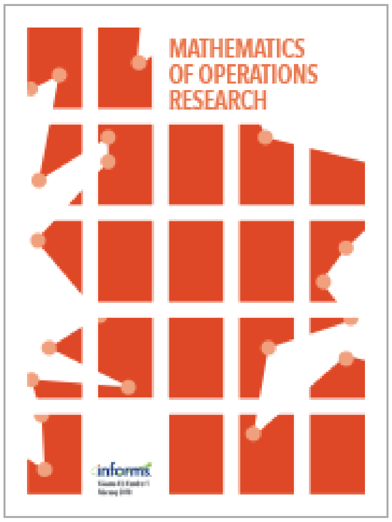

Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Stationary Strategies for Recursive Games

Piercesare Secchi,

MATHEMATICS OF OPERATIONS RESEARCH
Vol 22, No. 2, May 1997
Printed in U S A

STATIONARY STRATEGIES FOR RECURSIVE GAMES

PIERCESARE SECCHI

We study two-person, zero-sum recursive matrix games framing them in the more general context of nonleavable games. Our aim is to extend a result of Orkin (1972) by proving that a uniformly ϵ-optimal stationary strategy is available to player I (player II) in any recursive game such that the set where the value of the game is strictly greater (less) than its utility is finite. The proof exploits some new connections between nonleavable and leavable games. In particular we find sufficient conditions for the existence of ϵ-optimal stationary strategies for player I in a leavable game and we use these as a basis for constructing uniformly ϵ-optimal stationary strategies for the recursive games with which we are concerned.

1. Introduction. Consonant with Maitra and Sudderth (1996) a nonleavable stochastic game is a two-person, zero-sum competition played in stages and defined by five elements: a countable nonempty set of states S, two finite sets of actions A and B for player I and II respectively, a law of motion q which assigns to each triple $(x, a, b) \in S \times A \times B$ a probability distribution on the subsets of S and a bounded, real valued function u defined on S and called the utility of the game. The game is played as follows. Given an initial state $x \in S$, player I chooses, possibly at random, an action $a_1 \in A$, while simultaneously player II chooses, possibly at random, an action $b_1 \in B$. The game then moves to a new state $X_1 \in S$ according to the probability distribution $q(\cdot|x, a_1, b_1)$; the new state X_1 is announced to the players along with the chosen actions a_1 and b_1 and the procedure is then iterated. This produces a random sequence of states $X_0 = x, X_1, X_2, \dots$ and the payoff from player II to player I is fixed to be the expected value of

$$u^* = \limsup_{n \to \infty} u(X_n).$$

The principal aim of this paper is to investigate the problem concerning the existence of ϵ-optimal stationary strategies for a player engaged in a nonleavable game with a set W of absorbing states such that u is identically zero on W^c. The game begins in a state $x \in S$ and continues until the process $X_0 = x, X_1, X_2, \dots,$ reaches a $Y \in W$ where the game terminates and player II pays to I the quantity $u(Y)$. If the game never reaches W, then the payoff from II to I is 0. We call a game like this recursive because, as it will be shown later, it is equivalent to a recursive matrix game.

Maitra and Sudderth (1992) proved that any nonleavable game has a value by exploiting the connections between these games and leavable games defined as two-person, zero-sum competitions where, in addition to the rules defining a nonleavable game, player I, but not player II, is allowed to choose a stop rule t and the payoff from II to I is the expected value of the utility of the state X_t reached by the game when it is stopped. We will follow their example by first concerning ourselves

Received June 30, 1994; revised May 30, 1995 and May 15, 1996.
AMS 1991 subject classification. Primary: 90D05.
OR/MS Index 1978 subject classification. Primary: Games/Recursive.
Key words. Two-person, zero-sum recursive matrix games, nonleavable games, leavable games, stationary strategies.

0364-765X/97/2202/0494/$05.00

with the problem relative to the existence of ϵ-optimal stationary strategies in leavable games and then by showing a connection between leavable and nonleavable games which is different from those already known.

Investigations about recursive games began with the seminal papers of Everett (1957) and Milnor and Shapley (1957) who showed that in certain cases they have a value. Orkin (1972) made the notion of recursive game precise and proved that, when the state space of the game is finite, both players have a uniformly ϵ-optimal stationary strategy. We will extend Orkin's result by proving that a uniformly ϵ-optimal stationary strategy is available to player I (player II) in any recursive game with countable state space S, if the subset of S where the value of the game is strictly greater (less) than u is finite. Our proof has the additional advantage of showing explicitly what the ϵ-optimal stationary strategies are.

In the preliminaries section, after setting notation and terminology which will follow that of Maitra and Sudderth (1996), we will recall a few relevant results about leavable and nonleavable games. Section 3 examines the question regarding the existence of ϵ-optimal stationary strategies for leavable games. The following section is entirely dedicated to the study of an auxiliary leavable game which shows some new connections between leavable and nonleavable games. These provide the basis for our main result which is presented in the last section of the paper.

2. Preliminaries. Let $Z = S \times A \times B$ and define $H = Z \times Z \times \cdots$ to be the space of histories or sequences $h = (z_1, z_2, \dots)$ of elements of Z. For $n = 1, 2, \dots,$ a partial history of length n is a sequence $p = (z_1, \dots, z_n)$ of n elements of Z.

Let $\mathscr{P}(A)$ be the collection of probability distributions on the subsets of A and set $x \in S$ to be the initial state of the game. Following the tradition of Dubins and Savage (1976) as well as Maitra and Sudderth (1996), we define a strategy α for player I to be a sequence $\alpha_0, \alpha_1, \dots$ such that $\alpha_0 \in \mathscr{P}(A)$ and, for $n = 1, 2, \dots, \alpha_n$ is a mapping which assigns to each partial history p of length n an element of $\mathscr{P}(A)$. A strategy α is called stationary if there exists a function μ from S to $\mathscr{P}(A)$ such that

$$\alpha_n(z_1, \dots, z_n) = \mu(x_n)$$

for every integer $n \geq 1$, and $z_1 = (x_1, a_1, b_1), \dots, z_n = (x_n, a_n, b_n)$. Strategies β for player II are defined in the same way with B in place of A.

Given the initial state $x \in S$, the law of motion q along with a strategy α for player I and a strategy β for player II determines a probability distribution $P_{x,\alpha,\beta}$ on the sigma-field of subsets of H generated by the coordinate functions

$$Z_n(z_1, z_2, \dots, z_n, \dots) = z_n$$

for $n = 1, 2, \dots$. The expected value of a bounded, Borel measurable function g from H to the reals will be indicated with $\int g\, dP_{x,\alpha,\beta}$ or $E_{x,\alpha,\beta}\, g$.

A family of strategies $\overline{\alpha}$ for player I is a mapping from S to the collection of all available strategies for player I; that is, for every $x \in S$, $\overline{\alpha}(x)$ is a strategy for player I. The family $\overline{\alpha}$ is said to be stationary if there exists a function μ from S to $\mathscr{P}(A)$ such that

$$\overline{\alpha}(x)_0 = \mu(x),$$

$$\overline{\alpha}(x)_n(z_1, \dots, z_n) = \mu(x_n)$$

for every integer n, $x \in S$ and $z_1 = (x_1, a_1, b_1), \dots, z_n = (x_n, a_n, b_n)$. In this case we write $\overline{\alpha} = \mu^{\infty}$. Analogous definitions and notations hold for player II.

In the discrete time stochastic control literature our stationary strategies are often called semi-stationary since they depend on the initial state of the game, whereas the name of stationary strategy is used for what we called stationary family of strategies. We emphasize that, according to our definitions, if, for every $x \in S$, $\alpha(x)$ is defined to be a stationary strategy for I, the resulting collection of strategies $\{\alpha(x): x \in S\}$ does not need to be a stationary family.

If α is a strategy and $p = (z_1, \dots, z_n)$ is a partial history of length n, the conditional strategy $\alpha[p]$ is defined by

$$\alpha[p]_0 = \alpha_n(p),$$

$$\alpha[p]_m(z'_1, \dots, z'_m) = \alpha_{n+m}(z_1, \dots, z_n, z'_1, \dots, z'_m)$$

for all $m \geq 1$ and $(z'_1, \dots, z'_m)$.

A stopping time t is a mapping from H to $\{0, 1, \dots\} \cup \{\infty\}$ such that if h' agrees with h in the first $t(h)$ coordinates, then $t(h') = t(h)$. A stopping time which is everywhere finite is called a stop rule. Let t be a stop rule and $p = (z_1, \dots, z_n)$ a partial history of length n; if $t(z_1, \dots, z_n, z'_1, z'_2, \dots) \geq n$ for every history $h' = (z'_1, z'_2, \dots)$, we define the conditional stop rule $t[p]$ by setting

$$t[p](z'_1, z'_2, \dots) = t(z_1, \dots, z_n, z'_1, z'_2, \dots) - n.$$

When t is a nonzero stop rule, define on H the function p_t by setting, for every history $h = (z_1, z_2, \dots)$, $p_t(h) = (z_1, \dots, z_{t(h)})$. If g is a bounded, Borel measurable function from H to the reals, we will make frequent use of the following conditioning formula,

$$E_{x,\alpha,\beta}\, g = \int \left\{ E_{X_{t(h)},\, \alpha[p_t(h)],\, \beta[p_t(h)]}(gp_t(h)) \right\} dP_{x,\alpha,\beta}(h), \tag{2.1}$$

where, for every $h \in H$, the function $gp_t(h)$ is defined by

$$gp_t(h)(z'_1, z'_2, \dots) = g(z_1, \dots, z_{t(h)}, z'_1, z'_2, \dots)$$

for all $(z'_1, z'_2, \dots) \in H$. The same formula holds even when t is a nonzero stopping time such that $P_{x,\alpha,\beta}[t < \infty] = 1$. Note that (2.1) is a version of the more familiar relation

$$E_{x,\alpha,\beta}\, g = E_{x,\alpha,\beta}\left[E_{x,\alpha,\beta}[g \mid p_t] \right].$$

Finally we introduce the one-day operator G defined, for any bounded, real-valued function ϕ on S and for every $x \in S$, by

$$(G\phi)(x) = \inf_{\nu} \sup_{\mu} E_{x,\nu,\mu}\, \phi, \tag{2.2}$$

where μ and ν range over $\mathscr{P}(A)$ and $\mathscr{P}(B)$ respectively and

$$E_{x,\mu,\nu}\, \phi = \sum_{a \in A} \sum_{b \in B} \sum_{x_1 \in S} \phi(x_1) q(x_1 \mid x, a, b) \mu\{a\} \nu\{b\}.$$

By von Neumann's Theorem (von Neumann and Morgenstern 1947), $(G\phi)(x)$ is the value of the one-day game $\mathscr{A}(\phi)(x)$ where x is the initial state, players I and II choose, possibly at random, actions $a \in A$ and $b \in B$ respectively, and the game moves according to the law of motion q to the new state X_1; finally II pays I the expected value of $\phi(X_1)$. The same result by von Neumann proves also the existence of optimal randomized actions $\mu \in \mathscr{P}(A)$ and $\nu \in \mathscr{P}(B)$ for players I and II respectively.

For every $x \in S$, let $\mathscr{N}(u)(x)$ be the nonleavable game where player I chooses a strategy α, player II chooses a strategy β and II pays I the quantity $E_{x,\alpha,\beta}u^*$. The upper and lower values of $\mathscr{N}(u)(x)$ are, respectively,

$$\overline{V}(x) = \inf_\beta \sup_\alpha E_{x,\alpha,\beta} u^*, \qquad \underline{V}(x) = \sup_\alpha \inf_\beta E_{x,\alpha,\beta} u^*.$$

If $\overline{V}(x) = \underline{V}(x)$, then the game $\mathscr{N}(u)(x)$ is said to have value $V(x) = \overline{V}(x) = \underline{V}(x)$; in this case, a strategy α is optimal (ϵ-optimal) for player I if

$$E_{x,\alpha,\beta} u^* \geq V(x) \qquad \left(E_{x,\alpha,\beta} u^* \geq V(x) - \epsilon\right),$$

for any strategy β of player II. By reversing the inequalities we obtain analogous definitions for optimal (ϵ-optimal) strategies for player II.

Maitra and Sudderth (1992) proved that any nonleavable game has a value V. Moreover they proved (Maitra and Sudderth 1996) that V solves the optimality equation

$$V = GV. \tag{2.3}$$

Leavable games are two-person, zero-sum competitions introduced by Maitra and Sudderth (1996) for the main reason that they are profitable instruments for studying nonleavable games to which they are closely related.

Given S, A, B, u and a law of motion q as before, an initial state $x \in S$ determines a leavable game $\mathscr{L}(u)(x)$ in which player I chooses a strategy α and a stop rule t, player II chooses a strategy $\beta \in \mathscr{B}$, and II pays I the quantity $E_{x,\alpha,\beta}u(X_t)$. Let $\overline{U}(x)$ and $\underline{U}(x)$ be the upper and lower values of $\mathscr{L}(u)(x)$; that is

$$\overline{U}(x) = \inf_\beta \sup_{\alpha,t} E_{x,\alpha,\beta} u(X_t), \qquad \underline{U}(x) = \sup_{\alpha,t} \inf_\beta E_{x,\alpha,\beta} u(X_t).$$

Then Maitra and Sudderth (1992) showed that, for every $x \in S$, $\overline{U}(x) = \underline{U}(x) = U(x)$ or, equivalently, that $\mathscr{L}(u)(x)$ has value $U(x)$. They also proved the following lemma which characterizes the function U.

2.4 Lemma. *The value function U for the leavable game $\mathscr{L}(u)$ solves the optimality equation*

$$U = u \vee GU,$$

and is the least, bounded, real valued function ϕ defined on S such that

$$\text{(a) } \phi \geq u \quad \textit{and} \quad \text{(b) } G\phi \leq \phi.$$

In order to introduce a similar characterization for the value V of a nonleavable game $\mathscr{N}(u)$ we need to define an operator T which maps any bounded, real valued function u defined on S to the bounded, real valued function Tu defined, for every $x \in S$, by

$$Tu(x) = (GU)(x),$$

where U is the value of the leavable game $\mathscr{L}(u)$. Then Maitra and Sudderth (1992) proved the following lemma which characterizes the function V.

2.5 LEMMA. *The value function V for the nonleavable game $\mathscr{N}(u)$ is the largest, bounded, real valued function ϕ defined on S such that*

$$T(u \wedge \phi) = \phi.$$

We conclude these preliminaries by recalling a formula from Sudderth (1971). Let $x \in S$ and α and β be strategies for player I and II respectively. Then for any bounded, real valued function u defined on S,

$$E_{x,\alpha,\beta} u^* = \inf_s \sup_{t \geq s} E_{x,\alpha,\beta} u(X_t) \tag{2.6}$$

where s and t range over the set of stop rules.

3. Stationary strategies for leavable games. Let U be the value of the leavable game $\mathscr{L}(u)$ and assume that, in every state $x \in S$, II plays a randomized action $\nu(x) \in \mathscr{P}(B)$ which is optimal for him in the one-day game $\mathscr{A}(U)(x)$. Maitra and Sudderth (1992) proved that ν^∞ is an optimal stationary family for player II. In this section we prove a theorem which implies that, for any given initial state of $\mathscr{L}(u)$, there is an ϵ-optimal stationary strategy available to player I. When the state space of the game is finite our methods yield a stationary family which is uniformly ϵ-optimal for I.

On the state space S, define a sequence $\{U_n\}$ of real valued functions by setting $U_0 = u$ and, for $n = 1, 2, \dots$,

$$U_n = u \vee GU_{n-1}.$$

As Maitra and Sudderth (1992) have shown, U_n is the value of the n-day leavable game $\mathscr{L}_n(u)$ which has the same rules as $\mathscr{L}(u)$ except that player I must now choose a stop rule $t \leq n$. They also proved the following theorem which states that, for n large, U_n approximates the value U of $\mathscr{L}(u)$.

3.1 THEOREM. *For every $x \in S$,*

$$U(x) = \lim_{n \to \infty} U_n(x).$$

The next theorem, which is highly reminiscent of an analogous result proved by Sudderth (1969) for leavable gambling problems, states that, for every $\epsilon > 0$ and natural number m, player I has a stationary family uniformly ϵ-optimal for the m-day leavable game $\mathscr{L}_m(u)$.

Given an integer m and a quantity $\epsilon > 0$ set

$$0 = \epsilon_0 < \epsilon_1 < \cdots < \epsilon_m = \frac{\epsilon}{2}.$$

For every $x \in S$, define,

$$\phi(x) = \max_{0 \leq i \leq m} [U_i(x) - \epsilon_i].$$

Note that $\phi \geq u$. If $\phi(x) > u(x)$, let $k = k(x)$ be such that $\phi(x) = U_k(x) - \epsilon_k$ and set $\mu(x) \in \mathscr{P}(A)$ to be optimal for I in the one-day game $\mathscr{A}(U_{k-1})(x)$. If $\phi(x) = u(x)$

let $\mu(x) = \delta(a)$, the point mass at an action $a \in A$. Finally, as a function of m and ϵ, define the stationary family $\bar{\alpha} = \bar{\alpha}_{(m,\epsilon)} = \mu^{\infty}$.

3.2 THEOREM. *For every $x \in S$, there is a stop rule $t_\epsilon = t_\epsilon(x)$ such that*

$$E_{x,\bar{\alpha}(x),\beta} u(X_{t_\epsilon}) \geq U_m(x) - \epsilon \tag{3.3}$$

for any strategy β of player II.

PROOF. Let

$$\zeta = \frac{1}{2} \min_{1 \leq i \leq m} [\epsilon_i - \epsilon_{i-1}].$$

If $\phi(x) > u(x)$,

$$\begin{aligned} E_{x,\mu(x),\nu}\phi &\geq E_{x,\mu(x),\nu} U_{k-1} - \epsilon_{k-1} \\ &\geq GU_{k-1}(x) - \epsilon_{k-1} \\ &= U_k(x) - \epsilon_{k-1} \\ &\geq U_k(x) - \tfrac{1}{2}(\epsilon_k - \epsilon_{k-1}) - \epsilon_{k-1} \\ &= \phi(x) + \tfrac{1}{2}(\epsilon_k - \epsilon_{k-1}) \\ &\geq \phi(x) + \zeta, \end{aligned} \tag{3.4}$$

for any $\nu \in \mathscr{P}(B)$.

Fix now $x \in S$ and a strategy β for player II. Set

$$\eta_0 = E_{x,\mu(x),\beta_0}\phi - \phi(x),$$

and, for $n = 1, 2 \dots$, define,

$$\eta_n(z_1, \dots, z_n) = E_{x_n,\mu(x_n),\beta_n(z_1,\dots,z_n)}\phi - \phi(x_n)$$

for all $z_1 = (x_1, a_1, b_1), \dots, z_n = (x_n, a_n, b_n)$. Because of (3.4), $\eta_n(z_1, \dots, z_n) \geq \zeta > 0$, if $\phi(x_n) > u(x_n)$.

Now define a sequence of random variables $\{M_n\}$ by setting $M_0 = \phi(x)$ and, for $n = 1, 2, \dots$,

$$M_n = \phi(X_n) - \sum_{i=0}^{n-1} \eta_i(Z_1, \dots, Z_i).$$

It is easy to check that $\{M_n\}$ is a martingale with respect to $P_{x,\bar{\alpha}(x),\beta}$.

Set

$$\tau = \inf\{n \geq 0 \colon \phi(X_n) = u(X_n)\}.$$

For any given integer n, $\tau \wedge n$ is a bounded stop rule and the $P_{x,\bar{\alpha}(x),\beta}$-integral of $M_{\tau \wedge n}$ is well defined because ϕ is bounded. Therefore the Optional Sampling

Theorem of Doob (1953) guarantees that

$$E_{x,\bar{\alpha}(x),\beta} M_{\tau \wedge n} = \phi(x). \tag{3.5}$$

In particular, since $\eta_i(Z_1, \dots, Z_i) > 0$ if $i < \tau$, $\phi(X_{\tau \wedge n}) \geq M_{\tau \wedge n}$ and equation (3.5) implies that

$$E_{x,\bar{\alpha}(x),\beta} \phi(X_{\tau \wedge n}) \geq \phi(x). \tag{3.6}$$

Set $l = \sup_{x \in S} |\phi(x)|$ and $n \geq 1$. If $\phi(x) = u(x)$, $P_{x,\bar{\alpha}(x),\beta}[\tau \geq n] = 0$ whereas, if $\phi(x) > u(x)$,

$$\begin{aligned} 2l &\geq E_{x,\bar{\alpha}(x),\beta} \phi(X_{\tau \wedge n}) - \phi(x) \\ &= E_{x,\bar{\alpha}(x),\beta} \sum_{i=0}^{\tau \wedge n-1} \eta_i(Z_1, \dots, Z_i) \\ &\geq E_{x,\bar{\alpha}(x),\beta} I(\tau \geq n) \sum_{i=0}^{n-1} \eta_i(Z_1, \dots, Z_i) \\ &\geq n\zeta P_{x,\bar{\alpha}(x),\beta}[\tau \geq n], \end{aligned}$$

where the first equality is true because of equation (3.5) while the last two inequalities hold because $\eta_i(Z_1, \dots, Z_i) \geq \zeta > 0$ if $i < \tau$. Therefore, for $n = 1, 2, \dots$,

$$P_{x,\bar{\alpha}(x),\beta}[\tau \geq n] \leq \frac{2l}{n\zeta}$$

and thus

$$\begin{aligned} E_{x,\bar{\alpha}(x),\beta} \phi(X_{\tau \wedge n}) - E_{x,\bar{\alpha}(x),\beta} u(X_{\tau \wedge n}) &= E_{x,\bar{\alpha}(x),\beta} I(\tau > n)[\phi(X_{\tau \wedge n}) - u(X_{\tau \wedge n})] \\ &\leq 2l P_{x,\bar{\alpha}(x),\beta}[\tau \geq n] \\ &\leq \frac{4l^2}{n\zeta}. \end{aligned} \tag{3.7}$$

Choose n such that $4l^2/n\zeta \leq \epsilon/2$ and set $t_\epsilon = \tau \wedge n$. Equations (3.6) and (3.7) imply that

$$E_{x,\bar{\alpha}(x),\beta} u(X_{t_\epsilon}) \geq E_{x,\bar{\alpha}(x),\beta} \phi(X_{t_\epsilon}) - \frac{\epsilon}{2} \geq \phi(x) - \frac{\epsilon}{2} \geq U_m(x) - \epsilon,$$

and this proves (3.3). □

3.8 REMARK. The same result also holds for Borel stochastic games; for their definition and properties see Maitra and Sudderth (1993). The proof can be modified to handle measurability problems. □

Two notable corollaries follow easily from Theorem 3.1 and Theorem 3.2.

3.9 COROLLARY. *For any given initial state $x \in S$ and $\epsilon > 0$, player* I *has an ϵ-optimal stationary strategy for $\mathscr{L}(u)(x)$.*

However, player I need not have an optimal stationary strategy. In fact not even the existence of an optimal strategy is guaranteed for her as it was shown with an example by Kumar and Shiau (1981).

3.10 COROLLARY. *Let $\epsilon > 0$. If the sequence $\{U_n\}$ converges uniformly to U, player* I *has a uniformly ϵ-optimal stationary family for $\mathscr{L}(u)$.*

Therefore a uniformly ϵ-optimal stationary family is always available to I for leavable games with finite state space. The same result is not true when the state space of the game is countably infinite, as it was proved by Nowak and Raghavan (1991) with the next example.

3.11 EXAMPLE. Let $S = \{-1, 0, 1, 2, \dots\}$, $A = B = \{0, 1\}$, and u be the indicator function of $\{0\}$. Define the law of motion q by

$$q(-1|-1, a, b) = q(0|0, a, b) = 1$$

for every $a \in A$, $b \in B$, and, for $x = 1, 2, \dots$ set

$$q(x-1|x, 0, 0) = q(x+1|x, 0, 1) = q(0|x, 1, 1) = q(-1|x, 1, 0) = 1.$$

Obviously $U(-1) = 0$, $U(0) = 1$ whereas, for $x = 1, 2, \dots,$

$$U(x) = \frac{x+2}{2x+2} > \frac{1}{2}. \tag{3.12}$$

An interesting proof of (3.12), consonant with the present description of leavable stochastic games, can be found in Maitra and Sudderth (1996).

Fix $\epsilon > 0$ sufficiently small. Then there is no stationary family which is uniformly ϵ-optimal for player I. In fact, let $\overline{\alpha} = \mu^\infty$ where

$$\mu(x) = \theta(x)\delta(1) + [1 - \theta(x)]\delta(0)$$

for $x = 1, 2, \dots$. Consider the following two cases.

First assume that $\sum_{x=1}^{\infty} \theta(x) < \infty$. If β is a strategy for player II which uses action 1 at every state, then, for $x \geq 1$ and any stop rule t,

$$\begin{aligned}
&E_{x, \overline{\alpha}(x), \beta} u(X_t) \\
&\qquad \leq P_{x, \overline{\alpha}(x), \beta}[\text{reach}\{0\}] \\
&\qquad = \theta(x) + [1 - \theta(x)]\theta(x+1) + [1 - \theta(x)][1 - \theta(x+1)]\theta(x+2) + \cdots \\
&\qquad \leq \theta(x) + \theta(x+1) + \theta(x+2) \cdots.
\end{aligned}$$

Therefore, for any stop rule t,

$$\lim_{x \to \infty} E_{x, \overline{\alpha}(x), \beta} u(X_t) = 0,$$

and thus $\overline{\alpha}(x)$ cannot guarantee a return of $U(x) - \epsilon$ for x large.

Conversely, in the second case, assume that $\sum_{x=1}^{\infty} \theta(x) = \infty$. Let now β be the strategy for player II which uses action 0 at every state. Then, for $x \geq 1$ and any stop rule t,

$$\begin{aligned}
E_{x, \overline{\alpha}(x), \beta} u(X_t) &\leq P_{x, \overline{\alpha}(x), \beta}[\text{reach}\{0\}] \\
&= [1 - \theta(x)][1 - \theta(x-1)] \cdots [1 - \theta(1)] \\
&\leq \exp\left[-\sum_{i=1}^{x} \theta(i)\right].
\end{aligned}$$

So again, for any stop rule t,

$$\lim_{x\to\infty} E_{x,\bar{\alpha}(x),\beta}\, u(X_t) = 0. \quad \square$$

4. An auxiliary game. Before tackling the problem concerning the existence of ϵ-optimal stationary strategies for recursive games, we need to show a connection between leavable and nonleavable games which is slightly different from those presented by Maitra and Sudderth (1996).

Given S, A, B, u and a law of motion q as before, let V be the value function of the nonleavable game $\mathscr{N}(u)$.

4.1 LEMMA. *V is the value of the leavable game $\mathscr{L}(u \wedge V)$.*

PROOF. Let U' be the value of $\mathscr{L}(u \wedge V)$. Since $V \geq u \wedge V$ and $GV = V$, by Lemma 2.4 $V \geq U'$. But

$$V = T(u \wedge V) = GU' \leq U',$$

the first equality being true because of Lemma 2.5, the second one by definition of T and the last inequality because of Lemma 2.4. Therefore $V = U'$. $\square$

If $x \in S$ is the initial state of $\mathscr{L}(u \wedge V)$ and $V(x) \leq u(x)$, an optimal strategy for player I is to stop the game immediately and get the payoff $V(x)$ from II. On the contrary, if $V(x) > u(x)$, it is advisable for I to play the game for one day at least. This motivates the introduction of an auxiliary leavable game with the same state and action spaces as $\mathscr{N}(u)$, and with law of motion and utility defined as functions of q, u and the value V of $\mathscr{N}(u)$.

Let

$$L = \{x \in S\colon V(x) > u(x)\},$$

and, for every $a \in A$ and $b \in B$, set

$$\tilde{q}(\cdot|x,a,b) = \begin{cases} q(\cdot|x,a,b) & \text{if } x \in L, \\ \delta(x) & \text{otherwise}, \end{cases}$$

where $\delta(x)$ is the point mass at x. Define $\tilde{\mathscr{L}}(u \wedge V)$ to be the leavable game with law of motion $\tilde{q}$ and utility $u \wedge V$. The tilde superscript will be used for probabilities and expectations relative to the law of motion $\tilde{q}$; accordingly, $\tilde{G}$ will indicate the one day operator relative to $\tilde{q}$ whereas G is the one day operator relative to q.

4.2 LEMMA. *V is the value of the leavable game $\tilde{\mathscr{L}}(u \wedge V)$.*

PROOF. Let $\tilde{U}$ be the value of $\tilde{\mathscr{L}}(u \wedge V)$. Since $V \geq u \wedge V$ and $\tilde{G}V = V$, by Lemma 2.4 $V \geq \tilde{U}$. To prove the reversed inequality note that $\tilde{U} \geq u \wedge V$. If we show that $G\tilde{U} \leq \tilde{U}$, then Lemma 2.4 and Lemma 4.1 imply that $\tilde{U} \geq V$. Consider two cases.

If $V(x) > u(x)$, then

$$(G\tilde{U})(x) = (\tilde{G}\tilde{U})(x) \leq \tilde{U}(x).$$

If $V(x) \leq u(x)$, then

$$(G\tilde{U})(x) \leq (GV)(x) = V(x) = (u \wedge V)(x) = \tilde{U}(x)$$

where the first inequality is true because it has already been shown that $\tilde{U} \le V$, while the last equality follows by the definition of the law of motion $\tilde{q}$.

Therefore $G\tilde{U} \le \tilde{U}$ and this proves that $V = \tilde{U}$. □

With the next lemma we show that, when L is finite, for every $\eta > 0$ there is a stationary family $\overline{\alpha}_\eta$, constructed following the procedure described before Theorem 3.2, which is uniformly η-optimal for player I in $\tilde{\mathscr{L}}(u \wedge V)$.

4.3 LEMMA. *Assume that L is finite and let $\eta > 0$. Then player* I *has a stationary family $\overline{\alpha}_\eta$ such that, for every $x \in S$, there is a stop rule $t_\eta = t_\eta(x)$ for which*

$$\tilde{E}_{x,\,\overline{\alpha}_\eta(x),\,\beta}(u \wedge V)(X_{t_\eta}) \ge V(x) - \eta, \tag{4.4}$$

for any strategy β of player II.

PROOF. Define iteratively a sequence of functions $\{\tilde{U}_n\}$ by setting, $\tilde{U}_0 = u \wedge V$ and, for $n = 1, 2, \dots,$

$$\tilde{U}_n = (u \wedge V) \vee \tilde{G}\tilde{U}_{n-1}.$$

Theorem 3.1 states that the value V of $\tilde{\mathscr{L}}(u \wedge V)$ is the limit of the sequence $\{\tilde{U}_n\}$. Moreover, since $\tilde{U}_n(x) = V(x)$ for every $x \in L^c$ and $n = 0, 1, 2, \dots,$ the sequence $\{\tilde{U}_n\}$ converges uniformly to V, because L is finite. The lemma now follows from Corollary 3.10. □

The previous lemmas apply to any nonleavable game $\mathscr{N}(u)$. In the rest of this section we will restrict our attention to recursive games; this means that the law of motion q of $\mathscr{N}(u)$ is such that every state of a given subset W of S is absorbing and the utility function u is identically zero on W^c. To avoid trivialities, assume that u is not identically zero on S and that the sets W and $L = \{x \in S\colon V(x) > u(x)\}$ are both nonempty. Note that $L \subseteq W^c$ since $V(x) = u(x)$ for all $x \in W$.

Let

$$\zeta = \inf_{x \in L} V(x).$$

Since V is strictly positive on L, $\zeta > 0$ when L is finite. Now define the stopping times

$$\tau = \inf\{n \ge 0\colon X_n \in W\} \quad \text{and} \quad \xi = \inf\{n \ge 0\colon V(X_n) \le u(X_n)\}$$

with the usual convention that $\inf\{\varnothing\} = \infty$. Note that, $\xi \le \tau$ since $V = u$ on W.

Let $\tilde{\mathscr{L}}(u \wedge V)$ be the auxiliary game relative to the recursive game $\mathscr{N}(u)$ and t_η be the stop rule appearing in the statement of Lemma 4.3; from now on we will assume $t_\eta \le \xi$. In fact, if $x \in S$ is the initial state of the game and $\overline{\alpha}_\eta$ and t_η are a stationary strategy and a stop rule such that (4.4) is true, then, by conditioning to $p_{\xi \wedge t_\eta}$ and applying formula (2.1), we get

$$\begin{aligned} V(x) - \eta &\le \tilde{E}_{x,\,\overline{\alpha}_\eta(x),\,\beta}(u \wedge V)(X_{t_\eta}) \\ &= \int \tilde{E}_{X_{\xi \wedge t_\eta},\,\overline{\alpha}_\eta(X_{\xi \wedge t_\eta}),\,\beta[p_{\xi \wedge t_\eta}]}(u \wedge V)(X_{t_\eta[p_{\xi \wedge t_\eta}]})\, d\tilde{P}_{x,\,\overline{\alpha}_\eta(x),\,\beta} \\ &= \int_{\{\xi \le t_\eta\}} (u \wedge V)(X_\xi)\, d\tilde{P}_{x,\,\overline{\alpha}_\eta(x),\,\beta} + \int_{\{\xi > t_\eta\}} (u \wedge V)(X_{t_\eta})\, d\tilde{P}_{x,\,\overline{\alpha}_\eta(x),\,\beta} \\ &= \tilde{E}_{x,\,\overline{\alpha}_\eta(x),\,\beta}(u \wedge V)(X_{\xi \wedge t_\eta}), \end{aligned}$$

for any strategy β of II. Therefore (4.4) still holds if t_η is replaced by the stop rule $t_\eta \wedge \xi$.

4.5 LEMMA. *Let $0 < \eta < \zeta$ and assume that L is finite. Then, for all $x \in L$,*

$$\tilde{P}_{x,\bar{\alpha}_\eta(x),\beta}\left[\tau \le t_\eta\right] \ge \frac{\zeta - \eta}{\sup_{x \in S}|u(x)|} > 0, \tag{4.6}$$

for any strategy β of player II.

PROOF. Apply formula (2.1) by conditioning to $p_{\tau \wedge t_\eta}$. Then, for every $x \in L$,

$$\begin{aligned}
0 < \zeta - \eta \le V(x) - \eta &\le \tilde{E}_{x,\bar{\alpha}_\eta(x),\beta}(u \wedge V)\left(X_{t_\eta}\right)\\
&= \int \tilde{E}_{X_{\tau\wedge t_\eta},\bar{\alpha}_\eta(X_{\tau\wedge t_\eta}),\beta[p_{\tau\wedge t_\eta}]}(u \wedge V)\left(X_{t_\eta[p_{t_\eta\wedge\tau}]}\right) d\tilde{P}_{x,\bar{\alpha}_\eta(x),\beta}\\
&= \int_{\{\tau \le t_\eta\}} u(X_\tau)\, d\tilde{P}_{x,\bar{\alpha}_\eta(x),\beta} + \int_{\{\tau > t_\eta\}} (u \wedge V)\left(X_{t_\eta}\right) d\tilde{P}_{x,\bar{\alpha}_\eta(x),\beta}\\
&\le \int_{\{\tau \le t_\eta\}} u(X_\tau)\, d\tilde{P}_{x,\bar{\alpha}_\eta(x),\beta}\\
&\le \left\{\sup_{x \in S}|u(x)|\right\}\tilde{P}_{x,\bar{\alpha}_\eta(x),\beta}\left[\tau \le t_\eta\right],
\end{aligned}$$

for any β of II. Note that the inequality next to the last holds because $(u \wedge V)(X_{t_\eta}) \le 0$ if $t_\eta < \tau$. Since u is a non-null function, (4.6) is thus proved. □

An immediate consequence of the lemma is that, by means of a stationary family $\bar{\alpha}_\eta$, player I can reach for sure the set where $V \le u$ which contains W.

4.7 COROLLARY. *Assume that L is finite and let $0 < \eta < \zeta$. Then, for all $x \in S$,*

$$\tilde{P}_{x,\bar{\alpha}_\eta(x),\beta}[\xi < \infty] = 1,$$

for any strategy β of II.

PROOF. Let $x \in S$ be the initial state of $\tilde{\mathscr{L}}(u \wedge V)$ and fix a strategy β for II. If $x \in L^c$, then $\xi = \underline{0}$. If $x \in L$ and player I uses the stationary family $\bar{\alpha}_\eta$, because of (4.6) there is a positive probability that the set W will be reached by time t_η. This probability is bounded below by a strictly positive constant which does not depend on x. Thus the process of states $X_0 = x, X_1, X_2, \ldots$ will leave for sure the set L if player I incessantly uses $\bar{\alpha}_\eta$. □

In the terminology of gambling theory, the following theorem states that player I has a stationary family which is almost thrifty for the auxiliary game $\tilde{\mathscr{L}}(u \wedge V)$ when the set L is finite.

4.8 THEOREM. *Assume that L is finite and let $\epsilon > 0$. Then there is an $\eta \in (0, \zeta \wedge \epsilon)$ such that, for all $x \in S$,*

$$\tilde{E}_{x,\bar{\alpha}_\eta(x),\beta}V(X_\xi) \ge V(x) - \epsilon, \tag{4.9}$$

for any strategy β of player II.

PROOF. Set $M = \sup_{x \in S}|u(x)|$ and let $\eta \in (0, \zeta \wedge \epsilon)$ be sufficiently small so that the following inequality is satisfied:

$$\eta\left(2 + \frac{\log(\eta) - \log M}{\log(M - \zeta + \eta) - \log M}\right) < \epsilon.$$

For all $x \in L^c$, (4.9) is obvious since $\xi = \underline{0}$. So fix $x \in L$ and a strategy β for player II. For simplicity of notation in the rest of the proof we will write $\tilde{P}$ for the probability $\tilde{P}_{x, \bar{\alpha}_\eta(x), \beta}$ and $\tilde{E}$ for the corresponding expectation.

Define a sequence $\{s_n\}$ of stop rules by setting, for every history $h = (z_1, z_2, \dots) = ((x_1, a_1, b_1), (x_2, a_2, b_2), \dots)$,

$$s_0(h) = t_\eta(x)(h)$$

and, for $n = 0, 1, \dots,$

$$s_{n+1}(h) = \begin{cases} s_n(h) + t_\eta(X_{s_n(h)})(z_{s_n(h)+1}, z_{s_n(h)+2}, \dots) & \text{if } V(X_{s_n(h)}) > u(X_{s_n(h)}), \\ s_n(h) & \text{if } V(X_{s_n(h)}) \le u(X_{s_n(h)}). \end{cases}$$

Note that

$$\tilde{P}[s_1 > s_0] \le \tilde{P}[s_0 < \tau] \le 1 - \frac{\zeta - \eta}{M} < 1,$$

where the second inequality holds because of Lemma 4.5. By induction on n,

$$\tilde{P}[s_n > s_{n-1}] \le \left(1 - \frac{\zeta - \eta}{M}\right)^n.$$

Choose now an integer k such that,

$$M\left(1 - \frac{\zeta - \eta}{M}\right)^{k+1} \le \eta < M\left(1 - \frac{\zeta - \eta}{M}\right)^k.$$

Note that the choice of k does not depend on $\tilde{P}$. Now, by (4.4) and conditioning iteratively on $p_{s_{k-1}}, p_{s_{k-2}}, \dots, p_{s_0}$, we obtain

$$\begin{aligned} \tilde{E}_{x, \bar{\alpha}_\eta(x), \beta}(u \wedge V)(X_{s_k}) &= \int \tilde{E}_{X_{s_{k-1}}, \bar{\alpha}_\eta(X_{s_{k-1}}), \beta[p_{s_{k-1}}]}(u \wedge V)(X_{s_k[p_{s_{k-1}}]})\, d\tilde{P} \\ &\ge \tilde{E}_{x, \bar{\alpha}_\eta(x), \beta} V(X_{s_{k-1}}) - \eta \\ &\ge \tilde{E}_{x, \bar{\alpha}_\eta(x), \beta}(u \wedge V)(X_{s_{k-1}}) - \eta \\ &\ge V(x) - (k+1)\eta. \end{aligned}$$

Thus, by conditioning on p_ξ,

$$
\begin{aligned}
V(x) - (k+1)\eta &\le \tilde{E}_{x,\bar{\alpha}_\eta(x),\beta}(u \wedge V)(X_{s_k}) \\
&= \int_{\{\xi \le s_k\}} V(X_\xi)\, d\tilde{P} + \int_{\{\xi > s_k\}} (u \wedge V)(X_{s_k})\, d\tilde{P} \\
&= \int_{\{\xi \le s_k\}} V(X_\xi)\, d\tilde{P} \\
&= \tilde{E}_{x,\bar{\alpha}_\eta(x),\beta} V(X_\xi) - \int_{\{\xi > s_k\}} V(X_\xi)\, d\tilde{P} \\
&\le \tilde{E}_{x,\bar{\alpha}_\eta(x),\beta} V(X_\xi) + M\tilde{P}[s_k < \xi] \\
&\le \tilde{E}_{x,\bar{\alpha}_\eta(x),\beta} V(X_\xi) + M\tilde{P}[s_{k+1} > s_k] \\
&\le \tilde{E}_{x,\bar{\alpha}_\eta(x),\beta} V(X_\xi) + \eta.
\end{aligned}
$$

The second equality is true because $(u \wedge V)(X_{s_k}) = 0$ if $s_k < \xi$, the second inequality holds because $\sup_{x \in S}|V(x)| \le M$ and, finally, the next to the last inequality is true because $t_\eta(X_{s_k})(z_{s_k+1}, z_{s_k+2}, \dots) > 0$ if $s_k < \xi$ since $\eta < \zeta = \inf_{x \in L} V(x)$ so that, for all $x \in L$, $V(x) - \eta > u(x) = 0$. It follows that

$$\tilde{E}_{x,\bar{\alpha}_\eta(x),\beta} V(X_\xi) \ge V(x) - (k+2)\eta.$$

But, because of the way η and k were chosen, $(k+2)\eta \le \epsilon$ so that the theorem is proved. □

4.10 Remark. Since $q(\cdot|x, a, b) = \tilde{q}(\cdot|x, a, b)$ for every $x \in L$, $a \in A$, $b \in B$, the results of Corollary 4.7 and Theorem 4.8 are still true when $\tilde{P}$ and $\tilde{E}$ are replaced by P and E. That is, when L is finite, for every $\epsilon > 0$ there is an $\eta \in (0, \zeta \wedge \epsilon)$ such that, for all $x \in S$,

$$P_{x,\bar{\alpha}_\eta(x),\beta}[\xi < \infty] = 1, \tag{4.11}$$

and

$$E_{x,\bar{\alpha}_\eta(x),\beta} V(X_\xi) \ge V(x) - \epsilon \tag{4.12}$$

for any strategy β of II. □

5. Recursive games. A recursive matrix game, as defined by Orkin (1972), consists of a finite collection of finite matrices $S_0 = \{\Gamma^1, \Gamma^2, \dots, \Gamma^n\}$; associated with each matrix Γ^k is a finite set of real numbers $R_k = \{e_1^k, e_2^k, \dots, e_{m(k)}^k\}$. For $k = 1, \dots, n$, the entry (i, j) of the matrix Γ^k is a probability distribution p_{ij}^k on $S_0 \cup R_k$. The game starts at some matrix, say Γ^k. Player I chooses, possibly at random, a row i of the matrix and player II chooses, possibly at random, a column j. An element of $S_0 \cup R_k$ is then selected according to the distribution p_{ij}^k. If an element $e_l^k \in R_k$ is selected, the game terminates and player II pays to I the quantity e_l^k. If a matrix Γ^l is selected, then the game continues as before from Γ^l with no payoff from II to I. When the play goes on infinitely long the payoff from II to I is 0.

As was pointed out by Maitra and Sudderth (1996), it is not difficult to see that a recursive matrix game is equivalent to a nonleavable game. In fact set

$$W = R_1 \cup R_2 \cup \cdots \cup R_n,$$

and $S = S_0 \cup W$. Assume that all the matrices in S_0 have the same number m of rows and r of columns; if this is not the case add rows and columns in a way that does not affect the game. Let $A = \{1, \dots, m\}$ be the action space for player I and $B = \{1, \dots, r\}$ the action space for player II. Define the law of motion q by setting, for every $a \in A$ and $b \in B$,

$$q(\cdot|x, a, b) = \begin{cases} p_{ab}^k & \text{if } x = \Gamma^k \in S_0, \\ \delta(x) & \text{if } x \in W, \end{cases}$$

where $\delta(x)$ is the point mass at x. Finally, set

$$u(x) = \begin{cases} x & \text{if } x \in W, \\ 0 & \text{if } x \notin W. \end{cases}$$

The space S, the action sets A and B, the law of motion q and the utility u define a nonleavable game $\mathscr{N}(u)$, which we called recursive for reasons now evident. Let us indicate the value of this game with V as before.

In this section we will show that in any recursive game such that the set $L = \{x \in S\colon V(x) > u(x)\}$ is finite, for every $\epsilon > 0$, there is a uniformly ϵ-optimal stationary family available to player I. Because of the inherent symmetry of any recursive game, this will also prove that a uniformly ϵ-optimal stationary family is available to player II if it is finite the subset of S where the value of the game is strictly less than its utility.

Given $\eta > 0$ and a recursive game $\mathscr{N}(u)$ with value V, consider the auxiliary leavable game $\tilde{\mathscr{L}}(u \wedge V)$ introduced in the previous section and the η-optimal stationary family $\overline{\alpha}_\eta$ whose existence is guaranteed by Lemma 4.3 when L is finite. Let μ_η be the map from S to $\mathscr{P}(A)$ which defines $\overline{\alpha}_\eta$: that is assume $\overline{\alpha}_\eta = \mu_\eta^\infty$. Now define μ to be a function which maps every $x \in S$ to an element of $\mathscr{P}(A)$ which is optimal for player I in the one-day game $\mathscr{A}(V)(x)$. Finally, define on S the function μ'_η by setting

$$\mu'_\eta(x) = \begin{cases} \mu_\eta(x) & \text{if } x \in L, \\ \mu(x) & \text{otherwise}. \end{cases}$$

The next theorem states that, when L is finite, for every $\epsilon > 0$ there is a stationary family $\overline{\alpha}'_\eta = \mu'^\infty_\eta$ which is uniformly ϵ-optimal for player I. An informal description of $\overline{\alpha}'_\eta$ is the following: player I acts almost optimally in the leavable game $\tilde{\mathscr{L}}(u \wedge V)$ whenever the current state of the game is such that $V > u$ whereas player I conserves V by playing optimally in the one-day game $\mathscr{A}(V)$ when the current state of the game is such that $V \leq u$.

5.1 THEOREM. *Assume that L is finite and let $\epsilon > 0$. Then there is an $\eta > 0$ such that, for all $x \in S$,*

$$E_{x, \bar{\alpha}'_\eta(x), \beta} u^* \geq V(x) - \epsilon,$$

for any strategy β of player II.

PROOF. We will consider two cases.

First assume that L is empty. Then, for every $x \in S$, $V(x) \leq u(x)$ and $\bar{\alpha}'_\eta = \mu^\infty$ for any $\eta > 0$. Therefore, for all $x \in S$,

$$E_{x, \bar{\alpha}'_\eta(x), \beta} u(X_t) \geq E_{x, \bar{\alpha}'_\eta(x), \beta} V(X_t) \geq V(x)$$

for any stop rule t and any strategy β of player II. The second inequality follows by the Optional Sampling Theorem of Doob (1953) and the fact that the process $\{V(X_n)\}$ is a uniformly bounded submartingale with respect to $P_{x, \bar{\alpha}'_\eta(x), \beta}$. Apply now formula (2.6) and obtain, for all $x \in S$,

$$E_{x, \bar{\alpha}'_\eta(x), \beta} u^* \geq V(x)$$

for any strategy β of player II. This shows that $\bar{\alpha}'_\eta$ is optimal for I and proves the theorem in this case.

Now assume that L is nonempty but finite. As before, set

$$\zeta = \inf_{x \in L} V(x)$$

and $M = \sup_{x \in S} |u(x)|$. Let

$$\tau = \inf\{n \geq 0: X_n \in W\}, \qquad \xi = \inf\{n \geq 0: V(X_n) \leq u(X_n)\}$$

and define two sequences $\{t_n\}$ and $\{s_n\}$ of stopping times by setting $t_0 = \xi$, and, for any $n \geq 0$,

$$s_n = \inf\{n \geq t_n: V(X_n) > u(X_n)\}, \qquad t_{n+1} = \inf\{n \geq s_n: V(X_n) \leq u(X_n)\}$$

with the usual convention that $\inf\{\emptyset\} = \infty$.

By Remark 4.10, for any $\theta \in (0, \zeta)$ there is an $\eta \in (0, \theta)$ such that, for all $x \in S$,

$$E_{x, \bar{\alpha}'_\eta(x), \beta} V(X_\xi) \geq V(x) - \theta \tag{5.2}$$

and

$$P_{x, \bar{\alpha}'_\eta(x), \beta}[\xi < \infty] = 1 \tag{5.3}$$

for any strategy β of player II. Choose θ and $\eta < \theta$ such that

$$\theta\left(6 + \frac{\log \theta - \log M}{\log(M - \zeta + \eta) - \log M}\right) < \epsilon.$$

Now fix $x \in S$ and a strategy β for player II. We will begin by assuming that $x \in L$. For the sake of simplicity write $\bar{\alpha}'$ for $\bar{\alpha}'_\eta$, P for the probability $P_{x, \bar{\alpha}'_\eta(x), \beta}$ and E for the expected value computed according to P.

Notice first that

$$P[s_0 < \infty] \leq 1 - \frac{\zeta - \eta}{M}.$$

In fact,

$$\frac{\zeta - \eta}{M} \leq P(\tau \leq t_\eta) \leq P(\tau = \xi) \leq P(s_0 = \infty).$$

The first inequality follows from Lemma 4.5, the following one holds because $t_\eta \leq \xi \leq \tau$ while the last inequality is true since the game terminates as soon as W is reached. By induction on n and conditioning on $p_{s_{n-1}}$ one can show that, for any $n \geq 1$,

$$P(s_n < \infty) \leq \left(1 - \frac{\zeta - \eta}{M}\right)^{n+1}.$$

Let k be such that,

$$M\left(1 - \frac{\zeta - \eta}{M}\right)^{k+1} \leq \theta < M\left(1 - \frac{\zeta - \eta}{M}\right)^{k},$$

and observe that the choice of k does not depend on P.

Now set $V^* = \limsup_{n \to \infty} V(X_n)$. The aim of what follows is to show that

$$EV^* \geq V(x) - (k+3)\theta. \tag{5.4}$$

Notice that, for $n = 0, 1, 2, \ldots, t_n \leq s_n \leq t_{n+1}$. Let N be an integer and, for every $h \in H$, set

$$\lambda_N(h) = \inf\{n \geq 0 \colon t_n(h) + N < s_n(h)\}$$

with the usual convention that $\inf\{\varnothing\} = \infty$. Although λ_N need not be a stopping time, it is easy to check that $t_{\lambda_N} + N$ is a stopping time if we set $t_\infty = \infty$. Define now a sequence of stopping times $\{m_n\}$ by setting, for any $j = 0, 1, 2, \ldots,$

$$m_{2j} = t_j \wedge (t_{\lambda_N} + N)$$

and

$$m_{2j+1} = s_j \wedge (t_{\lambda_N} + N).$$

For all $n \geq 0$, the stopping time m_n is finite almost surely with respect to P. Intuitively, whenever the process $X_0 = x, X_1, X_2, \ldots$ is in L, it leaves L with probability one because of Corollary 4.7, whereas, if the process is in L^c, then either it reaches L before N days or it is definitely stopped at the Nth day.

For $j = 0, 1, 2, \ldots,$ set

$$Y_{2j} = V(X_{m_{2j}}) + j\theta$$

and

$$Y_{2j+1} = V(X_{m_{2j+1}}) + j\theta.$$

The sequence $\{Y_n\}$ is a submartingale with respect to P. This happens because the strategy $\overline{\alpha}'(x)$ is such that the expected value of V decreases by at most θ whenever the process $X_0 = x, X_1, X_2, \ldots$ leaves the set L, whereas V is conserved outside L. More precisely: for $n = 0, 1, 2, \ldots, E(|Y_n|) < \infty$, since V is bounded, and Y_n is measurable with respect to the sigma-field generated by $X_1, \ldots, X_{m_n}$. Also, if $j = 1, 2, \ldots,$

$$
\begin{aligned}
E&\left[Y_{2j}|X_1,\ldots,X_{m_{2j-1}}\right]\\
&= E\left[V\left(X_{t_j\wedge(t_{\lambda_N}+N)}\right)|X_1,\ldots,X_{m_{2j-1}}\right] + j\theta\\
&= E_{X_{m_{2j-1}},\overline{\alpha}'(X_{m_{2j-1}}),\beta[p_{m_{2j-1}}]}V\left(X_{(t_j\wedge(t_{\lambda_N}+N))[p_{m_{2j-1}}]}\right) + j\theta\\
&= I[\lambda_N \le j-1]E_{X_{t_{\lambda_N}+N},\overline{\alpha}'(X_{t_{\lambda_N}+N}),\beta[p_{t_{\lambda_N}+N}]}V(X_0)\\
&\qquad + I[\lambda_N \ge j]E_{X_{s_{j-1}},\overline{\alpha}'(X_{s_{j-1}}),\beta[p_{s_{j-1}}]}V(X_\xi) + j\theta\\
&\ge I[\lambda_N \le j-1]V\left(X_{t_{\lambda_N}+N}\right) + I[\lambda_N \ge j]\left[V\left(X_{s_{j-1}}\right) - \theta\right] + j\theta\\
&\ge V\left(X_{s_{j-1}\wedge(t_{\lambda_N}+N)}\right) + (j-1)\theta\\
&= Y_{2j-1}.
\end{aligned}
$$

The third equality above holds because $s_{j-1} < \infty$ if $\lambda_N \ge j$, while the first inequality follows by (5.2). On the other hand, if $j = 0, 1, 2, \ldots,$

$$
\begin{aligned}
E&\left[Y_{2j+1}|X_1,\ldots,X_{m_{2j}}\right]\\
&= E\left[V\left(X_{s_j\wedge(t_{\lambda_N}+N)}\right)|X_1,\ldots,X_{m_{2j}}\right] + j\theta\\
&= I[\lambda_N \le j-1]V\left(X_{t_{\lambda_N}+N}\right) + I[\lambda_N \ge j]E_{X_{t_j},\overline{\alpha}'(X_{t_j}),\beta[p_{t_j}]}V\left(X_{(s_j\wedge(t_{\lambda_N}+N))[p_{t_j}]}\right) + j\theta\\
&\ge I[\lambda_N \le j-1]V\left(X_{t_{\lambda_N}+N}\right) + I[\lambda_N \ge j]V\left(X_{t_j}\right) + j\theta\\
&= Y_{2j},
\end{aligned}
$$

where the second equality above holds because, if $\lambda_N \ge j$, then t_j is finite except on a set of probability zero, whereas the following inequality is true since, given $X_1, \ldots, X_{t_j}$, the process $V(X_{t_j+1}), V(X_{t_j+2}), \ldots$ is a submartingale up to time $(s_j \wedge (t_{\lambda_N} + N))[p_{t_j}] \le N$, so that Doob's Optional Sampling Theorem applies (Doob 1953).

Theorem 4.8 and the fact that $\{Y_n\}$ is a submartingale imply that

$$V(x) - \theta \le EV(X_\xi) = EY_0 \le EY_{2k+1} = EV\left(X_{s_k\wedge(t_{\lambda_N}+N)}\right) + k\theta.$$

To prove (5.4), let $N \to \infty$ and apply Fatou's Lemma obtaining

$$\begin{aligned}
&V(x) - (k+1)\theta \\
&\quad \le \limsup_{N\to\infty} EV\left(X_{s_k \wedge (t_{\lambda_N}+N)}\right) \\
&\quad \le E\left\{\limsup_{N\to\infty} V\left(X_{s_k \wedge (t_{\lambda_N}+N)}\right)\right\} \\
&\quad = \int_{\{s_k=\infty\}}\left[\limsup_{N\to\infty} V\left(X_{s_k \wedge (t_{\lambda_N}+N)}\right)\right] dP + \int_{\{s_k<\infty\}}\left[\limsup_{N\to\infty} V\left(X_{s_k \wedge (t_{\lambda_N}+N)}\right)\right] dP \\
&\quad \le \int_{\{s_k=\infty\}} V^* \, dP + \theta \\
&\quad \le EV^* + 2\theta.
\end{aligned}$$

Note that, given $\{s_k = \infty\}$, with probability one, $\lambda_N \le k$ and $s_k \wedge (t_{\lambda_N} + N) = t_{\lambda_N} + N$. This implies that, given $\{s_k = \infty\}$, with probability one, $\limsup_{N\to\infty} V(X_{s_k \wedge (t_{\lambda_N}+N)}) = V^*$ and explains part of the next to the last inequality above. The other part is true because V is bounded by M and $MP(s_k < \infty) \le \theta$.

Since $u \ge V$ on L^c, $u^* \ge V^*$ when the process of states $X_0 = x, X_1, X_2, \ldots$ returns to the set L only a finite number of times. Therefore

$$\begin{aligned}
Eu^* &= \int_{\{s_k=\infty\}} u^* \, dP + \int_{\{s_k<\infty\}} u^* \, dP \\
&\ge \int_{\{s_k=\infty\}} V^* \, dP - \theta \\
&\ge EV^* - 2\theta \\
&\ge V(x) - (k+5)\theta.
\end{aligned}$$

However k and θ were chosen in such a way that $(k+5)\theta \le \epsilon$. Thus we just proved that, for all $x \in L$,

$$E_{x, \bar{\alpha}'_\eta(x), \beta} u^* \ge V(x) - \epsilon, \tag{5.5}$$

for any strategy β of player II.

Equation (5.5) is true even when $x \notin L$. In fact it is not difficult to show that, for every $x \in L^c$ and $n \ge 0$,

$$P_{x, \bar{\alpha}'(x), \beta}(s_n < \infty) \le \left(1 - \frac{\zeta - \eta}{M}\right)^n$$

for any strategy β of player II. Therefore, by letting s_{k+1} take the place of s_k, the argument above can be easily modified for the purpose of proving that, for all $x \in L^c$,

$$E_{x, \bar{\alpha}'(x), \beta} u^* \ge V(x) - (k+6)\theta$$

for any strategy β of player II. But $(k+6)\theta \le \epsilon$ so that the proof of the theorem is now complete. □

The previous theorem shows that player I has a uniformly ϵ-optimal stationary family of strategies when the subset of S where $V > u$ is finite. Given the inherent symmetry of any recursive game, by interchanging the roles of the players the theorem also proves the existence of a uniformly ϵ-optimal stationary family for player II when the subset of S where $V < u$ is finite.

Before concluding the section, we note that, when the set $L \doteq \{x \in S: V(x) > u(x)\}$ is countably infinite, Theorem 5.1 is no longer true. In fact the game of Nowak and Raghavan introduced in Example 3.11 is easily seen to be a recursive game where the set L coincides with the set of all strictly positive integers. We have already shown that for this game there is no uniformly ϵ-optimal stationary family for player I, although, for any given initial state of the game, there is an ϵ-optimal stationary strategy available to her. We still do not know if this last fact is more generally true for any recursive game.

6. Acknowledgments. This paper is part of my Ph.D. thesis at the University of Minnesota. I am much indebted to my advisor William Sudderth for his precious help and I want to thank him for the stimulating discussions we had during the preparation of this work. I want also to thank an anonymous referee whose suggestions contributed to make this paper better.

References

Doob, J. L. (1953). *Stochastic Processes*. Wiley, New York.

Dubins, L. E., L. J. Savage (1976). *Inequalities for Stochastic Processes. How to Gamble if you Must*, Dover, New York.

Everett, H. (1957). Recursive games, Dresher, M. et al., ed., *Contributions to the Theory of Games*, III, volume 39 of *Annals of Mathematical Studies*, pp. 47–78. Princeton University Press, New Jersey.

Kumar, P. R., T. H. Shiau (1981). Existence of value and randomized strategies in zero-sum discrete time stochastic dynamic games, *Siam J. Control and Optim.* **19** 617–634.

Maitra, A., W. Sudderth (1992). An operator solution of stochastic games, *Israel J. Math.* **78** 33–49.

____, ____ (1993). Borel stochastic games with lim sup payoff, *Ann. Probab.* **21** 861–865.

____, ____ (1996). *Discrete Gambling and Stochastic Games*, Springer-Verlag, New York.

Milnor, J., L. S. Shapley (1957). On games of survival, Dresher, M. et al., ed., *Contributions to the Theory of Games*, III, volume 39 of *Annals of Mathematical Studies*, Princeton University Press, New Jersey.

Nowak, A. S., T. E. S. Raghavan (1991). Positive stochastic games and a theorem of Ornstein, Raghavan, T. E. S. et al., ed., *Stochastic Games and Related Topics*, Kluwer Academic Press, The Netherlands, 127–134.

Orkin, M. (1972). Recursive matrix games, *J. Appl. Probab.* **9** 813–820.

Sudderth, W. (1969). On the existence of good stationary strategies, *Trans. Amer. Math. Soc.* **135** 399–414.

____ (1971). On measurable gambling problems, *Ann. Math. Statist.* **42** 260–269.

von Neumann, J., O. Morgenstern (1947). *Theory of Games and Economic Behavior*, 2nd ed. Princeton University Press, New Jersey.

P. Secchi: Dipartimento di Economia Political Metodi Quantitativi, Università di Pavia, Via San Felice 5, I-27100 Pavia, Italy; e-mail: secchi@ipv36.unipv.it