Int J Game Theory (2000) 29:309–325

International Journal of Game Theory

Two-person repeated games with finite automata

Abraham Neyman[1], **Daijiro Okada**[2]

[1] Institute of Mathematics, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, ISRAEL and SUNY at Stony Brook, Stony Brook, NY 11794-4384, USA
(email: aneyman@math.huji.ac.il)
[2] Department of Economics, SUNY at Stony Brook, Stony Brook, NY 11794-4384, USA
(email: dokada@ms.cc.sunysb.edu)

Received February 1997/revised version March 2000

Abstract. We study two-person repeated games in which a player with a restricted set of strategies plays against an unrestricted player. An exogenously given bound on the complexity of strategies, which is measured by the size of the smallest automata that implement them, gives rise to a restriction on strategies available to a player.

We examine the asymptotic behavior of the set of equilibrium payoffs as the bound on the strategic complexity of the restricted player tends to infinity, but sufficiently slowly. Results from the study of zero sum case provide the individually rational payoff levels.

JEL classification: C73, C72.

Key words: repeated games, finite automata

---

1. Introduction

The objects of this study is two-person non-zero sum repeated games in which there is a bound on the complexity of strategies for only one of the players. Throughout the paper player 1 will be the restricted player. We employ automata to represent repeated game strategies. The complexity of a strategy is then defined to be the smallest number of states of an automaton required to implement it.

Specific models of repeated games studied are (1) **The finitely repeated game** $G^n(m(n))$ and (2) **The λ-discounted game** $G_\lambda(m(\lambda))$. Here, $m(\cdot)$ is the bound on the number of states of automata available to player 1 and this is a function of the number of repetitions, n, or the discount factor, λ. We examine the set of limit points of equilibrium payoffs of $G^n(m(n))$ (resp. $G_\lambda(m(\lambda))$) as

$m(n) \to \infty$ $(n \to \infty)$ (resp. $m(\lambda) \to \infty$ $(\lambda \to 1)$). The particular case considered in this paper is when $m(\cdot)$ grows "sufficiently slowly"[1] Formally, we will examine the cases when

$$\lim_{n\to\infty} \frac{m(n) \log m(n)}{n} = 0, \tag{1.1}$$

and

$$\lim_{\lambda\to 1}(1-\lambda)m(\lambda)\log m(\lambda) = 0. \tag{1.2}$$

We will show that, under these conditions on $m(\cdot)$, the Hausdorff limit of the set of equilibrium payoffs, $\mathrm{Lim}_{n\to\infty}\ E(m(n))$ and $\mathrm{Lim}_{\lambda\to 1}\ E(m(\lambda))$, exist and they coincide with the set of the feasible payoffs above certain individually rational levels. For player 1 this level will be his *maxmin* payoff in the one-shot game where *max* ranges over his pure actions and *min* ranges over player 2's pure actions, and for player 2 it will be her *minmax* of the one-shot game where *min* ranges over player 1's pure actions and *max* ranges over her own pure actions. The determination of these individually rational levels, under the conditions (1.1) and (1.2), can be provided by our analysis of the zero sum case, Neyman and Okada (1999). In this paper, however, we will explicitly construct player 2's strategy that effectively punishes the restricted player 1. This will provide a simpler proof of a result originally proved in Neyman and Okada (1999) using the concept of entropy. See also Neyman and Okada (2000) for an alternative proof.

It will be seen that the equilibria that we construct are in pure strategies and the equilibrium paths are cyclic. Our result for finitely repeated games implies in particular that, in finitely repeated prisoner's dilemma, the friendly, or nearly friendly, outcomes can be achieved in an equilibrium when there is a bound on the strategic complexity on only one of the players. In addition, Folk-Theorem type results like ours have an implication that in non-zero sum games, being restricted in terms of strategic possibility is not necessarily detrimental even against a powerful unrestricted player. Only the punishment will be severe for the restricted player.

Related literature includes Neyman (1999) and Papadimitriou and Yannakakis (1994) which contain several results on the asymptotic behavior of the set of equilibrium payoffs of two-person finitely repeated games when there are bounds on the strategic complexity for both players. These results encompass Neyman (1985)'s justification of cooperation in finitely repeated prisoners' dilemma. For example, the main theorem of Neyman (1999) states that if the two bounds on the size of automata are subexponential as a function of min{*the number of repetitions*, *the larger bound*}, then the asymptotic folk theorem is obtained. More precisely, let G be a two-person game in strategic form. Denote by v^i the *minmax* payoff for player i where *min* ranges over the other player's mixed actions and *max* ranges over i's own pure ac-

[1] For example, the condition (1.1) holds for all functions $m(n) = n^a$ where $0 < a < 1$, while it is violated for $m(n) = n$. In addition, the function $m(n) = n/\log n$, for which $m(n)/n \to 0$ $(n \to \infty)$ but $m(n)/n^a \to \infty$ $(n \to \infty)$ for all $0 < a < 1$, violates (1.1). The authors thank a referee for this comment.

tions. Let $G^n(m_1(n), m_2(n))$ be the n-fold repetitions of G in which player i's strategies are restricted to those implementable by automata of size at most $m_i(n)$, a function of n. If the sequence of triples $(n, m_1(n), m_2(n))_{n=1}^{\infty}$ satisfies the conditions $\min\{m_1(n), m_2(n)\} \to \infty$ $(n \to \infty)$ and

$$\lim_{n\to\infty} \frac{\log(\max\{m_1(n), m_2(n)\})}{\min\{n, m_1(n), m_2(n)\}} = 0,$$

then the set of equilibrium payoff vectors of $G^n(m_1(n), m_2(n))$ converges to the set of payoff vectors which are feasible and give player i at least v^i. Zemel (1989) contains results similar to Neyman (1985) but using modified finite automata which can send messages, in addition to those conveyed through the actions taken, during the play.

Similar results have been obtained for other classes of repeated games. Ben-Porath (1993) studies the undiscounted infinitely repeated games with finite automata. Let $G = ((A_i)_{i=1}^N, (r_i)_{i=1}^N)$ be an N-person game in strategic form, and

$$v^i = \min_{q \in \times_{j \neq i} \Delta(A^j)} \max_{a^i \in A^i} E_q(r(a^i, b))$$

and

$$w^i = \max_{p \in \Delta(A^i)} \min_{b \in \times_{j \neq i} A^i} E_p(r(a^i, b))$$

where $\Delta(X)$ denotes the set of probability distributions on a set X and E_μ denotes the expectation with respect to the probability μ. Consider the infinitely repeated game in which player i has a complexity bound $m_i(\kappa)$, parameterized by positive integer κ, with $m_1(\kappa) \leq \cdots \leq m_N(\kappa)$ and $m_1(\kappa) \to \infty(\kappa \to \infty)$. Denote this game by $G^\infty(m_1(\kappa), \ldots, m_N(\kappa))$ and the set of its equilibrium payoffs by $E^\infty(m_1(\kappa), \ldots, m_N(\kappa))$. One of his results asserts that if

$$\lim_{\kappa\to\infty} \frac{\log m_N(\kappa)}{m_1(\kappa)} = 0,$$

then,

(i) the set of feasible payoffs which give each player i at least v^i is included in $\liminf_{\kappa\to\infty} E^\infty(m_1(\kappa), \ldots, m_N(\kappa))$ and
(ii) $\limsup_{\kappa\to\infty} E^\infty(m_1(\kappa), \ldots, m_N(\kappa))$ is included in the set of feasible payoffs which give each player i at least w^i.

Note that $v^i \geq w^i$. For two-person games, we have $v^i = w^i$. Hence one can conclude that $\lim_{\kappa\to\infty} E^\infty(m_1(\kappa), m_2(\kappa))$ exists and it coincides with the set of equilibrium payoffs of the infinitely repeated game without complexity bound (the Folk Theorem). This result crucially depends on the study of the two-person zero sum case which provides the individually rational levels v^i and w^i. The exact asymptotics, i.e., the limit, if exists, of $E^\infty(m_1(\kappa), \ldots, m_N(\kappa))$ for N-person case is not known. See Section 4 of Neyman (1997). Lehrer (1988) contains a similar result for two-person games with bounded recall. Also see Lehrer (1994) for N-person case with bounded recall.

The main contribution of this paper is a determination of individually rational payoff levels together with the construction of equilibria which have certain robustness properties and can be applied to a wider variety of conditions on the order of magnitude of the complexity bound $m(\cdot)$.

The next section introduces the model of repeated games and finite automata. In Section 3 we will construct player 2's strategy which will be used to punish player 1 in equilibria constructed in the subsequent chapters. The results on the asymptotics of the set of equilibrium payoff vectors are presented in Section 4 (the finitely repeated games) and Section 5 (the discounted games). Section 6 concludes the paper.

2. Repeated games and automata

Let $G = (A, B, h, k)$ be a two-person game in strategic form where A and B are finite sets of actions, and, $h : A \times B \to \mathbb{R}$ and $k : A \times B \to \mathbb{R}$ are the payoff functions of player 1 and 2, respectively. We call G the stage game. Throughout the paper we will assume without loss of generality that all payoffs are nonnegative, i.e., $h(a, b) \geq 0$ and $k(a, b) \geq 0$ for all $(a, b) \in A \times B$. Denote the maxmin value of the stage game for player 1 in pure actions by h_* and the minimax value for player 2 in pure actions by k^*, i.e.,

$$h_* = \max_{a \in A} \min_{b \in B} h(a, b) \quad \text{and} \quad k^* = \min_{a \in A} \max_{b \in B} k(a, b).$$

Also set $\|h\| = \max_{a,b} |h(a, b)|$ and $\|k\| = \max_{a,b} |k(a, b)|$.

Given $G = (A, B, h, k)$ we next describe a new game in which G is played repeatedly (with complete information and standard signaling).

For each positive integer n, let S_n (resp. T_n) be the set of mappings from $(A \times B)^{n-1}$ to A (resp. to B) where $(A \times B)^0 = \{\phi\}$. A pure strategy of player 1 (resp. player 2) is an element of $S = \times_n S_n$ (resp. $T = \times_n T_n$). Equivalently, S (resp. T) is the set of all mappings on the set of all finite histories $\bigcup_{n \geq 1} (A \times B)^{n-1}$ to A (resp. B). A mixed strategy of player 1 (resp. player 2) is a probability distribution on S (resp. T). The sets of mixed strategies are denoted by $\Delta(S)$ and $\Delta(T)$.

Every pair of pure strategies (s, t) induces a play $\omega(s, t) = (\omega_\ell(s, t))_{\ell=1}^\infty \in (A \times B)^\infty$ where $\omega_\ell(s, t)$ is defined inductively as

$$\omega_\ell(s, t) = (a_\ell, b_\ell) = \begin{cases} (s_1(\phi), t_1(\phi)) & \text{for } \ell = 1 \\ (s_\ell(\omega_1, \ldots, \omega_{\ell-1}), t_\ell(\omega_1, \ldots, \omega_{\ell-1})) & \text{for } \ell > 1 \end{cases}$$

Accordingly, every pair (σ, τ) of mixed strategies induces a random play $\omega(\sigma, \tau) = (\omega_\ell(\sigma, \tau))_{\ell=1}^\infty$. We denote the corresponding probability distribution on the set of plays $(A \times B)^\infty$ by $P_{\sigma,\tau}$ and the expectation with respect to $P_{\sigma,\tau}$ by $E_{\sigma,\tau}$.

For each positive integer n we define the n-average payoff function of player 1, $h_n : S \times T \to \mathbb{R}$, by $h_n(s, t) = (1/n) \sum_{\ell=1}^n h(\omega_\ell(s, t))$. Also, for each $\lambda \in [0, 1)$ we define the λ-discounted payoff function of player 1, $h_\lambda : S \times T \to \mathbb{R}$ by $h_\lambda(s, t) = (1 - \lambda) \sum_{\ell=1}^\infty \lambda^{\ell-1} h(\omega_\ell(s, t))$. The n-average and the λ-discounted payoff functions of player 2, k_n and k_λ, are similarly defined. The bilinear extensions of h_n, h_λ, k_n and k_λ to $\Delta(S) \times \Delta(T)$ are denoted by the same symbols. Thus, for example, $h_n(\sigma, \tau) = E_{\sigma,\tau}[(1/n)\sum_{\ell=1}^{n} h(a_\ell, b_\ell)]$.

In this paper we study two classes of repeated games differentiated by their payoff functions.

Finitely Repeated Game $G^n = (S, T, h_n, k_n)$

The λ-Discounted Game $G_\lambda = (S, T, h_\lambda, k_\lambda)$

If two pure strategies of a player induce the same play against any pure strategy of the other player, they are said to be *equivalent*. For example, player 1's pure strategies s and s' are equivalent if $\omega_\ell(s, t) = \omega_\ell(s', t)$ for all pure strategy t of player 2 and all stages $\ell = 1, 2, \ldots$. Extending this notion to mixed strategies, we say that two strategies of a player are equivalent if, against any strategy of the other player, they induce the same probability over the plays of a repeated game.

Given the stage game $G = (A, B, h, k)$, an automaton of player 1 is defined by a four-tuple $M = \langle Q, q_1, f, g \rangle$. The first component Q is a set of *states*, and $q_1 \in Q$ is an *initial state*. The third component is an *action function*, $f : Q \to A$, and the last component is a *transition function*, $g : Q \times B \to Q$. By the *size* of an automaton we mean the cardinality of the set of its states, $|Q|$.

An automaton M plays a repeated game as follows. At each stage n it takes an action prescribed by f for the current state, say q_n, i.e., $f(q_n)$; it is set for q_1 at the first stage. Then it changes its state to q_{n+1} specified by g as a function of the current state q_n and player 2's action b_n, that is, $q_{n+1} = g(q_n, b_n)$.

Every automaton M induces a pure strategy s for player 1 in a repeated game in the following manner. First, for any sequence of player 2's actions $b_1, \ldots, b_n$ $(n \geq 2)$, define an extension of the transition function inductively by

$$g(q, b_1, \ldots, b_n) = g(g(q, b_1, \ldots, b_{n-1}), b_n).$$

Then for any history $\omega = ((a_1, b_1), \ldots, (a_n, b_n)) \in (A \times B)^n$, set

$$s(\omega) = f(g(q, b_1, \ldots, b_n))$$

which is an action taken at stage $n + 1$ $(n \geq 1)$. At the first stage, $s(\phi) = f(q_1)$.

Also, for every pure strategy $s \in S$ in a repeated game, there is an automaton that induces an equivalent strategy. If s is equivalent to a pure strategy induced by an automaton, we say that s is *implementable* by that automaton.

The size of the smallest automaton that implements a pure strategy serves as a measure of complexity of that strategy. To be more precise, for a given $s = (s_n) \in S$, we say that a finite history $\omega = ((a_1, b_1), \ldots, (a_\ell, b_\ell))$ is compatible with s if $a_n = s_n((a_1, b_1), \ldots, (a_{n-1}, b_{n-1}))$ for every $n = 1, \ldots, \ell$. Also, for an arbitrary finite history ω of length ℓ, define the induced strategy $s|\omega = ((s|\omega)_n)$ by $(s|\omega)_n(\omega') = s_{\ell+n}(\omega\omega')$ for each $\omega' \in (A \times B)^{n-1}$ where $\omega\omega'$ is the concatenation of ω and ω'. The number of distinct, or nonequivalent, strategies induced by s and finite histories compatible with s can be considered as a measure of complexity of implementing s. Indeed, it can be shown that *the size of the smallest automaton that implements s equals the number of the equivalence classes of* $\{(s|\omega)|\omega$ *is a finite history compatible with* $s\}$. Kalai and Stanford (1988) provides an analogous result for the full automata whose transition depends on the player's own action as well as the actions of the other players.

Henceforth, by the complexity of a pure strategy s, we mean the size of the smallest automaton that implements s. For each positive integer m, we denote by $S(m)$ the subset of S consisting of those pure strategies of player 1 whose complexity is at most m.

3. Individually rational payoff levels

We will present in this section a result which will be utilized in deriving much of the subsequent results. The situation under consideration is the one in which player 1 is restricted to a finite set of pure strategies. The nature of this set is arbitrary. In particular, it may contain pure strategies which cannot be implemented by any finite automata.

Theorem 3.1. *For every finite subset* S' *of* S *there exists* $\hat{t} \in T$ *such that for all* $s \in S'$

$$\textbf{(i)}\quad h_n(s,\hat{t}) \leq h_* + \frac{\|h\| \log_2|S'|}{n} \quad \textit{for all } n = 1, 2, \ldots,$$

and

$$\textbf{(ii)}\quad h_\lambda(s,\hat{t}) \leq h_* + (1-\lambda)\|h\| \log_2|S'| \quad \textit{for all } \lambda \in [0,1).$$

Proof: For each finite history $\omega = (\omega_1, \ldots, \omega_\ell)$, where $\omega_j = (a_j, b_j)$, let $S'(\omega)$ be the set of strategies in S' that are compatible with ω, i.e.,

$$S'(\omega) = \{s \in S' \mid s(\emptyset) = a_1, \text{ and } s(\omega_1, \ldots, \omega_{j-1}) = a_j \text{ for all } j = 2, \ldots, \ell\}.$$

For each $a \in A$ let $S'(\omega, a)$ be the set of strategies in $S'(\omega)$ that takes the action a at the history ω, i.e.,

$$S'(\omega, a) = \{s \in S'(\omega) \mid s(\omega) = a\}.$$

Clearly, if $a \neq a'$, then $S'(\omega, a)$ and $S'(\omega, a')$ are disjoint, and $\bigcup_{a \in A} S'(\omega, a) = S'(\omega)$.

Let $a(\omega)$ be an action of player 1 such that $|S'(\omega, a(\omega))| \geq |S'(\omega, a)|$ for all $a \in A$. Notice that if $a \neq a(\omega)$, then $|S'(\omega, a)|$ is at most one half of $|S'(\omega)|$: otherwise, $|S'(\omega, a(\omega))| + |S'(\omega, a)| \geq 2|S'(\omega, a)| > |S'(\omega)|$, a contradiction. This implies that for every $(a, b) \in A \times B$ with $a \neq a(\omega)$, if $\omega' = (\omega_1, \ldots, \omega_\ell, (a, b))$, then

$$|S'(\omega')| \leq \frac{|S'(\omega)|}{2}.$$

Define $\hat{t} \in T$ by $\hat{t}(\omega) \in \operatorname{argmin}_{b \in B} h(a(\omega), b)$. Take $s \in S'$ arbitrarily and let $(\omega_1, \omega_2, \ldots)$ be the play generated by $(s, \hat{t})$, $\omega_j = (a_j, b_j)$. Denote $\omega^\ell =$

$(\omega_1, \ldots, \omega_\ell)$. Of course, s is compatible with ω^ℓ for every ℓ, i.e., $s \in S(\omega^\ell)$. Therefore for all n,

$$|S'| 2^{-\sum_{\ell=1}^{n} I(a_\ell \neq a(\omega^{\ell-1}))} \geq |S(\omega^n)| \geq 1$$

where I is the indicator function. This implies that $\sum_{\ell=1}^{n} I(a_\ell \neq a(\omega^{\ell-1})) \leq \log_2 |S'|$, that is, the number of stages at which player 1's action differs from $a(\omega^\ell)$ is at most $\log_2 |S'|$.

Now let $\theta = (\theta_1, \theta_2, \ldots)$ be a nonincreasing sequence of nonnegative real numbers such that $\sum_{\ell=1}^{\infty} \theta_\ell = 1$. Define $h_\theta : S \times T \to \mathbb{R}$ by $h_\theta(s, t) = \sum_{\ell=1}^{\infty} \cdot \theta_\ell h(\omega_\ell(s, t))$. Take $s \in S'$. Then, since

$$h(\omega_\ell) \leq h_* I(a_\ell = a(\omega^{\ell-1})) + \|h\| I(a_\ell \neq a(\omega^{\ell-1}))$$

for every $\ell = 1, 2, \ldots$, we have

$$\begin{aligned} h_\theta(s, \hat{t}) &\leq \sum_{\ell=1}^{\infty} \theta_\ell (h_* I(a_\ell = a(\omega^{\ell-1})) + \|h\| I(a_\ell \neq a(\omega^{\ell-1}))) \\ &\leq h_* + \|h\| \sum_{\ell=1}^{\infty} \theta_\ell I(a_\ell \neq a(\omega^{\ell-1})) \\ &\leq h_* + \|h\| \sum_{\ell=1}^{\infty} \theta_1 I(a_\ell \neq a(\omega^{\ell-1})) \\ &\leq h_* + \theta_1 \|h\| \log_2 |S'|. \end{aligned}$$

(Recall our assumption $h(a, b) \geq 0$.) Note that if $\theta_\ell = 1/n$ for $\ell = 1, \ldots, n$ and $\theta_\ell = 0$ for $\ell > n$, then $h_\theta = h_n$. Hence **(i)**. Also, if $\theta_\ell = (1 - \lambda)\lambda^{\ell-1}$, $\ell = 1, 2, \ldots$, then $h_\theta = h_\lambda$. This proves **(ii)**. Q.E.D.

Remark 3.1: If s and s' in S' are equivalent, then for every finite history ω and every action $a \in A$, either both s and s' are in $S'(\omega, a)$ or neither is in $S'(\omega, a)$. Therefore one can replace $\log|S'|$ in the statement of Theorem 3.1 by $\log|S'/\sim|$ where $S'/\sim$ is the set of equivalence classes of S'.

Remark 3.2: Let $S_1, S_2, \ldots$ be a nondecreasing sequence of finite subsets of S. If $\log|S_n|/n \to 0$ as $n \to \infty$, then Theorem 3.1 **(i)** implies that for every $\varepsilon > 0$, there is n_0 such that for each $n \geq n_0$ there is $t \in T$ for which $\max_{s \in S_n} h_n(s, t) \leq h_* + \varepsilon$. Similar result is obtained from Theorem 3.1 **(ii)** for λ-discounted payoff by replacing the sequence $S_1, S_2, \ldots$ by a net $(S_\lambda | 0 \leq \lambda < 1)$ and the condition $\log|S_n|/n \to 0$ as $n \to \infty$ by $(1 - \lambda) \log|S_\lambda| \to 0$ as $\lambda \to 1$.

4. Finitely repeated game $G^n(m(n))$

In this section we study the modified version of the finitely repeated game, $G^n(m(n)) = (S(m(n)), T, h_n, k_n)$. The bound on the complexity of player 1's strategy, $m(\cdot)$, is a function of the number of repetitions n. Player 1 is allowed to use a mixed strategy provided that its support lies in $S(m(n))$. He can also use a behavioral strategy $s : \bigcup_{\ell \geq 0}(A \times B)^{\ell} \to \Delta(A)$ as long as it is equivalent to a mixed strategy with the support in $S(m(n))$.

A simple counting shows that the number of finite automata of size m is at most m^{Cm} for some positive constant C. Thus the number of equivalence classes of $S(m)$ is also bounded by m^{Cm}. The next lemma follows from Theorem 3.1 **(i)**, Remark 3.1 and Remark 3.2.

Lemma 4.1. *Suppose that* $m(n)\log m(n)/n \to 0$ *as* $n \to \infty$. *Then for every* $\varepsilon > 0$, *there is* n_0 *such that for each* $n > n_0$ *there is* $t \in T$ *such that*

$$h_n(s,t) \leq h_* + \varepsilon \quad \textit{for all } s \in S(m(n)).$$

Remark 4.1: As an immediate corollary of this lemma, we obtain the following result concerning the asymptotics of the value of two-person zero sum repeated games with finite automata which was proved in Neyman and Okada (1998a).[2] Consider $G = (A, B, h, k)$ to be a two-person zero sum game, i.e., $k = -h$. Denote the value of $G^n(m(n))$ by $V^n(m(n))$.

Corollary 4.1. *If* $m(n)\log m(n)/n \to 0$ *as* $n \to \infty$, *then* $V^n(m(n)) \to h_*$ *as* $n \to \infty$.

Denote by $E^n(m(n))$ the set of (Nash) equilibrium payoff vectors of $G^n(m(n))$. The next theorem provides an asymptotics of $E^n(m(n))$ when $m(n)$ grows sufficiently slowly. The convergence of sets is with respect to the Hausdorff topology in $\mathbb{R}^2$. To state the theorem formally we need some more notation. Let F be the convex hull of the set of payoffs feasible in pure actions of the stage game, that is, $F = \mathrm{Co}\{(h(a,b), k(a,b)) \mid (a,b) \in A \times B\}$, and let

$$\tilde{F} = \{(x,y) \in F \mid x \geq h_*, y \geq k^*\}.$$

The set $\tilde{F}$ is nonempty. For example, let $a^* \in \mathrm{argmax}_{a \in A}[\min_{b \in B} h(a,b)]$ and $b^* \in \mathrm{argmax}_{b \in B}\, k(a^*, b)$. Then it is easily seen that $h(a^*, b^*) \geq h_*$ and $k(a^*, b^*) \geq k^*$. Note that the point (h_*, k^*) does not necessarily belong to F and thus it may not belong to $\tilde{F}$. For example, for the 2×2 stage game

	L	R
T	0, 0	1, 2
B	2, 1	0, 0

we have $(h_*, k^*) = (0,1)$ and $F = \mathrm{Co}\{(0,0), (1,2), (2,1)\}$. So $(h_*, k^*) \notin F$. ($\tilde{F} = \mathrm{Co}\{(\frac{1}{2}, 1), (1,2), (2,1)\}$) See Figure 1.

Theorem 4.1. *If* $m(n) \to \infty$, $m(n)\log m(n)/n \to 0$ *as* $n \to \infty$, *and if there is* $(x,y) \in \tilde{F}$ *with* $x > h_*$, *then* $E^n(m(n)) \to \tilde{F}$ *as* $n \to \infty$.

[2] The previous proof utilized the notion of entropy.

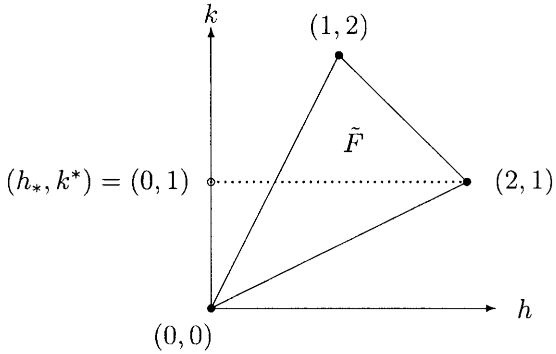


Fig. 1. (h_*, k^*) may not be in F

As demonstrated in the next example, the conclusion of the theorem fails if the condition on $\tilde{F}$ is not satisfied.

Example (Neyman (1999)). Consider the 2×2 stage game given below.

	L	R
T	1, 3	0, 4
B	1, 1	1, 0

Observe that $h_* = k^* = 1$ and $\tilde{F} = \{(1, y) \mid 1 \le y \le 3\}$. For this game $E^n(m(n)) = \{(1, 1)\}$ for every n regardless of $m(n)$. To see this first note that player 1 must receive 1 at every stage in any equilibrium path, and he can guarantee 1 with an automaton of size 1 (play B at every stage).

Suppose that, in some equilibrium (σ, τ) of $G^n(m(n))$, player 2 received more than 1. Then (T, L) must be played with a positive probability at some stage on the equilibrium path. Let $\tilde{n}$ be the last stage at which (T, L) is played with a positive probability, i.e.,

$$\tilde{n} = \max\{n' \mid 1 \le n' \le n,\ P_{\sigma,\tau}((a_{n'}, b_{n'}) = (T, R)) > 0\}.$$

Define $\tilde{\tau} = (\tilde{\tau}_\ell) \in \Delta(T)$ as

$$\tilde{\tau}_\ell(\omega) = \begin{cases} \tau_\ell(\omega) & \text{for } 1 \le \ell \le \tilde{n} - 1 \\ R & \text{for } \ell = \tilde{n} \\ L & \text{for } \ell > \tilde{n} \end{cases}$$

Then it is easily verified that $k_n(\sigma, \tilde{\tau}) > k_n(\sigma, \tau)$, contradicting to the supposition that (σ, τ) is an equilibrium. □

We now turn to the proof of Theorem 4.1. Given a point $z \in \mathbb{R}^2$ and a nonempty compact set $Z \subset \mathbb{R}^2$, define $d(z, Z) = \min_{z' \in Z} \|z - z'\|$. Since $\tilde{F}$ and $E^n(m(n))$ are nonempty compact subsets of F which is compact, the conclusion of the theorem, $E^n(m(n)) \to \tilde{F}$, is equivalent to

$$\overline{\mathrm{Lim}}_{n\to\infty} E^n(m(n)) = \underline{\mathrm{Lim}}_{n\to\infty} E^n(m(n)) = \tilde{F}$$

where

$$\overline{\mathrm{Lim}}_{n\to\infty} E^n(m(n)) = \{z \mid \forall\varepsilon > 0, \forall n', \exists n \geq n' \text{ such that } d(z, E^n(m(n))) < \varepsilon\},$$

and

$$\underline{\mathrm{Lim}}_{n\to\infty} E^n(m(n)) = \{z \mid \forall\varepsilon > 0, \exists n_0 \text{ such that } \forall n \geq n_0, d(z, E^n(m(n))) < \varepsilon\}.$$

We will establish the identity of the three sets through a pair of claims. Note that the first claim requires neither $m(n) \to \infty$ $(n \to \infty)$ nor the condition on $\tilde{F}$ present in the statement of Theorem 4.1.

Claim 4.1. *If $m(n) \log m(n)/n \to 0$ as $n \to \infty$, then $\overline{\mathrm{Lim}}_{n\to\infty} E^n(m(n)) \subset \tilde{F}$.*

Proof: Obviously, $E^n(m(n))$ is included in the set of payoff vectors achieved by mixed strategies which in turn is a subset of F. As the set F is closed, $\overline{\mathrm{Lim}}_{n\to\infty} E^n(m(n)) \subset F$.

Take $(x, y) \in \overline{\mathrm{Lim}}_{n\to\infty} E^n(m(n))$. First, the fact that player 1 can guarantee h_* at every stage using a pure action $a^* \in \mathrm{argmax}_{a\in A}[\min_{b\in B} h(a, b)]$ shows that $x \geq h_*$. Next if G is zero sum so that $h = -k$, then

$$h_* = \max_{a\in A} \min_{b\in B} -k(a, b) = -\min_{a\in A} \max_{b\in B} k(a, b) = -k^*.$$

It follows from Lemma 4.1 that for every $\varepsilon > 0$, there is n_0 such that for each $n \geq n_0$ player 2 has a pure strategy t such that

$$k_n(s, t) \geq k^* - \varepsilon \quad \text{for all } s \in S(m(n)).$$

Therefore, for every $n \geq n_0$, player 2 must receive at least $k^* - \varepsilon$ in any equilibrium of $G_n(m(n))$. This implies that $y \geq k^*$. Q.E.D.

Claim 4.2. *If $m(n) \to \infty$, $m(n) \log m(n)/n \to 0$ as $n \to \infty$, and if there is $(x, y) \in \tilde{F}$ with $x > h_*$, then $\tilde{F} \subset \underline{\mathrm{Lim}}_{n\to\infty} E^n(m(n))$.*

Proof: First, we deal with the case in which there is (x, y) in $\tilde{F}$ with $x > h_*$ and $y > k^*$. To show $\tilde{F} \subset \underline{\mathrm{Lim}}_{n\to\infty} E_n(m(n))$ it suffices to show that, for every $\delta > 0$, the set $\tilde{F}_\delta = \{(x, y) \in \tilde{F} \mid x > h_* + \delta, y > k^* + \delta\}$ is contained in $\underline{\mathrm{Lim}}_{n\to\infty} E_n(m(n))$.

Let $K = \max\{\|h\|, \|k\|\}$. Since we have assumed that the payoffs are nonnegative, it follows that $|r(a, b) - r(a', b')| \leq K$ for all $r = h, k$ and (a, b), $(a', b') \in A \times B$. Fix a $\delta > 0$ for which $\tilde{F}_\delta \neq \emptyset$ and take $(x, y) \in \tilde{F}_\delta$. Let $\varepsilon > 0$ be sufficiently small so that $\varepsilon < \min\{1, K/4\}$, $x > h_* + 2\varepsilon$, and $y > k^* + 2\varepsilon$.

Let $(a_i, b_i) \in A \times B$, $i = 1, 2, 3$, be such that (x, y) is a convex combination of $(h(a_i, b_i), k(a_i, b_i))$, $i = 1, 2, 3$. Thus there are $\alpha_i \geq 0$, $i = 1, 2, 3$, such that $\alpha_1 + \alpha_2 + \alpha_3 = 1$ and $(x, y) = \sum_{i=1}^{3} \alpha_i(h(a_i, b_i), k(a_i, b_i))$. Assume that $k(a_1, b_1) \le k(a_2, b_2) \le k(a_3, b_3)$ and, without loss of generality, $\alpha_3 > 0$. Let d be a sufficiently large positive integer so that, by setting $d_1 = [\alpha_1 d]$, $d_2 = [\alpha_2 d]$, $d_3 = d - (d_1 + d_2)$, and $(\bar{x}, \bar{y}) = (1/d) \sum_{i=1}^{3} d_i(h(a_i, b_i), k(a_i, b_i))$, the following inequalities hold:

$$\|(\bar{x}, \bar{y}) - (x, y)\| < \frac{\varepsilon}{2}, \tag{4.1}$$

and

$$d_3(\bar{y} - k^*) > K. \tag{4.2}$$

Note that $(\bar{x}, \bar{y})$ converges to (x, y) as d tends to infinity and thus (4.1) holds for a sufficiently large d. Also, since $\alpha_3 > 0$ implies that $d_3 \to \infty$ as $d \to \infty$, and since $\bar{y} > k^*$, (4.2) holds for a sufficiently large d.

Let $b_4 \in B$ be a best response of player 2 to the action a_3 of player 1. Define a sequence of action pairs of length d, $\xi = (\xi_1, \dots, \xi_d)$, by

$$\xi_j = \begin{cases} (a_1, b_1) & \text{for } j = 1, \dots, d_1 \\ (a_2, b_2) & \text{for } j = d_1 + 1, \dots, d_1 + d_2 \\ (a_3, b_3) & \text{for } j = d_1 + d_2 + 1, \dots, d. \end{cases}$$

(Recall that $d_1 + d_2 + d_3 = d$) Define a sequence of action pairs $\omega = (\omega_1, \dots, \omega_n) \in (A \times B)^n$ as follows. Let $q = [n/d]$. In the last d stages, ω coincides with ξ up to the one stage before the end and then finishes with (a_3, b_4):

$$(\omega_{n-d+1}, \dots, \omega_n) = (\xi_1, \dots, \xi_{d-1}, (a_3, b_4)),$$

From the stage $n - qd + 1$ up to $n - d$, ξ is repeated $q - 1$ times:

$$(\omega_{n-qd+1}, \dots, \omega_{n-d}) = \text{“}(\xi_1, \dots, \xi_d) \text{ repeated } q - 1 \text{ times.”}$$

Finally, in the first $n - qd$ stages, the tail part of ξ is played:

$$(\omega_1, \dots, \omega_{n-qd}) = (\xi_{(q+1)d-n+1}, \dots, \xi_d).$$

Notice that $(\omega_1, \dots, \omega_{n-1})$ is d-periodic. Clearly, for every $p = 1, \dots, n - 1$,

$$\sum_{\ell=p+1}^{n} h(\omega_\ell) \ge (n - p)\bar{x} - dK \tag{4.3}$$

The assumption $k(a_1, b_1) \le k(a_2, b_2) \le k(a_3, b_3)$ and the choice of b_4 imply that for every $p < n$,

$$\sum_{\ell=p+1}^{n} k(\omega_\ell) \ge (n - p)\bar{y}. \tag{4.4}$$

Define a pair of pure strategies $(\tilde{s}, \tilde{t})$ so that (i) they follow the path ω as long as the other player does so, and (ii) if player 2 deviated from ω, then $\tilde{s}$ takes a pure action $\tilde{a} \in \operatorname{argmin}_{a\in A}[\max_{b\in B} k(a,b)]$ at every stage afterward, while (iii) if player 1 deviated from ω for the first time at stage p, then $\tilde{t}$ starts playing a pure strategy $\hat{t}$ such that for every $s \in S(m(n))$ and $\ell = 1, 2, \ldots,$

$$h_\ell(s, \hat{t}) \le h_* + \frac{\|h\| \log|S(m(n))|}{\ell}.$$

Theorem 3.1 ensures the existence of such strategy.

The strategy $\tilde{s}$ is implementable by an automaton of size $d+1$: d states for playing the cycle phase of ω and one for the punishment.[3] So, if n is large enough so that $m(n) > d$, then $\tilde{s} \in S(m(n))$. Since the play induced by $(\tilde{s}, \tilde{t})$ is ω, we have

$$\|(h_n(\tilde{s}, \tilde{t}), k_n(\tilde{s}, \tilde{t})) - (\bar{x}, \bar{y})\| < \frac{dK}{n} < \frac{\varepsilon}{2} \quad \text{for } n > \frac{2dK}{\varepsilon}.$$

It follows from (4.1), using the triangle inequality, that $(h_n(\tilde{s}, \tilde{t}), k_n(\tilde{s}, \tilde{t}))$ is within ε of (x, y) for sufficiently large n.

Next we will show that no unilateral deviation from $(\tilde{s}, \tilde{t})$ leads to a strict improvement of the payoff. We start with player 2. Take $t \in T$. Assume that the strategy t deviates from the play ω at stage p. If $p \le n - d_3$, then the inequalities (4.2) and (4.4) imply that

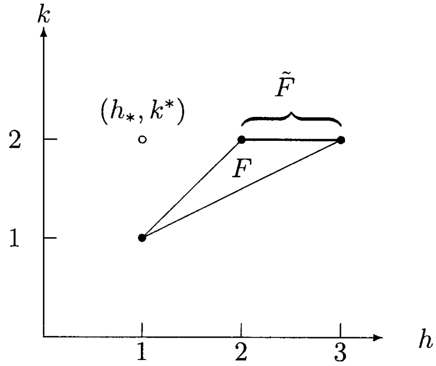
$$\begin{aligned} n(k_n(\tilde{s}, \tilde{t}) - k_n(\tilde{s}, t)) &\ge -K + (n-p)(\bar{y} - k^*) \\ &\ge -K + d_3(\bar{y} - k^*) > 0, \end{aligned}$$

while if $n - d_3 < p \le n$, recalling the choice of b_4,

$$\begin{aligned} n(k_n(\tilde{s}, \tilde{t}) - k_n(\tilde{s}, t)) &\ge (n-p)k(a_3, b_3) + k(a_3, b_4) \\ &\quad - (k(a_3, b_4) + (n-p)k^*) \\ &= (n-p)(k(a_3, b_3) - k^*) > 0. \end{aligned}$$

Thus we conclude that player 2 cannot benefit from a deviation from ω at any stage. Let us turn to player 1. Take $s \in S(m(n))$ and suppose that $(s, \tilde{t})$ resulted in player 1's deviation from the path ω. The fact that $s \in S(m(n))$ implies that such deviation must occur in the first $m(n)$ repetitions of the cycle $\omega_1, \ldots, \omega_d$, and hence in the first $m(n)d$ stages of the repeated game. Thus assume that the deviation occurred at stage $p \le m(n)d$. Let $(\omega'_1, \ldots, \omega'_p)$ be the play induced by $(s, \tilde{t})$ up to stage p and set $s' = (s|\omega'_1, \ldots, \omega'_p)$. Then by the construction of $\tilde{t}$, we have, recalling that the payoffs are assumed to be nonnegative,

[3] Although $\omega_n \ne \xi_d$, player 1's action in ω_n, a_3, is the same as his action in ξ_d.



1, 1	3, 2
2, 2	1, 1

$$(h_*, k^*) = (1, 2)$$
$$F = \mathrm{Co}\{(1,1), (2,2), (3,2)\}$$
$$\tilde{F} = \mathrm{Co}\{(2,2), (3,2)\}$$

Fig. 2. $\tilde{F} = \{(x, k^*) | h_0 \le x \le h_1\}$, $h_* \le h_0 < h_1$

$$h_n(s, \tilde{t}) \le \frac{pK}{n} + \left(1 - \frac{p}{n}\right) h_{n-p}(s', \hat{t})$$
$$\le \frac{pK}{n} + h_* + \frac{\|h\| \log|S(m(n))|}{n}$$

which would be less than $h_* + \varepsilon$ if, e.g.,

$$n \ge \frac{2}{\varepsilon} \max\{m(n)dK, \|h\| \log|S(m(n))|\}. \tag{4.5}$$

Our assumption on the order of magnitude of $m(n)$ guarantees (4.5) to hold for all sufficiently large n. Since $x > h^* + 2\varepsilon$ and $h_n(\tilde{s}, \tilde{t})$ is within ε of x, we have $h_n(s, \tilde{t}) < h_n(\tilde{s}, \tilde{t})$. We have thus shown that $(\tilde{s}, \tilde{t})$ is an equilibrium of $G^n(m(n))$ with a payoff vector within ε of our target payoff vector (x, y) provided that n is large enough.

Next assume $\tilde{F} = \{(x, k^*) | h_0 \le x \le h_1\}$ where $h_* \le h_0 \le h_1$ and $h_* < h_1$. See Figure 2 for example. In this case there are two action pairs (a, b) and (a', b') such that $k(a, b) = k^*$, $h(a, b) \le h_0$ and $(h(a', b'), k(a', b')) = (h_1, k^*)$. Take $(x, k^*) \in \tilde{F}$ with $x > h_* + 2\varepsilon$ where $\varepsilon > 0$ is sufficiently small. Let d be a sufficiently large positive integer and let $\xi = (\xi_1, \ldots, \xi_d) \in (A \times B)^d$ be such that (i) $\xi_j = (a, b)$ or (a', b'), and (ii) $|(1/d) \sum_{j=1}^{d} h(\xi_j) - x| < \varepsilon$. Let n be a sufficiently large positive integer and define the path $\omega = (\omega_1, \ldots, \omega_n) \in (A \times B)^n$ by $(\omega_1, \ldots, \omega_{n-qd}) = ((a', b'), \ldots, (a', b'))$ and $(\omega_{n-qd+1}, \ldots, \omega_n) =$ "ξ repeated q times" where $q = [n/d]$. Let $\tilde{s}$ be player 1's pure strategy that takes the action ω_ℓ^1 at stage ℓ regardless of the past history. Let $\tilde{t}$ be player 2's pure strategy that follows the path ω as long as player 1 does so, and if player 1 deviated from ω at stage ℓ for the first time, it immediately reverts to $\hat{t}$ against $S(m(n))$ as in the previous case.

Since player 2 receives k^* at every stage when $(\tilde{s}, \tilde{t})$ is played and it is the highest payoff for her in the stage game, she has no incentive to deviate from ω at any stage. An argument similar to the first case shows that player 1 cannot benefit from the deviation from ω, provided that n is sufficiently large.

Thus $(\tilde{s}, \tilde{t})$ is an equilibrium of $G^n(m(n))$ with payoff within ε of (x, k^*) for sufficiently large n. This completes the proof. Q.E.D.

Claims 4.1, 4.2 and the fact $\underline{\mathrm{Lim}}_{n\to\infty} E^n(m(n)) \subset \overline{\mathrm{Lim}}_{n\to\infty} E^n(m(n))$ establish Theorem 4.1.

5. The λ-discounted game $G_\lambda(m(\lambda))$

In this section we will study a modified version of the λ-discounted game, $G_\lambda(m(\lambda)) = (S(m(\lambda)), T, h_\lambda, k_\lambda)$. We consider the bound on the complexity of player 1's strategies to be a function of the discount factor λ such that $m(\lambda) \to \infty$ as $\lambda \to 1$. The next lemma is an analog of Lemma 4.1.

Lemma 5.1. *Suppose that* $(1-\lambda)m(\lambda)\log m(\lambda) \to 0$ *as* $\lambda \to 1$. *Then for every* $\varepsilon > 0$, *there is* λ_0 *such that for each* $\lambda \in [\lambda_0, 1)$ *there is* $t \in T$ *such that*

$$h_\lambda(s,t) \le h_* + \varepsilon \quad \textit{for all } s \in S(m(\lambda)).$$

Corollary 5.1. *Let* G *be a two-person zero sum game and* $V_\lambda(m(\lambda))$ *be the value of* $G_\lambda(m(\lambda))$. *If* $(1-\lambda)m(\lambda)\log m(\lambda) \to 0$ *as* $\lambda \to 1$, *then* $V_\lambda(m(\lambda)) \to h_*$ *as* $\lambda \to 1$.

Let us denote by $E_\lambda(m(\lambda))$ the set of (Nash) equilibrium payoff vectors of $G_\lambda(m(\lambda))$. The next theorem is an analogue of Theorem 4.1.

Theorem 5.1. *If* $(1-\lambda)m(\lambda)\log m(\lambda) \to 0$ *as* $\lambda \to 1$ *and if there is* $(x, y) \in \tilde{F}$ *with* $x > h_*$ *or* $y > k^*$, *then* $E_\lambda(m(\lambda)) \to \tilde{F}$ *as* $\lambda \to 1$.

Proof: Define the sets $\overline{\mathrm{Lim}}_{\lambda\to 1} E_\lambda(m(\lambda))$ and $\underline{\mathrm{Lim}}_{\lambda\to 1} E_\lambda(m(\lambda))$ similarly to $\overline{\mathrm{Lim}}_{n\to\infty} E^n(m(n))$ and $\underline{\mathrm{Lim}}_{n\to\infty} E^n(m(n))$. An argument similar to the proof of Claim 4.1 together with Lemma 5.1 shows that $\overline{\mathrm{Lim}}_{\lambda\to 1} E_\lambda(m(\lambda)) \subset \tilde{F}$. Below we will show that $\tilde{F} \subset \underline{\mathrm{Lim}}_{\lambda\to 1} E_\lambda(m(\lambda))$.

First, assume that there is a point (x, y) in $\tilde{F}$ such that $x > h_*$ and $y > k^*$. As in the proof of Claim 4.2, fix $\delta > 0$ with $\tilde{F}_\delta \neq \emptyset$ and take $(x, y) \in \tilde{F}_\delta$. Let $\varepsilon > 0$ be sufficiently small so that $x > h_* + 4\varepsilon$ and $y > k^* + 4\varepsilon$. Let d be a sufficiently large positive integer and let $\xi = (\xi_1, \ldots, \xi_d) \in (A \times B)^d$ be a finite sequence of action pairs such that

$$\left\| \frac{1}{d}\sum_{j=1}^{d}(h(\xi_j), k(\xi_j)) - (x, y) \right\| < \frac{\varepsilon}{2}. \tag{5.1}$$

Define a play $\omega = (\omega_1, \omega_2, \ldots)$ by $\omega_\ell = \xi_j$ if $\ell = j \pmod d$. That is, ω consists of repetitions of the finite cycle ξ. For each positive integer p let

$$(x_p(\lambda), y_p(\lambda)) = (1-\lambda)\sum_{\ell=p}^{\infty} \lambda^{\ell-p}(h(\omega_\ell), k(\omega_\ell)),$$

and set $(x(\lambda), y(\lambda)) = (x_1(\lambda), y_1(\lambda))$. As $(\omega_\ell)_{\ell=1}^\infty$ is d-periodic, $(x_p(\lambda), y_p(\lambda))$ converges to $(1/d)\sum_{j=1}^d (h(\xi_j), k(\xi_j))$ as $\lambda \to 1$ for each p. This convergence is uniform in p. So we can take λ sufficiently close to 1 so that for every $p = 1, 2, \ldots,$

$$\left\| (x_p(\lambda), y_p(\lambda)) - \frac{1}{d}\sum_{j=1}^{d}(h(\xi_j), k(\xi_j)) \right\| < \frac{\varepsilon}{2}. \tag{5.2}$$

Now we describe the equilibrium strategies $\tilde{s} \in S$ and $\tilde{t} \in T$. Player 1's strategy $\tilde{s}$ follows the play ω as long as player 2 does so. If player 2 deviated from ω at some stage, then from the next stage on $\tilde{s}$ takes a pure action $\tilde{a} \in \operatorname{argmin}_{a\in A}[\max_{b\in B} k(a,b)]$. Player 2's strategy $\tilde{t}$ also follows ω as long as player 1 does so. If player 1 deviated from ω at some stage, then at the next stage $\tilde{t}$ starts playing the pure strategy $\hat{t}$ constructed in the proof of Theorem 3.1 against player 1's strategy set $S(m(\lambda))$.

The strategy $\tilde{s}$ is implementable by an automaton of size at most $d+1$. So for λ sufficiently close to 1 so that $m(\lambda) > d$, we have $\tilde{s} \in S(m(\lambda))$. Note that $(h_\lambda(\tilde{s}, \tilde{t}), k_\lambda(\tilde{s}, \tilde{t})) = (x(\lambda), y(\lambda))$. Thus, by (5.1) and (5.2), the strategy pair $(\tilde{s}, \tilde{t})$ yields a payoff vector within ε of (x, y).

Take $s \in S(m(\lambda))$ and let $(\omega'_\ell)_{\ell=1}^\infty$ be the play induced by $(s, \tilde{t})$. Then, either $\omega_\ell = \omega'_\ell$ for all ℓ or there is the smallest p such that $\omega_p \neq \omega'_p$. (Note that both (ω_ℓ) and (ω'_ℓ) are deterministic plays.) In the latter case, Lemma 5.1 implies that

$$h_\lambda(s, \tilde{t}) < (1-\lambda)\left(\sum_{\ell=1}^{p-1}\lambda^{\ell-1}h(\omega_\ell) + \lambda^{p-1}h(\omega'_p) + \frac{\lambda^p}{1-\lambda}(h_* + \varepsilon)\right).$$

It follows from (5.1), (5.2), and the assumption $x > h_* + 4\varepsilon$, that, for λ sufficiently close to 1, $(h_* + \varepsilon) - x_{p+1}(\lambda) < -2\varepsilon$. Since $\lambda/(1-\lambda) \to \infty$ as $\lambda \to 1$, we have $K + (\lambda/(1-\lambda))((h_* + \varepsilon) - x_p(\lambda)) < -\varepsilon$ for λ sufficiently close to 1 and hence

$$\begin{aligned} h_\lambda(s, \tilde{t}) - h_\lambda(\tilde{s}, \tilde{t}) &= h_\lambda(s, \tilde{t}) - x(\lambda) \\ &\le (1-\lambda)\lambda^{p-1}\left(K + \frac{\lambda}{1-\lambda}((h_* + \varepsilon) - x_{p+1}(\lambda))\right) \\ &\le (1-\lambda)\lambda^{p-1}(-\varepsilon) < 0. \end{aligned}$$

If λ is sufficiently close to 1 so that $(1-\lambda)dK \le \lambda^d\varepsilon$, then player 2 would have no incentive to deviate from ω at any stage. Indeed, if player 2 deviated from ω for the first time in the p-th cycle, then the gain within the p-th cycle from the deviation is at most $(1-\lambda)\lambda^{(p-1)d}dK$ while the loss she incurs from the punishment is at least $\lambda^{pd}\varepsilon$.

Thus we have shown that, for all λ sufficiently close to 1, $(\tilde{s}, \tilde{t})$ is an equilibrium of $G_\lambda(m(\lambda))$ that yields a payoff vector within ε of (x, y).

Now suppose that $\tilde{F} = \{(h_*, y) | k_0 \le y \le k_1\}$ where $k^* \le k_0 \le k_1$ and $k^* < k_1$. Then there are action pairs (a, b) and (a', b') such that $h(a, b) = h_*$,

$k(a,b) \le k_0$ and $(h(a',b'),k(a',b')) = (h_*,k_1)$. For a given payoff vector (x,y) in $\tilde{F}$ with $y > k^*$, define $\xi = (\xi_1,\ldots,\xi_d)$ as in (5.1) except that $\xi_j = (a,b)$ or (a',b'). Let ω be the cyclic play with the cycle ξ. Define a strategy pair $(\tilde{s},\tilde{t})$ as follows: $\tilde{s}$ is the same as above, and $\tilde{t}$ follows ω regardless of the history. Note that player 1 receives h_* at every stage on the play ω and the assumption on $\tilde{F}$ implies that this is the highest payoff for him in the stage game. Thus player 1 cannot benefit by deviating from ω. The same argument as above shows that player 2 has no incentive to deviate from ω provided that λ is sufficiently close to 1. The proof for the case $\tilde{F} = \{(x,k^*)|h_0 \le x \le h_1\}$ with $h_* \le h_0 \le h_1$ and $h_* < h_1$ is similar and we omit it. Q.E.D.

For the following 2×2 game

	L	R
T	0, 1	1, 0
B	1, 0	0, 1

$(h_*,k^*) = (0,1)$ and hence $\tilde{F} = \{(0,1)\}$. It is clear, however, that player 1 must receive a strictly positive payoff in any equilibrium of the λ-discounted game. Thus one cannot dispense with the condition on $\tilde{F}$ in the statement of Theorem 5.1.

6. Concluding remarks

In the proof of Claim 4.2, we argued that player 1's deviation from equilibrium path, if any, must occur in a very early stage of the repeated game and so there are enough stages after the deviation so that player 2's punishment is effective. For this we only needed the condition on the order of magnitude of the complexity bound, $m(n) = o(n)$, which is weaker than our original condition $m(n)\log m(n) = o(n)$. The latter condition is needed to determine the individually rational levels, or rather, we know the individually rational levels only under this condition on $m(n)$. Suppose that we obtained a result that, for a particular sequence $(m(n))_{n=1}^{\infty}$ with $m(n) = o(n)$,

$$\lim_{n\to\infty} V^n(m(n)) = \mathrm{Val}(G) = \min_{\beta\in\Delta(B)} \max_{a\in A} h(a,\beta)$$

for every two-person zero sum game $G = (A,B,h)$. (See Neyman (1997) Conjecture 1 and 2.) Then an essentially the same proof as in Theorem 4.1 shows that, for such sequence $(m(n))_{n=1}^{\infty}$,

$$E^n(m(n)) \to \left\{(x,y)\in F \,\middle|\, x \ge \min_{\beta\in\Delta(B)} \max_{a\in A} h(a,\beta),\ y \ge \min_{\alpha\in\Delta(A)} \max_{b\in B} k(\alpha,b)\right\}$$

as $n \to \infty$, provided that there is a feasible payoff vector in which player 1 receives strictly more than $\min_{\beta\in\Delta(B)} \max_{a\in A} h(a,\beta)$. Similar argument holds for the discounted games.

References

Ben-Porath E (1993) Repeated games with finite automata. Journal of Economic Theory 59:17–32

Kalai E, Stanford W (1988) Finite rationality and interpersonal complexity in repeated games. Econometrica 56:397–410

Lehrer E (1988) Repeated games with stationary bounded recall strategies. Journal of Economic Theory 46:130–144

Lehrer E (1994) Finitely many players with bounded recall in infinitely repeated games. Games and Economic Behavior 7:390–405

Neyman A (1985) Bounded complexity justifies cooperation in the finitely repeated prisoner's dilemma. Economics Letters 19:227–229

Neyman A (1997) Cooperation, repetition, and automata. In Cooperation: Game-Theoretic Approaches, ed. by S. Hart, and A. Mas-Colell, vol. 155 of NATO ASI-Seies F, pp. 233–255. Springer Verlag

Neyman A (1999) Finitely repeated games with finite automata. Mathematics of Operations Research 23:513–552

Neyman A, Okada D (1999) Strategic entropy and complexity in repeated games. Games and Economic Behavior 29:191–223

Neyman A, Okada D (2000) Repeated games with bounded entropy. Games and Economic Behavior 30:228–247

Papadimitriou CH, Yannakakis M (1994) On complexity as bounded rationality: Extended abstract. In STOC 94, pp. 726–733, Montreal, Quebec, Canada

Zemel E (1989) Small talk and cooperation: A note on bounded rationality. Journal of Economic Theory 49:1–9