UNIFORM EQUILIBRIUM: MORE THAN TWO PLAYERS

EILON SOLAN
Northwestern University
Evanston, USA
Tel Aviv University
Tel Aviv, Israel

Abstract. The basic question addressed in this chapter is: Does every multi-player stochastic game (with finite state and action spaces) admit a uniform equilibrium payoff? To this day, no counterexample has been found. A positive answer has been given for several special classes, including zero-sum stochastic games [9], two-player non-zero-sum absorbing games [28] and two-player non-zero-sum stochastic games [26]. For multi-player games, existence of stationary equilibrium profiles has been proven for irreducible games [14],[3] and of "almost" stationary equilibrium profiles for games with additive rewards and additive transitions as well as for games with perfect information [25]. In this chapter I review recent results for games with more than two players.

1. Definitions

Recall that a *multi-player stochastic game* is given by (a) a finite set S of states, (b) a finite set I of players, (c) for every player $i \in I$, a finite set A^i of actions, set $A = \times_{i\in I} A^i$, (d) a payoff function $r : S \times A \to \mathbb{R}^I$, and (e) a transition rule $p : S \times A \to \Delta(S)$, where $\Delta(S)$ is the set of probability distributions over S.

We denote by z_n and a_n the state at stage n and the vector of actions chosen by the player at that stage.

A *strategy* for player i is a function $\sigma^i : \cup_{n\in\mathbb{R}} (S \times A)^{n-1} \to \Delta(A^i)$. A profile $\sigma = (\sigma^i)_{i\in I}$ is a vector of strategies, one for each player.

Every profile σ and every initial state z define a probability distribution over the space of plays $(S \times A)^{\mathbb{R}}$. The corresponding expectation operator is $\mathbf{E}_{z,\sigma}$.

For every discount factor $\lambda \in (0,1]$, every profile σ and every initial state z, the *expected λ-discounted payoff* is

$$\gamma_\lambda(z,\sigma) = \mathbf{E}_{z,\sigma}[\lambda \sum_{n\in\mathbb{R}} (1-\lambda)^{n-1} r(z_n,a_n)].$$

For every $N \in \mathbb{R}$, the *expected N-stage payoff* is

$$\gamma_N(z,\sigma) = \mathbf{E}_{z,\sigma}[\frac{1}{N}\sum_{n=1}^{N} r(z_n,a_n)].$$

Definition 1 *A vector $g \in \mathbb{R}^{S\times I}$ is a* uniform equilibrium payoff *if for every $\epsilon > 0$ there exist $\lambda_0 \in (0,1]$, $n_0 \in \mathbb{R}$ and a profile σ that satisfy for every player $i \in I$, every strategy σ'^i of player i, every initial state $z \in S$, every $\lambda \in (0,\lambda_0)$ and every $n \ge n_0$,*

$$\gamma^i_\lambda(z,\sigma) + \epsilon \ge g^i_z \ge \gamma^i_\lambda(z,\sigma^{-i},\sigma'^i) - \epsilon, \text{ and} \tag{1}$$
$$\gamma^i_n(z,\sigma) + \epsilon \ge g^i_z \ge \gamma^i_n(z,\sigma^{-i},\sigma'^i) - \epsilon. \tag{2}$$

A profile σ that satisfies (1) and (2) is a uniform ϵ-equilibrium.

In words, for every $\epsilon > 0$ there is a profile σ which is an ϵ-equilibrium in every discounted game, provided the discount factor is sufficiently small, and in every finite-stage game, provided it is sufficiently long.

The main question we address here is the following.

Question: Does every multi-player stochastic game admit a uniform equilibrium payoff?

Though in recent years existence of a uniform equilibrium payoff has been established in several classes of multi-player stochastic games, this question is still open.

2. An Example

Let us begin with an example of a three-player game, studied by Flesch et al. [5].

C

	C	Q
C	$0,0,0$	$0,1,3$ *
Q	$1,3,0$ *	$1,0,1$ *

Q

	C	Q
C	$3,0,1$ *	$1,1,0$ *
Q	$0,1,1$ *	$0,0,0$ *

In this game player 1 chooses a row, player 2 a column, and player 3 a matrix. An asterisked entry means that the entry is absorbing with probability 1. The non-asterisked entry is non-absorbing. The payoff in each entry is either the non-absorbing payoff or the absorbing payoff, depending on whether the entry is non-absorbing or absorbing.

Flesch et al. [5] proved that the game admits no stationary equilibrium (or stationary ϵ-equilibrium), and that the following cyclic strategy profile is an equilibrium:

1. At the first stage, the players play $(\frac{1}{2}C + \frac{1}{2}Q, C, C)$.
2. At the second stage, the players play $(C, \frac{1}{2}C + \frac{1}{2}Q, C)$.
3. At the third stage, the players play $(C, C, \frac{1}{2}C + \frac{1}{2}Q)$.
4. Afterwards, the players play cyclically those three mixed-action combinations, until absorption occurs.

If the players follow this profile then their expected payoff g satisfies: $g = \frac{1}{2}(1,3,0) + \frac{1}{4}(0,1,3) + \frac{1}{8}(3,0,1) + \frac{1}{8}g$, hence $g = (1,2,1)$.

Let us verify that this profile is an equilibrium.Since both the payoffs in this game and the profile we defined are cyclic, the conditional expected payoff for the players is $(1,1,2)$ if the realized action of player 1 at the first stage is C. Since the payoff for the players is $(1,3,0)$ if the realized action of player 1 at the first stage is Q, player 1 is indifferent between his two actions at the first stage. Moreover, the conditional expected payoff of player 2 is $\frac{1}{2} \times 1 + \frac{1}{2} \times 0 = \frac{1}{2}$ if he plays Q at the first stage, and $\frac{1}{2} \times 1 + \frac{1}{2} \times 3 = 2$ if he plays C, and the conditional expected payoff of player 3 is $\frac{1}{2} \times 1 + \frac{1}{2} \times 1 = 1$ if he plays Q at the first stage, and $\frac{1}{2} \times 2 + \frac{1}{2} \times 0 = 1$ if he plays C. Thus, if everyone follows this profile from the second stage on, no one can profit by deviating at the first stage. Since both the profile and the payoff structure are cyclic, similar analysis holds for all stages. Therefore, no player can profit by deviating in any finite number of stages. Since the profile is absorbing given *any* unilateral deviation, it follows that no player can profit by any type of deviation.

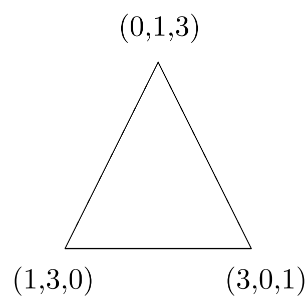

Figure 1: The payoff space

We shall now see a geometric presentation of this result (which was suggested by Nicolas Vieille). We are looking for an equilibrium payoff in the convex hull of $\{(1,3,0),(0,1,3),(3,0,1)\}$. In particular, this means that at most one player plays Q at every stage. Indeed, the sum of the coordinates of any vector in this convex hull is 4, while if at least two players play Q at the same stage, the sum of the coordinates of the corresponding absorbing payoff is at most 2. The convex hull is depicted in Figure 1.

Let us draw the indifference lines of the players; that is, line i includes all payoffs where player i receives 1. Each indifference line divides the convex hull into two halves: the payoffs that are "good" for the player, and the payoffs that are "bad" for the player. The diagram looks as follows.

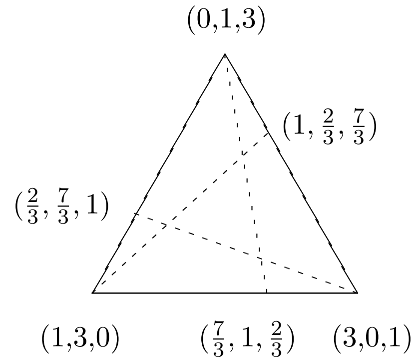

Figure 2: The payoff space with indifference lines

Assume that an equilibrium profile is given and that player 1 plays at the first stage the action Q with probability $p \in (0,1)$. This means that both the equilibrium payoff and the continuation payoff are *on the indifference line of player 1.* The probability of player 1 playing Q determines the distance between those two points. Similarly, if player i plays at stage n the action Q with probability strictly between 0 and 1, then both the payoff conditional on nonabsorption before stage n and the payoff conditional on nonabsorption before stage $n+1$ are on the indifference line of player i.

Thus, the continuation payoff at stage n must be on the edges of the dashed triangle in Figure 2, and any point on its three edges is an equilibrium payoff (this was proven for this game by Flesch et al. [5]). It turns out that the three extreme points of the dashed triangle in Figure 2 are $(1,2,1)$, $(1,1,2)$ and $(2,1,1)$.

For general payoff structure (with the same absorbing structure), as long as the intersection of the three "good" halves (the dashed triangle in Figure 2) is nonempty, there exists an equilibrium. Later we will see that if this intersection is empty, there exists a stationary equilibrium.

One might wonder whether this argument holds for four-player games as well. Unfortunately, the answer is negative. For four-player games we do not necessarily have such a cycle; an example is given in [21].

3. Three-Player Absorbing Games

In the present section we discuss the following generalization of the result of Flesch et al. [5].

Theorem 1 *(Solan [15]) Every three-player absorbing game admits a uniform equilibrium payoff.*

Proof. The basic idea is to follow Vrieze and Thuijsman's [28] proof for two-player absorbing games (see also [24]). Two difficulties that require special attention will arise.

Recall that $r^i(a)$ is the non-absorbing (daily) payoff to player i if the action profile a is played, $r^i_\star(a)$ is the absorbing payoff, and $p(a)$ is the probability that the game is absorbed if the action profile a is played. See [24] for a formal definition of an absorbing game.

Let v^i be the minmax value of player i. As proven in [12], v^i exists and is the limit of the λ-discounted minmax value of player i as λ goes to 0.

Step 1: Simple absorbing structure and low non-absorbing payoff.

We first deal with games that have the absorbing structure as in the above example (that is, each player has two actions, and only one entry is non-absorbing). Moreover, we assume that the non-absorbing payoffs are always below the minmax value; that is, $r^i(a) \leq v^i$ for every action profile a.

Let x_λ be a stationary λ-discounted equilibrium with a corresponding payoff $g_\lambda = \gamma_\lambda(x_\lambda)$. Using the algebraic approach [11], we assume that x_λ and g_λ are Puiseux functions of λ. Let x_0 and g_0 be the corresponding limits as $\lambda \to 0$.

Step 1a: x_0 is absorbing.

As in Vrieze and Thuijsman [28] (see also [24]), if x_0 is absorbing, then it induces an "almost" stationary equilibrium that yields a payoff g_0.

Step 1b: x_0 is non-absorbing (that is, $x_0 = (C, C, C)$).

As in Vrieze and Thuijsman [28] (see also [24]), for every λ we have

$$g_\lambda = p(x_\lambda) r_\star(x_\lambda) + (1 - p(x_\lambda)) \left(\lambda r(x_\lambda) + (1-\lambda) g_\lambda \right).$$

Solving this equation, we get that the λ-discounted payoff is a convex combination

$$g_\lambda = \alpha_\lambda r(x_\lambda) + (1 - \alpha_\lambda) r_\star(x_\lambda),$$

where

$$\alpha_\lambda = \frac{\lambda(1 - p(x_\lambda))}{p(x_\lambda) + \lambda(1 - p(x_\lambda))}. \tag{3}$$

Taking the limit as $\lambda \to 0$ gives

$$g_0 = \alpha_0 r(x_0) + (1 - \alpha_0) \lim_{\lambda \to 0} r_\star(x_\lambda), \tag{4}$$

where $\alpha_0 = \lim_{\lambda\to 0} \alpha_\lambda$.

If $\alpha_0 = 1$, then $g_0 = r(x_0)$. In this case, x_0 induces an "almost" stationary equilibrium. Indeed, if player i can profit by deviating, then by continuity arguments this deviation is profitable also against x_λ^{-i} in the λ-discounted game, for λ sufficiently small. We assume now that $\alpha_0 < 1$.

Since $r^i(x_0) \leq v^i = \lim_{\lambda\to 0} v^i_\lambda \leq \lim_{\lambda\to 0} g^i_\lambda = g^i_0$ and $\alpha_0 < 1$, (4) implies that

$$r(x_0) \leq g_0 \leq \lim_{\lambda\to 0} r_\star(x_\lambda). \tag{5}$$

Since x_0 is non-absorbing, $\lim_{\lambda\to 0} r_\star(x_\lambda)$ is in the convex hull of the three entries neighboring the non-absorbing entry. Equation (5) implies that if player 1 plays Q with positive probability under x_λ, for every λ sufficiently small, then $\lim_{\lambda\to 0} r^1_\star(x_\lambda) \geq g^1_0 = \lim_{\lambda\to 0} g^1_\lambda = r^1_\star(Q, C, C)$. Similar inequalities hold for the other two players. In particular, the dashed triangle in Figure 2 is nonempty, as $\lim_{\lambda\to 0} r_\star(x_\lambda)$ is in this triangle, and we can construct a cyclic equilibrium, as in the example.

Step 2: General non-absorbing payoffs (with the special absorbing structure).

Note that if the payoff in the non-absorbing entry is good, that is, if $r^1(C, C, C) \geq r^1_\star(Q, C, C)$, $r^2(C, C, C) \geq r^2_\star(C, Q, C)$ and $r^3(C, C, C) \geq r^3_\star(C, C, Q)$, then $r(C, C, C)$ is an equilibrium payoff that corresponds to the stationary strategy profile (C, C, C).

Define an auxiliary game $\tilde{G}$ where the daily payoff for player i is $\tilde{r}^i(x) = \min\{r^i(x), v^i\}$ if the *mixed action profile* x is played. Formally, for every stationary profile x, the λ-discounted payoff in $\tilde{G}$ is given by

$$\tilde{\gamma}^i_\lambda(x) = \lambda \mathbf{E}_x \left[\sum_{t=1}^{\infty} (1-\lambda)^{t-1} \left(\tilde{r}^i(x) 1_{t<t_\star} + r^i_\star(x) 1_{t\geq t_\star} \right) \right],$$

where $t_\star$ is the stage of absorption.

Since $\tilde{r}^i$ is continuous over the strategy space, the λ-discounted minmax value of player i in $\tilde{G}$, denoted by $\tilde{v}^i_\lambda$, exists. One can prove that the sequence $(\tilde{v}^i_\lambda)_\lambda$ converges to v^i, the minmax value of player i in the original game.

The function $\tilde{r}^i$ is quasi-concave and continuous; hence $\tilde{G}$ has a λ-discounted stationary equilibrium. The function $\tilde{r}$ is semialgebraic; hence one can choose for every λ a λ-discounted stationary equilibria x_λ in $\tilde{G}$ such that the mapping $\lambda \mapsto x_\lambda$ is a Puiseux function.

We now repeat the same analysis as in Step 1. Denote $g_\lambda = \tilde{\gamma}^i_\lambda(x_\lambda)$, and let $x_0 = \lim_{\lambda\to 0} x_\lambda$ and $g_0 = \lim_{\lambda\to 0} g_\lambda$.

If x_0 is absorbing, then it induces an "almost" stationary equilibrium, as in Step 1a. Thus, we assume that x_0 is non-absorbing. Then

$$g_0 = \alpha_0 \tilde{r}(x_0) + (1-\alpha_0) \lim r_\star(x_\lambda), \tag{6}$$

where $\alpha_0 = \lim_{\lambda\to 0}\alpha_\lambda$ and α_λ is given in (3). If $\alpha_0 = 1$ then $g_0 = \tilde{r}(x_0) \leq r(x_0)$, and, as in Step 1b, supplementing x_0 with threat strategies yields that $r(x_0)$ is a uniform equilibrium payoff. Otherwise, $\tilde{r}(x_0) \leq v = \lim_{\lambda\to 0} v_\lambda \leq \lim_{\lambda\to 0} g_\lambda = g_0$ and therefore by (6) $g_0 \leq \lim_{\lambda\to 0} r_\star(x_\lambda)$. Thus, the intersection in Figure 2 is nonempty, and there exists a cyclic equilibrium.

Step 3: General three-player absorbing game.

It is convenient here to view the absorbing game as a stochastic game with initial state z_0, such that all other states are absorbing.

We define the auxiliary game as in Step 2, and choose a Puiseux function $\lambda \mapsto x_\lambda$, where x_λ is a stationary equilibrium in the game $\tilde{G}$. We define $x_0 = \lim_{\lambda\to 0} x_\lambda$ and $g_0 = \lim_{\lambda\to 0}\tilde{\gamma}_\lambda(x_\lambda)$. Equation (6) holds in this more general setup.

If x_0 is absorbing, then as in Step 1a it induces an "almost" stationary equilibrium that yields the players an expected payoff g_0. Assume now that x_0 is non-absorbing. If $\alpha_0 = 1$, then as in Step 1b $r(x_0)$ is a uniform equilibrium payoff.

Since x_0 is non-absorbing, the non-absorbing state forms a weak communicating set under x_0. In [20] we defined exit distributions from communicating sets. Let $\mathcal{Q}(x_0)$ be the set of exit distributions. We denote by $\mathcal{Q}_i(x_0)$ the set of unilateral exits of player i (in [20] we had only two players) and by $\mathcal{Q}_0(x_0)$ the set of joint exits (exits that require perturbations of at least two players). Any exit $Q = \sum_{l\in L}\eta_l P_l \in \mathcal{Q}(x_0)$ can be decomposed (not necessarily uniquely) to a sum $Q = \sum_i \sum_{l\in L_i}\eta_l P_l + \sum_{l\in L_0}\eta_l P_l$, where $P_l \in Q_i(x_0)$ for $l \in L_i$, $i = 0,1,2,3$.

Since the setup is of absorbing games, an exit yields a terminal payoff. Define the set of *terminal payoffs* w.r.t. x_0 by

$$\mathcal{E}(x_0) = \{P \cdot r_\star \mid P \in \mathcal{Q}(x_0)\}.$$

It is easy to see that

$$\lim_{\lambda\to 0} r_\star(x_\lambda) \in \mathcal{E}(x_0).$$

Lemma 5.2 in [20] translates to

Lemma 1 *Let $Q = \sum_l \eta_l P_l \in \mathcal{Q}(x_0)$ with a decomposition (L_i), and let $g = \sum_l \eta_l P_l \cdot r_\star$. Let $\gamma : S \to \mathbb{R}^N$ coincide with $r_\star$ on $S \setminus z_0$ and $\gamma(z_0) = g$. If the following conditions hold*

1. *for every $l \in L_i$, $P_l \cdot \gamma^i_\star = g^i$;*
2. *for every player i and action a^i, $p(\cdot \mid x_0^{-i}, a^i)\gamma^i \leq g^i$;*

then g is an equilibrium payoff.

The last step of the proof is a combinatorial lemma that shows that if there is no cyclic equilibrium, then there exists $Q \in \mathcal{Q}(x_0)$ that satisfies Lemma 1. ∎

The same technique, without the need of the combinatorial lemma, proves the following result. An absorbing game is a *team game* if the players are divided into two teams, and the players in each team have the same payoffs (both absorbing and non-absorbing).

Theorem 2 *(Solan [16]) Every absorbing team game admits a uniform equilibrium payoff.*

Definition 2 *A strategy profile σ is* (x, ϵ)-perturbed *if it has the following structure:*

1. *Every player is checked by a statistical test.*
2. *As long as no player fails the statistical test, the mixed action profile prescribed by σ is ϵ-close to x (in the supremum topology).*
3. *The first player who fails the statistical test is punished with an ϵ-minmax profile forever.*

An ϵ-equilibrium profile σ is perturbed *if there exists a stationary profile x such that σ is (x, ϵ)-perturbed.*

In all classes of non-zero-sum stochastic games seen thus far where the uniform equilibrium is known to exist, there are ϵ-equilibrium profiles that are perturbed.

The importance of having a perturbed equilibrium is the relative simplicity of the ϵ-equilibrium profiles, as well as the method of the approach one should take to prove existence. Most of the proofs we have seen take a sequence of stationary equilibria in ϵ-approximating games that converge when ϵ goes to 0. Mertens and Neyman [9], Vrieze and Thuijsman [28] and Vieille [27] consider the discounted game, Solan [15] considers the discounted version of a variation of the game, and Flesch et al. [4], Vieille [26] and Solan [17] consider approximating games where the players have constrained strategy spaces. In the ϵ-equilibrium profile the players play mainly the limit stationary strategy, and perturb to other actions with small probability. In particular, the ϵ-equilibrium profile is perturbed.

Solan and Vieille [21] constructed a four-player game that has no perturbed equilibrium payoff. In particular, this example hints that the classical approach may not work in general.

4. Quitting Games

We look for a new approach to deal with multi-player games. Here we consider a simple class of games, namely quitting games. A *quitting game* is an absorbing game where each player has two actions: *continue* and *quit*. If everyone continues the game continues to the next stage (terminating with probability 0) and the daily payoff is 0. If at least one player quits, the game terminates with probability 1. The three-player game studied in Section 1 is a quitting game.

Theorem 3 *(Solan and Vieille [21]) Every quitting game that satisfies the following two conditions*

1. *if a single player quits, he receives 1;*
2. *if player i quits with some other players, he receives at most 1;*

admits a subgame-perfect ϵ-equilibrium payoff. Moreover, there is a cyclic ϵ-equilibrium strategy profile, but the length of the cycle can depend on ϵ.

Proof. The approach taken in the proof of Theorem 3 is different from the classical one. Instead of defining the best-reply correspondence, and looking for a fixed point, we look for a sequence $g_1, g_2, \ldots$ of payoff vectors such that g_k is an equilibrium payoff in the one-shot game with continuation payoff g_{k+1}. Denote by x_k the corresponding equilibrium strategy profile in this one-shot game. One can verify that if such a sequence exists, and if the sequence $(x_1, x_2, \ldots)$ is terminating with probability 1, then the profile $(x_1, x_2, \ldots)$ is an equilibrium of the quitting game.

Unfortunately, such a sequence may fail to exist; see Solan [19].

For every payoff vector $w \in \mathbb{R}^I$, let $G(w)$ be the one-shot game derived from the quitting game with continuation payoff w; that is, we replace the non-absorbing entry by an absorbing entry with absorbing payoff w. Let

$$W = \{w \in [-\rho, \rho]^I \mid w^i \leq 1 \text{ for at least one } i\}$$

where ρ is the maximal payoff in the game (in absolute values).

We denote by $\langle G(w), x\rangle^i$ the payoff of player i in the one-shot game $G(w)$ if the mixed action profile x is played.

Definition 3 *The action a^i is an* ϵ-best reply *against x^{-i} if*

$$\langle G(w), x^{-i}, a^i\rangle \geq \max_{b^i} \langle G(w), x^{-i}, b^i\rangle - \epsilon.$$

The mixed action profile x is an ϵ-equilbirium *of $G(w)$ if for every i and every $a^i \in \mathrm{supp}(x^i)$, a^i is an ϵ-best reply against x^{-i}.*

Lemma 2 *For every $w \in W$, the game $G(w)$ possesses a $\rho\epsilon$-equilibrium that is absorbing with probability at least ϵ. Moreover, the corresponding $\rho\epsilon$-equilibrium payoff is in W.*

Proof. Let x be an equilibrium payoff in $G(w)$. If x prescribes all players to continue, then the corresponding equilibrium payoff is w. In particular, there exists a player i with $w^i \leq 1$, and since w is an equilibrium payoff $w^i = 1$. If x does not prescribe all players to continue, then by the second condition there exists a player i who plays in x a fully mixed action, and therefore is indifferent between his two actions. Let $y^j = x^j$ for all $j \neq i$, and $y^i = \min\{x^i + \epsilon, 1\}$ otherwise (increase the probability to quit by ϵ). The profile y is a $\rho\epsilon$-equilibrium that is absorbing with probability at least ϵ. Moreover, the corresponding payoff is in W, since the expected payoff of player i is at most 1. ∎

Define the correspondence $\phi : W \to W$ as follows.For every $w \in W$, $\phi(w)$ is the set of all payoff vectors $\langle G(w), x\rangle$, where x is a $\rho\epsilon$-equilibrium which is absorbing with probability of at least ϵ; that is, under x absorption occurs at every stage with probability at least ϵ. ϕ is uppersemicontinuous, and by Lemma 2 it has nonempty values.

Lemma 3 *For every uppersemicontinuous correspondence ϕ with nonempty values from a compact set W into itself there exists a sequence $w_1, w_2, \ldots$ such that for each i, $w_i \in \phi(w_{i+1})$.*

It is straightforward to generate a sequence $w_1, w_2, \ldots$ such that $w_{i+1} \in \phi(w_i)$, but for our purposes the sequence should satisfy $w_i \in \phi(w_{i+1})$.

Proof. Define $W_0 = W$ and $W_{i+1} = \phi(W_i)$. Since W is compact and ϕ uppersemicontinuous with nonempty values, W_i is compact. By induction, $W_{i+1} \subseteq W_i$; hence $W_\infty = \cap W_i \neq \emptyset$.

Let $w_1 \in W_\infty$. For each i choose a sequence $w_1 = w_i^1, w_i^2, \ldots, w_i^i$ such that $w_i^j \in \phi(w_i^{j+1})$. By taking a subsequence, assume that $w_\infty^j = \lim_{i\to\infty} w_i^j$ exists for all j. By uppersemicontinuity, the sequence $(w_\infty^j)_j$ satisfies the lemma. ∎

We have generated a sequence of continuation payoffs (w_i) such that w_i is a payoff that corresponds to a $\rho\epsilon$-equilibrium in $G(w_{i+1})$ in which the per-stage probability of absorption is at least ϵ. Let x_i be the corresponding $\rho\epsilon$-equilibrium. The profile $\mathbf{x} = (x_1, x_2, \ldots)$ is a natural candidate for an ϵ'-equilibrium in the quitting game. However, if players follow $\mathbf{x}$ then at *every* stage each player may profit $\rho\epsilon$ by deviating. How do we know that these small profits do not aggregate?

One can now prove that either $(x_1, x_2, \ldots)$ is an ϵ'-equilibrium, or there exists a player i such that if i quits alone, all players get at least 1; hence a stationary ϵ-equilibrium exists. ∎

Problem. Can one bound the length of the cycle?

Problem. Does the result hold for general quitting games (without the two assumptions)?

5. Correlated Equilibrium

Correlation devices were introduced by Aumann [1], [2] for games in normal form. A correlation device chooses a private signal for every player before the start of play, and sends to each player the signal chosen for him. Each player can base his choice of an action on the private signal that he has received.

For multi-stage games, various generalizations of correlation devices have been introduced. The most general device receives at every stage a private message from each player and sends in return a private signal to each player (called a "communication device" [6], [7], [10], [8].) The most restrictive device, as in Aumann's definition, sends one signal before the start of play, and no signals are sent once play begins (called a "correlation device").

In between there are devices that base their choice on past signals that were sent, but not on past play (called an "autonomous correlation device" [6]), and devices that base their choice only on the current state (and not even on past signals) (called a "weak correlation device" [13]).

Here we concentrate on two types of correlation devices: (i) stationary devices, which choose at every stage a signal according to the same probability distribution, independently of any data, and (ii) autonomous devices, which base their choice of new signal on previous signals, but not on any other information.

Theorem 4 *(Solan and Vieille [22]) Every multi-player stochastic game admits a uniform correlated equilibrium payoff, using an autonomous correlation device. The equilibrium path is sustained using threat strategies, but punishment occurs only if a player disobeys the recommendation of the device.*

A stronger result is possible for positive recursive games (i.e., recursive games where the payoff in absorbing states is non-negative for all players).

Theorem 5 *(Solan and Vieille [22]) If the game is positive and recursive, then the correlation device can be taken to be stationary.*

The proofs utilize various methods that we have already seen, and one new idea. They are divided into two steps. First we construct a "good" strategy profile σ; meaning, a strategy profile that yields all players a high payoff, and no player can profit by a unilateral deviation that is followed by an indefinite punishment.

The construction of the "good" strategy profile uses the method of Mertens and Neyman [9] for Theorem 4, and a variant of the method of Vieille [26] for Theorem 5.

Second, we follow Solan [18] and define a correlation device that *mimics* that strategy profile: the device chooses a pure action profile according to the probability distribution given by the strategy profile, and recommends each player to play "his" action in this action profile. More formally, at each stage n, for every history h of length n and for every player i, the device chooses an action $a^i(h) \in A^i$ according to $\sigma^i(h)$. All choices are made independently of each other. The device then sends to each player i the collection of signals $(\sigma^i(h))_{h\in(S\times A)^{n-1}\times S}$. The players, who observe past play, know which history h has been realized, and each player i plays the action $a^i(h)$ that was recommended for him. To deter deviations, the device reveals to all players what its recommendations were in the previous stage. This way, a deviation is detected immediately, and can be punished. In particular, the device that we construct is not canonical (see [7]).

Problem. Does any stochastic game (or positive recursive game) admit a correlated equilibrium payoff, i.e, where the device sends only one signal before the start of play?

Solan and Vohra [23] answered this question affirmatively for absorbing games.

References

1. Aumann, R.J. (1974) Subjectivity and correlation in randomized strategies, *Journal of Mathematical Economics* **1**, 67–96.
2. Aumann, R.J. (1987) Correlated equilibrium as an expression of Bayesian rationality, *Econometrica* **55**, 1–18.
3. Federgruen, A. (1978) On N-person stochastic games with denumerable state space, *Advances in Applied Probability* **10**, 452–471.
4. Flesch, J., Thuijsman, F. and Vrieze, O.J. (1996) Recursive repeated games with absorbing states, *Mathematics of Operations Research* **21**, 1016–1022.
5. Flesch, J., Thuijsman, F. and Vrieze, O.J. (1997) Cyclic Markov equilibria in stochastic games, *International Journal of Game Theory* **26**, 303–314.
6. Forges, F. (1986) An approach to communication equilibria, *Econometrica* **54**, 1375–1385.
7. Forges, F. (1988) Communication equilibria in repeated games with incomplete information, *Mathematics of Operations Research* **13**, 77–117.
8. Mertens, J.-F. (1994) Correlated and communication equilibria, in J.-F. Mertens and S. Sorin (eds.), *Game-Theoretic Methods in General Equilibrium Analysis*, Kluwer Academic Publishers, Dordrecht, pp. 243–248.
9. Mertens, J.-F. and Neyman, A. (1981) Stochastic games, *International Journal of Game Theory* **10**, 53–66.
10. Myerson, R.B. (1986) Multistage games with communication, *Econometrica* **54**, 323-358.
11. Neyman, A. (2003) Real algebraic tools in stochastic games, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Kluwer Academic Publishers, Dordrecht, Chapter 6, pp. 57–75.
12. Neyman, A. (2003) Stochastic games: Existence of the minmax, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Kluwer Academic Publishers, Dordrecht, Chapter 11, pp. 173–193.
13. Nowak, A.S. and Raghavan, T.E.S. (1992) Existence of stationary equilibria with symmetric information for discounted stochastic games, *Mathematics of Operations Research* **17**, 519–526.
14. Sobel, M.J. (1971) Noncooperative stochastic games, *The Annals of Mathematical Statistics* **42**, 1930–1935.

15. Solan, E. (1999) Three-person absorbing games, *Mathematics of Operations Research* **24**, 669–698.
16. Solan, E. (2000) Absorbing team games, *Games and Economics Behavior* **31**, 245–261.
17. Solan, E. (2000) Stochastic games with two non-absorbing states, *Israel Journal of Mathematics* **119**, 29–54.
18. Solan, E. (2001) Characterization of correlated equilibria in stochastic games, *International Journal of Game Theory* **30**, 259–277.
19. Solan, E. (2001) The dynamics of the Nash correspondence and n-player stochastic games, *International Game Theory Review* **3**, 291–300.
20. Solan, E. (2003) Perturbations of Markov chains with applications to stochastic games, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Kluwer Academic Publishers, Dordrecht, Chapter 17, pp. 265–280.
21. Solan, E. and Vieille, N. (2001) Quitting games, *Mathematics of Operations Research* **26**, 265–285.
22. Solan, E. and Vieille, N. (2002) Correlated equilibrium in stochastic games, *Games and Economic Behavior* **38**, 362–399.
23. Solan, E. and Vohra, R.V. (2002) Correlated equilibrium payoffs and public signalling in absorbing games, *International Journal of Game Theory* **31**, 91-121.
24. Thuijsman, F. (2003) Repeated games with absorbing states, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Kluwer Academic Publishers, Dordrecht, Chapter 13, pp. 205–213.
25. Thuijsman, F. and Raghavan, T.E.S. (1997) Perfect information stochastic games and related classes, *International Journal of Game Theory* **26**, 403–408.
26. Vieille, N. (2000) Equilibrium in 2-person stochastic games II: The case of recursive games, *Israel Journal of Mathematics* **119**, 93–126.
27. Vieille, N. (2000) Small perturbations and stochastic games, *Israel Journal of Mathematics* **119**, 127–142.
28. Vrieze, O.J. and Thuijsman, F. (1989) On equilibria in repeated games with absorbing states, *International Journal of Game Theory* **18**, 293–310.