Available online at www.sciencedirect.com
SCIENCE DIRECT®

Stochastic Processes and their Applications 112 (2004) 319–329

stochastic processes and their applications

www.elsevier.com/locate/spa

Zero-sum dynamic games and a stochastic variation of Ramsey's theorem

Eran Shmaya[a], Eilon Solan[a,b,*,1]

[a] School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel
[b] Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208-2001, USA

Received 11 July 2003; received in revised form 23 February 2004; accepted 2 March 2004

Abstract

We show how a stochastic variation of a Ramsey's theorem can be used to prove the existence of the value, and to construct ε-optimal strategies, in two-player zero-sum dynamic games that have certain properties.

Keywords: Dynamic games; Ramsey's theorem; Value; Optimal strategies

1. Introduction

Competitive interaction between two players is quite common, and it is desirable to know whether such interaction has a value. That is, whether there is some quantity such that for every $\varepsilon > 0$ the maximizing player can guarantee receiving, on average, at least this quantity (up to ε), and the minimizing player can guarantee paying, on average, no more than this quantity (up to ε). Once the value exists, finding ε-optimal strategies for the two players (which guarantee that they receive at least the value, or pay no more than the value, up to ε) is also desirable.

When the interaction lasts for a single stage, or for a bounded number of stages, existence of the value is usually proven using a fixed-point argument, and hinges on the continuity of the payoff in the strategies of the players.

* Corresponding author. Tel.: +97236409635; fax: +97236409357.
E-mail addresses: gawain@post.tau.ac.il (E. Shmaya), e-solan@kellogg.northwestern.edu, eilons@post.tau.ac.il (E. Solan).
[1] The research of the second author was supported by the Israel Science Foundation (Grant No. 69/01-1).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2004.03.001

When the duration of the interaction is long but not known in advance, it is convenient to assume that the interaction lasts for countably many stages (see Aumann and Maschler, 1995, p. 143, and Neyman and Sorin, 2001, for justification). However, in this formulation, the payoff is often not continuous in the strategies of the players, and therefore one cannot use standard fixed-point theorems to prove the existence of the value. Various techniques have been employed in the literature to handle this problem (see, e.g., Mertens and Neyman, 1981; Maitra and Sudderth, 1993, 1998; Nowak, 1985), but it seems that each technique can be applied to only some models, under special conditions, or else is not constructive.

Here we present another tool for proving the existence of the value in infinite-stage competitive interactions, or two-player zero-sum dynamic games. We show how a stochastic variation of Ramsey's theorem[2] can be used to reduce the analysis of the infinite-stage interaction to the analysis of finite-stage interactions.

To exhibit the new technique, we apply it to the following generalization of stopping games (see Dynkin, 1969). At the outset of the game, the state of the world is chosen according to some known probability distribution, but is not revealed to the players. At every stage of the game, the players gain some information about the state of the world; both receive the same information. Then each player chooses an action. The pair of actions, together with the state of the world, determines a probability of termination, and a terminal payoff if the game terminates at that stage. If the game never terminates, the payoff to both players is 0.

The goal of the maximizing player is to maximize the expected payoff, and the goal of the minimizing player is to minimize this quantity.

The paper is arranged as follows. In Section 2 we formally present the model and the main result, stating that in our model the value exists. In Section 3 we state the stochastic variation of Ramsey's theorem that we use, and we apply it to our model to exhibit the new technique. Further discussion appears in Section 4.

2. The model and the main result

We consider infinite-stage dynamic games in discrete time that are given by:

- A probability space $(\Omega, \mathscr{F}, \mathbf{P})$ that captures the uncertainty of the state of the world. We denote by $\mathbf{E}$ the expectation w.r.t. $\mathbf{P}$.
- A filtration $(\mathscr{F}_n)_{n\in\mathbf{N}}$ that describes the information available to both players at stage n.
- Two measurable spaces $(A, \mathscr{A})$ and $(B, \mathscr{B})$ of *actions* for the two players.
- For each $n \in \mathbf{N}$, $\mathscr{F}_n \otimes \mathscr{A} \otimes \mathscr{B}$-measurable functions $p_n : \Omega \times A \times B \to [0,1]$ and $g_n : \Omega \times A \times B \to [-1,1]$; p_n indicates the probability of termination, while g_n indicates the terminal payoff.

[2] Ramsey (1930) proved that for every coloring of the complete infinite graph by finitely many colors there is a complete infinite monochromatic sub-graph.

The game is played as follows. At the outset, a state of the world $\omega \in \Omega$ is chosen according to the probability measure $\mathbf{P}$. At every stage n, the players independently and simultaneously choose actions $a_n \in A$ and $b_n \in B$. These choices must be measurable with regard to their information, namely, $\mathscr{F}_n$ and previously played actions. The game terminates with probability $p_n(\omega, a_n, b_n)$, and the terminal payoff is $g_n(\omega, a_n, b_n)$. The game continues to stage $n+1$ with probability $1 - p_n(\omega, a_n, b_n)$.

Our model is a generalization of stopping games (see, e.g., Dynkin, 1969; Rosenberg et al., 2001 or Touzi and Vieille, 2002). It is also closely related to general stochastic games (see, e.g., Nowak, 1985 or Maitra and Sudderth, 1998).

For every measurable space M, we denote by $\mathscr{P}(M)$ the space of probability distributions over M.

The space of infinite plays is $(A \times B)^{\mathbf{N}} \times \Omega$. We equip it with the product σ-algebra $(\mathscr{A} \otimes \mathscr{B})^{\mathbf{N}} \otimes \mathscr{F}$. We denote by $\mathscr{G}_n = (\mathscr{A} \times \mathscr{B})^{n-1} \otimes \mathscr{F}_n$ the σ-algebra that represents the information available to the players at stage n. It is convenient to consider $\mathscr{F}_n$ as a sub-σ-algebra of $\mathscr{G}_n$.

A *strategy* $\sigma = (\sigma_n)_{n \in \mathbf{N}}$ of player 1 is a collection of functions such that $\sigma_n : (A \times B)^{\mathbf{N}} \times \Omega \to \mathscr{P}(A)$ is $\mathscr{G}_n$-measurable, for every $n \in \mathbf{N}$. Strategies τ of player 2 are defined analogously.

Every pair (σ, τ) of strategies, together with $\mathbf{P}$, naturally defines a probability distribution over $(A \times B)^{\mathbf{N}} \times \Omega$. The corresponding expectation operator is denoted by $\mathbf{E}_{\sigma,\tau}$.

Denote by θ the stage of termination, so that $\theta = +\infty$ if termination never occurs. The distribution of θ is formally given by

$$\begin{aligned}
&\mathbf{P}(\theta = 1)(\omega, a, b) = p_1(\omega, a, b),\\
&\mathbf{P}(\theta \leqslant k)(\omega, a_1, b_1, \ldots, a_k, b_k)\\
&\quad = \mathbf{P}(\theta \leqslant k-1)(\omega, a_1, b_1, \ldots, a_{k-1}, b_{k-1})\\
&\qquad + (1 - \mathbf{P}(\theta \leqslant k-1)(\omega, a_1, b_1, \ldots, a_{k-1}, b_{k-1})) \times p_k(\omega, a_k, b_k).
\end{aligned}$$

For every pair (σ, τ) of strategies, the expected payoff is

$$\gamma(\sigma, \tau) = \mathbf{E}_{\sigma,\tau}[\mathbf{1}_{\{\theta < +\infty\}} g_\theta(a_\theta, b_\theta)],$$

where $\mathbf{1}$ is the indicator function.

The goal of player 1 is to maximize the expected payoff, while the goal of player 2 is to minimize this quantity.

Definition 1. If the equality

$$\sup_\sigma \inf_\tau \gamma(\sigma, \tau) = \inf_\tau \sup_\sigma \gamma(\sigma, \tau) \tag{1}$$

holds, then the common value is the *value* of the game. Given $\varepsilon \geqslant 0$, every strategy σ of player 1 that attains the supremum on the left-hand side of (1) up to ε is ε-*optimal* for player 1. Every strategy τ of player 2 that attains the infimum on the right-hand side of (1) up to ε is ε-*optimal* for player 2.

Our main result is the following

Theorem 1. *If A, B are compact metric spaces, and the functions $g_n(\omega,\cdot,\cdot)$ and $p_n(\omega,\cdot,\cdot)$ are continuous for each $\omega\in\Omega$ and every $n\in\mathbf{N}$, the game has a value.*

Our main contribution is not in the technical result, but in the new technique that we use for the proof.

We now briefly compare our result to the existing literature. Under some regularity assumptions on $(\Omega,\mathscr{F},\mathbf{P})$, the model we consider is a class of general stochastic games. Maitra and Sudderth (1993, 1998) proved the existence of the value in a fairly general setup of stochastic games.

Maitra and Sudderth (1993), using the operator approach and transfinite induction, proved that certain measurable stochastic games admit a value, and both players have universally measurable ε-optimal strategies, for every $\varepsilon>0$. Relative to this result, our contribution is that we give a constructive argument for the existence of ε-optimal strategies, which are also uniformly ε-optimal in the sense defined below (see Section 4).

Maitra and Sudderth (1998), using the fact that every Borel game is solvable, proved that finitely additive stochastic games admit a value, and both players have finitely-additive ε-optimal strategies for every $\varepsilon>0$ (see also Martin, 1998). Thus, relative to this result, our contribution is that both players have σ-additive uniformly ε-optimal strategies for every $\varepsilon>0$, rather than finitely additive ε-optimal strategies.

Rosenberg et al. (2001), using the technique of vanishing discount factors, proved that when A and B are finite, the value exists.

3. The proof

3.1. A stochastic variation of Ramsey's theorem

Ramsey (1930) proved that for every function c that attaches an element $c(k,l)\in C$, where C is a finite set, to every two non-negative integers $k<l$ there is an increasing sequence of integers $k_1<k_2<\cdots$ such that $c(k_1,k_2)=c(k_i,k_j)$ for every $i<j$.

We are going to attach an $\mathscr{F}_n$-measurable function $c_{n,\tau}$, whose range is some finite set C, to every non-negative integer n and every stopping time τ. We also impose a consistency requirement: if $\tau_1=\tau_2$ on an $\mathscr{F}_n$-measurable set F, then $c_{n,\tau_1}=c_{n,\tau_2}$ on F. Under these conditions, a weaker conclusion than that of Ramsey's theorem can be derived: for every $\varepsilon>0$ there exists an increasing sequence of stopping times $v_1<v_2<\cdots$ such that $\mathbf{P}(c_{v_1,v_2}=c_{v_2,v_3}=c_{v_3,v_4}=\cdots)>1-\varepsilon$.[3]

Let $(\Omega,\mathscr{F},\mathbf{P})$ be a probability space, and let $(\mathscr{F}_n)$ be a filtration. A stopping time v (to the filtration $(\mathscr{F}_n)_{n\in\mathbf{N}}$) is a function $v:\Omega\to\mathbf{N}$ such that the set $\{v=n\}$ is $\mathscr{F}_n$-measurable for every $n\in\mathbf{N}$. For every $A,B\in\mathscr{F}$, *A holds on B* if and only if $\mathbf{P}(A^c\cap B)=0$.

[3] For every two bounded stopping times $v<v'$, the $\mathscr{F}_v$-measurable function $c_{v,v'}$ is defined by $c_{v,v'}(\omega)=c_{v(\omega),v'}(\omega)$.

Definition 2. An *NT-function* is a function that assigns to every integer $n \geqslant 0$ and every bounded stopping time v an $\mathscr{F}_n$-measurable r.v. that is defined over the set $\{v > n\}$. We say that an NT-function f is *C-valued*, for some set C, if the r.v. $f_{n,v}$ is C-valued, for every $n \geqslant 0$ and every bounded stopping time v.

Definition 3. An NT-function f is *consistent* if for every $n \geqslant 0$, every $\mathscr{F}_n$-measurable set F, and every two bounded stopping times v_1, v_2, we have

$$v_1 = v_2 > n \text{ on } F \text{ implies } f_{n,v_1} = f_{n,v_2} \text{ on } F.$$

When f is an NT-function, and $v_1 < v_2$ are two bounded stopping times, we denote $f_{v_1,v_2}(\omega) = f_{v_1(\omega),v_2}(\omega)$. Thus f_{v_1,v_2} is an $\mathscr{F}_{v_1}$-measurable r.v.

The following theorem was proved by Shmaya and Solan (2002, Theorem 4.3).

Theorem 2. *For every finite set* C, *every* C*-valued consistent NT-function* f, *and every* $\varepsilon > 0$, *there exists a sequence of bounded stopping times* $1 \leqslant v_1 < v_2 < v_3 < \cdots$ *such that*

$$\mathbf{P}(f_{v_1,v_2} = f_{v_2,v_3} = f_{v_3,v_4} = \cdots) > 1 - \varepsilon.$$

3.2. Application to games

Let $\Gamma = (\Omega, \mathscr{F}, \mathbf{P}, A, B, (p_n, g_n))$ be a dynamic game.

For every two bounded stopping times $v_1 < v_2$, and every $\mathscr{F}_{v_2}$-measurable function h, let $\Gamma(v_1, v_2, h)$ be the two-player zero-sum game that starts at stage v_1 and, if not terminated earlier, terminates at stage v_2 with terminal payoff h. We do not introduce a new concept of a strategy in $\Gamma(v_1, v_2, h)$. Rather, we take the strategy space in $\Gamma(v_1, v_2, h)$ to coincide with that of Γ, and we use conditional expectation on the event $\{\theta \geqslant v_1\}$.

The following standard lemma states that $\Gamma(v_1, v_2, h)$ admits a value (see, e.g., Nowak, 1985, Theorem 5.2). It follows using backward induction from Sion's (1958) minimax theorem and a measurable selection theorem (e.g., Kuratowski and Ryll-Nardzewski, 1965).

Lemma 1. *Let* $v_1 < v_2$ *be bounded stopping times, and* h *an* $\mathscr{F}_{v_2}$*-measurable function such that* $\|h\|_\infty \leqslant 1$. *Under the assumptions of Theorem* 1, *there exists an* $\mathscr{F}_{v_1}$*-measurable function* $v(\Gamma(v_1, v_2, h))$, *and a pair* (σ^*, τ^*) *of strategies such that for every pair* (σ', τ') *of strategies*,

$$\mathbf{E}_{\sigma^*,\tau'}[\mathbf{1}_{\{v_1 \leqslant \theta < v_2\}} g_\theta(a_\theta, b_\theta) + \mathbf{1}_{\{v_2 \leqslant \theta\}} h \mid \mathscr{G}_{v_1}] \geqslant v(\Gamma(v_1, v_2, h))\mathbf{1}_{\{v_1 \leqslant \theta\}}$$

and

$$\mathbf{E}_{\sigma',\tau^*}[\mathbf{1}_{\{v_1 \leqslant \theta < v_2\}} g_\theta(a_\theta, b_\theta) + \mathbf{1}_{\{v_2 \leqslant \theta\}} h \mid \mathscr{G}_{v_1}] \leqslant v(\Gamma(v_1, v_2, h))\mathbf{1}_{\{v_1 \leqslant \theta\}}. \tag{2}$$

Actually, there are optimal strategies $\sigma^* = (\sigma_n^*)$ and $\tau^* = (\tau_n^*)$ such that σ_n^* and τ_n^* are $\mathscr{F}_n$-measurable, rather than $\mathscr{G}_n$-measurable (that is, the actions chosen at each stage

n do not depend on previously chosen actions). Moreover, one can verify that (2) still holds if we replace σ^* in Lemma 1 by any strategy σ such that, for every $n \in \mathbf{N}$, $\sigma_n = \sigma_n^*$ on $\{v_1 \leqslant n < v_2\}$.

The following lemma summarizes simple monotonicity and continuity properties of the value operator.

Lemma 2. (a) *If* $\|h\|_\infty \leqslant 1$ *then* $\|v(\Gamma(v_1, v_2, h))\|_\infty \leqslant 1$.

(b) *If* $v_1 < v_2 < v_3$, *then* $v(\Gamma(v_1, v_2, v(\Gamma(v_2, v_3, h)))) = v(\Gamma(v_1, v_3, h))$.

(c) *If* $F \in \mathscr{G}_{v_1}$ *and* $h \leqslant h'$ *on* F, *then* $v(\Gamma(v_1, v_2, h)) \leqslant v(\Gamma(v_1, v_2, h'))$ *on* F.

(d) *If* $F \in \mathscr{G}_{v_1}$ *and the sequence* (h_n) *converges pointwise to* h *on* F, *then*

$$\lim_{n\to\infty} v(\Gamma(v_1, v_2, h_n)) = v(\Gamma(v_1, v_2, h)) \quad \text{on } F.$$

Set $C = \{\text{‘+’}, \text{‘−’}\}$, and define a C-valued NT-function c as follows. For every $n \in \mathbf{N}$ and every stopping time v,

$$c(n, v) = \begin{cases} \text{‘+’} & \text{if } v(\Gamma(n, v, 0)) > 0, \\ \text{‘−’} & \text{if } v(\Gamma(n, v, 0)) \leqslant 0. \end{cases}$$

Lemma 2(c) implies that c is a consistent NT-function.

Finally, fix, once and for all $\varepsilon > 0$. By Theorem 2 there exists an increasing sequence (v_k) of stopping times such that

$$\mathbf{P}(c(v_1, v_2) = c(v_2, v_3) = c(v_3, v_4) = \cdots) > 1 - \varepsilon. \tag{3}$$

3.3. An auxiliary game

For every $k \in \mathbf{N}$ define

$$E_k^+ = \{c(v_1, v_2) = \text{‘+’ and } c(v_k, v_{k+1}) = \text{‘−’}\}$$

and

$$E_k^- = \{c(v_1, v_2) = \text{‘−’ and } c(v_k, v_{k+1}) = \text{‘+’}\}.$$

By (3), $\mathbf{P}(\bigcup_{k\in\mathbf{N}} (E_k^+ \cup E_k^-)) < \varepsilon$. Moreover, E_k^+ and E_k^- are in $\mathscr{F}_{v_k}$.

We now define an auxiliary game Γ'_ε, which is similar to Γ, except that it has a different payoff function $(g'_n)_{n\in\mathbf{N}}$, that is defined as follows:

$$g'_n(\omega) = \begin{cases} 0 & \omega \in E_k^+ \cup E_k^- \text{ for some } k \leqslant n, \\ g_n(\omega) & \text{otherwise.} \end{cases}$$

Thus, whenever $c(v_k, v_{k+1}) \neq c(v_1, v_2)$, we set the payoff to be 0 from stage v_k and onwards.

Denote by $\gamma'(\sigma, \tau)$ the expected payoff under the pair of strategies (σ, τ) in Γ'_ε. Since $\mathbf{P}(g_n \neq g'_n \text{ for some } n \in \mathbf{N}) \leqslant \mathbf{P}\left(\bigcup_{k\in\mathbf{N}} (E_k^+ \cup E_k^-)\right) < \varepsilon$, and since payoffs are bounded by 1, for every pair of strategies (σ, τ) one has $|\gamma(\sigma, \tau) - \gamma'(\sigma, \tau)| < \varepsilon$.

3.4. Sufficiency of the analysis of the auxiliary game

The following lemma asserts that if for every $\varepsilon > 0$ there are 3ε-optimal strategies in Γ'_ε, then the original game admits a value.

Lemma 3. *If for every $\varepsilon > 0$ there exist $V_\varepsilon \in [-1, 1]$ and a pair $(\sigma_\varepsilon, \tau_\varepsilon)$ of strategies that satisfy $\inf_{\tau'} \gamma'(\sigma_\varepsilon, \tau') \geqslant V_\varepsilon - 3\varepsilon$ and $\sup_{\sigma'} \gamma'(\sigma', \tau_\varepsilon) \leqslant V_\varepsilon + 3\varepsilon$, then $V := \lim_{\varepsilon\to 0} V_\varepsilon$ exists, and is the value of Γ.*

Observe that we do not require that V_ε is the value of Γ'_ε, or, for that matter, that the games $(\Gamma'_\varepsilon)_\varepsilon > 0$ have values.

Proof. Let V be any accumulation point of the sequence $(V_\varepsilon)_{\varepsilon>0}$ as ε goes to 0. Since $|\gamma(\sigma, \tau) - \gamma'(\sigma, \tau)| < \varepsilon$, the assumptions imply that $\inf_{\tau'} \gamma(\sigma_\varepsilon, \tau') \geqslant V_\varepsilon - 4\varepsilon$ and $\sup_{\sigma'} \gamma(\sigma', \tau_\varepsilon) \leqslant V_\varepsilon + 4\varepsilon$.

Therefore, for every δ there is $\varepsilon > 0$ sufficiently small such that $\inf_{\tau'} \gamma(\sigma_\varepsilon, \tau') \geqslant V_\varepsilon - 4\varepsilon \geqslant V - \delta$ and $\sup_{\sigma'} \gamma(\sigma', \tau_\varepsilon) \leqslant V_\varepsilon + 4\varepsilon \leqslant V + \delta$. In particular, V is the value of Γ. □

Thus, our goal is to find $V_\varepsilon \in [-1, 1]$ and to construct a pair (σ, τ) of strategies such that $\inf_{\tau'} \gamma'(\sigma, \tau') \geqslant V_\varepsilon - 3\varepsilon$ and $\sup_{\sigma'} \gamma'(\sigma', \tau) \leqslant V_\varepsilon + 3\varepsilon$.

In Section 3.5 we define V_ε, in Section 3.6 we define σ, and in Section 3.7 we prove that $\inf_{\tau'} \gamma'(\sigma, \tau') \geqslant V_\varepsilon - 3\varepsilon$. The construction of τ, and the proof that $\sup_{\sigma'} \gamma'(\sigma', \tau) \leqslant V_\varepsilon + 3\varepsilon$, is analogous to that of σ, and hence omitted.

3.5. Properties of the coloring

By construction of Γ'_ε, if $v(\Gamma'_\varepsilon(v_1, v_2, 0)) > 0$ then $v(\Gamma'_\varepsilon(v_k, v_{k+1}, 0)) \geqslant 0$ for every $k \in \mathbf{N}$, whereas if $v(\Gamma'_\varepsilon(v_1, v_2, 0)) \leqslant 0$ then $v(\Gamma'_\varepsilon(v_k, v_{k+1}, 0)) \leqslant 0$ for every $k \in \mathbf{N}$.

Let $D_+ = \{v(\Gamma'_\varepsilon(v_1, v_2, 0)) > 0\} \in \mathscr{F}_{v_1}$, and $D_- = \{v(\Gamma'_\varepsilon(v_1, v_2, 0)) \leqslant 0\} \in \mathscr{F}_{v_1}$. Plainly, (D_+, D_-) is a partition of Ω.

On D_+, $v(\Gamma'_\varepsilon(v_k, v_{k+1}, 0)) \geqslant 0$ for every $k \in \mathbf{N}$. By Lemma 2(b, c),

$$v(\Gamma'_\varepsilon(v_k, v_{l+1}, 0)) = v(\Gamma'_\varepsilon(v_k, v_l, v(\Gamma'_\varepsilon(v_l, v_{l+1}, 0)))) \geqslant v(\Gamma'_\varepsilon(v_k, v_l, 0)) \quad \text{on } D_+.$$

Similarly,

$$v(\Gamma'_\varepsilon(v_k, v_{l+1}, 0)) \leqslant v(\Gamma'_\varepsilon(v_k, v_l, 0)) \quad \text{on } D_-.$$

In particular, for every fixed $k \in \mathbf{N}$, the sequence $(v(\Gamma'_\varepsilon(v_k, v_l, 0)))_{l>k}$ is a sequence of $\mathscr{F}_{v_k}$-measurable functions, which is non-decreasing on D_+ and non-increasing on D_-. Therefore, this sequence has a limit h^*_k, which is $\mathscr{F}_{v_k}$-measurable.

Applying Lemma 2(b, d), we get

$$\begin{aligned} h^*_k &= \lim_{l\to\infty} v(\Gamma'_\varepsilon(v_k, v_l, 0)) \\ &= \lim_{l\to\infty} v(\Gamma'_\varepsilon(v_k, v_{k+1}, v(\Gamma'_\varepsilon(v_{k+1}, v_l, 0)))) \\ &= v(\Gamma'_\varepsilon(v_k, v_{k+1}, h^*_{k+1})). \end{aligned} \tag{4}$$

Set $V_\varepsilon = \mathbf{E}[v(\Gamma'_\varepsilon(1, v_1, h^*_1))]$.

3.6. Definition of a strategy σ

Choose $l \in \mathbf{N}$ sufficiently large such that

$$\mathbf{P}(D_+ \cap \{v(\Gamma'_\varepsilon(v_1, v_l, 0)) \geqslant h_1^* - \varepsilon\}) > \mathbf{P}(D_+) - \varepsilon. \tag{5}$$

For every $k \in \mathbf{N}$ choose an optimal strategy σ_k for player 1 in the game $\Gamma'_\varepsilon(v_k, v_{k+1}, 0)$, and an optimal strategy σ_k^* in the game $\Gamma'_\varepsilon(v_k, v_{k+1}, h_{k+1}^*)$. Choose an optimal strategy $\sigma_{1,l}$ for player 1 in the game $\Gamma'_\varepsilon(v_1, v_l, 0)$, and, finally, an optimal strategy σ_0 in the game $\Gamma'_\varepsilon(1, v_1, h_1^*)$.

Recall that D_+ and D_- are $\mathscr{F}_{v_1}$-measurable. Define a strategy σ for player 1 as follows:

- σ follows σ_0 up to stage v_1.
- If $\omega \in D_-$, σ follows σ_k^* between stages v_k and v_{k+1}, for every $k \in \mathbf{N}$.
- If $\omega \in D_+$, σ follows $\sigma_{1,l}$ between stages v_1 and v_l. Then, for every $k \geqslant l$, σ follows σ_k between stages v_k and v_{k+1}.

3.7. The strategy σ is 3ε-optimal

Let τ' be an arbitrary strategy of player 2. We prove that $\gamma'(\sigma, \tau') \geqslant V_\varepsilon - 3\varepsilon$. For convenience, set $r_\theta = g'_\theta(a_\theta, b_\theta)$ if $\theta < +\infty$ and $r_\theta = 0$ if $\theta = +\infty$. This is the terminal payoff in the game.

Since $\sigma_{1,l}$ is optimal in $\Gamma'_\varepsilon(v_1, v_l, 0)$,

$$\mathbf{E}_{\sigma,\tau'}[\mathbf{1}_{\{v_1 \leqslant \theta < v_l\}} r_\theta \mid \mathscr{G}_{v_1}] \geqslant v(\Gamma'_\varepsilon(v_1, v_l, 0))\mathbf{1}_{\{v_1 \leqslant \theta\}} \quad \text{on } D_+. \tag{6}$$

Since for every $k \in \mathbf{N}$, σ_k is optimal in $\Gamma'_\varepsilon(v_k, v_{k+1}, 0)$,

$$\mathbf{E}_{\sigma,\tau'}[\mathbf{1}_{\{v_k \leqslant \theta < v_{k+1}\}} r_\theta \mid \mathscr{G}_{v_k}] \geqslant v(\Gamma'_\varepsilon(v_k, v_{k+1}, 0))\mathbf{1}_{\{v_k \leqslant \theta\}} \geqslant 0 \quad \text{on } D_+.$$

Taking the conditional expectation w.r.t. $\mathscr{G}_{v_1}$, and summing over $k \geqslant l$, gives us

$$\mathbf{E}_{\sigma,\tau'}[\mathbf{1}_{\{v_l \leqslant \theta\}} r_\theta \mid \mathscr{G}_{v_1}] \geqslant 0 \quad \text{on } D_+. \tag{7}$$

From (6) and (7) we have

$$\mathbf{E}_{\sigma,\tau'}[\mathbf{1}_{\{v_1 \leqslant \theta\}} r_\theta \mid \mathscr{G}_{v_1}] \geqslant v(\Gamma'_\varepsilon(v_1, v_l, 0))\mathbf{1}_{\{v_1 \leqslant \theta\}} \quad \text{on } D_+.$$

By taking the expectation and using (5), we obtain

$$\begin{aligned}\mathbf{E}_{\sigma,\tau'}[\mathbf{1}_{D_+\cap\{v_1 \leqslant \theta\}} r_\theta] &\geqslant \mathbf{E}[\mathbf{1}_{D_+\cap\{v_1 \leqslant \theta\}} v(\Gamma'_\varepsilon(v_1, v_l, 0))] \\ &\geqslant \mathbf{E}[\mathbf{1}_{D_+\cap\{v_1 \leqslant \theta\}} h_1^*] - 3\varepsilon. \end{aligned} \tag{8}$$

Since for every $k \in \mathbf{N}$, σ_k^* is optimal in $\Gamma'_\varepsilon(v_k, v_{k+1}, h_{k+1}^*)$, and by the recursive relation (4),

$$\begin{aligned}&\mathbf{E}_{\sigma,\tau'}[\mathbf{1}_{\{v_k \leqslant \theta < v_{k+1}\}} r_\theta + \mathbf{1}_{\{v_{k+1} \leqslant \theta\}} h_{k+1}^* \mid \mathscr{G}_{v_k}] \\ &\quad \geqslant \mathbf{1}_{\{v_k \leqslant \theta\}} v(\Gamma'_\varepsilon(v_k, v_{k+1}, h_{k+1}^*)) = \mathbf{1}_{\{v_k \leqslant \theta\}} h_k^* \quad \text{on } D_-. \end{aligned} \tag{9}$$

Then taking the conditional expectation w.r.t. $\mathcal{G}_{v_{k-1}}$ in (9), and adding it to (9) where k has been replaced by $k-1$, we get

$$\begin{aligned}
&\mathbf{E}_{\sigma,\tau'}[\mathbf{1}_{\{v_{k-1}\leqslant\theta<v_{k+1}\}}r_\theta+\mathbf{1}_{\{v_{k+1}\leqslant\theta\}}h^*_{k+1}\mid\mathcal{G}_{v_{k-1}}]\\
&\quad\geqslant\mathbf{E}_{\sigma,\tau'}[\mathbf{1}_{\{v_{k-1}\leqslant\theta<v_{k+1}\}}r_\theta+\mathbf{1}_{\{v_{k+1}\leqslant\theta\}}h^*_{k}\mid\mathcal{G}_{v_{k-1}}]\\
&\quad\geqslant\mathbf{1}_{\{v_{k-1}\leqslant\theta\}}h^*_{k-1}\quad\text{ on } D_-.
\end{aligned}$$

Continuing inductively one obtains for every $k\in\mathbf{N}$

$$\mathbf{E}[\mathbf{1}_{\{v_1\leqslant\theta<v_k\}}r_\theta+\mathbf{1}_{\{v_k\leqslant\theta\}}h^*_k\mid\mathcal{G}_{v_1}]\geqslant\mathbf{1}_{\{v_1\leqslant\theta\}}h^*_1\quad\text{ on } D_-.$$

Since $h^*_k\leqslant 0$ for every k on D_-, it follows by taking the expectation that

$$\begin{aligned}
\mathbf{E}_{\sigma,\tau'}[\mathbf{1}_{D_-\cap\{v_1\leqslant\theta<v_k\}}r_\theta]&\geqslant\mathbf{E}_{\sigma,\tau'}[\mathbf{1}_{D_-\cap\{v_1\leqslant\theta<v_k\}}r_\theta+\mathbf{1}_{D_-\cap\{\theta\geqslant v_k\}}h^*_k]\\
&\geqslant\mathbf{E}[\mathbf{1}_{D_-\cap\{v_1\leqslant\theta\}}h^*_1].
\end{aligned}$$

By the bounded convergence theorem, we deduce that

$$\mathbf{E}_{\sigma,\tau'}[\mathbf{1}_{D_-\cap\{v_1\leqslant\theta\}}r_\theta]\geqslant\mathbf{E}[\mathbf{1}_{D_-\cap\{v_1\leqslant\theta\}}h^*_1]. \tag{10}$$

By (8) and (10),

$$\mathbf{E}_{\sigma,\tau'}[\mathbf{1}_{\{v_1\leqslant\theta\}}r_\theta]\geqslant\mathbf{E}[\mathbf{1}_{\{v_1\leqslant\theta\}}h^*_1]-3\varepsilon. \tag{11}$$

Since σ_0 is optimal in the game $\Gamma'_\varepsilon(1,v_1,h^*_1)$,

$$\mathbf{E}_{\sigma,\tau'}[\mathbf{1}_{\{\theta<v_1\}}r_\theta+\mathbf{1}_{\{v_1\leqslant\theta\}}h^*_1]\geqslant\mathbf{E}[v(\Gamma'_\varepsilon(1,v_1,h^*_1))]=V_\varepsilon. \tag{12}$$

By (11) and (12),

$$\gamma'(\sigma,\tau')=\mathbf{E}_{\sigma,\tau'}[r_\theta]\geqslant V_\varepsilon-3\varepsilon. \tag{13}$$

4. Further discussion

Here we discuss the assumptions that our proof hinges on, as well as further topics.

Our argument relies on the assumption that the evolution of the game is independent of the actions chosen by the players. That is, the players have no control over the information they receive during the game, but only over the probability of termination and the terminal payoff. It is most desirable to extend our technique to the case where players do influence their information.

Another aspect that we critically need is that the information be symmetric: both players should have the same information at every stage. If this is not the case, then the value need not exist (see Laraki, 2000, for an example.) It is interesting to know under which informational structure the value still exists.

The strategy σ that we constructed in Section 3.6 is uniform in the following sense. There is $N\geqslant 0$ such that for every $n\geqslant N$

$$\inf_\tau\mathbf{E}_{\sigma,\tau}[\mathbf{1}_{\{\theta\leqslant n\}}g_\theta(a_\theta,b_\theta)]\geqslant V_\varepsilon-5\varepsilon.$$

That is, the strategy σ is $5\varepsilon+|V-V_\varepsilon|$-optimal in every finite-stage interaction, provided the interaction is sufficiently long.

The proof relies on the observation that if for some bounded stopping time v, where $v \geqslant v_k$ for k sufficiently large, the expected payoff under (σ, τ) up to stage v is significantly different from V_ε, then the probability of termination between stages v_k and v must be bounded away from 0. Therefore, such an event can occur only finitely many times. Details are standard and omitted.

We assumed that the functions $p_n(\omega, \cdot, \cdot)$ and $g_n(\omega, \cdot, \cdot)$ are continuous for every $\omega \in \Omega$. However, all we need is that for every $\mathscr{F}_{n+1}$-measurable function f, the one-stage game $\Gamma(n, n+1, f)$ with terminal payoff f admits a value. More formally, we now present a more general version of the one-shot game.

Definition 4. Let $(\Omega, \mathscr{F}, \mathbf{P})$ be a probability space, X and Y be two measurable sets of strategies, and $u : \Omega \times X \times Y \to [-1, 1]$ be a measurable payoff function. The game $(\Omega, \mathscr{F}, \mathbf{P}, X, Y, u)$ admits a *value* if there exists an $\mathscr{F}$-measurable function $v : \Omega \to [-1, 1]$, and, for every $\varepsilon > 0$, there exist $\mathscr{F}$-measurable functions $x : \Omega \to X$ and $y : \Omega \to Y$, such that

$$\sup_{x' \in X} u(\omega, x', y(\omega)) - \varepsilon \leqslant v(\omega) \leqslant \inf_{y' \in Y} u(\omega, x(\omega), y') + \varepsilon, \qquad \mathbf{P}\text{-a.e.}$$

The proof of the following extension of Theorem 1 follows the same lines as the proof we presented.

Theorem 3. *Let $\Gamma = (\Omega, \mathscr{F}, \mathbf{P}, (\mathscr{F}_n), A, B, (g_n, p_n))$ be an infinite-stage dynamic game. Assume that for every $n \in \mathbf{N}$ and every $\mathscr{F}_n$-measurable function $h : \Omega \to [-1, 1]$, the one-shot game $(\Omega, \mathscr{F}_n, \mathbf{P}, \mathscr{P}(A), \mathscr{P}(B), u)$ admits a value, where $u(\omega, x, y) = \int_A \int_B p_n(a, b) g_n(a, b) + (1 - p_n(a, b)) h \, \mathrm{d}x(a) \, \mathrm{d}y(b)$. Then the game Γ admits a value.*

Theorem 2 can be applied to prove the existence of an equilibrium in two-player non-zero-sum stopping games in discrete time (see Shmaya and Solan, 2002). However, whereas for non-zero-sum finite-stage games a lot of structure is needed to ensure the existence of an equilibrium that satisfies certain desirable properties, for zero-sum games the technique works in a much more general setup.

References

Aumann, R.J., Maschler, M.B., 1995. Repeated Games with Incomplete Information. MIT Press, Cambridge, MA.

Dynkin, E.B., 1969. Game variant of a problem on optimal stopping. Soviet Math. Dokl. 10, 270–274.

Kuratowski, K., Ryll-Nardzewski, C., 1965. A general theorem on selectors. Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys. 13, 397–403.

Laraki, R., 2000. Jeux répétés à information incomplète: approche variationnelle. Thèse de Doctorat de l'Univérsité Paris 6, France.

Maitra, A., Sudderth, W., 1993. Borel stochastic games with limsup payoff. Ann. Probab. 21, 861–885.

Maitra, A., Sudderth, W., 1998. Finitely additive stochastic games with borel measurable payoffs. Int. J. Game Theory 27, 257–267.

Martin, D.A., 1998. The determinacy of blackwell games. J. Symbolic Logic 63, 1565–1581.

Mertens, J.F., Neyman, A., 1981. Stochastic Games. Int. J. Game Theory 10, 53–66.

Neyman, A., Sorin, S., 2001. Zero-sum two-person repeated games with public uncertain duration process. Discussion Paper 259, Center for Rationality, The Hebrew University of Jerusalem.

Nowak, A.S., 1985. Universally measurable strategies in zero-sum stochastic games. Ann. Probab. 13, 269–287.

Ramsey, F.P., 1930. On a problem of formal logic. Proc. London Math. Soc. 30, 264–286.

Rosenberg, D., Solan, E., Vieille, N., 2001. Stopping games with randomized strategies. Probab. Theory Related Fields 119, 433–451.

Shmaya, E., Solan, E., 2002. Two-player non-zero-sum stopping games in discrete time. Ann. Probab., to appear.

Sion, M., 1958. On general minimax theorems. Pacific J. Math. 8, 171–176.

Touzi, N., Vieille, N., 2002. Continuous-time Dynkin games with mixed strategies. SIAM J. Control Optim. 41, 1073–1088.