# On the existence of easy initial states for undiscounted stochastic games

Tijs, S.H.; Vrieze, O.J.

*Published in:*
Mathematics of Operations Research

*Publication date:*
1986

*Document Version*
Peer reviewed version

# ON THE EXISTENCE OF EASY INITIAL STATES FOR UNDISCOUNTED STOCHASTIC GAMES

S.H. TIJS and O.J. VRIEZE
Catholic University of Nijmegen
The Netherlands

This paper deals with undiscounted infinite stage two-person zero-sum stochastic games with finite state and action spaces. It was recently shown that such games possess a value. But in general there are no optimal strategies. We prove that for each player there exists a non-empty set of easy initial states, i.e. starting states for which the player possesses an optimal stationary strategy. This result is proved with the aid of facts derived by Bewley and Kohlberg for the limit discount equation for stochastic games.

AMS 1980 subject classification. Primary 90D15, Secondary 93C30.

Key words. Stochastic game, Undiscounted payoff criterion, limit discount equation, optimal stationary strategy, easy initial state.

## 1. INTRODUCTION

In this paper we restrict our attention to stochastic games with finite state and action spaces. The theory of such stochastic games started in 1953 with the fundamental paper of Shapley [6]. He proved, that such a game possesses a value and that there exists for each player a stationary strategy, which is optimal for each initial state, under the condition that the payoffs are discounted. The theory of undiscounted stochastic games started with a paper of Gillette [3] in 1957. For a long time, it was an open question, whether such undiscounted games possess a value. This question was answered in the affirmative, recently by Monash [5] and Mertens and Neyman [4].

In a nice paper [2] in 1968, Blackwell and Ferguson studied an example of an undiscounted stochastic game (the big match), and showed that no optimal strategy exists for one of the players. Even if one wants to play $\varepsilon$-optimal ($\varepsilon > 0$), in general one has to use complicated history dependent strategies. So the question arose, whether there are perhaps certain initial states for which playing optimal or $\varepsilon$-optimal can be accomplished with a simple stationary strategy. With this question we deal in this paper. Let us call a state an *easy initial state for player* $i \in \{1,2\}$, if there exists a stationary strategy, such that this strategy is optimal for the player, if a play of the game starts in that state. Then Shapley's result can be reformulated as follows: for a discounted stochastic game each initial state is easy for each player and there is a stationary strategy for each player, which is optimal for each initial state. The example of Blackwell and Ferguson learns us that for undiscounted games the situation is not so simple.

The main result of this paper is that we show that for each stochastic game, for each player there exists a non-empty subset of easy initial states and a stationary strategy for the player, which is optimal for each play starting in that subset of states. These subsets are related to the main part of the unique solution of the limit discount equation of the stochastic game.

The organization of this paper is as follows. In section 2 we give the necessary facts about stochastic games, about the limit discount equation for such games and we formulate our main result in theorem 2.1. In section 3 this theorem is proved with the aid of some lemmas. We conclude that section with some examples.

## 2. THE MAIN THEOREM

In the following, $\Gamma = \langle S, \{A_{is} : i \in \{1,2\}, s \in S\}, r, p\rangle$ is a fixed zero-sum two-person stochastic game with finite state space $S = \{1,2,\ldots,z\}$, with for each player $i \in \{1,2\}$ in each state $s \in S$ available a finite action space $A_{is}$, and with reward function r and transition probability map p. Hence, $r(s,a_1,a_2)$ is the direct reward and $p(t|s,a_1,a_2)$ is the probability that the system is in the next stage in state t, given that the system is now in state s and given that player i, $i = 1,2$, chooses action $a_i \in A_{is}$.

We are interested in the undiscounted infinite stage case i.e. we assume that there are an infinite number of stages $n = 0,1,2,\ldots$ where the players choose actions and we assume that a payoff stream $(\pi_0,\pi_1,\pi_2,\ldots)$ for player 1, is evaluated by the expression

$$\liminf_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} \pi_n \tag{2.1}$$

where $\pi_n$ is the (expected) payoff to player 1, made by player 2 at stage n.

In the following we denote by F the ordered field of real Puiseux series, consisting of power series of the form

$$\sum_{k=-\infty}^{N} a_k \theta^{k/L} \tag{2.2}$$

where N is an arbitrary integer, L is an arbitrary positive integer, $a_k$ is a real number for each k, and where the series $\sum_{k=-\infty}^{N} a_k t^{k/L}$ converges for all sufficiently large real numbers t. Addition and multiplication in F are defined in an obvious way and the order is defined as follows.

$$\sum_{k=-\infty}^{N} a_k \theta^{k/L} \geq 0 \text{ iff } a_N \geq 0 \text{ and for all } k < N: \quad a_k \geq 0 \text{ if } a_{k+1} = a_{k+2} = \ldots = a_N = 0 \tag{2.3}$$

The z-fold Cartesian product of F with itself is denoted by $F^z$. The ordered subfield of F of expressions of the form (2.2), where in each expression L is equal to a fixed positive integer M, is denoted by $F_M$.

For our study the *limit discount equation* for the stochastic game Γ, introduced by Bewley and Kohlberg [1], plays an important role. It is the equation in $F^z$, given by

$$x = \text{val} \langle r + (1+\theta^{-1})^{-1} Px \rangle \tag{2.4}$$

which is an abbreviation of a system of z equations in F, for each state one, namely

$$x_s = \text{val} \langle r(s,.,.) + (1+\theta^{-1})^{-1} \sum_{t \in S} p(t|s,.,.)x_t \rangle \text{ for each } s \in S \tag{2.5}$$

where the right hand term of (2.5) is the value of the (mixed extension of the) matrix game $M_s(x)$, with pure action spaces $A_{1s}$ and $A_{2s}$ for the row and column player, respectively, and with

$$r(s,a_1,a_2) + (1+\theta^{-1})^{-1} \sum_{t \in S} p(t|s,a_1,a_2)x_t \in F \tag{2.5}$$

the payoff made to player 1 by player 2, if player 1 chooses pure action $a_1$ and player 2 action $a_2$.

Bewley and Kohlberg [1] proved that the system of equations (2.5) possesses a unique solution $V = (V(1),V(2),\dots,V(z)) \in F^z$. Furthermore, they proved that there exists a positive integer M (dependent on $\Gamma$) such that $V(s) \in F_M$ and for each $s \in S$, V(s) is of the form

$$V(s) = \sum_{k=-\infty}^{M} v_k(s)\theta^{k/M}. \tag{2.7}$$

Since the entries in the matrix $M_S(V)$ are then also elements of the ordered field $F_M$, it follows from a well-known result of H. Weyl [7], that player 1 possesses an optimal mixed action $\tilde{\rho}(s)$, which chooses for each $a \in A_{1s}$ with probability $\tilde{\rho}(s,a) \in F_M$ this action. Hence, $\tilde{\rho}(s,a)$ is of the form

$$\tilde{\rho}(s,a) = \sum_{u=0}^{\infty} \tilde{\rho}_{-u}(s,a)\theta^{-u/M} \tag{2.8}$$

where for each $s \in S$ and each $u \in \mathbb{N}$

$$\sum_{a \in A_{1s}} \tilde{\rho}_0(s,a) = 1, \quad \sum_{a \in A_{1s}} \tilde{\rho}_{-u}(s,a) = 0 \tag{2.9}$$

and where for each $s \in S$, $a \in A_{1s}$ and $u \in \mathbb{N}$

$$\tilde{\rho}_0(s,a) \geq 0 \tag{2.10}$$

$$\tilde{\rho}_{-u}(s,a) \geq 0 \text{ if } \tilde{\rho}_{-u+1}(s,a) = \tilde{\rho}_{-u+2}(s,a) = \dots = \tilde{\rho}_0(s,a) = 0 \tag{2.11}$$

Let $\tilde{\rho}_0 = (\tilde{\rho}_0(1),\dots,\tilde{\rho}_0(s),\dots,\tilde{\rho}_0(z))$ be the stationary strategy for player 1 in the stochastic game $\Gamma$, which takes action $a \in A_{1s}$ with probability $\tilde{\rho}_0(s,a) \in \mathbb{R}$, if the game is in state s.

Note that $\tilde{\rho}_0(s,a) = \lim_{t\to\infty} \tilde{\rho}_t(s,a)$ for all $s \in S$ and $a \in A_{1s}$, where $\tilde{\rho}_t(a,s)$ is the expression which we obtain by replacing $\theta$ by the real number t in (2.8). In [1] it was proved that for large real members t, the stationary strategy $\tilde{\rho}_t$ is optimal for player 1 in the discounted stochastic game with discount factor $(1+t^{-1})^{-1}$. So $\tilde{\rho}_0$ can be seen as the limit of optimal stationary strategies for player 1 in $\beta$-discounted games for $\beta \uparrow 1$.

We will show in section 3 that this stationary strategy $\tilde{\rho}_0$ is an optimal stationary strategy for all plays of $\Gamma$, starting in a non-empty subset $S^*$. Here $S^*$ is defined as follows. Let

$$\sum_{k=1}^{M} b_k \theta^{k/M} := \max_{s \in S} \left( \sum_{k=1}^{M} v_k(s)\theta^{k/M} \right) \tag{2.12}$$

(the maximum w.r.t. the order in F). Then

$$S^* := \{s \in S: \sum_{k=1}^{M} b_k \theta^{k/M} = \sum_{k=1}^{M} v_k(s)\theta^{k/M}\} \tag{2.13}$$

If we call in the expressions (2.7), the part $\sum_{k=1}^{M} v_k(s)\theta^{k/M}$, corresponding to the positive powers of $\theta$, the *main part* of V(s), then we can say that $S^*$ consists precisely of those states in S with maximal main part. Theorem 2.1 below says that $S^*$ consists of easy initial states for player 1.

An important fact, recently proved by Mertens and Neyman [4], will also be used in the following, namely that the undiscounted stochastic game $\Gamma$ with evaluation rule as in (2.1) possesses a value and that this value equals the leading coefficient $v_M(s)$ in (2.7), if the initial state is s.

We formulate now our main result. The proof will be given in section 3.

THEOREM 2.1. Notations as above.

The states in $S^*$ are easy initial states for player 1 in the stochastic game $\Gamma$. $\tilde{\rho}_0$ is an optimal stationary strategy for player 1 in $\Gamma$, if the initial state is an element of $S^*$.

Of course, a similar theorem can be formulated for player 2. Easy initial states for player 2 are elements of $S^{**}$, where $S^{**}$ consists of those states s, for which the main part of V(s) is minimal.

## 3. SOME LEMMAS AND THE PROOF OF THEOREM 2.1.

Let us take an arbitrary, but from now on, fixed stationary strategy $\sigma = (\sigma(1),\sigma(2),\ldots,\sigma(z))$ for player 2 in the stochastic game $\Gamma$. Then for all $s \in S$, $a \in A_{2s}$

$$\sigma(s,a) \in \mathbb{R}_+ , \quad \sum_{a \in A_{2s}} \sigma(s,a) = 1 \tag{3.1}$$

Since, for each $s \in S$, $\sigma(s)$ is a mixed action for player 2 in the matrix game $M_s(V)$, and $\tilde{\rho}(s)$ is an optimal mixed action for player 1 in that game with value V(s), we have for each $s \in S$:

$$-V(s) + R(s) + (1+\theta^{-1})^{-1} \sum_{t \in S} P(s,t)\, V(t) \geq 0 \tag{3.2}$$

in which

$$R(s) = \sum_{u=0}^{\infty} R_{-u}(s)\theta^{-u/M} \tag{3.3}$$

is the expected direct reward in state s, and

$$P(s,t) = \sum_{u=0}^{\infty} P_{-u}(s,t)\theta^{-u/M} \tag{3.4}$$

is the expected probability that the system jumps to $t \in S$, if in state s, player 1 uses the mixed action $\tilde{\rho}(s)$ and player 2 uses $\sigma(s)$. For the coefficients in (3.3) and (3.4) we have for all $u \in \{0,1,2,\ldots\}$

$$R_{-u}(s) = \sum_{a_1 \in A_{1s}} \sum_{a_2 \in A_{2s}} r(s,a_1,a_2)\, \tilde{\rho}_{-u}(s,a_1)\sigma(s,a_2) \tag{3.5}$$

$$P_{-u}(s,t) = \sum_{a_1 \in A_{1s}} \sum_{a_2 \in A_{2s}} p(t|s,a_1,a_2)\tilde{\rho}_{-u}(s,a_1)\sigma(s,a_2) \tag{3.6}$$

For further use, we note that, in view of (2.9), (2.10), (2.11) and (3.1), for all $u \in \mathbb{N}$:

$$\sum_{t\in S} P_0(s,t) = 1, \quad \sum_{t\in S} P_{-u}(s,t) = 0 \tag{3.7}$$

$$P_0(s,t) \geq 0 \tag{3.8}$$

$$P_{-u}(s,t) \geq 0 \text{ if } P_{-u+1}(s,t) = P_{-u+2}(s,t) = \ldots = P_0(s,t) = 0 \tag{3.9}$$

The left hand side of inequality (3.2) is a Puiseux series of the form

$$\sum_{k=-\infty}^{M} c_k(s)\theta^{k/M} \tag{3.10}$$

for which, for each $k < M$ we have:

$$c_M(s) \geq 0 \text{ and } c_k(s) \geq 0 \text{ if } c_M(s) = c_{M-1}(s) = \ldots = c_{k+1}(s) = 0 \tag{3.11}$$

In the following we are especially interested in the expressions for the coefficients corresponding to non-negative powers. For them we obtain, in view of (3.2), (3.3), (3.4) and (2.7), for $k \in \{M,M-1,\ldots,2,1\}$:

$$c_k(s) = -v_k(s) + \sum_{t\in S}\sum_{u=0}^{M-k} P_{-u}(s,t)v_{k+u}(t) \tag{3.12}$$

$$c_0(s) = -v_0(s)+R_0(s)+\sum_{t\in S}\sum_{u=0}^{M} P_{-u}(s,t)v_u(t)-\sum_{t\in S} P_0(s,t)v_M(t) \tag{3.13}$$

The following subsets of S will play a role. For $k \in \{1,2,\ldots,M\}$ let

$$S_k := \{s \in S: c_k(s) = c_{k+1}(s) = \ldots = c_M(s) = 0\} \tag{3.14}$$

$$T_k := \{s \in S: v_k(s) = b_k(s), v_{k+1}(s) = b_{k+1}(s),\ldots,v_M(s) = b_M(s)\} \tag{3.15}$$

(Here $c_k(s)$, $v_k(s)$ and $b_k(s)$ are the coefficients, occurring in (3.10), (2.7) and (2.12)).

Define also $T_{M+1} := S$, $S_{M+1} := S$. We note that

$$S^* = T_1 \tag{3.16}$$

$$S_{k-1} \subset S_k,\ T_{k-1} \subset T_k \text{ for } k \in \{2,\ldots,M,M+1\} \tag{3.17}$$

From (3.11) we obtain

$$c_{k-1}(s) \geq 0 \text{ for } k \geq 1 \text{ and } s \in S_k, \; c_{k-1}(s) > 0 \text{ for } k \geq 2 \text{ and } s \in S_k \backslash S_{k-1} \quad (3.18)$$

and from (2.12), for $k \in \{M+1, M, \ldots, 2\}$:

$$v_{k-1}(s) \leq b_{k-1} \text{ for } s \in T_k, \; v_{k-1}(s) < b_{k-1} \text{ for } s \in T_k \backslash T_{k-1} \quad (3.19)$$

For $s \in T_M$, we have, by (3,11), (3.12), (3.15), (3.19) and (3.7):

$$\begin{aligned} 0 \leq c_M(s) &= -v_M(s) + \sum_{t \in S} P_0(s,t)\, v_M(t) \\ &\leq b_M(-1 + \sum_{t \in S} P_0(s,t)) = 0 \end{aligned} \quad (3.20)$$

This implies, using (3.19) and (3.7):

$$c_M(s) = 0 \qquad \text{for all } s \in T_M \quad (3.21)$$

$$P_0(s,t) = 0 \qquad \text{for all } s \in T_M, \; t \in S \backslash T_M \quad (3.22)$$

LEMMA 3.1. For each $k \in \{M, M-1, \ldots, 2, 1\}$, $(Z_k)$ and $(Y_k)$ hold, where $(Z_k)$ is the property

$$T_k \subset S_k \quad (3.23)$$

and $(Y_k)$ the property:

$$P_{-u}(s,t) = 0 \text{ for all } s \in T_k, \; t \in S \backslash T_{k+u}, \; u \in \{0, 1, \ldots, M-k\} \quad (3.24)$$

PROOF. In view of (3.21) and (3.22), $(Z_M)$ and $(Y_M)$ hold. So the proof of the lemma will be finished if, for each $k \in \{M, M-1, \ldots, 2\}$, we can show that $(Z_k)$ and $(Y_k)$ imply $(Z_{k-1})$ and $(Y_{k-1})$. Hence, suppose that $(Z_k)$ and $(Y_k)$ hold for a $k \in \{M, M-1, \ldots, 2\}$. If we prove that for each $s \in T_{k-1}$:

$$c_{k-1}(s) = 0 \quad (3.25)$$

$$P_0(s,t) = 0 \text{ if } t \in T_k \backslash T_{k-1} \quad (3.26)$$

$$P_{-u}(s,t) = 0 \text{ if } t \in T_{k+u} \backslash T_{k+u-1} \text{ and } u \in \{1, 2, \ldots, M-k+1\} \quad (3.27)$$

then we can combine $(Z_k)$ and (3.25), using (3.17) and conclude that $(Z_{k-1})$ holds; and we can combine $(Y_k)$, (3.26) and (3.27) to conclude that $(Y_{k-1})$ holds. So the only things to prove are (3.25)-(3.27). Take $s \in T_{k-1}$. Observe, that $s \in T_{k-1} \subset T_k \subset S_k$ implies with (3.18), that

$$c_{k-1}(s) \geq 0 \tag{3.28}$$

By $(Y_k)$ and (3.17) we have for $u \in \{1,2.,,,.M-k+1\}$, $v < u$, $t \in T_{k+u} \setminus T_{k-1+u}$:

$$P_{-v}(s,t) = 0 \tag{3.29}$$

This implies, in view of (3.9)

$$P_{-u}(s,t) \geq 0 \text{ for all } t \in T_{k+u} \setminus T_{k-1+u} \tag{3.30}$$

Furthermore, by $(Y_k)$ and (3.17)

$$P_{-u}(s,t) = 0 \text{ if } t \notin T_{k+u} \tag{3.31}$$

So (3.7) and (3,31) imply for $u \in \{1,2,\ldots,M-k+1\}$

$$\sum_{t \in T_{k+u} \setminus T_{k-1+u}} P_{-u}(s,t) = -\sum_{t \in T_{k-1+u}} P_{-u}(s,t) \tag{3.32}$$

By (3.12)

$$c_{k-1}(s) = -v_{k-1}(s) + \sum_{t \in S} \sum_{u=0}^{M-k+1} P_{-u}(s,t) v_{k-1+u}(t) \tag{3.33}$$

If we put

$$d_0(s) := -v_{k-1}(s) + \sum_{t \in S} P_0(s,t) v_{k-1}(t)$$
$$d_u(s) := \sum_{t \in S} P_{-u}(s,t) v_{k-1+u}(t) \text{ for } u \in \{1,2,\ldots,M-k+1\} \tag{3.34}$$

Then, by (3.28) and (3.33):

$$0 \leq c_{k-1}(s) = \sum_{u=0}^{M-k+1} d_u(s) \tag{3.35}$$

In view of $s \in T_{k-1} \subset T_k$, (3.15), $(Y_k)$, (3.19) and (3.7)

$$d_0(s) = -b_{k-1} + \sum_{t \in T_k} P_0(s,t)v_{k-1}(t)$$
$$\leq b_{k-1}(-1 + \sum_{t \in T_k} P_0(s,t)) = 0 \tag{3.36}$$

and in view of (3.19)

$$d_0(s) = 0 \text{ iff (3.26) holds} \tag{3.37}$$

For $u \in \{1,2,\ldots,M-k+1\}$ we have by $(Y_k)$, (3.15), (3.32), (3.19) and (3.30)

$$d_u(s) = \sum_{t \in T_{k+u}} P_{-u}(s,t)\ v_{k+u-1}(t) =$$
$$= \sum_{t \in T_{k+u} \setminus T_{k+u-1}} P_{-u}(s,t)v_{k+u-1}(t) + \sum_{t \in T_{k+u-1}} P_{-u}(s,t)b_{k+u-1}$$
$$= \sum_{t \in T_{k+u} \setminus T_{k+u-1}} P_{-u}(s,t)(v_{k+u-1}(t)-b_{k+u-1}) \leq 0 \tag{3.38}$$

and

$$d_u(s) = 0 \text{ iff (3.27) holds.} \tag{3.39}$$

Now (3.35), (3.36) and (3.38) imply (3.25) and

$$d_u(s) = 0 \text{ for each } u \in \{0,1,\ldots,M-k+1\} \tag{3.40}$$

It follows from (3.40), (3.37) and (3.39) that (3.26) and (3.27) hold. So we have proved (3.25) - (3.27) which finishes this proof. □

The foregoing lemma will be repeatedly used in the proof of

LEMMA 3.2. For each $s \in S^*$, we have

$$0 \leq -v_0(s) + \sum_{t \in S^*} P_0(s,t)v_0(t) + R_0(s)-b_M \tag{3.41}$$

For each $s \in S^*$ and $t \in S \setminus S^*$: $P_0(s,t) = 0$. (3.42)

PROOF. (3.42) is in view of (3.16) already proved in lemma 2.1. (Cf. (3.24) with $k=1$ and $u=0$.) Take $s \in S^*$. By $(Y_1)$ of lemma 2.1: $s \in S^* = T_1 \subset S_1$. So by (3.18) and (3.13):

$$0 \leq c_0(s) = -v_0(s)+R_0(s) + \sum_{t \in S} \sum_{u=0}^{M} P_{-u}(s,t)v_u(t) - \sum_{t \in S} P_0(s,t)v_M(t) \tag{3.43}$$

This implies that we have shown that (3.41) holds, if we prove

$$\sum_{t\in S}\sum_{u=1}^{M} P_{-u}(s,t)v_u(t) \le 0 \tag{3.44}$$

and

$$\sum_{t\in S} P_0(s,t)v_M(t) = b_M \tag{3.45}$$

By lemma 2.1 and (3.42) $\sum_{t\in S} P_0(s,t)v_M(t) = \sum_{t\in T_1} P_0(s,t)v_M(t) = b_M$, so (3.45) holds. Take $u \in \{1,2,\ldots,M\}$. For each $v \in \{0,1,\ldots,u-1\}$, we have in view of lemma 2.1:

$$P_{-u}(s,t) = 0 \quad \text{if } t \notin T_{1+u} \tag{3.46}$$

$$P_{-v}(s,t) = 0 \quad \text{if } t \in T_{1+u}\backslash T_u \tag{3.47}$$

So by (3.9) and (3.47):

$$P_{-u}(s,t) \ge 0 \quad \text{if } t \in T_{1+u}\backslash T_u \tag{3.48}$$

By (3.7) and (3.46):

$$\sum_{t\in T_{1+u}\backslash T_u} P_{-u}(s,t) = -\sum_{t\in T_u} P_{-u}(s,t) \tag{3.49}$$

But then, by (3.46), (3.49), (3.48) and (2.12), for each $u \in \{1,\ldots,M\}$:

$$\sum_{t\in S} P_{-u}(s,t)v_u(t) = \sum_{t\in T_{1+u}\backslash T_u} P_{-u}(s,t)(v_u(t)-b_u) \le 0 \tag{3.50}$$

Now (3.50) implies (3.44) and this finishes the proof. □

For notational conveniences, we now suppose w.l.o.g. that $S^* = \{1,2,\ldots,z^*\}$ with $z^* \le z$. Let $P^*$ be the $z^* \times z^*$-submatrix $[P_0(s,t)]_{s=1,t=1}^{z^*\ \ z^*}$ in the left upper corner of the stochastic $z \times z$-matrix $P_0 = [P_0(s,t)]_{s=1,t=1}^{z\ \ z}$. Then, by (3.42) of lemma 3.2 and (3.7), this submatrix $P^*$ is also a probability matrix. This implies that the Cesaro-limit

$$Q^* := \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} (P^*)^n$$

of $P^*$, is the $z^* \times z^*$-submatrix in the left upper corner of the Cesaro-limit $Q_0$ of $P_0$.

If we denote for an element $x = (x_1, x_2, \ldots, x_z) \in \mathbb{R}^z$, the element $(x_1, x_2, \ldots, x_{z^*}) \in \mathbb{R}^{z^*}$, by $\hat{x}$, then, obviously we have

$$(Q_0 x)_s = (Q^* \hat{x})_s \quad \text{for all } s \in \{1,2,\ldots,z^*\} \tag{3.51}$$

Now we give the

PROOF OF THEOREM 2.1. Suppose that player 1 uses stationary strategy $\tilde{\rho}_0$ in the undiscounted game $\Gamma$ and player 2 the stationary strategy $\sigma$. Then the limit average expected payoff for player 1 is given by the formula

$$(Q_0 R_0)_s = \lim_{N \to \infty} \frac{1}{N+1} \sum_{n=0}^{\infty} P_0^n(s,t) R_0(t)$$

if s is the initial state. So, by (3.51), for $s \in S^*$ this payoff equals $(Q^* \hat{R}_0)_s$. From (3.41) in lemma 3.2 we derive

$$0 \leq -\hat{v}_0 + P^* \hat{v}_0 + \hat{R}_0 - b_M 1_{z^*} \tag{3.52}$$

where $1_{z^*} \in \mathbb{R}^{z_*}$ is the vector, for which all coordinates are equal to 1. Since $Q^*(0) = 0$, $Q^* 1_{z^*} = 1_{z^*}$, $Q^* P^* = Q^*$ and $Q^*$ is monotone and linear, we may conclude from (3.52):

$$0 \leq -Q^* \hat{v}_0 + Q^* P^* \hat{v}_0 + Q^* \hat{R}_0 - b_M 1_{z^*} = Q^* \hat{R}_0 - b_M 1_{z^*} \tag{3.53}$$

So for $s \in S^*$:

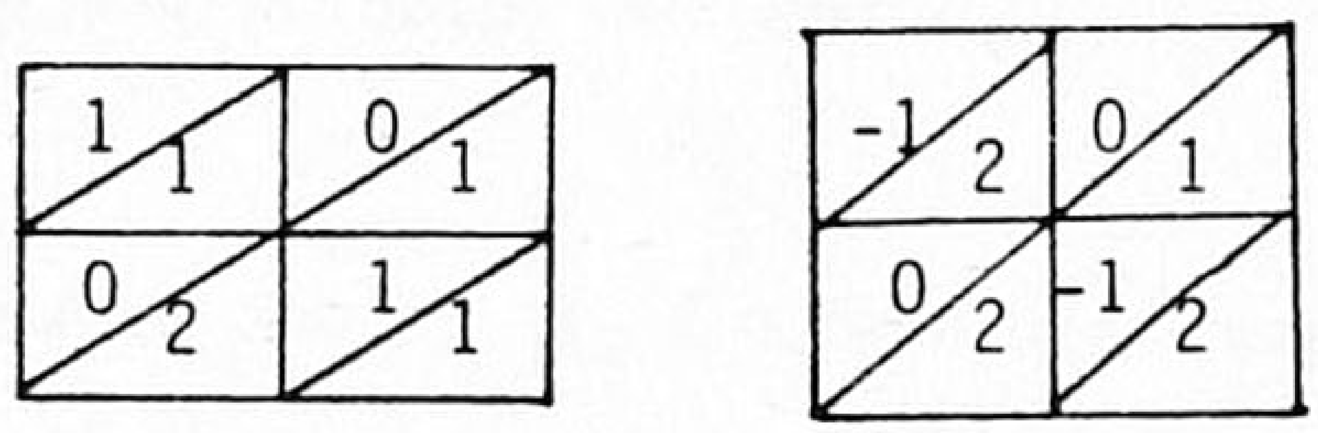
$$(Q_0 R_0)_s = (Q^* \hat{R}_0)_s \geq b_M(s) \tag{3.54}$$

and by the mentioned result of Mertens and Neyman [4], the right member of (3.54) is the value of the undiscounted stochastic game $\Gamma$, if the initial state is s. So we have proved that player 1, using $\tilde{\rho}_0$, can obtain at least the value, if the initial state is in $S^*$, against each stationary strategy of player 2. Since it is well-known that player 2 cannot do better with arbitrary (history dependent) strategies, if player 1 uses $\tilde{\rho}_0$, we may conclude that $\tilde{\rho}_0$ is optimal for player 1 for each initial state $s \in S^*$. This finishes the proof of theorem 2.1.

□

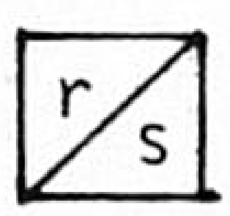
We conclude with two examples. In the first example the set of easy initial states of player 1 (2) coincides with $S^*$ ($S^{**}$) and $S = S^* \cup S^{**}$, $S^* \cap S^{**} = \emptyset$. In the second example $S^*$ ($S^{**}$) is a proper subset of the set of easy initial states of player 1 (2) and $S^* \cup S^{**} \neq S$.

EXAMPLE 3.3. Let $\Gamma$ be the game described by




State 1 State 2

where the notation



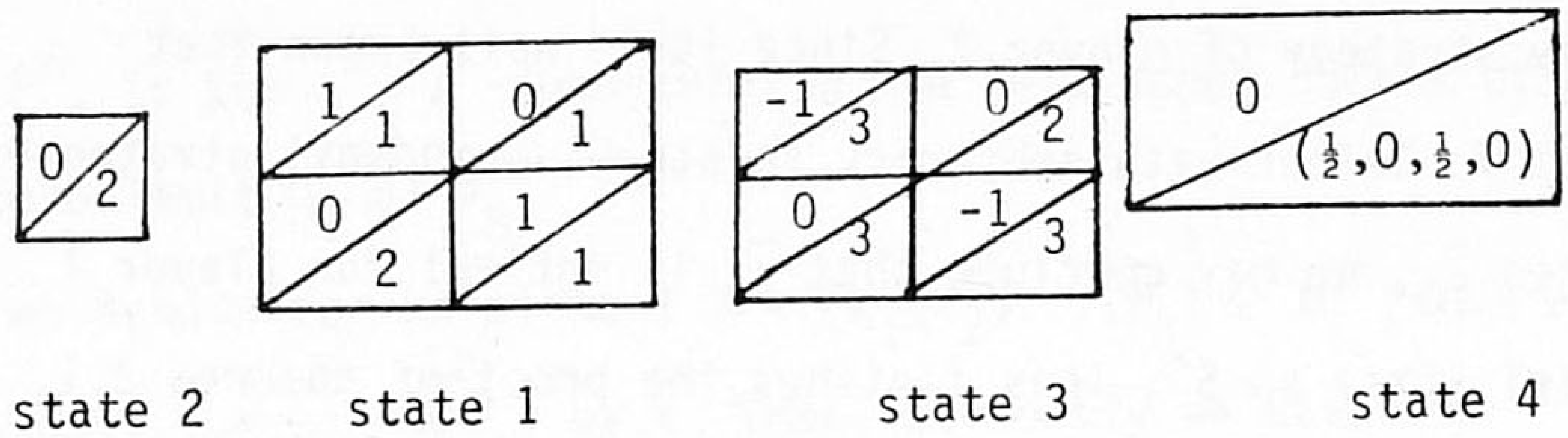

means that if the players choose the row and column corresponding to this box, then player 2 pays the amount r to player 1 and the state in the next stage is s. Then $V = (V(1),V(2)) = (V(1),-V(1))$ with

$$V(1) = \tfrac{1}{2}\sqrt{2\theta+1} - \tfrac{1}{2} = 0\theta + \tfrac{1}{2}\sqrt{2}\,\theta^{\frac{1}{2}} - \tfrac{1}{2} + \tfrac{1}{8}\sqrt{2}\,\theta^{-\frac{1}{2}} + \ldots$$

Hence, $S^* = \{1\}$, $S^{**} = \{2\}$. The value of the undiscounted game is (0,0) and it is clear that state 1 is not an easy initial state for player 2 and that state 2 is not easy for player 1.

For the strategy $\tilde{\rho}_0$, we have that $\tilde{\rho}_0(1,1) = 1$. This strategy is optimal for player 1 if the play starts in state 1.

EXAMPLE 3.4. Let $\Gamma$ be the game described by




In this game, if the system is in state 4, then it jumps to states 1 and 3 with probability $^1/_2$.

Now

$$V = (V(1),0,-V(1),0) \quad \text{with}$$

$$V(1) = (1+\theta^{-1})(\sqrt{\theta+1}-1) = 0\theta + \theta^{\frac{1}{2}} - 1 + \ldots$$

This implies that $S^* = \{1\}$, $S^{**} = \{3\}$ and that the value of the undiscounted game is (0,0,0,0).

State 4 is not easy for both players and state 1 (3) is not an easy initial state for player 2 (1), while state 2 is easy for both players.

## References

[1] Bewley, T. and Kohlberg, E. (1976). The Asymptotic Theory of Stochastic Games. *Math. Oper. Res.* 1 197-208.

[2] Blackwell, D. and Ferguson, T.S. (1968). The Big Match. *Ann. Math. Statist.* 39 159-168.

[3] Gillette, D. (1957). Stochastic Games with Zero Stop Probabilities. *Contributions to the Theory of Games*, Vol.III (*Annals of Mathematics Studies*, no.39). Princeton University Press, Princeton, N.J., 179-187.

[4] Mertens, J.-F. and Neyman, A. (1980). *Stochastic Games*. Report No. 22/80, Inst. for Advanced Studies. The Hebrew University of Jerusalem, Israel.

[5] Monash, C.A. (1980). *Stochastic Games. The minimax theorem.* Ph.D. dissertation, Department of Mathematics of the Harvard University, Cambridge, Mass.

[6] Shapley, L.S. (1953). Stochastic Games. *Proc. Nat. Acad. Sci. U.S.A.* 39 1095-1100.

[7] Weyl, H. (1950). Elementary Proof of the Minimax Theorem. *Contributions to the Theory of Games*, Vol. I (*Annals of Mathematics Studies*, No.24). Princeton University Press, Princeton, N.J., 19-25.