\Pi_2,p_1,p_2>$, in which Π_1 and Π_2 are non-empty sets and $p_1: \Pi_1 \times \Pi_2 \to \mathbb{R}$, $p_2: \Pi_1 \times \Pi_2 \to \mathbb{R}$ are real-valued functions on $\Pi_1 \times \Pi_2$. A point $(\pi_1^*,\pi_2^*) \in \Pi_1 \times \Pi_2$ is called an *equilibrium point* of the game $<\Pi_1,\Pi_2,p_1,p_2>$ if

$$p_1(\pi_1^*,\pi_2^*) = \max_{\pi_1 \in \Pi_1} p_1(\pi_1,\pi_2^*),\ p_2(\pi_1^*,\pi_2^*) = \max_{\pi_2 \in \Pi_2} p_2(\pi_1^*,\pi_2);$$

and is called an ε-*equilibrium point* ($\varepsilon > 0$) if

$$p_1(\pi_1^*,\pi_2^*) \geq \sup_{\pi_1 \in \Pi_1} p_1(\pi_1,\pi_2^*) - \varepsilon,\ p_2(\pi_1^*,\pi_2^*) \geq \sup_{\pi_2 \in \Pi_2} p_2(\pi_1^*,\pi_2) - \varepsilon.$$

The set of equilibrium points of $<\Pi_1,\Pi_2,p_1,p_2>$ is denoted by $E(p_1,p_2)$ and the set of ε-equilibrium points by $E^{\varepsilon}(p_1,p_2)$.

For fixed Π_1,Π_2, let $B(\Pi_1,\Pi_2)$ be the metric space of pairs (p_1,p_2) of bounded realvalued functions on $\Pi_1 \times \Pi_2$, provided with the metric d defined by:

$$d((p_1,p_2),(p_1',p_2')) := \max\{\|p_1-p_1'\|,\|p_2-p_2'\|\}$$

for all $(p_1,p_2) \in B(\Pi_1,\Pi_2)$ and $(p_1',p_2') \in B(\Pi_1,\Pi_2)$. Let $BE(\Pi_1,\Pi_2)$ be the subset of $B(\Pi_1,\Pi_2)$, consisting of those pairs (p_1,p_2) for which $E^{\varepsilon}(p_1,p_2) \neq \emptyset$ for all $\varepsilon > 0$.

The following two theorems are extensions of theorems 3.6 and 3.7 in TIJS [10], pp.99-100.

THEOREM 5.2. *Let* $\varepsilon \geq 0$, $\delta \geq 0$, $(p_1,p_2) \in B(\Pi_1,\Pi_2)$, $(p_1',p_2') \in B(\Pi_1,\Pi_2)$ *and* $d((p_1,p_2),(p_1',p_2')) \leq \delta$. *Then* $E^{\varepsilon}(p_1,p_2) \subset E^{\varepsilon+2\delta}(p_1',p_2')$.

PROOF. Let $(\pi_1^*,\pi_2^*) \in E^{\varepsilon}(p_1,p_2)$. Then for each $\pi_1 \in \Pi_1$

$$p_1'(\pi_1,\pi_2^*) \le p_1(\pi_1,\pi_2^*) + \delta \le p_1(\pi_1^*,\pi_2^*) + \varepsilon + \delta \le p_1'(\pi_1^*,\pi_2^*) + \varepsilon + 2\delta$$

and, analogously, for each $\pi_2 \in \Pi_2$:

$$p_2'(\pi_1^*,\pi_2) \le p_2'(\pi_1^*,\pi_2^*) + \varepsilon + 2\delta.$$

Hence $(\pi_1^*,\pi_2^*) \in E^{\varepsilon+2\delta}(p_1',p_2')$. □

THEOREM 5.3. $BE(\Pi_1,\Pi_2)$ *is a closed subset of* $B(\Pi_1,\Pi_2)$.

PROOF. Suppose that $(\bar{p}_1,\bar{p}_2)$ is an element of the closure of $BE(\Pi_1,\Pi_2)$, and let $\varepsilon > 0$. Then we can take $(p_1,p_2) \in BE(\Pi_1,\Pi_2)$ such that $d((p_1,p_2),(\bar{p}_1,\bar{p}_2)) < \frac{1}{4}\varepsilon$. Take $(\pi_1,\pi_2) \in E^{\frac{1}{2}\varepsilon}(p_1,p_2) \neq \emptyset$. Then, in view of theorem 5.2, $(\pi_1,\pi_2) \in E^{\varepsilon}(\bar{p}_1,\bar{p}_2) \neq \emptyset$. Hence $(\bar{p}_1,\bar{p}_2) \in BE(\Pi_1,\Pi_2)$ and we may conclude that $BE(\Pi_1,\Pi_2)$ is closed. □

Now let Π_1 and Π_2 be topological spaces. Put $CBE(\Pi_1,\Pi_2) := \{(p_1,p_2) \in \in BE(\Pi_1,\Pi_2) \mid p_1$ and p_2 are continuous functions$\}$.

THEOREM 5.4. *Let* Π_1 *and* Π_2 *be compact metric spaces. Then*

(1) $E(p_1,p_2) \neq \emptyset$ *for each* $(p_1,p_2) \in CBE(\Pi_1,\Pi_2)$.

(2) $(p_1,p_2) \mapsto E(p_1,p_2)$ *is an upper semicontinuous multifunction from* $CBE(\Pi_1,\Pi_2)$ *into* $\Pi_1 \times \Pi_2$.

(3) $(p_1,p_2) \mapsto E^{\varepsilon}(p_1,p_2)$ *is upper semicontinuous for each* $\varepsilon > 0$.

PROOF. (a) We note that (1) follows from

$$E(p_1,p_2) = \bigcap_{\varepsilon>0} E^{\varepsilon}(p_1,p_2)$$

where $E^{\varepsilon}(p_1,p_2)$ is, for each $\varepsilon > 0$, a non-empty closed subset of the compact set $\Pi_1 \times \Pi_2$.

(b) Let $\varepsilon \ge 0$. Let (p_1^1,p_2^1), (p_1^2,p_2^2), (p_1^3,p_2^3), ... be a sequence in $CBE(\Pi_1,\Pi_2)$ converging to (p_1,p_2). Let $(\pi_1^n,\pi_2^n) \in E^{\varepsilon}(p_1^n,p_2^n)$ for each $n \in \mathbb{N}$ and suppose that $\lim_{n\to\infty} \pi_1^n = \pi_1^*$, $\lim_{n\to\infty} \pi_2^n = \pi_2^*$. If we can show that $(\pi_1^*,\pi_2^*) \in E^{\varepsilon}(p_1,p_2)$, then we have proved (2) and (3). Put $\delta_n := = d((p_1^n,p_2^n), (p_1,p_2))$ for each $n \in \mathbb{N}$. Then, by theorem 5.2,

$(\pi_1^n,\pi_2^n) \in E^{\varepsilon+2\delta_n}(p_1,p_2)$ for each $n \in \mathbb{N}$. Because $\lim_{n\to\infty} \delta_n = 0$ (and $E^{\lambda}(p_1,p_2) \subset E^{\mu}(p_1,p_2)$ if $0 \le \lambda \le \mu$), we may conclude that for each fixed $k \in \mathbb{N}$ and n sufficiently large: $(\pi_1^n,\pi_2^n) \in E^{\varepsilon+k^{-1}}(p_1,p_2)$. But then $(\pi_1^*,\pi_2^*) \in$ $\in E^{\varepsilon+k^{-1}}(p_1,p_2)$, since $E^{\varepsilon+k^{-1}}(p_1,p_2)$ is closed and $(\pi_1^*,\pi_2^*) = \lim_{n\to\infty} (\pi_1^n,\pi_2^n)$. So

$$(\pi_1^*,\pi_2^*) \in E^{\varepsilon}(p_1,p_2) = \bigcap_{k\in\mathbb{N}} E^{\varepsilon+k^{-1}}(p_1,p_2). \quad \square$$

REMARKS

5.5.1. Theorem 5.4 can be extended to the class of games, which is studied in VRIEZE [12], namely the class of mixed extensions of the games $<\{A_i | i \in I\},\{p_i | i \in I\}>$, where I is a countable set of players, where the action space A_i of player i is a compact topological space, satisfying the first axiom of countability and where the payoff function p_i: $X_{i\in I} A_i \to \mathbb{R}$ for player i is a continuous function with respect to the product topology.

5.5.2. In a similar way to that in which we have extended results obtained for the zero-sum game in normal form to the zero-sum discounted stochastic game in section 3, we could extend some of the results of this section (e.g. theorem 5.4) to the general sum discounted stochastic game.

REFERENCES

[1] BERGE, C., *Espaces topologiques*, Dunod, Paris (1959).

[2] BLACKWELL, D., *Discounted dynamic programming*, Annals of Math. Stat. 36, 226-235 (1965).

[3] BOHNENBLUST, H.F., S. KARLIN & L.S. SHAPLEY, *Solutions of discrete two-person games*, Annals of Mathematics Studies 24, 51-72 (1950).

[4] KRABS, W., *Stetige Abänderung der Daten bei nichtlinearer Optimierung und ihrer Konsequenzen*, Operations Research Verfahren 25, 93-113 (1977).

[5] PARTHASARATHY, K.R., *Probability measures on metric spaces*, Acad. Press, New York-London (1967).

[6] PARTHASARATHY, T., *Discounted, positive and non-cooperative stochastic games*, Int. Journal of Game Theory 2, 25-37 (1973).

[7] PARTHASARATHY, T. & M. STERN, *Markov games- a survey*, University of Illinois, Chicago (1976).

[8] SCHWEITZER, P.J., *Perturbation theory and finite Markov chains*, J. Appl. Probability, 5, 401-413 (1968).

[9] SHAPLEY, L.S., *Stochastic games*, Proc. Nat. Acad. Sci. 39, 1095-1100 (1953).

[10] TIJS, S.H., *Semi-infinite and infinite matrix games and bimatrix games*, Ph.D. dissertation, Department of Mathematics, Catholic University, Nijmegen, The Netherlands (1975).

[11] TIJS, S.H., *Semi-infinite programs and semi-infinite matrix games*, report 7630, Catholic University, Nijmegen, The Netherlands (1976).

[12] VRIEZE, O.J., *Non-cooperative countable-person games with compact action spaces*, report BW 65/76, Mathematical Centre, Amsterdam (1976).

[13] VRIEZE, O.J., *The stochastic non-cooperative countable-person game with countable state space and compact action spaces under the discounted payoff criterion,* report BW 66/76, Mathematical Centre, Amsterdam (1976).

[14] WHITT, W., *Continuity of Markov processes and dynamic programs*, Yale University, New Haven (1975).