

The Folk Theorem for Finitely Repeated Games with Mixed Strategies

Olivier Gossner

February 1994

REVISED VERSION

Abstract

This paper proves a Folk Theorem for finitely repeated games with mixed strategies. To obtain this result, we first show a similar property for finitely repeated games with terminal payoffs.

I wish to thank J. Abdou, who introduced me to this field of research, and S. Sorin for his constant help, support and fruitful discussions. My special thanks go to Amrita Dhillon and Denis Gromb for very helpful discussions.

Olivier Gossner, L. S. T. A. , Université Paris 6,
4 Place Jussieu, 75005 Paris, FRANCE
e-mail: gossner@dma.ens.fr

1 Introduction

The perfect Folk Theorem (Aumann and Shapley [1], Rubinstein [10]) states that any payoff that is feasible and better than the minimax payoff for any player in a one-shot game is a subgame perfect equilibrium payoff of the corresponding infinitely repeated game with standard signaling and without discounting. Since then, results have been proved for different structures of repeated games: mainly discounted or finitely repeated games. But one should also distinguish results that rely on the assumption that players use only pure strategies, and results that allow them to use mixed strategies. We will refer to the former ones as Folk Theorems in pure strategies, as opposed to Folk Theorems in mixed strategies.

In a one-shot game, the difference between pure and mixed strategies relies essentially on the structure of the strategy spaces (convex in one case but usually not in the other). In a repeated game, the signaling structures are radically different in both models. If we assume that the mixed actions of the players are revealed after each round, we are basically in a model with an extended set of pure strategies. On the other hand, with mixed strategies, only the action corresponding to a realization of their mixed moves is announced. In this case, it is therefore impossible to force some player to use a mixed move by threatening him with a punishment if he disobeys. Since most proofs of Folk Theorems rely on some similar arguments of threat and reward, proofs of Folk Theorems in pure strategies don't easily extend to mixed strategies.

The Folk Theorem by Aumann and Shapley [1], Rubinstein [10] holds both in pure and in mixed strategies. Fudenberg and Maskin [4] and [5] show that both a Folk Theorem in pure strategies and a Folk Theorem in mixed strategies hold for discounted infinitely repeated games (see also Neyman [9] and Sorin [12]). Benoit

and Krishna [2] obtain a Folk Theorem in pure strategies for finitely repeated games under some assumptions on the one-shot game, and exhibit some finitely repeated games for which it fails. Overlapping generations games –OLGs hereafter–, inspired by Samuelson’s economic model, form another important class of games. Kandori [7], and Smith [11] in a more general case, proved a Folk Theorem in pure strategies for OLGs. For the study of OLGs, Kandori introduced finitely repeated games with terminal payoffs which are interesting from a technical point of view. These are finitely repeated games in which players receive an additional payoff which depends on the history of the game.

The scope of this paper is to extend the Folk Theorem for finitely repeated games from pure strategies to mixed strategies. To do this, we first prove a Folk Theorem in mixed strategies for finitely repeated games for terminal payoffs. In section 2, we recall the models of finitely repeated games and finitely repeated games with terminal payoffs. We also review the results of Benoit and Krishna [2] and Kandori [7] for these games. The Folk Theorem in mixed strategies for finitely repeated games with terminal payoffs is presented in section 3. In section 4, we show how the terminal payoffs can be constructed as equilibrium payoffs of a subgame of the repeated game. Using this, we finally prove that a Folk Theorem in mixed strategies holds for finitely repeated games.

2 The models

2.1 The one-shot game

Let G be an I -player normal form game, the set of players being represented by $I = \{1, \dots, I\}$, with payoff function $g : \prod_i A^i \rightarrow \mathbb{R}^I$. A^i is player i ’s set of pure strategies and is finite, and $g^i(a)$ (the i -th component of $g(a)$) is his payoff when

the action profile is a .

Players may use mixed strategies, and the simplex $\Delta(A^i) = S^i$ represents the set of mixed strategies for player i . $g : \prod_i S^i \rightarrow \mathbb{R}^I$ will also denote the canonical extension of g , thus $g(s^1, \dots, s^I)$ is the expected payoff when each player i uses the mixed strategy s^i .

We will use the notations $A = \prod_i A^i$, $A^{-i} = \prod_{j \neq i} A^j$, and similarly for S , S^{-i} . For every player i , we select a minimax in mixed strategies against player i for players other than i , $m_i^{-i} \in S^{-i}$, and a best response m_i^i against m_i^{-i} , so that $m_i = (m_i^{-i}, m_i^i) \in S$.

Without loss of generality, we normalize the payoffs so that for every i , $g^i(m_i) = 0$. Let the convex hull of the set of feasible payoffs be $F = \text{co} \{g(a), a \in A\}$. $V = F \cap \mathbb{R}_+^I$ is the (closed) set of payoffs that are individually rational in mixed strategies and feasible. $C = \|G\| = \max_{i,a} |g^i(a)|$ is the norm of G . For $X \in \mathbb{R}^I$ and $\rho \in \mathbb{R}_+$, we will denote by $\mathcal{B}(X, \rho)$ the closed ball with center X and radius ρ .

G being fixed, G^* represents the I -player normal form game in which players' pure strategy sets are S^i , with payoff function g . In G^* , players use only pure strategies, that are the original mixed strategies of G .

2.2 Finitely repeated games

Given a game G as above, the finitely repeated game $G(T)$ is defined as the game G being repeated T times, where players are observing at each stage the action profiles of G . A history at stage $t \in \{1, \dots, T\}$ is an element of H_t , where H_t is the set of t -tuples of elements of A if $t > 0$, and $H_0 = \{\emptyset\}$ for convenience. A strategy of player i is a sequence $\sigma^i = (\sigma_t^i)_{1 \leq t \leq T}$ of mappings $\sigma_t^i : H_{t-1} \rightarrow S^i$. If the history at stage T is (a_1, \dots, a_T) the (average) payoff vector in $G(T)$ is:

$$g_T(a_1, \dots, a_T) = \frac{1}{T} \sum_{t=1}^T g(a_t)$$

An I -tuple of strategies $\sigma = (\sigma^1, \dots, \sigma^I)$ induces a probability distribution on the set of histories at stage T , and hence on the payoff vectors. Recall that σ is a Nash equilibrium of $G(T)$ if for every player i , the strategy σ^i maximizes his expected payoff in $G(T)$ given other players' strategies. Also σ is a subgame perfect equilibrium of $G(T)$ when for all $h \in H_t$, σ_h is a Nash equilibrium of $G(T-t)$, where $\sigma_h(h')$ is σ applied to the history h followed by h' . Let $E(T)$ denote the set of (average expected) subgame perfect equilibrium payoffs of $G(T)$.

$G^*(T)$ represents the game G^* repeated T times. In this game a history at time $t > 0$ is a sequence of t elements of S . The set of subgame perfect equilibrium payoffs of $G^*(T)$ will be denoted by $E^*(T)$.

Note that since every player must get at least his minimax payoff in any equilibrium, $E(T)$ and $E^*(T)$ are subsets of \mathbb{R}_+^I . It is obvious that $E(T)$ and $E^*(T)$ are also subsets of F , therefore $E(T) \subset V$ and $E^*(T) \subset V$. Two Folk Theorems, one proved by Benoit and Krishna [2], and the other which is the scope of this paper, give general conditions under which the converse inclusions are also true when T goes to infinity.

The following result deals with the case of pure strategies:

Theorem 1 (Benoit and Krishna, 1985) *Consider $G^*(T)$ and assume that:*

(i) *For every player i , there exists two Nash equilibria e^i and f^i of G such that*

$$g^i(e^i) > g^i(f^i)$$

(ii) *$\dim F = I$*

then

$$\forall \varepsilon > 0, \exists \bar{T} \in \mathbb{N}, \forall T \geq \bar{T}, \forall v \in V, \mathcal{B}(v, \varepsilon) \cap E^*(T) \neq \emptyset$$

Another way to state this Theorem is to say that, under assumptions (i) and (ii), the limit for the Hausdorff topology of the set of equilibrium payoffs of $G^*(T)$ as T tends to infinity is the whole space V .

The scope of this paper is to prove that Theorem 1 can be extended to the case where the players use mixed strategies.

Theorem A (Main Theorem) *Assume (i) and (ii), then:*

$$\forall \varepsilon > 0, \exists \bar{T} \in \mathbb{N}, \forall T > \bar{T}, \forall v \in V, \mathcal{B}(v, \varepsilon) \cap E(T) \neq \emptyset$$

2.3 Finitely repeated games with terminal payoffs

Let \mathcal{W} be a nonempty subset of \mathbb{R}^I and ω a function from H_T to \mathcal{W} . The finitely repeated game with terminal payoffs $G(T, \mathcal{W}, \omega)$ is defined as being the game G repeated T times in which players receive an additional payoff at the end of the T stages. This terminal payoff is given by ω applied to the history at stage T . \mathcal{W} is called the terminal payoffs set and ω is the terminal payoffs function. If the history is (a_1, \dots, a_T) at stage T , the (average) payoff vector in $G(T, \mathcal{W}, \omega)$ is:

$$g_{T, \mathcal{W}, \omega}(a_1, \dots, a_T) = \frac{1}{T} \sum_{t=1}^T g(a_t) + \frac{1}{T} \omega(a_1, \dots, a_T)$$

The set $E(T, \mathcal{W}, \omega)$ of subgame perfect equilibrium payoffs of $G(T, \mathcal{W}, \omega)$ is defined in the usual sense. Since our terminal payoffs function ω will vary, we are interested in the study of $E(T, \mathcal{W}) = \cup_{\omega} E(T, \mathcal{W}, \omega)$, that we call by extension the set of subgame perfect equilibrium payoffs of $G(T, \mathcal{W})$. A history h_T will be called an equilibrium path of $G(T, \mathcal{W})$ if it is an equilibrium path of $G(T, \mathcal{W}, \omega)$ for some ω .

When the one-shot game is G^* , we also define the finitely repeated game with terminal payoffs $G^*(T, \mathcal{W}, \omega^*)$ (ω is a function from H_T^* to \mathcal{W}) and the sets of

equilibrium payoffs $E^*(T, \mathcal{W}, \omega)$ and $E^*(T, \mathcal{W}) = \cup_w E(T, \mathcal{W}, \omega)$.

We present here a version of Kandori's Folk Theorem [7] for finitely repeated games with terminal payoffs played in pure strategies.

Theorem 2 (Kandori, 1992) *If $F \cap \mathbb{R}_{++}^I$ is non empty, for any $\varepsilon > 0$, there exists a finite set $\mathcal{W} \subset \mathbb{R}^I$ and $T_0 \in \mathbb{N}$ such that:*

$$\forall T > T_0, \forall v \in V, \mathcal{B}(v, \varepsilon) \cap E^*(T, \mathcal{W}) \neq \emptyset$$

In the next section, we extend this result to mixed strategies, and use some particular spaces of terminal payoffs.

3 A Folk Theorem for finitely repeated games with terminal payoffs

In this section, we prove a ‘‘robust’’ Folk Theorem for finitely repeated games with terminal payoffs with mixed strategies, where in fact, the points in the terminal payoffs set may vary up to some ρ without affecting the results.

3.1 Statement of the result

We denote by REW (like rewards) the subset $\{0, 1\}^I \cup \{(-1, -1, \dots, -1)\}$ of \mathbb{R}^I . For $(W, \rho) \in \mathbb{R}_+^2$, $\Theta^{W, \rho}$ is the correspondence from REW to \mathbb{R}^I with compact values defined by $\Theta^{W, \rho}(r) = \mathcal{B}(Wr, \rho)$. A selection θ of $\Theta^{W, \rho}$ picks for every $r \in REW$ a point $\theta(r)$ in the ball $\Theta^{W, \rho}(r)$.

Theorem B *If $F \cap \mathbb{R}_{++}^I$ is non empty,*

$$\forall \rho \geq 0, \forall \varepsilon > 0, \exists W_0 \in \mathbb{R}, \exists T_0 \in \mathbb{N},$$

$$\forall T \geq T_0, \forall W \geq W_0, \forall v \in V, \text{ and for every selection } \theta \text{ of } \Theta^{W, \rho}$$

$$\mathcal{B}(v, \varepsilon) \cap E(T, \theta(REW)) \neq \emptyset$$

Note that in Theorem B, as in Theorem 2, the set of terminal payoffs does not depend on T . The game can be repeated as long as we want without increasing the size of the terminal payoffs set.

As a Corollary we get an equivalent to Theorem 2 in mixed strategies:

Corollary 1 *If $F \cap \mathbb{R}_{++}^I$ is non empty, for every $\varepsilon > 0$, there exists a finite subset \mathcal{W} of \mathbb{R}^I and T_0 in \mathbb{N} such that for every $T \geq T_0$ and every v in V :*

$$\mathcal{B}(v, \varepsilon) \cap E(T, \mathcal{W}) \neq \emptyset$$

We will first prove Theorem B for a fixed $v \in V$. Since $F \cap \mathbb{R}_{++}^I$ is non empty and convex, v can be approximated by payoffs that are in $F \cap \mathbb{R}_{++}^I$. With the notation $g(lA) = \frac{1}{l}\{\sum_{k=1}^l g(a_k), (a_1, \dots, a_l) \in H_l\}$, $l \in \mathbb{N}$, the set $\cup_l g(lA)$ is dense in F and \mathbb{R}_{++}^I is open, thus $\cup_l g(lA) \cap \mathbb{R}_{++}^I$ is dense in $F \cap \mathbb{R}_{++}^I$. Therefore to prove the theorem for v , it is enough to prove that for every sequence of action profiles $\tilde{a} = (a_1, a_2, \dots, a_l) \in H_l$ such that $\frac{1}{l}\sum_{k=1}^l g(a_k) \gg 0$, the path consisting of repeated cycles of this sequence is an equilibrium path of $G(T, \theta(REW))$ for any selection θ of $\Theta^{W,\rho}$, when W and T are big enough.

For a such sequence and a given selection, we will now construct equilibrium strategies having the following properties:

- Players conform with the path above when the history is consistent with it.
- If player i deviates “long enough” before the end of the game, other players will punish him, i.e. they will use for some number P of stages a strategy near m_i^{-i} ; that is what we will refer to as a punishing period.
- The players who effectively punish (effective punishers) will receive a reward as terminal payoff.

- If any player deviates in the last stages of the game, the terminal payoff will be bad for every player.

The elements of $\theta(\{0, 1\}^I)$ are payoffs that will reward players who were effective punishers. $\theta(-1, \dots, -1)$ is a collective punishment to prevent from “late” deviations.

The major difficulty one encounters when dealing with mixed strategies is to determine whether a player is an effective punisher or not. This will depend on the observation of action profiles issued during the punishing period, and thus we introduce the “test functions” $\alpha_{\eta}^{i,j}$ as below.

3.2 The test functions

Suppose that after some period of the game, we want to check if players have used some mixed strategies “near” a fixed strategy in $s \in S$. The history during the considered period gives us some probability distribution on the action profiles, and one could simply compare this distribution with the distribution induced by s on A . Here we want to determine what players j have used the mixed strategy s^j , therefore we introduce a different test for each player by the following way:

For a history of the game during some period be expressed as a t -tuple $h = (h_{t_0+1}, \dots, h_{t_0+t}) \in H_t$, let $n(a)$ be the number of occurrences of $a \in A$. Also $n(a^{-j})$ is the number of times that players others than j issued the action $a^{-j} \in A^{-j}$, $n(a^{-j}) = \sum_{b^j \in A^j} n(a^{-j}, b^j)$. A distance between player j 's observed strategy during some history h and the repetition of s^j is defined by:

$$D^j(h, s^j) = \frac{1}{t} \sum_{a \in A} |n(a) - n(a^{-j})P(s^j = a^j)|$$

where $P(s^j = a^j)$ represents the probability for player j to issue the action a^j while

using the strategy s^j .

After a punishing period, we want to compare each player's strategy and a minimax strategy, so when the history during a punishing period against player i is $h \in H_P$, some test functions are given for $\eta > 0$ by:

$$\alpha_\eta^{i,j} = \begin{cases} 1 & \text{if } D^j(h, m_i^j) < \eta \\ 0 & \text{otherwise} \end{cases}$$

An effective punisher j against i will be a punisher so that $\alpha_\eta^{i,j} = 1$. The two following lemmata show that if P is large enough, and η small enough, when all punishers are effective the deviator receives an average payoff which is less than the payoff in the normal path. Conversely, players who play the minimax strategy are usually effective punishers.

Lemma 3.1 *For every $\varepsilon > 0$, and $\eta > 0$, there exists P_0 in \mathbb{N} such that if $P \geq P_0$, if any player $j \neq i$ uses the strategy m_i^j during P stages where the history is $h \in H_P$, the probability that $\alpha_\eta^{i,j}(h) = 0$ is less than ε whatever are the strategies used by players other than j .*

Proof:

This proof uses the approachability theory, cf Blackwell [3], or Mertens, Sorin and Zamir [8].

For any player j fixed, we consider the game \tilde{G} with two players I and II, strategy sets A^j for player I and A^{-j} for player II, and vector payoff $\tilde{g}(a)$ in \mathbb{R}^A with 1 in its component indexed by a and 0 elsewhere. For $h = (h_1, h_2, \dots, h_t) \in H_t$, $\bar{x}_t = \bar{x}_t(h) = \frac{1}{t} \sum_{k=1}^t \tilde{g}(h_k)$. $\bar{x}_{t,a}$ is the component of \bar{x}_t indexed by a . Note that:

$$D^j(h, s^j) = \sum_{a \in A} |\bar{x}_{t,a} - \sum_{b^j \in A^j} \bar{x}_{t,(a^{-j}, b^j)} P(s^j = a^j)|$$

Thus we can write D^j as a function from $[0, 1]^A \times S^j$ to \mathbb{R} defined by the above formula, that we will still denote by D^j .

When player I uses the strategy m_i^j , the convex hull $R(m_i^j)$ of the set of points $\{\tilde{g}(s^{-j}, m_i^j), s^{-j} \in A^{-j}\}$ is equal to the set of all $\bar{x} \in [0, 1]^A$ such that $D^j(\bar{x}, m_i^j) = 0$. The approachability theory tells us that this set is approachable for I by using the constant strategy m_i^j . This means that for all $\varepsilon_1 > 0$, there exists an integer P_0 such that for every strategy of player II, the probability $P(\sup_{t \geq P_0} \delta_t \leq \varepsilon_1)$ is greater than $1 - \varepsilon_1$, where δ_t is the distance between \bar{x}_t and $R(m_i^j)$.

In particular, since the function $D^j(\cdot, m_i^j)$ is continuous, for ε_1 sufficiently small, the probability that $D^j(\bar{x}_t, m_i^j) \geq \eta$ is smaller than ε .

Lemma 3.2 *For every $\varepsilon > 0$, there exists $\eta > 0$ such that for every i , and every $h_t = (h_1, h_2, \dots, h_t) \in H_t$ if for $j \neq i$ $\alpha_\eta^{i,j}(h) = 1$ then:*

$$\frac{1}{t} \sum_{k=1}^t g^i(h_k) < \varepsilon$$

Proof:

First we reorder the players such that the player to be punished is called player I . Consider an history $h = (h_1, h_2, \dots, h_t)$ and $\eta > 0$ such that for all $j \neq I$, $\alpha_\eta^{I,j}(h) = 1$. Then for every $a \in A$ and every $j \neq I$:

$$\frac{1}{t} |n(a) - n(a^{-j})P(m_I^j = a^j)| < \eta$$

In particular for every $a = (a^1, a^2, \dots, a^I) \in A$:

$$\frac{1}{t} |n(a) - \sum_{b^1 \in A^1} n(b^1, a^2, \dots, a^I)P(m_I^1 = a^1)| < \eta$$

Since for every $b^1 \in A^1$,

$$\frac{1}{t} |n(b^1, a^2, \dots, a^I) - \sum_{b^2 \in A^2} n(b^1, b^2, a^3, \dots, a^I) \mathbb{P}(m_I^2 = a^2)| < \eta$$

for every a we get:

$$\frac{1}{t} |n(a) - \sum_{(b^1, b^2) \in A^1 \times A^2} n(b^1, b^2, a^3, \dots, a^I) \mathbb{P}(m_I^1 = a^1) \mathbb{P}(m_I^2 = a^2)| < 2\eta$$

Repeating the same procedure $I - 1$ times leads to the formula:

$$\forall a \in A, \frac{1}{t} |n(a) - \sum_{(b^1, b^2, \dots, b^{I-1}) \in A^{I-1}} n(b^1, b^2, \dots, b^{I-1}, a^I) \prod_{j=1}^{I-1} \mathbb{P}(m_I^j = a^j)| < (I - 1)\eta$$

If we write $r(a^I) = \frac{1}{t} \sum_{b^{-I} \in A^{-I}} n(b^{-I}, a^I)$, we get that $\sum_{a^I \in A^I} r(a^I) = 1$, so that $\{r(a^I), a^I \in A^I\}$ defines a point $r^I \in S^I$ and:

$$\frac{1}{t} \sum_{k=1}^t g^I(h_k) = \frac{1}{t} \sum_{a \in A} n(a) g^I(a) \leq g^I(m_I^{-I}, r^I) + (I - 1)C\eta \leq (I - 1)C\eta$$

This proves that when η is small enough, if all punishers are effective punishers, the payoff of player I during the punishing period is less than ε .

3.3 Proof of Theorem B

Fix $v \in V$, $\rho \geq 0$, and $\tilde{a} = (a_1, \dots, a_l)$ such that $v' = \frac{1}{l} \sum_{t=1}^l g(a_t) \gg 0$. We will fix the parameters η , P and W_0 such that for any $W > W_0$, $T > P$ and any selection θ of $\Theta^{W, \rho}$, the following algorithm defines a terminal payoffs function ω and strategies that are a subgame perfect equilibrium of $G(T, \theta(REW, \omega))$, with equilibrium path repetitions of \tilde{a} .

Initialization:

Put $r_j = 0$ for all $j \in I$, and $t = 1$.

NORM (Normal path):

Play a_k at stage $t = k \pmod{l}$ until $t = T$, then go to **End**.

If player i deviates from **NORM** at stage $t_0 < T - P$, go to **P(i)**.

If any player deviates from **NORM** at stage $t_0 \geq T - P$, go to **LD**.

P(i) (Punishment of player i):

Play in G during P stages then redefine

$$r_j = \alpha_{\eta}^{i,j}(h_{t_0+1}, \dots, h_{t_0+P}) \text{ for all } j \neq i$$

and keep r_i unchanged.

Then go back to **NORM**.

LD (Late Deviations):

Redefine $r_j = -1$ for all $j \in I$, play in G until $t = T$ and go to **End**.

End:

Players receive $\theta(r_1, \dots, r_I)$ as terminal payoff.

We first choose η fixed by Lemma 3.2 so that for every i and every P , if for all $j \neq i$ $\alpha_{\eta}^{i,j}(h_1, \dots, h_P) = 1$ then:

$$\frac{1}{P} \sum_{t=1}^P g^i(h_t) < \frac{v'_i}{2}$$

Also let $\varepsilon_1 > 0$ and P_1 be so that for any $P > P_1$, and for each i :

$$\frac{1}{P}(2\rho + 2C) + 2I\varepsilon_1 C < \frac{v'_i}{2}$$

Lemma 3.1 gives for η and ε_1 a value P_0 , we now fix $P = kl$, $P > \max\{P_0, P_1\}$.

The algorithm does not define explicitly what the strategies of the players during a punishing period are, but we will prove that: if W is large enough, the strategy p_i^j of player j consisting of playing m_i^j repeatedly during the P stages of a punishing period of player i dominates any alternative σ^j strategy for which $P(\alpha_\eta^{i,j} = 0) > 2\varepsilon_1$.

In fact, playing a strategy that gives $P(\alpha_\eta^{i,j} = 0) > 2\varepsilon_1$ leads to a maximum expected payoff which is the sum of:

- PC during the punishing period.
- Some payoff, call it U^j , during the intermediate period between the end of the punishing period and the last stage of the game.
- The terminal payoff is at most $W + \rho$ if $\alpha_\eta^{i,j} = 1$, and at most of ρ if $\alpha_\eta^{i,j} = 0$, thus the expected terminal payoff is less than $(1 - 2\varepsilon_1)W + \rho$.

whereas using the strategy p_i^j gives as minimum expected payoff the sum of:

- $-PC$ during the punishing period.
- The same payoff U^j during the intermediate period.
- As above, we see that the expected terminal payoff is greater than $(1 - \varepsilon_1)W - \rho$.

Therefore the only condition that is needed is $W > W_0$ with:

$$W_0 \geq \frac{2CP + 2\rho}{\varepsilon_1}$$

We now check that for such a P , ε_1 , η , and W no player has incentive to deviate from **NORM** at stage $t_0 \leq T - P$.

A player who deviates from **NORM** at stage $t_0 \leq T - P$ can expect as payoff no more than the sum of:

- C maximal payoff during the stage of the deviation.
- During the punishing period, the payoff is $\sum_{t=0}^P g^i(h_{t_0+t}) < P \frac{v'_i}{2}$ if all punishers are effective punishers, which has a probability greater than $1 - 2I\varepsilon_1$ to occur, and if not the payoff during this period is less than $2CP$. Therefore an upper bound of the expected payoff is $2I\varepsilon_1CP + \frac{v'_i}{2}P$.
- Some payoff U^i during the intermediate period.
- The terminal payoff can raise from some w_i up to a maximum of $w_i + 2\rho$ due to a change of the parameter r .

By following the path **NORM** player i would receive:

- At least $-C$ at stage t_0 .
- Pv'_i during the punishing period.
- U^i during the intermediate period.
- w_i as terminal payoff.

Thus we need only to check that the following condition is satisfied:

$$C + 2I\varepsilon_1P + \frac{v'_i}{2}P + 2\rho \leq -C + v'_iP$$

The validity of the inequation above is a consequence of the definition of ε_1 and P .

To prevent any late deviation, i. e. a deviation occurring at a stage $t \geq T - P$, we assign to W_0 a value that is greater than $2PC + 2\rho$. For these values of the parameters, the algorithm above defines a subgame perfect equilibrium of $G(T, \theta(REW))$ with as equilibrium path repeated cycles of \tilde{a} .

Thus we proved so far that the theorem holds for any $v \in \cup_{l \in I} (lA) \cap \mathbb{R}_{++}^I$, and hence for any fixed $v \in V$. Now, see that for any $\rho \geq 0$ and $\varepsilon > 0$, since V is compact, it is included in a finite union of balls $\cup_{k \in K} \mathcal{B}(u_k, \frac{\varepsilon}{2})$, $u_k \in V$. We just pointed out that for every $k \in K$, there exists W_k , and T_k such that for $T > T_k$, and $W > W_k$, $\mathcal{B}(u_k, \frac{\varepsilon}{2}) \cap E(T, \theta(REW)) \neq \emptyset$ for any selection θ of $\Theta^{W, \rho}$. It is easy to check that if θ is a selection of $\Theta^{W, \rho}$ with $W \geq \max_k W_k$ and if $T \geq \max_k T_k$, for any $v \in V$ $\mathcal{B}(v, \varepsilon) \cap E(T, \theta(REW)) \neq \emptyset$.

This completes the proof of Theorem B

4 Proof of the Main Theorem

To prove Theorem A for a fixed game G , we first prove the existence of some T_1 such that $\text{co}(E(T_1))$ has dimension I (Lemma 4.1). Using this, we show how to construct T_2 and payoffs in $T_2E(T_2)$ that define a selection θ of $\Theta^{W, \rho}$ as in Theorem B (Lemma 4.2). Theorem A then appears as a consequence of the fact that the repeated game $G(T+T_2)$ can be viewed as the repeated game with terminal payoffs $\frac{T}{T+T_2}G(T, T_2E(T_2))$, where the factor $\frac{T}{T+T_2}$ is a renormalization factor due to different averaging of payoffs in $G(T+T_2)$ and in $G(T, T_2E(T_2))$.

Lemma 4.1 *Under hypothesis (i) and (ii), there exists $T_1 \in \mathbb{N}$ such that $\dim(\text{co } E(T_1)) = I$.*

Proof:

Let (A_0, A_1, \dots, A_I) be $I + 1$ action profiles of G so that $\dim(\text{co } \{g(A_i), i \in I\}) = I$, and let us consider the strategies defined for $j \in I$ by:

- Play A_j , then P times e^1 , then P times e^2, \dots , then P times e^I .
- If player i deviates at stage 1, play P times f^i instead of e^i later.

These strategies define a subgame perfect equilibrium of $G(IP + 1)$ when P is large enough, since the gain from deviating at first stage is compensated for by the loss due to the repetition of “bad” equilibria later.

Therefore for every k , $\frac{P}{IP+1} \sum_{i=1}^I g(e^i) + g(A_k) \in E(IP + 1)$, thus $\dim(\text{co } E(IP + 1)) = I$.

Lemma 4.2 *Under hypothesis (i) and (ii):*

$$\exists \rho_0 \geq 0, \forall W \in \mathbb{R}, \exists T_2 \in \mathbb{N}, \exists U \in \mathbb{R}^I$$

$$\forall r \in REW, \exists \theta(r) \in T_2 E(T_2), \theta(r) - U \in \mathcal{B}(Wr, \rho_0)$$

Proof:

Let T_1 be given by Lemma 4.2, and $(U_0, U_1, \dots, U_I) \in E(T_1)^{I+1}$ be such that $\dim(\text{co } (U_1 - U_0, \dots, U_I - U_0)) = I$, and let $V_i = U_i - U_0$, so $(V_i)_{i \in I}$ is a basis of \mathbb{R}^I , and put $\rho_0 = \sum_{i=1}^I \|V_i\|$.

For every $W \in \mathbb{R}$ and every $r \in REW$, there exists integer numbers $(\alpha_1(r), \dots, \alpha_I(r))$ such that:

$$\left\| \sum_{i=1}^I \alpha_i(r) V_i - Wr \right\| < \rho_0$$

In fact every ball of radius ρ_0 contains at least one point of the lattice of \mathbb{R}^I generated by $\{V_1, V_2, \dots, V_I\}$. Put $\alpha^0 = \inf_{i,r} \alpha_i(r)$, and $\beta_i(r) = \alpha_i(r) - \alpha_0$, therefore we have:

$$\forall r \in REW \quad \left\| \sum_{i=1}^I \beta_i(r) V_i - (Wr - \alpha^0 \sum_{i=1}^I V_i) \right\| < \rho_0$$

and let be $\gamma = \sup_r \sum_i \beta_i(r)$. For every $r \in REW$,

$$\left\| \sum_{i=1}^I \beta_i(r) U_i + (\gamma - \sum_{i=1}^I \beta_i(r)) U_0 - (Wr + \gamma U_0 - \alpha^0 \sum_{i=1}^I V_i) \right\| < \rho_0$$

Note that, if for $i \in \{1, 2\}$, σ_i is a subgame perfect equilibrium of $G(t_i)$ with vector payoff z_i , then the strategies (σ_1, σ_2) consisting of following σ_1 , then σ_2 define a subgame equilibrium of $G(t_1 + t_2)$ with vector payoff $z_1 + z_2$. Hence for every r in REW , $\sum_i \beta_i(r) U_i + (\gamma - \sum_i \beta_i(r)) U_0$ is in $\gamma T_1 E(\gamma T_1)$. This proves Lemma 4.2 with $T_2 = \gamma T_1$, $\theta(r) = \sum_i \beta_i(r) U_i + (\gamma - \sum_i \beta_i(r)) U_0$ and $U = \gamma U_0 - \alpha^0 \sum_i V_i$.

Proof of Theorem A:

Let be $\rho_0 \geq 0$ fixed by Lemma 4.2. For all $\varepsilon > 0$, let W_0 be given by Theorem B for ρ_0 and $\frac{\varepsilon}{2}$ ($F \cap \mathbb{R}_{++}^I$ is non empty from assumption (i)), then choose U and T_2 that fit Lemma 4.2 for W_0 .

Lemma 4.2 shows the existence of elements $\theta(r) \in T_2 E(T_2)$ that define a selection θ of $\Theta^{W_0, \rho_0} + U$. Since any translation of vector U of the terminal payoffs has no effect on the strategies of the players, from Theorem B we get the existence of T_0 such that:

$$\forall T > T_0, \forall v \in V, \mathcal{B}(v, \frac{\varepsilon}{2}) \cap E(T, \theta(REW)) \neq \emptyset$$

Therefore

$$\forall T > T_0, \forall v \in V, \mathcal{B}(v, \frac{\varepsilon}{2}) \cap E(T, T_2 E(T_2)) \neq \emptyset$$

Every subgame perfect equilibrium of $G(T, T_2 E(T_2))$ extends to a subgame perfect equilibrium of $G(T+T_2)$, moreover for every $T > 0$ $E(T+T_2) = \frac{T}{T+T_2} E(T, T_2 E(T_2))$. Since T_2 is constant and T goes to infinity, this implies the existence of \bar{T} such that:

$$\forall T > \bar{T}, \forall v \in V, \mathcal{B}(v, \varepsilon) \cap E(T + T_2) \neq \emptyset$$

which completes the proof of Theorem A.

5 Conclusion and possible extensions

Thus we proved a Folk Theorem in mixed strategies for finitely repeated games. In this theorem, the limit set of average equilibrium vector payoffs is V . We already noticed that $E(T)$ and $E^*(T)$ are always included in V . Therefore Theorem A gives a full characterization of $\lim_{T \rightarrow \infty} E(T)$ when (i) and (ii) hold. Theorem A may fail without the assumption (i) that implies the existence of at least two Nash equilibria of G , as the counterexample of the prisoner's dilemma shows. We also know that Theorem 1 fails without the "Full Dimensionality" assumption (ii) by an example given by Benoit and Krishna ([2], example 3.2). The method used to prove Theorem A might be extended to other classes of games, like games with signals, and can be used to prove a Folk Theorem in mixed strategies for Overlapping Generations Games [6].

References

- [1] Aumann, R. J. and L. S. Shapley (1976), ‘Long-term competition — A game theoretic analysis’, preprint.
- [2] Benoit, J. P. and V. Krishna (1985), “Finitely repeated games”, *Econometrica*, 53, 905-922.
- [3] Blackwell, D. (1956), “An analog of the minimax theorem for vector payoffs”, *Pacific Journal of Mathematics*, 6, 1-8.
- [4] Fudenberg, D. and E. Maskin (1986), “The Folk Theorem in repeated games with discounting and with incomplete information”, *Econometrica*, 54, 533-554.
- [5] Fudenberg, D. and E. Maskin (1991), “On the dispensability of public randomizations in discounted repeated games”, *Journal of Economic Theory*, 53, 428-438.
- [6] Gossner, O. (1993) “Overlapping Generations Games with Mixed Strategies”, mimeo.
- [7] Kandori, Y. (1992), “Repeated games played by overlapping generations of players”, *Review of Economic Studies*, 59, 81-92.
- [8] Mertens, J. F. , S. Sorin and S. Zamir (1992) “Repeated games”, book to appear.
- [9] Neyman, A. (1988), “Stochastic games”, mimeo.
- [10] Rubinstein, A. (1977), “Equilibrium in Supergames”, CRIMEGT RM 25.

- [11] Smith, L. (1989), “Folk Theorems in overlapping generations games”, *Games and Economic Behavior*, 4, 426-449.
- [12] Sorin, S. (1990), “Supergames”, in “Game theory and applications”, T. Ichiishi, A. Neyman and Y. Tauman eds. , Academic Press, 46-63.