Acknowledgment

First of all, I thank my wife, Alex, for listening and supporting in the last couple of years.
I deeply thank my advisor, Prof. Abraham Neyman, whose help, advice and supervision was invaluable.
I also thank Janos Flesch, with whom I discussed three-player games, three anonymous referees of *Mathematics of Operations Research*, whose comments substantially improved the presentation of the results in section 4, and Nicolas Vieille, who commented on the presentation of the results in section 5.
Finally, I would like to thank the *Center of Rationality and Interactive Decision Theory*, Sergiu Hart and Hana Shemesh. Their hospitality and support during the period this research took place is greatly acknowledged.

Contents

5 Recursive Games with the Absorbing Property 71

1 Introduction

The present monograph deals with uniform equilibria in stochastic games. In the Introduction we review informally the basic definitions and the known results in this topic. We begin by introducing the model, continue with the known results for the discounted and the uniform equilibria, and finally we review the main results and ideas presented in the monograph.

1.1 The Model

A stochastic game is played in stages. At every stage the game is in some state of the world, and each player, given the whole history (including the current state), chooses an action in his action space. The action combination that was chosen by all the players, together with the current state, determine the daily payoff that each player receives and the probability distribution according to which the new state of the game is chosen.

The results that appear in this monograph are for stochastic games where the number of players, states and available actions are finite.

One may consider a finite t-stage game — the game terminates after t stages, and the payoff for the players is their average payoff. Another standard version is the infinite discounted game, where the infinite sequence of payoffs $(r_1^i, r_2^i, \ldots)$ of player i is evaluated by the discounted sum

$$(1-\beta)\sum_{t=1}^{\infty}\beta^{t-1}r_t^i,$$

and $\beta \in [0,1)$ is the discount factor.

It is fairly easy to prove (i) the existence of a stationary equilibrium profile for the discounted case (stationary, in the sense that the mixed action that is chosen by each player at every stage depends only on the current state, rather than on the whole history), and (ii) the existence of an equilibrium profile (which usually depends on the stage of the game, as well as on the current state) for finite t-stage games. However, the equilibrium profiles in both cases usually depend on the exact discount factor or on the exact duration of the game; an equilibrium profile for one discount factor might yield some players a low payoff if the discount factor is slightly changed and the equilibrium profiles for finite t-stage games differ for every t.

A strategy profile is a *uniform ϵ-equilibrium* if no player can profit more than ϵ by deviating in *any* sufficiently long game, or in *any* discounted game, for discount factor sufficiently close to 1. Moreover, no player can profit more than ϵ in the infinite game as well (a precise definition is given in section 3.4).

Aumann and Maschler [2] mention several reasons to study the uniform equilibrium:

1. A uniform ϵ-equilibrium profile is an ϵ-equilibrium profile in any finite game whose duration is sufficiently long, as well as in any discounted game, when the discount factor is sufficiently close to 1. Thus, the uniform equilibrium can be used
    - if the game is "long", but its the exact duration is not known, or if the players are sufficiently patient, but the exact discount factor, is not known,
    - or when the players have bounded rationality, and taking the stage of the game into account is too complex.
2. Uniform ϵ-equilibrium profiles are usually simple to describe, and the players follow rules that make use only of fairly simple statistics of the past history. Thus, uniform equilibria lead to simple "rules of thumb" for the players.
3. To study optimal behavior for the players in games that continue indefinitely, and the payoff of any finite number of stages is negligible.
4. Since uniform ϵ-equilibrium profiles depend only on the structure of the game (rather than on its duration or the discount factor), one can compare optimal behavior of different games — how does an additional player changes the optimal behavior, or the addition of states or actions for the players.

In the following two subsections we introduce the discounted and the uniform equilibria, and the results concerning these two types of equilibria. The proofs in the monograph make use of the discounted equilibria, rather then equilibria in finite t-stage games. However, bearing in mind the results of Bewley and Kohlberg [4], one could use finite games instead of discounted games.

1.2 The Discounted Equilibrium

Let $0 \leq \beta < 1$. In the β-discounted game each player i evaluates a profile σ by

$$v_\beta^i(s,\sigma) = \mathbf{E}_{s,\sigma}\left((1-\beta)\sum_{t=1}^{\infty}\beta^{t-1}r_t^i\right)$$

where s is the initial state and r_t^i is the daily payoff that player i receives at stage t.

A strategy profile σ is a *β-discounted (Nash) equilibrium* if $v_\beta^i(s,\sigma) \geq v_\beta^i(s,\sigma^{-i},\tau^i)$ for every player i and every strategy τ^i of player i, where $\sigma^{-i} = (\sigma^j)_{j\neq i}$. The payoff vector $v_\beta(\cdot,\sigma)$ is a *β-discounted equilibrium payoff.* It is easy to verify that if the equilibrium payoff of a two-player zero-sum game exists, then it is unique. In this case, the unique equilibrium payoff of player 1 is the *discounted value* of the game.

Shapley [26] introduced the model of stochastic games and proved that every two-player zero-sum discounted stochastic game has a discounted value. Moreover, there are stationary equilibrium profiles. Fink [12] generalized this result for n-player stochastic games.

Bewley and Kohlberg [4] proved that the value of a two-player zero-sum discounted game, as a function of the discount factor, is a Puiseux function (that is, it has an expansion as a Laurent series in fractional powers). In particular, it follows that the limit of the discounted value, as the discount factor tends to 1, exists. Moreover, they proved that there exists a Puiseux function that assigns to every discount factor a stationary equilibrium profile in the corresponding discounted game. Following similar lines it can be proven (see, e.g., Mertens, Sorin and Zamir [21]) that for every n-player stochastic game there exists a Puiseux function that assigns for every discount factor an equilibrium payoff and an equilibrium strategy profile in the corresponding discounted game. As we will see, this result is used extensively for proving results on the uniform equilibrium payoff.

1.3 The Uniform Equilibrium

1.3.1 Uniform Equilibrium Payoff

A payoff vector $g = (g_s^i)$ is a *uniform ϵ-equilibrium payoff* if there exists a strategy profile σ and a finite horizon $t_e \in \mathbf{N}$ such that for every player i, every strategy τ^i of player i and every initial state s

- If the players follow the profile σ, then the expected average daily payoff for player i in every finite t stage game (for $t > t_e$), as well as the expected $\liminf$ of the average daily payoffs, is at least $g^i_s - \epsilon$.
- If the players follow the profile (σ^{-i}, τ^i), then the expected average payoff of player i in every finite t stage game (for $t > t_e$), as well as the expected $\limsup$ of the average daily payoffs, is at most $g^i_s + \epsilon$.

The strategy profile σ is an *ϵ-equilibrium profile.*

It can easily be proven that σ is also an ϵ-equilibrium profile in the discounted game, for discount factors sufficiently close to 1.

A payoff vector g is a *uniform equilibrium payoff* if it is a uniform ϵ-equilibrium payoff for every $\epsilon > 0$. If the game is two-player zero-sum and it has a uniform equilibrium payoff, then the uniform equilibrium payoff is unique. In this case the unique equilibrium payoff of player 1 is the (uniform) *value* of the game.

Many authors used the undiscounted equilibrium (the players evaluate a stream of payoffs by its $\liminf$, $\limsup$ or some other Banach limit) rather than the uniform equilibrium. Several results that were proved for undiscounted equilibria hold for uniform equilibria as well, with minor modifications in the proofs. In the sequel we mention these results as proved for uniform equilibria.

1.3.2 Special Classes of Stochastic Games

A state is *absorbing* if once it is reached, the probability to leave it, whatever the players play, is 0.

A *recursive game* is a stochastic game where the payoff for the players in all the non-absorbing states is identically 0, whatever actions the players play.

A *repeated game with absorbing states* is a stochastic game where all the states but one are absorbing.

A stochastic game is of *perfect information* if in every state only one player has a non-degenerate action space (that is, only one player has more than one possible action), and of *switching control* if in every state only one player controls the transitions.

An *irreducible game* is a stochastic game where the game reaches every state infinitely often, whatever the players play.

Stochastic team games are stochastic games where the players are divided into two teams, and the players of each team have the same payoff function. Since different players in the same team cannot correlate their actions, the strategy space of a team, viewed as a single player, is strictly larger than the product of the strategy spaces of the players of that team (provided that the team consists of at least two players). Therefore one cannot deduce trivially results on team games from the corresponding results on two-player games.

1.3.3 Zero-Sum Games

The study of the undiscounted evaluation has begun by Everett [10] and Gillette [15].

Everett [10] proved that for two-player zero-sum recursive games the value exists, and there are stationary ϵ-equilibrium profiles.

Gillette [15] studied games of perfect information and irreducible games, and proved the existence of stationary equilibrium profiles in both cases. Gillette introduced the following example of the "Big Match" which is a two-player zero-sum repeated game with absorbing states.

EXAMPLE 1 *The "Big Match"*

	L	R
T	$-1, 1$ *	$1, -1$ *
B	$1, -1$	$-1, 1$

Player 1 is the row player, while player 2 is the column player. The daily payoff is as indicated by the cell. An cell marked with an asterisk means that once this cell is reached, then the game moves to an absorbing state, where the payoff for the players is as indicated by the cell. If an unmarked cell is chosen, then the game remains at the same state.

Gillette proves that in this game there is no stationary equilibrium profile.
■

Nevertheless, Blackwell and Ferguson [5] proved that the value of the "Big Match" is 0 (that is, the value of the non-absorbing state), and they constructed ϵ-equilibrium profiles. In the ϵ-equilibrium profile, player 2 plays at every stage with equal probability his two actions, and player 1 plays a history-dependent strategy. An ϵ-equilibrium strategy that Blackwell and Ferguson suggest for player 1 is to play T at stage t with probability $\frac{1}{1/\epsilon+k_t}$, where k_t is the number of times that player 2 played R until stage t, minus the number of times that player 2 played L until stage t. It can easily be proven that a 0-equilibrium profile for player 1 does not exist, and there are neither stationary nor Markovian ϵ-equilibrium profiles.

Kohlberg [16] generalized the result of Blackwell and Ferguson, and proved that every zero-sum repeated game with absorbing states has a value.

Mertens and Neyman [20] generalized this result further, and proved that every two-player zero-sum stochastic game has a value. Moreover, they proved that the limit of the β-discounted value is equal to the uniform value. Their proof relies on the result of Bewley and Kohlberg [4] that was mentioned before, that the value of the β-discounted game is a Puiseux function in β.

1.3.4 Non Zero-Sum Games

The study of non zero-sum games turns out to be much more difficult. Sorin [30] introduced the following example of a two-player non zero-sum repeated game with absorbing states:

Example 2

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

The unique β-discounted equilibrium payoff of this game is $(1/2, 2/3)$, and the points on the interval $(1/2, 1) - (2/3, 2/3)$ are the only uniform equilibrium payoffs. Since the limit of the β-discounted equilibrium payoffs is not on this interval, it follows that the approach of Mertens and Neyman [20] cannot be used for non zero-sum games. ■

By studying Sorin's example, Vrieze and Thuijsman [36] were able to prove that every two-player repeated game with absorbing states has a uniform equilibrium payoff. Though stationary uniform ϵ-equilibrium profiles need not exist in two-player repeated games with absorbing states, one can prove that 'almost' stationary ϵ-equilibrium strategies do exist. A profile of strategies is *'almost' stationary* if it is given by a stationary profile and a statistical test; the players follow the stationary profile as long as no player fails the statistical test. Once a player has failed the test, he is punished with an ϵ-min-max strategy forever. (an ϵ-min-max strategy against player 1 is the strategy of player 2 in an ϵ-equilibrium profile of the zero-sum game that is defined by the payoffs of player 1).

Rogers [25] and Sobel [28] proved that stationary uniform equilibrium profiles exist in every irreducible game and Thuijsman and Raghavan [31] prove the same result for switching control games.

By carefully analyzing the proof of Vrieze and Thuijsman [36], Vieille [32] proved that every stochastic game with three states has a uniform equilibrium payoff.

Vieille [33, 34] proved that a uniform equilibrium payoff exists in every two-player (non zero-sum) stochastic game if and only if a uniform equilibrium payoff exists in every two player *positive recursive games with the absorbing property.* These games are recursive games where the payoff for player 2 in absorbing states is positive, and satisfy the following absorbing property: for every fully mixed stationary strategy of player 2, the game eventually reaches an absorbing state with probability 1, whatever player 1 plays. Existence of a uniform equilibrium payoff for this class of games was given only recently by Vieille [35].

The only result on stochastic games with more than two players, whose proof is different than the standard proofs for two-player games, was given by Flesch et al. [13], who introduced the following example of a three-player repeated game with absorbing states.

EXAMPLE 3

W

	L	*R*
T	$0, 0, 0$	$0, 1, 3$ *
B	$1, 3, 0$ *	$1, 0, 1$ *

E

	L	*R*
T	$3, 0, 1$ *	$1, 1, 0$ *
B	$0, 1, 1$ *	$0, 0, 0$ *

In this game at every stage player one chooses a row, player 2 chooses a column and player 3 chooses a matrix.

Flesch et al. proved that there is no stationary uniform ϵ-equilibrium profile in this game, but nevertheless a uniform equilibrium payoff does exist. The equilibrium strategy profile has a cyclic nature: the mixed action that player i plays at stage t is equal to the mixed action that player $i+1 \bmod 3$ plays at stage $t+1$. For more details, see section 4.2.5. ∎

1.4 Application

Stochastic games have many applications in economy, biology, political sciences and other fields. Bargaining between agents, interactions between different species, the behavior of political parties and political alliances can be modeled by stochastic games.

Shubik and Whitt [27] modeled an economy with one non-durable good, where the money is valued only as a means to obtain more real goods, as a stochastic game. The state variable in this case is the vector of amounts of money each player possesses.

Winston [38] gave a model of an arm race as a stochastic game. In his model there is a weapon development competition between two countries, and the difference between the level of development between the countries is the state variable.

Filar [11] studied a traveling inspector model: an inspector should inspect some facilities, who can profit by violating the regulations. The aim of the inspector is to minimize the cost for the society, due to inspection and violations. In this model the state variable is the current facility that the inspector inspects.

Levhari and Mirman [19], Amir [1], Dutta and Sundaram [9] and others consider a common property resource model. Two agents simultaneously exploit a productive asset (or resource). Any amount of the resource left over after consumption in a given period forms the investment for that period, and is transformed into the available stock for the next period through a production function.

Team games were studied in [6, 7, 8] for understanding interactions between players of the same team, and more precisely, the "free rider" phenomenon in team games.

Kohlberg and Zamir [17] proved that if every two player zero-sum repeated game with absorbing states has a uniform equilibrium payoff then every repeated game with symmetric incomplete information and deterministic signaling has a uniform equilibrium payoff. This result was generalized by Neyman and Sorin for existence of a uniform equilibrium payoff in n-player repeated games with symmetric incomplete information and deterministic signaling [23] and for non-deterministic signaling [24].

1.5 The Present Monograph

The goal of the present monograph is to shed more light on the existence of the uniform equilibrium payoff in n-player stochastic games, and on the structure of the uniform ϵ-equilibrium profiles.

We consider both n-player repeated games with absorbing states, and two-player stochastic games, and we prove existence of a uniform equilibrium payoff in classes of stochastic games, where existence was unknown before.

In this section we outline the main ideas of the monograph. We first concentrate on n-player repeated games with absorbing states, and then on two-player stochastic games.

1.5.1 Two-Player Repeated Games with Absorbing States

Since a basic idea of our approach was presented by Vrieze and Thuijsman [36] in their proof for existence of a uniform equilibrium payoff in two-player non zero-sum repeated games with absorbing states, we give a sketch of their proof.

Vrieze and Thuijsman consider a sequence of β-discounted equilibria in the game that converges to a limit as $\beta \to 1$, and they construct different types of uniform ϵ-equilibrium profiles according to various properties of this sequence. Denote by $x = (x^1, x^2)$ the limit of the β-discounted equilibrium profiles, and by $g = (g^1, g^2)$ the limit of the corresponding β-discounted equilibrium payoffs. Note that x can be viewed as a stationary profile as well.

Vrieze and Thuijsman prove that three cases can occur. (i) The stationary profile x is absorbing, and then, by adding threat strategies to x, one can construct a uniform ϵ-equilibrium profile. (ii) The stationary profile x is non-absorbing, but the expected non-absorbing payoff for the players if they follow the profile x is at least g. Then, by adding threat strategies to x, one can devise a uniform ϵ-equilibrium profile. (iii) The stationary profile

x is non-absorbing, but the non-absorbing payoff for one player, say player 1, if the players follow the profile x, is strictly less then g^1. In this case, as Vrieze and Thuijsman prove, player 2 has an action a^2 (or a *perturbation*) such that by playing the stationary profile (x^1, a^2) the game will be eventually absorbed, and the payoff for both players is at least g. Using x, a^2 and threat strategies, one can construct a uniform ϵ-equilibrium profile.

It turns out that for more than two players, if the first two cases do not occur then it might be the case that neither the payoff that the players receive by the stationary profile x nor any absorbing perturbation of any subset of the players (or a convex combination of some perturbations), yield all the players a payoff which is at least g.

1.5.2 n-Player Repeated Games with Absorbing States

In order to overcome this difficulty, we define an auxiliary game, where the non-absorbing payoff is bounded by the min-max value of the players in the original game. We prove that for every player, the discounted value of the auxiliary game converges, as the discount factor tends to 1, to his uniform value in the original game.

Consider a sequence of discounted equilibria in the auxiliary game. Denote by x the limit of the discounted stationary equilibrium profiles, and by g the limit of the corresponding discounted equilibrium payoffs. It turns out that the first two cases of Vrieze and Thuijsman yield an 'almost' stationary ϵ-equilibrium profile as above, and, if they do not hold, then there exists a *convex combination* of some perturbations that yields each player i a payoff which is at least g^i. When there are three players, or in team games, one can construct, using this convex combination, a uniform ϵ-equilibrium profile for every $\epsilon > 0$.

In all the uniform ϵ-equilibrium profiles that we construct, the players play mainly the limit of the discounted stationary equilibrium profiles, and perturb to other actions with a very small probability, while checking statistically whether the other players do not deviate. Once a deviation is detected, the deviator is punished with an ϵ-min-max profile forever. Such a profile is called a *perturbed* profile.

Unfortunately, our approach cannot be generalized for more than three players. In section 4.8 we give an example of a four player repeated game with absorbing states where there exists a sequence of discounted equilibrium profiles in the auxiliary game that converges to a limit, but one cannot construct a uniform ϵ-equilibrium profile where the players play mainly the limit mixed-action. Recently Solan and Vieille [29] found an example of a four player repeated game with absorbing states that has no perturbed uniform equilibrium payoff. It is currently not known whether a uniform equilibrium payoff exists in every n-player repeated game with absorbing states, for $n \geq 4$.

The uniform equilibrium payoff that we construct is not necessarily equal to g, the limit of the discounted equilibrium payoffs of the auxiliary game. The reason is that the discounted payoff is a convex combination of the non-absorbing payoff and the absorbing payoff, whereas the undiscounted payoff of an absorbing mixed-action combination depends only on the absorbing part.

For this reason we could not generalize the construction for stochastic games with more than one non-absorbing state — if there is an equality between the uniform equilibrium payoff and the limit of the discounted payoffs, then one can turn all non-absorbing states but one into absorbing states, which yield the players a payoff equal to this limit, and get a repeated game with absorbing states. By our result, this game has an equilibrium payoff, which would be equal (by the hypothesis) to the original limit of the discounted payoffs. Thus, one could construct an equilibrium payoff by working "state after state".

However, since the hypothesis is incorrect, and the limit of the discounted payoff is not necessarily equal to the constructed equilibrium payoff, such an approach fails.

1.5.3 Two-Player Recursive Games with the Absorbing Property

To overcome this problem we consider the undiscounted payoff instead of the discounted payoff, that is, a player evaluates a stationary profile x by

$$\mathbf{E}_{x,s}\left(\lim_{t\to\infty}\frac{1}{t}\sum_{j=1}^{t} r_j^i\right).$$

Since the undiscounted payoff is not continuous over the space of stationary strategies (with the maximum norm), we cannot use standard fixed point theorems.

In order to "make" the undiscounted payoff continuous, we use a result of Vieille [34]. As mentioned above, Vieille proved that for two player games, it is sufficient to prove the existence of an equilibrium payoff for positive recursive games with the absorbing property.

If we consider such a game, and restrict player 2 to a compact subset of the fully mixed stationary strategies, then the undiscounted payoff is continuous over the strategy space. We define ϵ-approximating games where player 2 must play every action with a positive probability, which is greater than some function of ϵ. We consider a sequence of stationary undiscounted equilibria in the ϵ-approximating games as $\epsilon \to 0$, and, as in Vrieze and Thuijsman [36], if there are at most two non-absorbing states then we construct, according to various properties of the sequence, uniform ϵ-equilibrium profiles.

Unfortunately, the uniform equilibrium payoff that we construct need not be equal to the limit of the sequence of the equilibrium payoffs of the ϵ-approximating games, hence we cannot generalize the proof for stochastic games with more than two non-absorbing states.

2 Preliminaries

2.1 On Puiseux Functions

Denote by $\mathcal{F}$ the collection of all *Puiseux series*, that is, the collection of all the formal sums $\sum_{k=K}^{\infty} a_k(1-\theta)^{k/M}$ where $K \in \mathbf{Z}$, $M \in \mathbf{N}$, $(a_k)_{k=K}^{\infty}$ are real numbers and there exists $\theta_0 \in (0,1)$ such that $\sum_{k=K}^{\infty} a_k(1-\theta)^{k/M}$ converges for every $\theta \in (\theta_0, 1)$.

We use θ both as an abstract symbol and as a real number. This dual use should not confuse the reader.

It is well known (see, e.g. Walker [37] or Bewley and Kohlberg [4]) that $\mathcal{F}$ is an ordered field, when addition and multiplication are defined in a similar way to the same operations on power series, and $\sum_{k=K}^{\infty} a_k(1-\theta)^{k/M} > 0$ if and only if $\sum_{k=K}^{\infty} a_k(1-\theta)^{k/M} > 0$ for every θ sufficiently close to 1.

We define the *degree* of any non-zero Puiseux series by:

$$\deg\left(\sum_{k=K}^{\infty} a_k(1-\theta)^{k/M}\right) \stackrel{\text{def}}{=} \frac{\min\{k \mid a_k \neq 0\}}{M}$$

and $deg(0) = \infty$.

Definition 2.1 *A function $\hat{f} : [0,1) \to \mathbf{R}$ is a* Puiseux function *if there exists a Puiseux series $\sum_{k=K}^{\infty} a_k(1-\theta)^{k/M}$ and $\theta_0 \in (0,1)$ such that $\hat{f}(\theta) = \sum_{k=K}^{\infty} a_k(1-\theta)^{k/M}$ for every $\theta \in (\theta_0,1)$.*

As a rule, Puiseux functions are denoted with a hat.

The *degree* of a Puiseux function is the degree of the corresponding Puiseux series, and the order on $\mathcal{F}$ induce an order on Puiseux functions.

Note that $\hat{f}(\theta) = o((1-\theta)^{\deg(\hat{f})-c})$ for every $c > 0$. If $\deg(\hat{f}) \geq 0$ then $\lim_{\theta\to 1} \hat{f}(\theta)$ is finite. In this case define $\hat{f}(1) \stackrel{\text{def}}{=} \lim_{\theta \to 1} \hat{f}(\theta)$. Note also that

$$\deg(\hat{f}\hat{g}) = \deg(\hat{f}) + \deg(\hat{g}). \tag{1}$$

Clearly we have:

Lemma 2.2 *Let $\hat{f}, \hat{g}$ be two Puiseux functions such that $\hat{f}, \hat{g} > 0$. $\lim_{\theta\to 1} \frac{\hat{f}(\theta)}{\hat{g}(\theta)} \in (0,\infty)$ if and only if $\deg(\hat{f}) = \deg(\hat{g})$, and $\lim_{\theta\to 1} \frac{\hat{f}(\theta)}{\hat{g}(\theta)} = 0$ if and only if $\deg(\hat{f}) > \deg(\hat{g})$.*

2.2 Semi-Algebraic Sets

DEFINITION 2.3 *For every* $d \geq 1$*, let* $\mathcal{C}_d$ *be the collection of all subsets of* $\mathbf{R}^d$ *of the form* $\{x \in \mathbf{R}^d \mid p(x) = 0\}$ *or* $\{x \in \mathbf{R}^d \mid p(x) > 0\}$*, where* p *is an arbitrary polynomial.*

A set $C \subseteq \mathbf{R}^d$ *is* semi-algebraic *if it is in the finitely generated algebra which is spanned by* $\mathcal{C}_d$*.*

By Theorem 8.14 in Forster [14] we have

LEMMA 2.4 *If the graph of a real valued function defined on* $(0, 1)$ *is a semi-algebraic set, then the function is a Puiseux function.*

By Lemma 2.4 and Theorem 2.2.1 in Benedetti and Risler [3] we have, by induction on d:

THEOREM 2.5 *Let* $C \subseteq \mathbf{R}^d$ *be a semi-algebraic set, whose projection over its first coordinate includes the interval* $(0, 1)$*. Then there exists a vector of Puiseux functions* $\hat{f} = (\hat{f}^i)_{i=1}^{d-1} : (0, 1) \to \mathbf{R}^{d-1}$ *such that* $(\theta, \hat{f}(\theta)) \in C$ *for every* $\theta \in (0, 1)$*.*

2.3 Notations

For every finite set I, we denote by $\Delta(I)$ the set of all probability distributions over I.

Let I be a finite set, $J \subseteq I$, $\mu \in \Delta(I)$ a probability distribution such that $\sum_{i \in J} \mu(i) > 0$, and $\mathcal{L} = \{L_1, \ldots, L_n\}$ a partition of J (that is, $\{L_j\}_{j=1}^n$ are disjoint sets, whose union in J). The *conditional probability* induced by μ over $\mathcal{L}$, $\mu_{\mathcal{L}}$, is a probability distribution over $\mathcal{L}$ that is defined by:

$$\mu_{\mathcal{L}}(L_j) = \frac{\sum_{i \in L_j} \mu(i)}{\sum_{i \in J} \mu(i)}.$$

For every $a, b \in \mathbf{R}^d$, $a \geq b$ if $a^i \geq b^i$ for every $i = 1, \ldots, d$, and $a > b$ if $a \geq b$ and $a \neq b$. Whenever we use a norm, it is the maximum norm. If $\| a - b \| \leq \epsilon$, we say that a is ϵ*-close* to b. For every $\epsilon > 0$ let

$$B(a, \epsilon) = \{a' \in \mathbf{R}^d \mid \ \| a - a' \| \leq \epsilon\}.$$

We identify each $a_0 \in A$ with $\alpha \in \Delta(A)$ that is defined by

$$\alpha^a = \begin{cases} 1 & a = a_0 \\ 0 & a \neq a_0 \end{cases}$$

By convention, a sum over an empty set of indices is 0.

3 Stochastic Games

3.1 The Model

A *stochastic game* is a 5-tuple $G = (N, S, (A^i)_{i \in N}, h, w)$ where

- N is a finite set of players.
- S is a finite set of states.
- For every $i \in N$, A^i is a finite set of actions available for player i in each state s. Denote $A = \times_{i \in N} A^i$.
- $h : S \times A \to \mathbf{R}$ is the daily payoff function, $h^i(s, a)$ being the daily payoff for player i in state s when the action combination a is played. Let $R \geq 1$ be a bound on $|h|$.
- $w : S \times A \to \Delta(S)$ is the transition function.

Note that since the state space is finite, the assumption that the available set of actions is independent of the state is not restrictive.

Let $H^n = S \times (A \times S)^n$ be the space of all histories of length n, $H_0 = \cup_{n \in \mathbf{N}} H^n$ be the space of all finite histories and $H = S \times (A \times S)^{\mathbf{N}}$ be the space of all infinite histories. For every finite history $h_0 \in H_0$, $L(h_0)$ is its length and $s_L(h_0)$ is its last stage.

We define a partial order on H_0. Let $h_0 = (s_0, a_1, s_1, \ldots, a_t, s_t)$ and $h'_0 = (s'_0, a'_1, s'_1, \ldots, a'_{t'}, s'_{t'})$ be two histories. $h'_0 \leq h_0$ if and only if $t' \leq t$ and $(s'_0, a'_1, \ldots, s'_{t'}) = (s_0, a_1, \ldots, s_{t'})$, that is, h'_0 is a beginning of h_0. Define $h'_0 < h_0$ if and only if $h'_0 \leq h_0$ and $h'_0 \neq h_0$. If $h_0 \in H_0$ and $h \in H$, we say that $h_0 < h$ if $h = (s_0, a_1, s_1, \ldots)$, $h_0 = (s'_0, a'_1, s'_1, \ldots, a'_t, s'_t)$ and $(s'_0, a'_1, \ldots, s'_t) = (s_0, a_1, \ldots, s_t)$, that is, h_0 is a beginning of h.

For every $i \in N$, let $X^i = \Delta(A^i)$, the set of all mixed-action combinations of player i. We denote $X = \times_{i \in N} X^i$, $X^{-i} = \times_{j \neq i} X^i$ and $X^L = \times_{i \in L} X^i$ for every $L \subseteq N$. The multi-dimensional extensions of h and w to X are denoted also by h and w.

Definition 3.1 *A* behavioral strategy *of player i is a function $\sigma^i : H_0 \to X^i$. A strategy σ^i is* stationary *if $\sigma^i(h_0)$ depends only on $s_L(h_0)$.*

A *strategy profile* (or simply a *profile*) is a vector of strategies, one for each player. Every profile σ and finite history h_0 induce a probability measure over H (equipped with the σ-algebra generated by all the finite cylinders). The probability measure is the measure that is induced by σ given that the history h_0 has occurred (regardless of the probability of h_0 under σ). We denote this measure by $\Pr_{h_0,\sigma}$, and expectation according to it by $\mathbf{E}_{h_0,\sigma}$. If $h_0 = (s)$ we denote the expectation by $\mathbf{E}_{s,\sigma}$.

For every strategy σ and finite history $h_0 = (s_0, a_1, s_1, \dots, a_t, s_t)$ we define the strategy $\sigma_{|h_0}$ by:

$$\sigma_{|h_0}(h'_0) = \sigma(s_0, a_1, s_1, \dots, a_t, s'_0, a'_0, \dots, a'_{t'}, s'_{t'})$$

where $h'_0 = (s'_0, a'_0, \dots, a'_{t'}, s'_{t'})$.

Every $x^i \in X^i$ can be viewed as a stationary strategy of player i. Every such strategy is identified with a vector in $\mathbf{R}^{|S| \cdot |A^i|}$

3.2 The Discounted Payoff

We denote by r^i_t the daily payoff that player i receives at stage t.

Let σ be a strategy profile, $s \in S$, $\beta \in [0, 1)$ and $i \in N$. The expected β-discounted payoff for player i if the initial state is s and the players follow the profile σ is given by:

$$v^i_\beta(s, \sigma) = (1 - \beta) \sum_{t=1}^{\infty} \beta^{t-1} \mathbf{E}_{s,\sigma} r^i_t,$$

where r^i_t is the payoff of player i at stage t.

Definition 3.2 *Let $i \in N$ and $\beta \in [0, 1)$. The vector $(c^i_s(\beta))_{s \in S}$ is* the β-discounted min-max value of player i *if the following two conditions hold:*

- *For every strategy profile σ^{-i} of players $N \setminus \{i\}$ there exists a strategy σ^i of player i such that $v^i_\beta(s, \sigma) \geq c^i_s(\beta)$ for every $s \in S$.*
- *There exists a strategy profile σ^{-i} of players $N \setminus \{i\}$ such that for every strategy σ^i of player i, $v^i_\beta(s, \sigma) \leq c^i_s(\beta)$ for every $s \in S$.*

Since the discounted payoff is continuous over the strategy space, and the setup is finite, it follows that the β-discounted min-max value exists.

Definition 3.3 *The strategy profile σ is a β-*discounted equilibrium *if for every player i, every strategy τ^i of player i and every initial state s,*

$$v_\beta^i(s, \sigma) \geq v_\beta^i(s, \sigma^{-i}, \tau^i).$$

*The payoff vector $(v_\beta(s, \sigma))_{s \in S}$ is a β-*discounted equilibrium payoff.

Fink [12] has proved the following:

Theorem 3.4 *For every $\beta \in [0, 1)$ there exists a β-discounted stationary equilibrium profile in every stochastic game.*

3.3 The Uniform MinMax Value

Definition 3.5 *Let $i \in N$. The vector $(c_s^i)_{s \in S} \in \mathbf{R}^S$ is the* uniform min-max value *of player i if for every $\epsilon > 0$ there exists $t_c \in \mathbf{N}$ and a profile σ_ϵ^{-i} of players $N \setminus \{i\}$ such that for every initial state $s \in S$:*

- *For every strategy σ^i of player i*

  $$\mathbf{E}_{s, \sigma_\epsilon^{-i}, \sigma^i} \left(\limsup_{t \to \infty} \frac{r_1^i + r_2^i + \cdots + r_t^i}{t} \right) \leq c_s^i + \epsilon$$

  and for every $t \geq t_c$

  $$\mathbf{E}_{s, \sigma_\epsilon^{-i}, \sigma^i} \left(\frac{r_1^i + r_2^i + \cdots + r_t^i}{t} \right) \leq c_s^i + \epsilon.$$

- *For every strategy profile σ^{-i} of players $N \setminus \{i\}$ there exists a strategy σ^i of player i such that*

  $$\mathbf{E}_{s, \sigma} \left(\liminf_{t \to \infty} \frac{r_1^i + r_2^i + \cdots + r_t^i}{t} \right) \geq c_s^i - \epsilon$$

  and for every $t > t_c$,

  $$\mathbf{E}_{s, \sigma} \left(\frac{r_1^i + r_2^i + \cdots + r_t^i}{t} \right) \geq c_s^i - \epsilon.$$

The profile σ_ϵ^{-i} is a uniform ϵ-min-max profile against player *i.*

LEMMA 3.6 *For every player i, the min-max value c^i exists. Moreover, $c^i = \lim_{\beta \to 1} c^i(\beta)$, the limit of the discounted min-max value of player i.*

This result was proved by Mertens and Neyman [20] for two-player stochastic games, and an unpublished proof of Neyman [22] that follows similar lines proves the result for n-player stochastic games.

3.4 The Uniform Equilibrium Payoff

DEFINITION 3.7 *The payoff vector $(g(s))_{s \in S} \in \mathbf{R}^{N \times S}$ is a* uniform ϵ-equilibrium payoff *if there exist a strategy profile σ_ϵ and $t_e \in \mathbf{N}$ such that for every player i, every strategy τ^i of player i, every initial state s and every $t \geq t_e$*

$$\mathbf{E}_{s,\sigma_\epsilon}\left(\frac{r_1^i + r_2^i + \cdots + r_t^i}{t}\right) \geq g^i(s) - \epsilon, \tag{2}$$

$$\mathbf{E}_{s,\sigma_\epsilon}\left(\liminf_{t\to\infty} \frac{r_1^i + r_2^i + \cdots + r_t^i}{t}\right) \geq g^i(s) - \epsilon, \tag{3}$$

$$\mathbf{E}_{s,\sigma_\epsilon^{-i},\tau^i}\left(\frac{r_1^i + r_2^i + \cdots + r_t^i}{t}\right) \leq g^i(s) + \epsilon \qquad \textit{and} \tag{4}$$

$$\mathbf{E}_{s,\sigma_\epsilon^{-i},\tau^i}\left(\limsup_{t\to\infty} \frac{r_1^i + r_2^i + \cdots + r_t^i}{t}\right) \leq g^i(s) + \epsilon. \tag{5}$$

The profile σ_ϵ is a uniform ϵ-equilibrium profile*. The payoff vector $g \in \mathbf{R}^{N \times S}$ is a* uniform equilibrium payoff *if it is an ϵ-equilibrium payoff for every $\epsilon > 0$.*

3.5 On Perturbed Equilibrium

In order to be able to implement the ϵ-equilibrium profile, we need the profile to be "simple". In example 1, one cannot construct ϵ-equilibrium profiles which are stationary or Markovian. For some classes of stochastic games existence of *'almost' stationary* ϵ-equilibrium profiles was established. Flesch et al. proved that in the game presented in example 3 there are no 'almost' stationary ϵ-equilibrium profiles.

In this section we define a broader class of profiles, called *perturbed* profiles. A profile in this class is given by a stationary strategy, small perturbations and statistical tests. The players play mainly the stationary strategies, but perturb to other actions with small probability. All along, the actions of each player are screened by a statistical test, and the first player who fails the test is punished forever.

We later prove that different classes of stochastic games admit equilibrium payoffs, whose corresponding ϵ-equilibrium profiles are perturbed.

The importance of perturbed strategies are that the distribution of the actions almost resembles stationary strategies. Thus, if the players are different species, and the actions stand for various types of this specie, then an equilibrium where at even stages half the population should consists of males and half of females, while at odd stages two thirds should be females, is clearly undesirable.

Note that perturbed equilibrium profiles may be complex: we say nothing about the complexity of the perturbations, or on the statistical test.

For every function $f : H_0 \to \{0, 1\}$, let $Z(f) \subseteq H_0$ be the set of all finite histories h_0 such that $f(h_0) = 0$, but $f(h'_0) = 1$ for every $h'_0 < h_0$.

Every function $f : H_0 \to \{0, 1\}$ can represent a statistical test — if $f(h_0) = 1$ then the player does not fail the test, while if $f(h_0) = 0$ then the player fails the test. $Z(f)$ is the set of all finite histories in which the player fails the test for the first time.

For every vector function $f : H_0 \to \{0, 1\}^N$ we define $Z(f) \subseteq H_0$ to be the set of all finite histories h_0 such that $h_0 \in Z(f^i)$ for some $i \in N$, but $h'_0 \notin Z(f^j)$ for every $h'_0 < h_0$ and every $j \in N$. For every $h_0 \in Z(f)$, the *index* of h_0 is the minimal i such that $h_0 \in Z(f^i)$.

Every vector function $f : H_0 \to \{0, 1\}^N$ can represent a vector of statistical tests, one test for each player. $Z(f)$ is the set of finite histories where a failure of some player is observed for the first time, and the index of the history is the identity of the first player who failed his test.

We denote by $G(f)$ the set of all finite histories h_0 such that $f^i(h'_0) = 1$ for every $i \in N$ and $h'_0 \leq h_0$. $G(f)$ is the set of "good" histories, where no failure of the test has ever occurred.

Define

$$G^\star(f) = \{h \in H \mid h_0 \in G(f) \qquad \forall h_0 < h\}$$

and

$$Z^\star(f) = H \setminus G^\star(f).$$

$G^\star(f)$ is the set of all *infinite* histories where no failure is detected along the whole game, and $Z^\star(f)$ is the set of all infinite histories where a failure of some player is detected at some point.

Formally, the new class of equilibrium payoffs is defined as follows:

DEFINITION 3.8 *Let $x \in X$ and $\epsilon > 0$. A profile σ is (x,ϵ)-*perturbed *if there exist*

- *a function $f : H_0 \to \{0,1\}^N$*
- *and for every $i \in N$, an ϵ-min-max strategy τ_ϵ^{-i} against player i*

such that

- *For every history $h_0 \in G(f)$ we have $\parallel \sigma(h_0) - x \parallel < \epsilon$.*
- *For every history $h_0 \in Z(f)$ with index i_0 we have $\sigma_{|h_0}^{-i_0} = \tau_\epsilon^{-i_0}$.*

Finally

DEFINITION 3.9 *Let $x \in X$. The payoff vector $g \in \mathbf{R}^{N\times S}$ is a* uniform (x,ϵ)-perturbed equilibrium payoff *if it is an ϵ-equilibrium payoff, and there exists an ϵ-equilibrium profile for g which is (x,ϵ)-perturbed. x is the* base *of the ϵ-equilibrium payoff. The payoff vector $g \in \mathbf{R}^{N\times S}$ is a* uniform x-perturbed equilibrium payoff *(or a* perturbed equilibrium payoff*) if it is an (x,ϵ)-perturbed equilibrium payoff for every $\epsilon > 0$.*

In all the classes of stochastic games for which existence of the undiscounted equilibrium payoff was proven, there exists a perturbed equilibrium payoff. These classes, apart of the classes that admit stationary equilibria, are two-player non zero-sum repeated games with absorbing states (Vrieze and Thuijsman [36]), and positive recursive games with the absorbing property (Vieille [35]). Recall that for two-player zero-sum stochastic games, there exists an ϵ-equilibrium profile σ such that $\parallel \sigma(h_0) - x \parallel < \epsilon$ for every finite history $h_0 \in H_0$, where x is a fixed mixed action combination (the limit of the discounted stationary equilibria). Moreover, the expected payoff that these profiles yield converge, as ϵ tends to 0, to the value of the game (Mertens and Neyman [20]).

However, Solan and Vieille [29] show that stochastic games in general need not admit perturbed equilibrium payoffs.

In the present paper we prove existence of a uniform perturbed equilibrium payoff in the following three classes of stochastic games: three-player repeated games with absorbing states, repeated team games with absorbing states and two player stochastic games with two non-absorbing states.

Since the results in this monograph refer to uniform equilibria, whenever we write equilibrium payoff, ϵ-equilibrium profiles and min-max value, we mean the uniform equilibrium payoff, uniform ϵ-equilibrium profile and uniform min-max value respectively. Whenever we refer to the discounted version of these notion, we explicitly mention the word discounted.

4 Repeated Games with Absorbing States

In this section we prove that every three-player repeated game with absorbing states, as well as every repeated team game with absorbing states, has an equilibrium payoff. We begin by introducing an equivalent formulation of repeated games with absorbing states (section 4.1). We then give five sets of sufficient conditions for existence of a perturbed equilibrium payoff (section 4.2). We derive some preliminary results in sections 4.3-4.6, including the definition of an auxiliary game (section 4.4), which is the core of the proof. Afterwards we prove that every three-player repeated game with absorbing states has an equilibrium payoff (section 4.7), and we explain why our approach cannot be used for games with more than three players (section 4.8). We finally prove that every repeated team game with absorbing states has an equilibrium payoff (section 4.9), and prove a geometric result, which is used along the section (section 4.10).

4.1 An Equivalent Formulation

Recall that a repeated game with absorbing state is a stochastic game where all the states but one are absorbing. Since every absorbing state is a standard repeated game, we can choose for each such state one equilibrium payoff, and assume that once this state is reached, all future payoffs are equal to this equilibrium payoff.

Thus, an equivalent representation of a repeated game with absorbing states is by a 5-tuple $G = (N, (A^i, h^i, u^i)_{i \in N}, w)$ where:

- N is a finite set of players.
- For every player $i \in N$, A^i is a finite set of pure actions available to player i. Denote $A = \times_{i \in N} A^i$.
- For every player i, $h^i : A \to \mathbf{R}$ is a function that assigns to each action combination $a \in A$ a non-absorbing payoff for player i.
- For every action combination $a \in A$, $w(a)$ is the probability that the game is absorbed if this action combination is played. If the game is absorbed, $u^i(a)$ is the payoff that player i receives, at each future stage.

The game is played as follows. At every stage each player chooses an action $a^i \in A^i$. If the game is not already absorbed, the players receive a daily payoff $h(a)$, where $a = (a^i)_{i \in N}$. With probability $w(a)$ the game is absorbed, and then the players receive a payoff $u(a)$ at each future stage. With probability $1 - w(a)$ the game continues.

Recall that the multi-linear extensions of h and w are also denoted by h and w. A mixed action combination $x \in X$ is *absorbing* if $w(x) > 0$ and *non-absorbing* otherwise. Define an extension of u to X by:

$$u^i(x) = \sum_{a \in A} \left(\prod_{i \in N} x^i_{a^i} \right) w(a) u^i(a) / w(x) \tag{6}$$

whenever x is absorbing and 0 otherwise. $u^i(x)$ is the expected absorbing payoff for player i given absorption occurs when the players play the mixed action combination x. Note that u is semi-algebraic, and that wu is multi-linear.

Since there is only one "interesting" state, we omit the state variable from all the entities that we have defined, and they all refer to the non-absorbing state as the initial state.

4.2 Different Types of Equilibria

In this section we present five sets of sufficient conditions for existence of a perturbed equilibrium payoff in a repeated game with absorbing states. The first three sets are for games with arbitrary number of players, while the fourth is given only for three-player games. The fifth set of sufficient conditions is a generalization of the third and fourth conditions for n-player games.

The first and second sufficient conditions and a degenerate version of the third were used by Vrieze and Thuijsman [36] for two-player games.

For every player i define the function $e^i : X \to \mathbf{R}$ by

$$e^i(x) = w(x) u^i(x) + (1 - w(x)) c^i.$$

$e^i(x)$ is the maximal payoff that player i can guarantee if at the current stage the players play the mixed-action combination x, and from the next stage player i is punished forever with an ϵ-min-max profile, for an arbitrary small

ϵ. Define

$$E^i(x) = \max_{a^i \in A^i} e^i(x^{-i}, a^i).$$

$E^i(x)$ is the maximal payoff that player i can guarantee by "deviating" when the mixed action combination x should be played, and then be punished with an ϵ-min-max strategy, for an arbitrary small ϵ.

4.2.1 The Structure of the Proofs

With every set of sufficient conditions we give an example that illustrates the corresponding ϵ-equilibrium profile, a formal definition of the ϵ-equilibrium profile, and a proof that the profile is indeed an ϵ-equilibrium.

Since the proofs in all the different cases have the same structure, we will now sketch the structure of the proofs.

In all the cases, we will be given a mixed action combination $x \in X$ and a payoff vector $g \in \mathbf{R}^N$ such that $g^i \geq E^i(x)$ for every player $i \in N$. We will fix $\epsilon > 0$ and proceed as follows.

Step a: Definition of a profile σ
We will define a profile σ such that $\| \sigma(h_0) - x \| \leq \epsilon$ for every finite history $h_0 \in H_0$. The profile σ will always be cyclic, but the length of the cycle may depend on ϵ, and its exact nature is derived from the conditions we impose. Therefore the limit $\lim_{t\to\infty}(r_1^i + \cdots + r_t^i)/t$ exists a.s. w.r.t. σ. We will define σ in such a way that

$$\left| \mathbf{E}_\sigma \left(\lim_{t\to\infty} (r_1^i + \cdots + r_t^i)/t \right) - g^i \right| < \epsilon. \tag{7}$$

Step b: Definition of a statistical test f
We will define a statistical test $f : H_0 \to \{0,1\}^N$. Recall that $G^\star(f)$ is the collection of all the infinite histories where the statistical test never fails, and $Z(f)$ is the collection of all the finite histories where the statistical test fails for the first time.

We will prove that for every profile τ such that $\Pr_\tau(G^\star(f)) > 0$, $\lim_{t\to\infty}(r_1^i + \cdots + r_t^i)/t$ exists a.s. w.r.t. τ conditional on $G^\star(f)$ and

$$\mathbf{E}_\tau \left(\limsup_{t\to\infty} (r_1^i + \cdots + r_t^i)/t \mid G^\star(f) \right) \leq g^i + \epsilon R. \tag{8}$$

That is, as long as no "deviation" is detected, the expected payoff for any player i cannot exceed $g^i + \epsilon R$.

Step c: False Detection of Deviation
Next we will prove that if the players follow σ, then the probability of false detection of deviation is smaller than ϵ, that is:

$$\Pr_\sigma(G^\star(f)) > 1 - \epsilon. \tag{9}$$

Step d: Definition of a Profile σ_f
For every profile σ and statistical test f, we define an (x, ϵ)-perturbed profile σ_f as follows. The players follow σ as long as no player fails the statistical test. The first player who fails the statistical test is punished with an ϵ-min-max profile forever. Formally, $\sigma_f(h_0) = \sigma(h_0)$ for every $h_0 \in G(f)$, and $\sigma_{f|h_0} = \tau_\epsilon^{-i}$ for every $h_0 \in Z(f)$ with index i, where τ_ϵ^{-i} is any ϵ-min-max strategy against player i.

In all the cases, we define σ and f in such a way that the profile σ_f is a $3\epsilon R$-equilibrium for g.

In the following steps we show how we will prove that σ_f is a $3\epsilon R$-equilibrium profile.

Step e: Eqs. (2) and (3) in Definition 3.7 hold
We shall now see that from the above steps it follows that Eqs. (2) and (3) in Definition 3.7 hold w.r.t. σ_f. Indeed, by (7), (9) and since the payoffs are bounded by $R \geq 1$ it follows that

$$\mathbf{E}_{\sigma_f}\left(\liminf_{t\to\infty}(r_1^i + \cdots + r_t^i)/t\right) \geq g^i - 2\epsilon R.$$

Assume that t_e is sufficiently large such that for every $t \geq t_e$

$$\mathbf{E}_\sigma\left((r_1^i + \cdots + r_t^i)/t\right) \geq g^i - 2\epsilon \qquad \forall i \in N. \tag{10}$$

It follows from (9) and (10) that for every $t \geq t_e$

$$E_{\sigma_f}\left((r_1^i + \cdots + r_t^i)/t\right) \geq g^i - 3\epsilon R \qquad \forall i \in N.$$

Step f: Eqs. (4) and (5) in Definition 3.7 hold
This part of the proof is different for each set of sufficient conditions. Since the limit $\lim_{t\to\infty}(r_1^i + \cdots + r_t^i)/t$ exists a.s. w.r.t. σ, there exists $t_e \in \mathbf{N}$ such that for every $t \geq t_e$

$$\mathbf{E}_\tau \left((r_1^i + \cdots + r_t^i)/t \mid G^\star(f) \right) \leq g^i + \epsilon. \tag{11}$$

By (8) and (11) it is left to verify that for every player i, every strategy τ^i of player i and every finite history $h_0 \in Z(f)$ with index i,

$$\mathbf{E}_{h_0,\sigma_f^{-i},\tau^i} \left(\limsup_{t\to\infty} (r_1^i + \cdots + r_t^i)/t \right) \leq g^i + \epsilon R$$

and for every $t \geq t_e$

$$\mathbf{E}_{h_0,\sigma_f^{-i},\tau^i} \left((r_1^i + \cdots + r_t^i)/t \right) \leq g^i + \epsilon R.$$

To summarize, for each set of sufficient conditions we need to define a profile σ that satisfies (7), a statistical test f that satisfies (8), and to prove (9) and Step f. Along the different proofs we will refer to these steps as our guidelines.

4.2.2 An 'Almost' Stationary Non-Absorbing Equilibrium

EXAMPLE 4

W

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

E

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

The min-max value of each player in this game is 0.

One equilibrium payoff in the game is $(1/2, 1/2, 1/2)$. Consider the non-absorbing stationary profile $(\frac{1}{2}T + \frac{1}{2}B, \frac{1}{2}L + \frac{1}{2}R, E)$. Player 1 is indifferent between his actions, while player 2 prefers to play always R, and player 3 prefers to play always W. Since the profile is non-absorbing, the players can conduct a statistical test and check whether player 2 plays each action with probability close to 1/2. If the distribution of the realized actions of player 2 is not sufficiently close to $(1/2, 1/2)$, then the other two players punish him with a min-max profile forever. A deviation of player 3 is detected immediately, and, given the game is not absorbed, can be punished with a min-max profile.

It is easy to verify that the stationary non-absorbing profile, supplemented with these threat strategies, is an ϵ-equilibrium profile for $h(x)$, where the ϵ comes from the probability of false detection of deviation in the statistical test. ■

Lemma 4.1 *Let $x \in X$ be a non-absorbing mixed action combination such that $h^i(x) \geq E^i(x)$ for every player $i \in N$. Then $h(x)$ is an equilibrium payoff.*

Note that by the assumption it follows that $h^i(x) \geq c^i$ for every player i.

Proof: Let $\epsilon > 0$ be fixed, and denote $g = h(x)$. Define a stationary profile σ by:

$$\sigma(h_0) = x \qquad \forall h_0 \in H_0.$$

It is clear that σ satisfies (7).

Define a statistical test as follows.

1. Each player i is checked if his strategy is compatible with σ (that is, he does not play an action outside $\text{supp}(x^i)$).
2. At every stage $t \geq t_1$ (where t_1 is defined below), each player i is checked whether the distribution of his realized actions is ϵ-close to x^i.

Formally, the statistical test is given by a function $f : H_0 \to \{0,1\}^N$ that is defined as follows. $f^i(a_1, a_2, \ldots, a_t) = 0$ if and only if $a^i_t \notin \text{supp}(x^i)$ *or* $t \geq t_1$ and $\left\| \frac{a^i_1 + \cdots + a^i_t}{t} - x^i \right\| > \epsilon$. It is clear that (8) holds.

The constant $t_1 \geq 1/\epsilon$ is chosen sufficiently large such that the probability of false detection of deviation is bounded by ϵ, that is, for every $i \in N$

$$\Pr\left(\| \bar{X}^i_t - x^i \| < \epsilon \quad \forall t > t_1 \right) > 1 - \epsilon/|N| \tag{12}$$

where $\bar{X}^i_t = \frac{1}{t}\sum_{j=1}^t X^i_j$ and $\{X^i_j\}$ are i.i.d. r.v. with distribution x^i. Hence (9) holds.

We shall now verify that (4) and (5) in Definition 3.7 hold. As remarked in Step f, we fix a player i, a strategy τ^i of player i and a history $h_0 \in Z(f)$ with index i.

We assume that t_1 is sufficiently large such that:

$$\frac{t_1 h^i(x) + t_c R}{t_1 + t_c} \leq h^i(x) + \epsilon \qquad \forall i \in N. \tag{13}$$

Let $t_2 \geq t_c$ be sufficiently large such that

$$\frac{t_1 R + t_2 c^i}{t_1 + t_2} \leq c^i + \epsilon \qquad \forall i \in N. \tag{14}$$

If $L(h_0) \leq t_1$ then by (14) and the condition, for every $t \geq t_1 + t_2$,

$$\mathbf{E}_{h_0,\sigma_f^{-i},\tau^i}\left(\frac{r_1^i + \cdots + r_t^i}{t}\right) \leq h^i(x) + 2\epsilon R \tag{15}$$

and

$$\mathbf{E}_{h_0,\sigma_f^{-i},\tau^i}\left(\liminf_{t\to\infty} \frac{r_1^i + \cdots + r_t^i}{t}\right) \leq h^i(x) + \epsilon R. \tag{16}$$

If, on the other hand, $L(h_0) > t_1$ then it follows by (13) and the condition that (15) and (16) hold.

Thus σ_f is a $3\epsilon R$-equilibrium profile for $h(x)$, and $h(x)$ is an x-perturbed equilibrium payoff. ■

4.2.3 An 'Almost' Stationary Absorbing Equilibrium

Consider the game in example 4. Another equilibrium payoff in this game is $(0, 1, 1)$, and an equilibrium profile is to play the stationary profile $(T, \frac{1}{2}L + \frac{1}{2}R, W)$, while checking for a deviation of players 1 and 3. Once a deviation is detected, the deviator is punished with a min-max strategy profile. Note that player 2 is indifferent between his actions, hence the fact that his deviations cannot be checked (since the game is absorbed after the first stage, whatever he plays) does not affect this equilibrium. ■

LEMMA 4.2 *Let $x \in X$ be an absorbing mixed action combination that satisfies the following two conditions:*

1. *$u^i(x) \geq E^i(x) \qquad \forall i \in N$.*
2. *$u^i(x) = u^i(x^{-i}, a^i)$ for every $i \in N$ and every $a^i \in \mathrm{supp}(x^i)$ such that $w(x^{-i}, a^i) > 0$.*

Then $u(x)$ is a perturbed equilibrium payoff.

Proof: Denote $g = u(x)$. We will consider the profile that was defined in the proof of Lemma 4.1, but assume that the constant t_1 is sufficiently large to satisfy an additional requirement. We will then prove, as in Lemma 4.1, that this profile is an ϵ-equilibrium for g.

It is clear that if the players follow the stationary profile σ that was defined in the proof of Lemma 4.1 then (7) holds, and that (8) holds as well.

Let $\eta > 0$ be fixed. Let $\epsilon \in (0, \eta)$ be sufficiently small such that every $y \in B(x, \epsilon)$ satisfies that $w(y) > w(x)/2$ and $\| u(y) - u(x) \| < \eta$.

Consider the statistical test f that was defined in the proof of Lemma 4.1, but assume that the constant t_1 is sufficiently large such that if no deviation is detected then absorption occurs before stage t_1 with probability greater than $1 - \epsilon$, that is, $(1 - w(x)/2)^{t_1} < \epsilon$.

By the choice of t_1, (9) holds.

Fix a player i, a strategy τ^i of player i and a history $h_0 \in Z(f)$ with index i. If $L(h_0) \leq t_1$ then by (14) and the conditions, for every $t \geq t_1 + t_2$ (where t_2 is defined in the proof of Lemma 4.1)

$$\mathbf{E}_{h_0, \sigma_f^{-i}, \tau^i} \left(\frac{r_1^i + \cdots + r_t^i}{t} \right) \leq \frac{t_1 u^i(x) + (t - t_1)(E^i(x) + \epsilon)}{t} \leq u^i(x) + \epsilon \quad (17)$$

and

$$\mathbf{E}_{h_0, \sigma_f^{-i}, \tau^i} \left(\liminf_{t \to \infty} \frac{r_1^i + \cdots + r_t^i}{t} \right) \leq u^i(x) + \epsilon. \quad (18)$$

Since

$$\Pr_\sigma \left(h \in H \mid \exists h_0 < h \text{ s.t. } h_0 \in Z_f \text{ and } L(h_0) \geq t_1 \right) < \epsilon$$

it follows that

$$\mathbf{E}_{\sigma_f^{-i}, \tau^i} \left(\frac{r_1^i + \cdots + r_t^i}{t} \right) \leq u^i(x) + 3\epsilon R$$

and

$$\mathbf{E}_{\sigma_f^{-i}, \tau^i} \left(\liminf_{t \to \infty} \frac{r_1^i + \cdots + r_t^i}{t} \right) \leq u^i(x) + 2\epsilon R.$$

Thus σ_f is a $3\epsilon R$-equilibrium profile for $u(x)$, and $u(x)$ is an x-perturbed equilibrium payoff. ■

4.2.4 Average of Perturbations

EXAMPLE 5

		W			E	
	L	C	R	L	C	R
T	$0,0,0$	$0,0,0$	$0,1,0$ *	$0,0,0$	$0,0,0$	
M	$0,0,0$	$1,4,1$ *		$3,0,1$ *		
B	$1,-2,3^*$					

The empty cells may be arbitrary (both the payoff and whether they are absorbing or not). Let R be the maximal payoff (in absolute values). We fix $\epsilon > 0$ and denote $\delta = \epsilon/R$. Note that the min-max value of players 1 and 2 is at most 1, and the min-max value of player 3 is at most 0.

One equilibrium payoff is $(1,4,1)$ and an ϵ-equilibrium profile for it is to play at every stage the mixed-action combination $((1-\delta)T+\delta M, (1-\delta)L+\delta C, W)$. Clearly if the players follow this profile then they receive the desired payoff, and no player can profit more than $\delta R = \epsilon$ by deviating.

Absorption occurs at every stage with probability δ^2, while perturbations of players 1 and 2 occur with probability δ. Therefore, if δ is sufficiently small, the players can conduct statistical tests to check whether player 1 plays the action M in frequency δ, and whether player 2 plays the action L with frequency δ. Though these tests are not necessary, since no player can profit by deviating, they can still be employed.

Another equilibrium payoff is $(2,2,1) = \frac{1}{2}(1,4,1) + \frac{1}{2}(3,0,1)$, and an ϵ-equilibrium profile for it is:

- At odd stages play $((1-\delta)T+\delta M, (1-\delta)L+\delta C, W)$.
- At even stages play $((1-\delta)T+\delta M, L, (1-\delta)W+\delta E)$.

If the players follow this profile then their expected payoff is approximately $(2,2,1)$. Since player 1 prefers absorption by (M,L,E), while player 2 prefers absorption by (M, C, W), the players should conduct statistical tests and check whether each of them follows this profile, and punish a deviator with an ϵ-min-max profile.

Yet a third equilibrium payoff is $(1,1,2) = \frac{1}{2}(1,4,1) + \frac{1}{2}(1,-2,3)$, and an ϵ-equilibrium profile for it is:

- At odd stages play $((1-\delta)T + \delta M, (1-\delta)L + \delta C, W)$.
- At even stages play $((1-\delta^2)T + \delta^2 B, L, W)$.

If the players follow this profile then their expected payoff is approximately $(1,1,2)$. In this case the players cannot check whether player 1 plays at even stages the action B in frequency δ^2, since once he plays this action the game terminates with probability $1-\delta$, and there is a probability of $1/2$ that he never plays B. Nevertheless player 1 has no incentive to deviate, and therefore such a check is not needed. However. the players do need to check whether player 2 perturbs at odd stages as he should.

Actually, every convex combination (g^1, g^2, g^3) of the four absorbing cells $(1,-2,3), (1,4,1), (0,1,0)$ and $(3,0,1)$ in which $g^1, g^2 \geq 1$, $g^3 \geq 0$ that satisfies:

- If $(1,-2,3)$ has a positive weight in this combination then $g^1 = 1$.
- If $(0,1,0)$ has a positive weight in this combination then $g^2 = 1$.

is an equilibrium payoff. ■

Definition 4.3 *Let $x \in X$ be a non-absorbing mixed-action combination and $L \subseteq N$. An action combination $b^L \in \times_{i \in L} A^i$ is an* absorbing neighbor of x by L *if*

- $w(x^{-L}, b^L) > 0$.
- $w(x^{-L'}, b^{L'}) = 0$ *for every strict subset L' of L.*

If $L = \{i\}$ then the absorbing neighbor is called a *single absorbing neighbor* of player i.

We denote by $\mathcal{B}(x)$ the set of all absorbing neighbors of x, and by $\mathcal{B}_i(x)$ the set of all single absorbing neighbors of player i. Note that $\mathcal{B}(x)$ is never empty (as long as there is an absorbing action combination), but $\mathcal{B}_i(x)$ may be empty.

LEMMA 4.4 *Let $x \in X$ be a non-absorbing mixed action combination. Let $\mu \in \Delta(\mathcal{B}(x))$ and denote $g = \sum_{b^L \in \mathcal{B}(x)} \mu(b^L) u(x^{-L}, b^L)$. Assume the following conditions hold:*

1. $g^i \geq E^i(x) \qquad \forall i \in N$.
2. $u^i(x^{-i}, a^i) = g^i$ *for every player* i *and every action* $a^i \in \mathcal{B}_i(x) \cap \mathrm{supp}(\mu)$.

Then g *is an equilibrium payoff.*

Proof: Let $\epsilon > 0$ be sufficiently small, $T \in \mathbf{N}$ sufficiently large, and $m : [1, \ldots, T] \to \mathrm{supp}(\mu)$ such that

$$\left| \frac{\#\{j \mid m(j) = b^L\}}{T} - \mu(b^L) \right| < \epsilon/3, \qquad \forall b^L \in \mathrm{supp}(\mu). \tag{19}$$

That is, m is a discrete approximation of μ. Extend the domain of m to $\mathbf{N}$ by $m(t) = m(t \bmod T)$ for every $t > T$. Let $L(t)$ be the set of players for which $m(t)$ is an absorbing neighbor of x.

In the sequel, $\delta \in (0, \epsilon)$ is sufficiently small, such that

$$(1-\delta)^T > 1 - \epsilon/3 \tag{20}$$

and $t_1, t_2 \in \mathbf{N}$ are sufficiently large. For every $b^L \in \mathrm{supp}(\mu)$, let $\delta(b^L) = (\delta / w(x^{-L}, b^L))^{1/|L|}$. Define a cyclic profile σ as follows:

- At stage t the players play the mixed action combination $(1-\delta(m(t)))x + \delta(m(t))(x^{-L(t)}, m(t))$.

If the players follow σ then the probability of absorption at each stage t is $\delta(m(t))^{|L(t)|} w(x^{-L(t)}, m(t)) = \delta$. Fix consecutive T stages. By (19) and (20), the probability that the game is absorbed by a neighbor $b^L \in \mathrm{supp}(\mu)$, given absorption occurs in these T stages, is ϵ-close to $\mu(b^L)$. It follows that (7) holds.

Let $\eta \in (0, \epsilon/3)$ be sufficiently small such that for every $b^L \in \mathrm{supp}(\mu)$ and $y \in B(x, \eta)$ we have $\| u(x^{-L}, b^L) - u(y^{-L}, b^L) \| < \epsilon$. Define a statistical test f as follows. Each player i is checked for the following:

1. Whether his realized actions are compatible with σ.

2. For every $b^L \in \mathrm{supp}(\mu)$ such that $i \notin L$, whether the distribution of his realized actions, restricted to stages j such that $m(j) = b^L$, is η-close to x^i. This check is done only after stage $t_1 T$.

3. For every $b^L \in \mathrm{supp}(\mu)$ such that $i \in L$, whether player i plays the action b^i during stages j such that $m(j) = b^L$ with probability $\delta(b^L)$ (that is, the realized probability p should satisfy $1 - \eta/|N| < p/\delta(b^L) < 1 + \eta/|N|$). This check is done only after stage $t_2 T$.

4. If $\mathrm{supp}(\mu) \subseteq \mathcal{B}_i(x)$, whether the game was absorbed before stage t_0 (where t_0 is defined below).

The second test checks whether the players play mainly the mixed action combination x, and the third test checks whether the players perturb to absorbing neighbors b^L such that $|L| \geq 2$ in the pre-specified frequencies. By the second condition, it is not necessary to check whether players perturb to single absorbing neighbors in the pre-specified frequencies. Moreover, such a check cannot be done. However, if *all* the absorbing neighbors in μ are single absorbing neighbors of player i, then it might be in the interest of player i *never* to perturb to his single absorbing neighbors, and to receive the non-absorbing payoff forever. The last test takes care of this type of deviation.

Formally, $f^i(a_1, \ldots, a_t) = 0$ if and only if at least one of the following holds:

- $i \notin L(t)$ and $a_t^i \notin \mathrm{supp}(x^i)$.
- $i \in L(t)$ and $a_t^i \notin \mathrm{supp}(x^i) \cup \{m^i(t)\}$.
- $t \geq t_1 T$, $i \notin L(t)$ and $\left\| x^i - \frac{\sum_{j<t \mid m(j)=m(t)} a_j^i}{t} \right\| > \eta$.
- $t \geq t_2 T$, $i \in L(t)$ and
$$\left| 1 - \#\{j < t \mid m(j) = m(t) \text{ and } a_j^i = m^i(t)\}/\delta(m(t)) \right| < \eta/|N|.$$
- If $\mathrm{supp}(\mu) \subseteq \mathcal{B}_i(x)$ and $t \geq t_0$, where $t_0 \in \mathbf{N}$ satisfies that $(1 - \delta(m(1)))^{t_0} < \epsilon$.

Note that if no deviation is ever detected, then the game is bounded to be eventually absorbed. By (19), (20), the second condition and the definitions of η and f it follows that (8) holds.

It is left to prove that (9) holds, and that no player can gain too much by deviating.

We claim that it is sufficient to prove the following:

a) If the game is absorbed before a deviation is detected, while player i plays an action $a^i \in \mathcal{B}_i(x) \cap \operatorname{supp}(\mu)$, then player i's expected payoff is at most $g^i + \epsilon$.

b) If player i deviates, by altering the probability in which he plays actions within $\operatorname{supp}(x^i)$, or the action $m^i(t)$ at stage t with $|L(t)| \geq 2$, then the probability of absorption before the statistical test is employed is at most ϵ.

c) By a detectable deviation no player can profit more than $2\epsilon R$.

d) The probability of false detection of deviation is bounded by ϵ (that is, (9) holds).

Indeed, (a)-(c) imply that Step f holds. Note that (a) holds by the second condition and the definition of σ, and (c) holds by the first condition and the definition of σ.

Let us now see how to choose the constants δ,t_1 and t_2 such that (b) and (d) will hold. To insure (d) we need for the second test that

$$\Pr\left(\| \bar{X}^i_t - x^i \| < \eta \quad \forall t > t_1 T\right) > 1 - \epsilon/2|N| \qquad \forall i \in N \tag{21}$$

where $\bar{X}^i_t = \frac{1}{t}\sum_{j=1}^t X^i_j$ and $\{X^i_j\}$ are i.i.d. r.v. with distribution x^i. For the third test we need that

$$\Pr\left(\| \bar{Y}^i_t/\delta(m(t)) - 1 \| < \eta \quad \forall t > t_2 T\right) > 1 - \epsilon/2|N| \qquad \forall i \in N \tag{22}$$

where $\bar{Y}^i_t = \frac{1}{t}\sum_{j=1}^t Y^i_j$ and $\{Y^i_j\}$ are i.i.d Bernoulli r.v. with $P(Y^i_j = 1) = \delta(m(t))$. To insure (b) we need for the first test that

$$(1-\delta)^{t_1 T} > 1 - \epsilon \tag{23}$$

and for the third test that

$$\left(1-\delta^{(|L(t)|-1)/|L(t)|}\right)^{t_2T} \geq (1-\delta^{1/2})^{t_2T} > 1-\epsilon \qquad \forall t=1,\ldots,T. \quad (24)$$

We claim that there exist δ, t_1 and t_2 such that (21), (22), (23) and (24) hold. We need the following lemma:

LEMMA 4.5 *Let* $\epsilon > 0$, $p = 1/n$ *for some* $n \in \mathbf{N}$ *and* $(X_t)_{t\in\mathbf{N}}$ *be i.i.d. Bernoulli random variables with* $P(X_t = 1) = p$. *There exists* $t_\star \in \mathbf{N}$ *such that*

$$P\left(\left|\frac{\sum_{j=1}^{t} X_j}{tp} - 1\right| < 2\epsilon \qquad \forall t > \frac{t_\star}{p}\right) > 1-\epsilon. \quad (25)$$

Proof: Let $\lambda \in (1, 1+\epsilon)$ and $t_\star = \lambda/\epsilon^3(\lambda-1)$. By Kolmogorov's Inequality (see, e.g., Lamperti [18], p. 46), for every $k \in \mathbf{N}$

$$\begin{aligned}\Pr\left(\max_{\lambda^k t_\star/p < t \leq \lambda^{k+1}t_\star/p}\left|\sum_{j=1}^{t}(X_j-p)\right| < \epsilon\lambda^{k+1}t_\star p\right) &\leq \frac{\lambda^{k+1}t_\star p(1-p)}{\epsilon^2\lambda^{2(k+1)}t_\star^2p^2} \\ &< \frac{1}{\epsilon^2\lambda^{k+1}t_\star p} \qquad (26)\end{aligned}$$

Summing (26) over all $k \geq 0$ yields

$$\Pr\left(\max_{t_\star/p<t}\left|\sum_{j=1}^{t}(X_j-p)\right| < 2\epsilon t\right) < \frac{\lambda}{\epsilon^2 t_\star(\lambda-1)} \leq \epsilon$$

and (25) follows. ■

Let $t_1 \in \mathbf{N}$ be sufficiently large to satisfy (21). Let $\rho_0 > 0$ be such that for every $\rho \in (0, \rho_0)$

$$(1-\rho)^{\rho^{-1/2}} = \left((1-\rho)^{1/\rho}\right)^{\rho^{1/2}} > 1-\epsilon. \quad (27)$$

Let $\delta \in (0, \rho_0^2)$ be sufficiently small such that (23) holds. Denote $t_2 = 1/T\delta^{1/4}$. We assume δ is sufficiently small such that $t_\star = t_2T$ satisfies Lemma 4.5. Since $\delta^{1/2} < \rho_0$, it follows by (27) that (24) holds, and since t_2T satisfies Lemma 4.5, (22) holds. ■

4.2.5 A Cyclic Equilibrium

Example 6

	W	
	L	R
T	0, 0, 0	0, 1, 3 *
B	1, 3, 0 *	1, 0, 1 *

	E	
	L	R
T	3, 0, 1 *	1, 1, 0 *
B	0, 1, 1 *	0, 0, 0 *

This game was studied by Flesch et al. (1997). The game is symmetric in the sense that for every player i and action combination $a = (a^1, a^2, a^3) \neq (T, L, W)$ we have

$$u^i(a^1, a^2, a^3) = u^{i+1 \bmod 3}(a^3, a^1, a^2),$$

where we identify $T = L = W$ and $B = R = E$.

Flesch et al. (1997) prove that $(1, 2, 1)$ is an equilibrium payoff. The equilibrium profile that Flesch et al. suggest is the following.

- At the first stage, the players play $(\frac{1}{2}T + \frac{1}{2}B, L, W)$.
- At the second stage, the players play $(T, \frac{1}{2}L + \frac{1}{2}R, W)$.
- At the third stage, the players play $(T, L, \frac{1}{2}W + \frac{1}{2}E)$.
- Afterwards, the players play cyclicly those three mixed-action combinations, until absorption occurs.

If the players follow this profile then their expected payoff is $(1, 2, 1)$, and it can easily be checked that no player has a profitable deviation.

Let $\epsilon > 0$. A more robust ϵ-equilibrium profile for this game, that does not depend on the payoffs of the cells $(T, R, E), (B, R, W)$ and (B, L, E) is the following. Let $\delta < \epsilon/R$ (where R is the maximal payoff in absolute values) and $n \in \mathbf{N}$ satisfy that $(1 - \delta)^n = 1/2$.

- The players play $((1 - \delta)T + \delta B, L, W)$ for n stages (thus, the overall probability to be absorbed by the action combination (B, L, W) is $1/2$).
- Then the players play $(T, (1 - \delta)L + \delta R, W)$ for n stages.

- Then the players play $(T, L, (1-\delta)W + \delta E)$ for n stages.
- Afterwards, the players play cyclicly those three phases, until absorption occurs.

■

Definition 4.6 *Let $a, b, c \in \mathbf{R}^3$. The three vectors (a, b, c) are* left-cyclic *if $b_1 > a_1 > c_1$, $c_2 > b_2 > a_2$ and $a_3 > c_3 > b_3$, and* right-cyclic *if $b_1 < a_1 < c_1$, $c_2 < b_2 < a_2$ and $a_3 < c_3 < b_3$. They are* cyclic *if they are either left-cyclic or right-cyclic. They are* positive cyclic *if they are cyclic and*

$$\det \begin{pmatrix} 0 & b_1 - a_1 & c_1 - a_1 \\ a_2 - b_2 & 0 & c_2 - b_2 \\ a_3 - c_3 & b_3 - c_3 & 0 \end{pmatrix} > 0.$$

Whenever we say that three vectors are cyclic, it should be understood that the first vector serves as the a in the above definition, the second vector serves as the b, and the third serves as the c.

Note that if (a, b, c) are left-cyclic, then $((a_1, a_3, a_2), (c_1, c_3, c_2), (b_1, b_3, b_2))$ are right-cyclic.

Lemma 4.7 *If (a, b, c) are positive right-cyclic vectors then the system of equations*

$$a_1 = \frac{\beta b_1 + (1-\beta)\gamma c_1}{\beta + (1-\beta)\gamma} \tag{28}$$

$$b_2 = \frac{\gamma c_2 + (1-\gamma)\alpha a_2}{\gamma + (1-\gamma)\alpha} \tag{29}$$

$$c_3 = \frac{\alpha a_3 + (1-\alpha)\beta b_3}{\alpha + (1-\alpha)\beta} \tag{30}$$

has a unique solution. Moreover, this solution satisfies $\alpha, \beta, \gamma \in (0, 1)$.

Proof: Assume w.l.o.g. that $a_1 = b_2 = c_3 = 0$. Since the vectors are cyclic, it follows that $a_2, a_3, b_1, b_3, c_1, c_2 \neq 0$. Note that every solution (α, β, γ) satisfies that $\alpha, \beta, \gamma \notin \{0, 1\}$.

We are going now to calculate β. By (29) and (30) we have

$$\alpha = \frac{-\gamma c_2}{(1-\gamma)a_2} = \frac{\beta b_3}{\beta b_3 - a_3}, \tag{31}$$

and by (28) we have

$$\gamma = \frac{-\beta b_1}{(1-\beta)c_1}. \tag{32}$$

Substituting (32) in (31) and dividing by β yields

$$\frac{b_1 c_2}{c_1 - \beta c_1 + \beta b_1} = \frac{b_3 a_2}{\beta b_3 - a_3}.$$

Hence

$$\beta = \frac{a_2 b_3 c_1 + a_3 b_1 c_2}{b_3 b_1 c_2 + a_2 b_3 c_1 - a_2 b_3 b_1}$$

is uniquely determined. Since (a, b, c) are right-cyclic it follows that the denominator is positive, while since they are positive cyclic, the numerator is also positive. Hence $\beta > 0$. To prove that $\beta < 1$ it is sufficient to prove that $b_3 b_1 c_2 - a_2 b_3 b_1 - a_3 b_1 c_2 > 0$, which holds since (a, b, c) are right-cyclic.

In a similar way we prove that $\alpha, \gamma \in (0, 1)$. ■

The following sufficient condition is given only for three-player repeated games with absorbing states, hence we assume that $N = \{1, 2, 3\}$.

Lemma 4.8 *Let $x \in X$ be a non-absorbing mixed action combination and for every $i \in N$ let $y^i \in X^i$ such that*

1) *For every $i \in N$, $w(x^{-i}, y^i) > 0$ and $u^i(x^{-i}, y^i) \geq E^i(x)$.*
2) *$(u(x^{-1}, y^1), u(x^{-2}, y^2), u(x^{-3}, y^3))$ are positive cyclic vectors.*
3) *For every player i and every action $a^i \in \operatorname{supp}(y^i)$, $w(x^{-i}, a^i) > 0$ and $u^i(x^{-i}, a^i) = u^i(x^{-i}, y^i)$.*

Then there exists an equilibrium payoff.

Proof: Assume w.l.o.g. that $(u(x^{-1}, y^1), u(x^{-2}, y^2), u(x^{-3}, y^3))$ are right-cyclic (otherwise, change the names of players 2 and 3, and recall the remark after Definition 4.6). By condition 2 and Lemma 4.7 there exist $\alpha, \beta, \gamma \in (0,1)$ such that

$$\begin{aligned} u^1(x^{-1}, y^1) &= \frac{\beta u^1(x^{-2}, y^2) + (1-\beta)\gamma u^1(x^{-3}, y^3)}{\beta + (1-\beta)\gamma}, & (33) \\ u^2(x^{-2}, y^2) &= \frac{\gamma u^2(x^{-3}, y^3) + (1-\gamma)\alpha u^2(x^{-1}, y^1)}{\gamma + (1-\gamma)\alpha} \quad \text{and} & (34) \\ u^3(x^{-3}, y^3) &= \frac{\alpha u^3(x^{-1}, y^1) + (1-\alpha)\beta u^3(x^{-2}, y^2)}{\alpha + (1-\alpha)\beta}. & (35) \end{aligned}$$

Let $\epsilon > 0$ be fixed. Let $\delta_1, \delta_2, \delta_3 \in (0, \epsilon)$ be sufficiently small and $n_1, n_2, n_3 \in \mathbf{N}$ satisfy the following

$$\begin{aligned} (1 - \delta_1 w(x^{-1}, y^1))^{n_1} &= 1 - \alpha \\ (1 - \delta_2 w(x^{-2}, y^2))^{n_2} &= 1 - \beta \\ (1 - \delta_3 w(x^{-3}, y^3))^{n_3} &= 1 - \gamma. \end{aligned} \tag{36}$$

Define a profile σ as follows:

- Phase 1: The players play the mixed action combination $(1-\delta_1)x + \delta_1(x^{-1}, y^1)$ for n_1 stages.
- Phase 2: The players play the mixed action combination $(1-\delta_2)x + \delta_2(x^{-2}, y^2)$ for n_2 stages.
- Phase 3: The players play the mixed action combination $(1-\delta_3)x + \delta_3(x^{-3}, y^3)$ for n_3 stages.
- The players repeat cyclicly these three phases until absorption occurs.

If the players follow σ then the probability that the game is absorbed during the first phase is $1-(1-\delta_1 w(x^{-1}, y^1))^{n_1} = \alpha$. Similarly, the probability that the game is absorbed during the second and third phases are β and γ respectively. Hence the game will be eventually absorbed with probability 1.

We first calculate the expected payoff for the players if they follow σ. The expected payoff for player 1 is, by (33),

$$\frac{\alpha u^1(x^{-1}, y^1) + (1-\alpha)\beta u^1(x^{-2}, y^2) + (1-\alpha)(1-\beta)\gamma u^1(x^{-3}, y^3)}{1 - (1-\alpha)(1-\beta)(1-\gamma)} = u^1(x^{-1}, y^1).$$

Moreover, for every $j \leq n_1$ his expected payoff given absorption has not occurred in the first j stages is $u^1(x^{-1}, y^1)$.

Similarly, the expected payoff of player 2 given absorption has not occurred during the first n_1 stages is $u^2(x^{-2}, y^2)$. By condition 2, $u^2(x^{-1}, y^1) > u^2(x^{-2}, y^2)$, and therefore the expected payoff of player 2 is

$$\alpha u^2(x^{-1}, y^1) + (1 - \alpha) u^2(x^{-2}, y^2) \geq u^2(x^{-2}, y^2).$$

Moreover, for every $j \leq n_1$ the expected payoff for player 2 given absorption has not occurred in the first j stages is at least $u^2(x^{-2}, y^2)$.

In a similar way, the expected payoff of player 3 given absorption has not occurred during the first $n_1 + n_2$ stages is $u^3(x^{-3}, y^3)$. Since the profile is cyclic, it follows that his expected payoff at the beginning of the game is $u^3(x^{-3}, y^3)$. By condition 2, $u^3(x^{-1}, y^1) < u^3(x^{-3}, y^3)$, and therefore for every $j \leq n_1$ his expected payoff given absorption has not occurred during the first j stages is at least $u^3(x^{-3}, y^3)$.

Denote $g = (u^1(x^{-1}, y^1), \alpha u^2(x^{-1}, y^1) + (1 - \alpha) u^2(x^{-2}, y^2), u^3(x^{-3}, y^3))$. Then (7) holds. Since $(u(x^{-1}, y^1), u(x^{-2}, y^2), u(x^{-3}, y^3))$ are right-cyclic and by the first assumption, it follows that

$$g^i \geq u^i(x^{-i}, y^i) \geq E^i(x) \qquad \forall i \in N.$$

Let $\eta \in (0, \epsilon)$ sufficiently small such that for every $z \in B(x, \eta)$ we have $w(z^{-i}, y^i) > w(x^{-i}, y^i)/2$ and $\parallel u(x^{-i}, y^i) - u(z^{-i}, y^i) \parallel < \epsilon$.

Define a statistical test f as follows. Each player i is checked for the following:

- Whether his realized action is compatible with σ.
- Let $t_0 = k(n_1 + n_2 + n_3)$ be the first stage of phase 1 at the $k + 1$st cycle. At each stage t such that $t_0 + t_1 \leq t \leq t_0 + n_1$ (where t_1 is defined below) players 2 and 3 are checked whether the distribution of their realized actions at stages $t_0, t_0 + 1, \ldots, t - 1$ is η-close to x^2 and x^3 respectively.

  Analogous checks are done in phases 2 and 3.

Formally, the statistical test is defined as follows. Let $t, k \in \mathbf{N}$ satisfy that $k(n_1 + n_2 + n_3) \leq t < k(n_1 + n_2 + n_3) + n_1$. $f^i(a_1, a_2, \ldots, a_t) = 0$ if and only if at least one of the following holds:

- $i \neq 1$ and $a_t^i \notin \text{supp}(x^i)$.
- $i = 1$ and $a_t^1 \notin \text{supp}(x^1) \cup \text{supp}(y^1)$.
- $i \neq 1$, $t \geq t_0 + t_1$, where $t_0 = k(n_1 + n_2 + n_3)$, and $\left\| \frac{\sum_{j=t_0}^{t-1} a_j^i}{t-t_0} - x^i \right\| > \eta$.

The function f is defined analogously for every t that satisfies $k(n_1 + n_2 + n_3) + n_1 \leq t < (k+1)(n_1 + n_2 + n_3)$ for some $k \in \mathbf{N}$. By condition 3 and the definitions of f and η, (8) holds.

We choose the various constants in the following way. Let $k_0 \in \mathbf{N}$ be sufficiently large such that if no deviation is detected in σ_f, and at least one of the players follows σ_f, then absorption occurs during the first k_0 cycles with probability greater than $1 - \epsilon/2$. Formally,

$$(1 - \alpha/2)^{k_0}, (1 - \beta/2)^{k_0}, (1 - \gamma/2)^{k_0} < \epsilon/2. \tag{37}$$

Let $t_1 \in \mathbf{N}$ be sufficiently large such that

$$\Pr\left(\| \bar{X}_t^i - x^i \| < \eta \quad \forall t > t_1 \right) > 1 - \epsilon/6k_0 \qquad \forall i \in N \tag{38}$$

where $\bar{X}_t^i = \frac{1}{t}\sum_{j=1}^t X_j^i$ and $\{X_j^i\}$ are i.i.d. r.v. with distribution x^i. By (37) and (38) it follows that (9) holds.

Let $\delta_1, \delta_2, \delta_3 > 0$ be sufficiently small such that

$$(1 - \delta_i)^{t_1} > 1 - \epsilon \qquad \forall i \in N. \tag{39}$$

Moreover, we choose $\{n_i\}$ and $\{\delta_i\}$ in such a way that (36) holds.

Let $t_3 = k_0(n_1 + n_2 + n_3)$. If no deviation is detected then absorption occurs before stage t_3 with probability greater than $1 - \epsilon/2$. Let $t_4 \in \mathbf{N}$ be sufficiently large such that

$$\frac{t_3 R + (t_4 - t_3)c^i}{t_4} \leq c^i + \epsilon \qquad \forall i \in N. \tag{40}$$

We will show that in phase 1 no player can deviate and profit more than $2\epsilon R$. The proofs for the other phases is analogous.

Fix a player i, a strategy τ^i of player i and a history $h_0 \in Z(f)$ with index i such that $L(h_0) < n_1$. If in h_0 player i fails the second test then by (39) the probability of absorption before the statistical test is employed is smaller than ϵ, and by the first condition and the choice of t_4, for every $t \geq t_4$:

$$\mathbf{E}_{h_0,\sigma_f^{-i},\tau^i}\left(\frac{r_1^i + \cdots + r_t^i}{t}\right) \leq \frac{t_3 R + (t-t_3)(E^i(x)+\epsilon)}{t} \leq g^i + 2\epsilon \tag{41}$$

and

$$\mathbf{E}_{h_0,\sigma_f^{-i},\tau^i}\left(\liminf_{t\to\infty}\frac{r_1^i + \cdots + r_t^i}{t}\right) \leq E^i(x)+\epsilon \leq g^i + \epsilon. \tag{42}$$

If, on the other hand, player i fails the first test, then by the first condition, Eqs. (41) and (42) hold.

Thus σ_f is a $2\epsilon R$-equilibrium profile for g, and g is an x-perturbed equilibrium payoff. ■

4.2.6 A Generalization

We now generalize the third and fourth sufficient conditions to a single sufficient condition, that holds for an arbitrary number of players.

Lemma 4.9 *Let x be a non-absorbing mixed-action combination. For every $j \in \mathbf{N}$, let $\mu_j \in \Delta(\mathcal{B}(x))$ and $\alpha_j \in (0,1]$ such that $\sum_{j\in\mathbf{N}} \alpha_j = \infty$. Let $(r_j)_{j\in\mathbf{N}}$ be the unique solution of the following system of linear equation:*

$$r_j = \alpha_j\left(\sum_{b^L\in\mathcal{B}(x)} \mu_j(b^L)u(x^{-L},b^L)\right) + (1-\alpha_j)r_{j+1} \qquad \forall j \in \mathbf{N}. \tag{43}$$

Assume that the following conditions hold:

1) *$r_j^i \geq E^i(x)$ for every $i, j \in \mathbf{N}$.*
2) *For every $j \in \mathbf{N}$ and every player $i \in N$ such that there exists a single absorbing perturbation $b^i \in \mathrm{supp}(\mu_j)$ of player i we have:*

$$r_{j+1}^i = u^i(x^{-i},b^i) = r_j^i.$$

Then for every $j \in \mathbf{N}$, r_j is an x-perturbed equilibrium payoff.

Note that since $\sum_{j\in\mathbf{N}} \alpha_j = \infty$ it follows that the system of linear equations (43) has a unique solution.

The ϵ-equilibrium profile is constructed in phases: in phase j, the players play an $\epsilon/2^j$-equilibrium profile as defined in the proof of Lemma 4.4 for μ_j, either until the game is absorbed, or a deviation is detected, or until the overall probability of absorption during phase j is at least $\alpha_j - \epsilon/2^j$. If the players follow this profile then the overall probability of false detection of deviation is bounded by ϵ, and since $\sum_{j\in\mathbf{N}} \alpha_j = \infty$, if no deviation is detected then the game will eventually be absorbed with probability 1. As in the proof of Lemma 4.4, no player can profit more than ϵ by deviating.

Note that this equilibrium can be generalized to equilibrium profiles which are *not* perturbed, by using at stage j a (possibly) different mixed-action combination x_j, instead of the same x always.

4.3 On the Discounted Game

Recall that $v^i_\beta(x)$ is the expected β-discounted payoff for player i if the players follow the stationary profile x.

In this section we recall a fundamental formula that is already derived by Vrieze and Thuijsman [36] for the discounted payoff that is used later, and we prove that for every $\epsilon > 0$ there exists an interval $(\beta(\epsilon), 1)$ such that for every mixed action x^{-i} of the players $N \setminus \{i\}$, player i has a reply which is ϵ-good in every β-discounted game (where $\beta \in (\beta(\epsilon), 1)$).

The function $v_\beta(x)$ satisfies the recursion formula:

$$v_\beta(x) = (1-\beta)h(x) + \beta w(x)u(x) + \beta(1-w(x))v_\beta(x). \tag{44}$$

Since h and uw are multi-linear, it follows that v_β is quasi-concave. Indeed, assume that for $x, y \in X$

$$\frac{(1-\beta)h(x) + \beta w(x)u(x)}{1-\beta+\beta w(x)}, \frac{(1-\beta)h(y) + \beta w(y)u(y)}{1-\beta+\beta w(y)} \geq c.$$

Then

$$\begin{aligned} & v_\beta(\lambda x + (1-\lambda)y) = \\ & = \frac{\lambda\left((1-\beta)h(x) + \beta w(x)u(x)\right) + (1-\lambda)\left((1-\beta)h(y) + \beta w(y)u(y)\right)}{\lambda(1-\beta+\beta w(x)) + (1-\lambda)(1-\beta+\beta w(y))} \\ & \geq c. \end{aligned} \tag{45}$$

The solution of (44) is:

$$v_\beta(x) = \frac{(1-\beta)h(x) + \beta w(x)u(x)}{1-\beta+\beta w(x)}.$$

Let

$$\alpha_\beta(x) = \frac{1-\beta}{1-\beta+\beta w(x)} \tag{46}$$

then

$$v_\beta(x) = \alpha_\beta(x)h(x) + (1-\alpha_\beta(x))u(x). \tag{47}$$

In other words, $v_\beta(x)$ is a convex combination of the non-absorbing payoff $h(x)$ and the absorbing payoff $u(x)$. The weight of each factor depends on the absorbing probability $w(x)$ and on the discount factor β.

$\alpha_\beta(x)$ is defined only for $\beta \in [0,1)$, but whenever $w(x) > 0$ we can define it continuously for $\beta = 1$.

If $w(x) = 0$ then $\alpha_\beta(x) = 1$ for every β, and therefore $v_\beta(x) = h(x)$ — if the strategy is non-absorbing then the expected discounted payoff for the players is equal to the non-absorbing payoff, whatever the discount factor is.

For every fixed x, if $w(x) > 0$ then $\lim_{\beta\to 1}\alpha_\beta(x) = 0$, which means that if x is absorbing then $\lim_{\beta\to 1} v_\beta(x) = u(x)$.

For every $\epsilon > 0$, define

$$\beta(\epsilon) = \inf\{\beta' \in [0,1) \mid |c^i(\beta) - c^i| \leq \epsilon \quad \forall \beta \in (\beta', 1)\}.$$

By Lemma 3.6, $\beta(\epsilon) < 1$ for every $\epsilon > 0$.

For every $\epsilon > 0$, every player $i \in N$, every action $a^i \in A^i$ and every stationary profile $x^{-i} \in X^{-i}$, let

$$\Gamma_\epsilon(x^{-i}, a^i) = \{\beta \in (\beta(\epsilon), 1) \mid v^i_\beta(x^{-i}, a^i) \geq c^i - \epsilon\}.$$

That is, the set of all discount factors β, such that a^i is an ϵ-good reply of player i against x^{-i} in the β-discounted game.

By the definition of $\beta(\epsilon)$,

$$\cup_{a^i\in A^i}\Gamma_\epsilon(x^{-i}, a^i) = (\beta(\epsilon), 1). \tag{48}$$

By (47), if $\Gamma_\epsilon(x^{-i}, a^i) \neq \emptyset$ then at least one of the following inequalities hold:

$$h^i(x^{-i}, a^i) \geq c^i - \epsilon \qquad \text{or} \tag{49}$$

$$u^i(x^{-i}, a^i) \geq c^i - \epsilon. \tag{50}$$

Clearly if for a given pair (x^{-i}, a^i) both (49) and (50) hold, then $\Gamma_\epsilon(x^{-i}, a^i) = (\beta(\epsilon), 1)$. If $\Gamma_\epsilon(x^{-i}, a^i) \neq \emptyset$, $w(x^{-i}, a^i) > 0$ and $h^i(x^{-i}, a^i) < c^i - \epsilon$ then $u^i(x^{-i}, a^i) > c^i - \epsilon$ and vice versa, if $u^i(x^{-i}, a^i) < c^i - \epsilon$ then $h^i(x^{-i}, a^i) > c^i - \epsilon$. By the continuity of $v_\beta(x)$, as a function of β for every fixed x, $\Gamma_\epsilon(x^{-i}, a^i)$ is relatively closed in $(\beta(\epsilon), 1)$.

LEMMA 4.10 *Let* $\epsilon > 0$, $i \in N$, $x^{-i} \in X^{-i}$ *and* $a^i \in A^i$. *The set* $\Gamma_\epsilon(x^{-i}, a^i)$ *is either empty, or has the form* $(\beta(\epsilon), \beta_1]$, $[\beta_2, 1)$ *or* $(\beta(\epsilon), 1)$.

Proof: For every fixed $x \in X$, the function

$$\alpha_\beta(x) = 1 - \frac{\beta w(x)}{1 - \beta + \beta w(x)}$$

is monotonic decreasing in β.

Assume that $\beta' \in \Gamma_\epsilon(x^{-i}, a^i)$. Since $\alpha_\beta(x^{-i}, a^i)$ is monotonic decreasing in β, if (49) holds then every $\beta \in (\beta(\epsilon), \beta')$ is also in $\Gamma_\epsilon(x^{-i}, a^i)$. Symmetrically, if (50) holds then every $\beta \in (\beta', 1)$ is also in $\Gamma_\epsilon(x^{-i}, a^i)$. Since $\Gamma_\epsilon(x^{-i}, a^i)$ is relatively closed in $(\beta(\epsilon), 1)$ the result follows. ■

Since $\lim_{\beta \to 1} \alpha_\beta(x) = 0$ whenever $w(x) > 0$ and by the continuity of w and u^i we get:

LEMMA 4.11 *Let* $\epsilon > 0$, $i \in N$, $x^{-i} \in X^{-i}$ *and* $a^i \in A^i$ *such that* $w(x^{-i}, a^i) > 0$. *If there exists* $\beta' \in (\beta(\epsilon), 1)$ *such that* $[\beta', 1) \subseteq \Gamma_\epsilon(x^{-i}, a^i)$ *then* $u^i(x^{-i}, a^i) \geq c^i - \epsilon$, *whereas if* $\Gamma_\epsilon(x^{-i}, a^i) = (\beta(\epsilon), \beta']$ *for some* $\beta' < 1$ *then* $u^i(x^{-i}, a^i) < c^i - \epsilon$.

The result that we need in the next section is the following.

LEMMA 4.12 *Let* $\epsilon > 0$, $i \in N$ *and* $x^{-i} \in X^{-i}$. *Assume that there is no action* $a^i \in A^i$ *such that* $\Gamma_\epsilon(x^{-i}, a^i) = (\beta(\epsilon), 1)$. *Then there exists* $y^i \in X^i$ *such that* $h^i(x^{-i}, y^i) = c^i - \epsilon$ *and* $u^i(x^{-i}, y^i) \geq c^i - \epsilon$.

Proof: Let $x^{-i} \in X^{-i}$ satisfy the assumptions. Since A^i is finite, by (48) and Lemma 4.10, there exist actions a^i and b^i of player i such that $\Gamma_\epsilon(x^{-i}, a^i) = (\beta(\epsilon), \beta_1]$, $\Gamma_\epsilon(x^{-i}, b^i) = [\beta_2, 1)$ and $\beta_1 \geq \beta_2$. Indeed, let a^i be an action that maximizes β_1 among all actions $a' \in A^i$ such that $\Gamma_\epsilon(a', x^{-i}) = (\beta(\epsilon), \beta_1]$, and let b^i be an action that minimizes β_2 among all actions $a' \in A^i$ such that $\Gamma_\epsilon(a', x^{-i}) = [\beta_2, 1)$. It follows from Lemma 4.10 and (48) that $\beta_1 \geq \beta_2$, as desired.

Note that $w(x^{-i}, a^i) > 0$, otherwise $h^i(x^{-i}, a^i) \geq c^i - \epsilon$, which implies that $\Gamma_\epsilon(x^{-i}, a^i) = (\beta(\epsilon), 1)$. Similarly, $w(x^{-i}, b^i) > 0$.

By Lemma 4.11, $u^i(x^{-i}, b^i) \geq c^i - \epsilon$ and $u^i(x^{-i}, a^i) < c^i - \epsilon$. By (47) and the assumption it follows that $h^i(x^{-i}, b^i) < c^i - \epsilon$ and $h^i(x^{-i}, a^i) > c^i - \epsilon$.

Let $q \in (0,1)$ solve the equation

$$qh^i(x^{-i}, a^i) + (1-q)h^i(x^{-i}, b^i) = c^i - \epsilon.$$

Let y^i be the mixed action of player i where he plays the action a^i with probability q and the action b^i with probability $1-q$.

Clearly

$$h^i(x^{-i}, y^i) = qh^i(x^{-i}, a^i) + (1-q)h^i(x^{-i}, b^i) = c^i - \epsilon \tag{51}$$

and

$$w(x^{-i}, y^i) = qw(x^{-i}, a^i) + (1-q)w(x^{-i}, b^i) > 0. \tag{52}$$

Since both $v^i_\beta(x^{-i}, a^i) \geq c^i - \epsilon$ and $v^i_\beta(x^{-i}, b^i) \geq c^i - \epsilon$ we get by the quasi-concavity of v^i_β that $v^i_\beta(x^{-i}, y^i) \geq c^i - \epsilon$. By (47), (51) and (52) we get that $u^i(x^{-i}, y^i) \geq c^i - \epsilon$, as desired. ■

4.4 The Auxiliary Game

In this section we associate with every repeated game with absorbing states G and a continuous quasi-concave function $\tilde{h} : X \to \mathbf{R}^N$ a new game $\tilde{G}(\tilde{h})$, similar to the original game, but with a different way to calculate the daily payoff — the daily payoff for the players in $\tilde{G}(\tilde{h})$ is given by $\tilde{h}$, instead of h. We then define a specific function $\tilde{h}$ and we prove that the β-discounted min-max value of each player in the $\tilde{G}(\tilde{h})$ exists, its limit, as β tends to 1, is equal to the min-max value of that player in the original game, and that stationary β-discounted equilibrium profiles in the auxiliary game exist.

Let $G = (N, (A^i, h^i, u^i)_{i \in N}, w)$ be a repeated game with absorbing states and $\tilde{h} : X \to \mathbf{R}^N$ be continuous and quasi-concave. The *auxiliary game* $\tilde{G}(\tilde{h})$, is played in the same way as G, but the non-absorbing payoff at stage

t is $\tilde{h}(x_t)$, where x^i_t is the mixed action that the strategy of player i indicates him to play at stage t.

We denote by $\tilde{v}_\beta(x)$ the expected β-discounted payoff in $\tilde{G}$ given a pure stationary profile x is played.

Formally, for every profile $\sigma : H_0 \to X$ and $\beta \in [0,1)$, the expected β-discounted payoff in $\tilde{G}$ is given by:

$$\tilde{v}_\beta(\sigma) = (1-\beta)\sum_{t=1}^{\infty} \beta^{t-1}\mathbf{E}_\sigma \left(\mathbf{1}_{t_\star \geq t}\tilde{h}(x_t) + \mathbf{1}_{t_\star < t}u(x_{t_\star})\right)$$

where $t_\star$ is the stage in which absorption occurs.

An equivalent definition for $\tilde{G}(\tilde{h})$ is that we consider a game with a larger strategy space. Each player chooses at every stage a *mixed action* rather than a pure action, and the non-absorbing payoff is given by the function $\tilde{h}$. The game $\tilde{G}$ is given by this larger game, when the players are restricted to *pure* strategies (that is, at every stage each player chooses one mixed action, and he cannot lotter between some mixed actions).

Example 7

Consider the following two-player zero-sum repeated game with absorbing states:

	y	$1-y$
x	$1,-1$	$0,0$ *
$1-x$	$-1,1$	$0,0$ *

The min-max value of both players in this game is 0. Any stationary profile (x,y) where $x \in [1/2,1]$ and $y=0$ is a discounted equilibrium profile (for every discount factor).

Define the function $\tilde{h} : X \to \mathbf{R}$ by

$$\tilde{h}^i(x) = \min\{h(x), 0\}$$

where $h(x)$ is the non-absorbing payoff. In the auxiliary version $\tilde{G}(\tilde{h})$ there is one more stationary discounted equilibrium, $(x = 1/2, y = 1/2)$, since if player 1 plays $x = 1$ then his expected payoff in the auxiliary game is 0. Note, however, that if player 1 plays $x = 0$ (while player 2 plays $y = 1/2$) then his expected payoff in the β-discounted auxiliary game is strictly less than 0, though $x = 0$ is in the support of his stationary β-discounted equilibrium strategy. ∎

The equivalent of (47) for the auxiliary game is:

$$\tilde{v}_\beta(x) = \alpha_\beta(x)\tilde{h}(x) + (1 - \alpha_\beta(x))u(x) \tag{53}$$

where α_β is given in (46). Since $\tilde{h}$ is quasi-concave and wu is multi-linear, it follows that $\tilde{v}_\beta$ is quasi-concave for every fixed β.

Since $\tilde{v}_\beta$ is continuous over X, it follows that $\tilde{c}^i_\beta$, the β-discounted min-max value of player i in the auxiliary game, exists. Note that, as for standard stochastic games, $\tilde{c}^i_\beta$ is the min-max value when the players can use only stationary strategies.

The result that we need for the auxiliary game in the sequel is the following:

Lemma 4.13 *For every player i, $\lim_{\beta\to 1} \tilde{c}^i(\beta) = c^i$.*

Proof: Define for every $i \in N$ a function $\tilde{h}^i : X \to \mathbf{R}$ by

$$\tilde{h}^i(x) = \min\{h^i(x), c^i\}.$$

Note that $\tilde{h}$ is continuous, quasi-concave and semi-algebraic.

It is clear that $\tilde{c}^i(\beta) \leq c^i(\beta)$ for every $\beta \in (0, 1)$, hence $\lim_{\beta\to 1} \tilde{c}^i(\beta) \leq \lim_{\beta\to 1} c^i(\beta) = c^i$.

For the opposite inequality, let $x^{-i} \in X^{-i}$ and $\epsilon > 0$. First we show that there exists a mixed action $y^i \in X^i$ of player i such that

$$\tilde{v}^i_\beta(x^{-i}, y^i) \geq c^i - \epsilon \qquad \forall \beta \in (\beta(\epsilon), 1). \tag{54}$$

Case 1: There is no action $a^i \in A^i$ of player i such that $\Gamma_\epsilon(x^{-i}, a^i) = (\beta(\epsilon), 1)$.
By lemma 4.12 there exists $y^i \in X^i$ such that $h^i(x^{-i}, y^i) = c^i - \epsilon$ and $u^i(x^{-i}, y^i) \geq c^i - \epsilon$. In particular by (53), $\tilde{v}^i_\beta(x^{-i}, y^i) \geq c^i - \epsilon$ for every $\beta \in (\beta(\epsilon), 1)$.

Case 2: There exists $a^i \in A^i$ such that $\Gamma_\epsilon(x^{-i}, a^i) = (\beta(\epsilon), 1)$.
By the definition of Γ_ϵ, $v^i_\beta(x^{-i}, a^i) \geq c^i - \epsilon$ for every $\beta \in (\beta(\epsilon), 1)$. Therefore, by (47), for every $\beta \in (\beta(\epsilon), 1)$

$$c^i - \epsilon \leq v^i_\beta(x^{-i}, a^i) = \alpha_\beta(x^{-i}, a^i)h^i(x^{-i}, a^i) + (1 - \alpha_\beta(x^{-i}, a^i))u^i(x^{-i}, a^i).$$

If $w(x^{-i}, a^i) = 0$ then $\alpha_\beta(x^{-i}, a^i) = 1$ for every $\beta \in (\beta(\epsilon), 1)$, and therefore $h^i(x^{-i}, a^i) \geq c^i - \epsilon$. In particular $\tilde{v}^i_\beta(x^{-i}, a^i) = \tilde{h}^i(x^{-i}, a^i) \geq c^i - \epsilon$. If, on the other hand, $w(x^{-i}, a^i) > 0$ then $\lim_{\beta \to 1} \alpha_\beta(x^{-i}, a^i) = 0$ and therefore $u^i(x^{-i}, a^i) \geq c^i - \epsilon$. It follows by (53) and the definition of $\tilde{h}$ that $\tilde{v}^i_\beta(x^{-i}, a^i) \geq c^i - \epsilon$, as desired.

Since for every $x^{-i} \in X^{-i}$ there exists $y^i \in X^i$ such that (54) holds, it follows that $\lim_{\beta \to 1} \tilde{c}^i(\beta) \geq c^i - \epsilon$. Since ϵ is arbitrary, the result follows. ■

Remark: Assume that $d^i \geq c^i$ for every $i \in N$. Define a function $\tilde{h}^i : X \to \mathbf{R}$ by

$$\tilde{h}^i(x) = \min\{h^i(x), d^i\}.$$

It then follows that the discounted min-max value of each player i in the auxiliary game defined by the new function $\tilde{h}$ converges to c^i, as the discount factor converges to 1.

Remark: Using a similar method to the one used by Mertens and Neyman [20] and Neyman [22], it is possible to prove that the undiscounted min-max value of each player in the auxiliary game exists, and is equal to his min-max value in the original game.

Remark: One could define the auxiliary game, by changing the values in the cells, instead of changing the multi-linear extension. However, in this case, Lemma 4.13 would not be true. Consider example 1. By re-defining the non-absorbing payoff in each cell to be the minimum of the original non-absorbing payoff and the min-max value (which is 0), we get the following game

	L	R
T	$-1, 1$ *	$1, -1$ *
B	$0, -1$	$-1, 0$

but the min max value of both players in this game is $-\frac{1}{3}$, which is less than their min max value in the original game.

Since $\tilde{v}_\beta$ is quasi concave and continuous for every fixed β, and X is compact, it follows that:

LEMMA 4.14 *For every $\beta \in [0, 1)$ there exists a stationary β-discounted equilibrium in $\tilde{G}$, i.e. there exists a mixed action combination $x \in X$ such that for every player i and mixed action $y^i \in X^i$ of player i,*

$$\tilde{v}^i_\beta(x) \geq \tilde{v}^i_\beta(x^{-i}, y^i).$$

4.5 Puiseux Stationary Profiles

DEFINITION 4.15 *Let I be a finite set. A vector $\hat{f} = (\hat{f}_i)_{i\in I}$ of Puiseux functions is called* a Puiseux probability distribution *if $\hat{f}_i \geq 0$ for every $i \in I$ and $\sum_{i\in I} \hat{f}_i = 1$.*

Clearly we have:

LEMMA 4.16 *Let I be a finite set and $\hat{f}$ a Puiseux probability distribution over I. Then $\hat{f}(1)$ is a probability distribution over I.*

DEFINITION 4.17 *A vector of Puiseux functions $\hat{x} = (\hat{x}^i_{a^i})_{i\in N, a^i\in A^i}$ is a* Puiseux stationary profile *if $\hat{x}^i_{a^i}$ is a Puiseux function for every $i \in N$ and $a^i \in A^i$, and $\hat{x}(\theta) = (\hat{x}^i_{a^i}(\theta))_{i\in N, a^i\in A^i}$ is a stationary profile for every $\theta \in (0, 1)$.*

The following lemma, which follows from Lemma 2.2, connects the degree of the probability of absorption to Equation (53).

LEMMA 4.18 *Let $\hat{x}$ be a Puiseux stationary profile.* $\deg(w(\hat{x})) = 1$ *if and only if* $\lim_{\theta\to 1} \alpha_\theta(\hat{x}(\theta)) \in (0, 1)$. $\deg(w(\hat{x})) < 1$ *if and only if* $\lim_{\theta\to 1} \alpha_\theta(\hat{x}(\theta)) = 0$. $\deg(w(\hat{x})) > 1$ *if and only if* $\lim_{\theta\to 1} \alpha_\theta(\hat{x}(\theta)) = 1$.

LEMMA 4.19 *Let $\hat{x}$ be a Puiseux stationary profile such that* $\lim_{\theta\to 1} \alpha_\theta(\hat{x}(\theta)) \leq 1$. *Let $i \in N$ and $a^i \in A^i$ such that* $\deg(w(\hat{x}^{-i}, a^i)) < \deg(w(\hat{x}))$. *Then*

$$\lim_{\theta\to 1} \tilde{v}_\theta(\hat{x}^{-i}(\theta), a^i) = \lim_{\theta\to 1} u(\hat{x}^{-i}(\theta), a^i).$$

Proof: By the assumption and Lemma 4.18 it follows that $\deg(w(\hat{x}^{-i}, a^i)) < \deg(w(\hat{x})) \leq 1$. By a second use of Lemma 4.18 it follows that $\lim_{\theta \to 1} \alpha_\theta(\hat{x}^{-i}(\theta), a^i) = 0$. The result follows by (53). ■

By (53) and since $\alpha_\beta(x)$ is continuous at $\beta = 1$ for any absorbing mixed action combination x it follows that:

COROLLARY 4.20 *If* $w(\hat{x}(1)) > 0$ *then*

$$\lim_{\theta \to 1} \tilde{v}^i_\theta(\hat{x}(\theta)) = u^i(\hat{x}(1)) = \frac{\sum_{a^i \in \mathrm{supp}(\hat{x}^i(1))} \hat{x}^i_{a^i}(1) w(\hat{x}^{-i}(1), a^i) u^i(\hat{x}^{-i}(1), a^i)}{w(\hat{x}(1))}.$$

A Puiseux stationary profile $\hat{x}$ is *absorbing* if for every θ sufficiently close to 1, $\hat{x}(\theta)$ is an absorbing stationary profile.

Let $\hat{x}$ be an absorbing Puiseux stationary profile. For every $L \subseteq N$ and $b^L \in A^L$ we define the Puiseux function $\hat{x}_{b^L}$ by:

$$\hat{x}_{b^L} = \Pi_{i \in L} \hat{x}^i_{b^i}.$$

Define the probability distribution $\mu_{\hat{x}}$ over $\mathcal{B}(\hat{x}(1))$ as follows:

$$\mu_{\hat{x}}(b^L) = \lim_{\theta \to 1} \frac{\hat{x}_{b^L}(\theta) w(\hat{x}^{-L}(1), b^L)}{\sum_{a^T \in \mathcal{B}(\hat{x}(1))} \hat{x}_{a^T}(\theta) w(\hat{x}^{-T}(1), a^T)}.$$

LEMMA 4.21 *If* $\hat{x}$ *is absorbing but* $\hat{x}(1)$ *is non-absorbing then*

$$\lim_{\theta \to 1} u(\hat{x}(\theta)) = \sum_{b^L \in \mathcal{B}(\hat{x}(1))} \mu_{\hat{x}}(b^L) u(\hat{x}^{-L}(1), b^L).$$

Proof: Define

$$\hat{x}^\star_{-L}(\theta) = \Pi_{i \notin L} \left(\sum_{a^i \in \mathrm{supp}(\hat{x}^i(1))} \hat{x}^i_{a^i}(\theta) \right).$$

$\hat{x}^\star_{-L}(\theta)$ is the probability that the realized action of every player $i \in N \setminus L$ is in $\mathrm{supp}(\hat{x}^i(1))$, given the players $N \setminus L$ play the mixed action combination $\hat{x}^{-L}(\theta)$.

For every mixed action combination x, let

$$\mathcal{B}^0(x) = \{b^L \in A^L \mid L \subseteq N, w(x^{-L}, b^L) > 0\}.$$

Note that $\mathcal{B}^0(x) \supseteq \mathcal{B}(x)$ for every $x \in X$. By the definition of $\mathcal{B}$, if $b^L \in \mathcal{B}^0(\hat{x}(1)) \backslash \mathcal{B}(\hat{x}(1))$ then there exists $a^T \in \mathcal{B}(\hat{x}(1))$ such that $\lim_{\theta \to 1} \frac{w(x^{-L}(\theta), b^L)}{w(x^{-T}(\theta), a^T)} = 0$.

By re-ordering the summation in (6)

$$u(\hat{x}(\theta)) = \sum_{b^L \in \mathcal{B}^0(\hat{x}(1))} \frac{\hat{x}_{b^L}(\theta) \hat{x}^\star_{-L}(\theta) w(\hat{x}^{-L}(1), b^L) u(\hat{x}^{-L}(1), b^L)}{\sum_{a^T \in \mathcal{B}^0(\hat{x}(1))} \hat{x}_{a^T}(\theta) \hat{x}^\star_{-T}(\theta) w(\hat{x}^{-T}(1), a^T)}.$$

Since $\hat{x}^\star_{-L}(1) = 1$ for every $L \subset N$, taking the limit as $\theta \to 1$, and using Lemma 2.2, yields

$$\begin{aligned} \lim_{\theta \to 1} u(\hat{x}(\theta)) &= \lim_{\theta \to 1} \sum_{b^L \in \mathcal{B}(\hat{x}(1))} \frac{\hat{x}_{b^L}(\theta) w(\hat{x}^{-L}(1), b^L) u(\hat{x}^{-L}(1), b^L)}{\sum_{a^T \in \mathcal{B}(\hat{x}(1))} \hat{x}_{a^T}(\theta) w(\hat{x}^{-T}(1), a^T)} \\ &= \sum_{b^L \in \mathcal{B}(\hat{x}(1))} \mu_{\hat{x}}(b^L) u(\hat{x}^{-L}(1), b^L) \end{aligned}$$

as desired. ■

4.6 A Preliminary Result

In this section we prove that there exists a mixed action combination that satisfies some "nice" properties. This mixed action combination is used in the following sections to construct perturbed equilibria in three player repeated games with absorbing states, and in repeated team games with absorbing states.

Lemma 4.22 *There exist $x \in X$, $g \in \mathbf{R}^N$, $\mu \in \Delta(\mathcal{B}(x))$ and $d : \cup_i A^i \to [0, \infty)$ that satisfy:*

a) $g^i \geq E^i(x)$ for every player $i \in N$.

b) At least one of the following holds:

  i) $w(x) = 0$ and $h(x) \geq g$.

  ii) $w(x) > 0$, $u(x) = g$ and $u^i(x^{-i}, a^i) = g^i$ for every player i and every action a^i such that $w(x^{-i}, a^i) > 0$.

  iii) 1) $w(x) = 0$.

2) $\sum_{b^L \in \mathcal{B}(x)} \mu(b^L) u(x^{-L}, b^L) \geq g$.

3) For every player i *and every action* $a^i \notin \mathrm{supp}(x^i)$

$$\sum_{b^L \ \mid \ i \in L, b^i = a^i} \mu(b^L) u^i(x^{-L}, b^L) = \left(\sum_{b^L \ \mid \ i \in L, b^i = a^i} \mu(b^L) \right) g^i.$$

4) If there is strict inequality in (b).iii(2) then $\sum_{i \in L} d(b^i) = 1$ *for every* $b^L \in \mathrm{supp}(\mu)$.

Note that the second condition in (b).ii follows from the third condition in (b).ii.

Condition (a) states that g is a "good" payoff vector, that is, no player can profit by deviating and being punished forever from the second stage. Condition (b).i states that x is non-absorbing and the non-absorbing payoff is "good" for all the players. Condition (b).ii states that x is absorbing, the absorbing payoff is "good" for all the players and the players are indifferent between their actions. Condition (b).iii states that x is non-absorbing, the average of the payoff in the absorbing neighbors of x is "good", while this average is equal to the average given any player i plays an action which is not one of his "main" actions. The last condition in (b).iii, which is technical, is essential for later results.

Proof: The idea is to consider a sequence of β-discounted stationary equilibria in the auxiliary game $\tilde{G}$, that converges to a limit. The values of x, g, μ and d are derived from this sequence, and conditions (a) and (b) are proven using various inequalities that discounted equilibria satisfy, and by taking the limit $\beta \to 1$.

Step 1: Definition of a Puiseux stationary profile $\hat{x}$

Consider the set

$$\mathcal{E} = \left\{ (\beta, x) \in (0,1) \times X \mid \tilde{v}^i_\beta(x) \geq \tilde{v}^i_\beta(x^{-i}, y^i) \qquad \forall i \in N, y^i \in X^i \right\}. \quad (55)$$

$(\beta, x) \in \mathcal{E}$ if and only if x is a β-discounted stationary equilibrium profile in $\tilde{G}$.

Note that $\mathcal{E}$ is semi-algebraic. Indeed, using the recursive formula of the discounted payoff (44) the condition $a = v_\beta(x)$ (as vectors) is a system of equations between polynomials in β, a and x. Therefore $\mathcal{E}$ is a projection of a semi-algebraic set, and in particular it is semi-algebraic (Benedetti and Risler [3]).

By Lemma 4.14 the projection of $\mathcal{E}$ over the first coordinate is the interval $(0,1)$. By Theorem 2.5 there exists a Puiseux stationary profile $\hat{x}$ such that $(\beta, \hat{x}(\beta)) \in \mathcal{E}$ for every $\beta \in (0,1)$.

Step 2: Definition of x, g, μ, d
Define $x = \hat{x}(1)$, $g = \lim_{\beta\to 1} \tilde{v}_\beta(\hat{x}(\beta))$, $\mu = \mu_{\hat{x}}$ and $d(a^i) = \deg(\hat{x}^i_{a^i})$ for every $a^i \in \cup_i A^i$.

Since $\tilde{v}^i_\beta(\hat{x}(\beta))$ is a Puiseux function, g is well defined. By Lemmas 3.6 and 4.13

$$g = \lim_{\beta\to 1} \tilde{v}_\beta(\hat{x}(\beta)) \geq \lim_{\beta\to 1} \tilde{c}(\beta) = c. \tag{56}$$

Step 3: Assertion (a) holds
Fix a player i and an action $a^i \in A^i$. If $w(\hat{x}^{-i}(1), a^i) = 0$ then (a) follows from (56). If, on the other hand, $w(\hat{x}^{-i}(1), a^i) > 0$ then $\lim_{\beta\to 1} \alpha_\beta(\hat{x}^{-i}(\beta), a^i) = 0$. Since $\hat{x}^i(\beta)$ is a best reply against $\hat{x}^{-i}(\beta)$, by (53) and the continuity of u at absorbing mixed action combinations, it follows that

$$g^i = \lim_{\beta\to 1} \tilde{v}^i_\beta(\hat{x}(\beta)) \geq \lim_{\beta\to 1} \tilde{v}^i_\beta(\hat{x}^{-i}(\beta), a^i) = \lim_{\beta\to 1} u^i(\hat{x}^{-i}(\beta), a^i) = u^i(\hat{x}^{-i}(1), a^i). \tag{57}$$

Assertion (a) follows from (56), (57) and the definition of E^i.

Step 4: The distinction between the possibilities in (b)
Denote $q = \lim_{\beta\to 1} \alpha_\beta(\hat{x}(\beta))$. Substituting $\hat{x}(\beta)$ instead of x in (53) and taking the limit as $\beta \to 1$ yields, by the continuity of $\tilde{h}$ and (56),

$$c \leq g = \lim_{\beta\to 1} \tilde{v}_\beta(\hat{x}(\beta)) = q\tilde{h}(\hat{x}(1)) + (1-q) \lim_{\beta\to 1} u(\hat{x}(\beta)). \tag{58}$$

If $q = 1$ then $h(\hat{x}(1)) \geq \tilde{h}(\hat{x}(1)) = g$ and $w(\hat{x}(1)) = 0$. Therefore (b).i holds.

If $w(\hat{x}(1)) > 0$ then $q = 0$ and the second claim in (b).ii holds. In Step 5 we prove that in this case (b).ii holds.

Otherwise $q < 1$ and $w(\hat{x}(1)) = 0$. Therefore $\hat{x}$ is absorbing. Since $\tilde{h}(x) \leq c$, it follows by (58) that $\lim_{\beta\to 1} u(\hat{x}(\beta)) \geq g$, and by Lemma 4.21, (b).iii(2) holds. We shall later see that in this case (b).iii holds.

Step 5: If $w(\hat{x}(1)) > 0$ then assertion (b).ii holds
Let $i \in N$ and $a^i \in \text{supp}(x^i)$ such that $w(x^{-i}, a^i) > 0$. Then $\lim_{\beta \to 1} \alpha_\beta(\hat{x}^{-i}(\beta), a^i) = 1$ and, by the optimality of $\hat{x}^i(\beta)$,

$$u^i(x^{-i}, a^i) = \lim_{\beta \to 1} \tilde{v}^i_\beta(\hat{x}^{-i}(\beta), a^i) \leq \lim_{\beta \to 1} \tilde{v}^i_\beta(\hat{x}(\beta)) = g^i. \tag{59}$$

By corollary 4.20 there is an equality in (59), as desired.

From now on we assume that $q < 1$ and $w(\hat{x}(1)) = 0$.

Step 6: Assertion (b).iii(3) holds
Let $i \in N$ be fixed. For every $a^i \in A^i$, denote $\mathcal{B}(\mu, a^i) = \{b^L \in \text{supp}(\mu) \mid i \in L, b^i = a^i\}$.

By Lemma 4.19 it follows that

$$\begin{aligned} g^i &= \lim \tilde{v}^i_\beta(\hat{x}(\beta)) \\ &= \alpha \lim_{\beta \to 1} \tilde{v}^i_\beta(\hat{x}^{-i}(\beta), \hat{x}^i(1)) + \sum_{a^i \notin \text{supp}(\hat{x}^i(1))} \alpha_{a^i} \lim_{\beta \to 1} \tilde{v}^i_\beta(\hat{x}^{-i}(\beta), a^i) \end{aligned} \tag{60}$$

where $\alpha_{a^i} = (1-q) \sum_{b^L \in \mathcal{B}(\mu, a^i)} \mu(b^L)$, and $\alpha = 1 - \sum_{a^i \notin \text{supp}(\hat{x}^i(1))} \alpha_{a^i}$. By the optimality of $\hat{x}^i(\beta)$ it follows that $\tilde{v}^i_\beta(\hat{x}^{-i}(\beta), a^i) = g^i$ for every $a^i \notin \text{supp}(\hat{x}^i(1))$.

Step 7: Assertion (b).iii(4)
If there is strict inequality in (b).iii(3) then $q \in (0, 1)$. Hence, by Lemma 4.18, $\deg(w(\hat{x}(\beta))) = 1$. Therefore, for every $b^L \in \text{supp}(\mu)$,

$$\sum_{i \in L} d(b^i) = \sum_{i \in L} \deg(\hat{x}^i_{b^i}) = \deg(w(\hat{x}(\beta))) = 1.$$

■

4.7 Three Players Repeated Games with Absorbing States

The main result of this section is:

THEOREM 4.23 *Every three-player repeated game with absorbing states has a perturbed equilibrium payoff.*

Recall that $\mathcal{B}_i(x)$ is the set of single absorbing neighbors of player i w.r.t. x.

For the proof we need the following lemma:

LEMMA 4.24 *Let* $x \in X$, $g \in \mathbf{R}^N$, $\mu \in \Delta(\mathcal{B}(x))$ *and* $d : \cup_i A^i \to [0, \infty)$ *satisfy condition 4.22(b).iii. At least one of the following hold.*

a) There exists a probability distribution $\nu \in \Delta(\mathcal{B}(x))$ *such that* $\text{supp}(\nu) \subseteq \text{supp}(\mu)$*, for every player i*

$$\sum_{b^L \in \text{supp}(\nu)} \nu(b^L) u^i(x^{-L}, b^L) \geq g^i \tag{61}$$

and for every player i such that $\mathcal{B}_i(x) \cap \text{supp}(\nu) \neq \emptyset$ *there is an equality in (61).*

b) For every player i there exists $y^i \in \Delta(\mathcal{B}_i(x))$ *such that* $\text{supp}(y^i) \subset \text{supp}(\mu)$ *and* $(u(x^{-1}, y^1), u(x^{-2}, y^2), u(x^{-3}, y^3))$ *are positive cyclic vectors.*

The proof of the lemma is postponed to section 4.10.

Proof of Theorem 4.23:
Theorem 4.23 is an immediate consequence of Lemmas 4.1, 4.2, 4.4, 4.8, 4.22 and 4.24.

Indeed, 4.22(b).i and 4.22(a) imply that the conditions of Lemma 4.2 hold. 4.22(b).ii and 4.22(a) imply that the conditions of Lemma 4.1 hold. 4.24(b) and 4.22(a) imply that the conditions of Lemma 4.4 hold, while 4.24(a) and 4.22(a) imply that the conditions of Lemma 4.8 hold. ■

4.8 More Than Three Players

Unfortunately, our approach cannot be generalized for more than three players. In our proof we construct for every $\epsilon > 0$ and a sequence of discounted equilibria in the auxiliary game that converges to a mixed action combination x, an (x, ϵ)-perturbed equilibrium profile. Such a construction need not be possible for games with more than three players, as can be seen by the following four-player game:

$$\begin{array}{cccc}
 & & \mathbf{4} & \\
 & \mathbf{2} & & \mathbf{2} \\
\mathbf{1} & \begin{array}{|c|c|}\hline 0,0,0,0 & 4,1,0,0^* \\ \hline 1,4,0,0^* & 1,1,0,0^* \\ \hline\end{array} & \mathbf{1} & \begin{array}{|c|c|}\hline 0,0,4,1^* & 0,1,0,1^* \\ \hline 1,0,0,1^* & 1,1,0,1^* \\ \hline\end{array} \\
\mathbf{3} & & & \\
\mathbf{1} & \begin{array}{|c|c|}\hline 0,0,1,4^* & 0,1,1,0^* \\ \hline 1,0,1,0^* & 1,1,1,0^* \\ \hline\end{array} & \mathbf{1} & \begin{array}{|c|c|}\hline 0,0,1,1^* & 0,1,1,1^* \\ \hline 1,0,1,1^* & 1,1,1,1^* \\ \hline\end{array}
\end{array}$$

In this game player 1 chooses a row, player 2 chooses a column, player 3 chooses either the top two matrices or the bottom two matrices, and player 4 chooses either the left two matrices or the right two matrices.

Note that the auxiliary game is essentially the same as the original game, and that there are the following symmetries in the payoff function: for every 4-tuple of actions (a, b, c, d) we have:

$$\begin{array}{rcll}
v^1(a,b,c,d) & = & v^2(b,a,d,c), & \\
v^3(a,b,c,d) & = & v^4(b,a,d,c) & \quad\text{and} \\
v^1(a,b,c,d) & = & v^3(c,d,a,b). &
\end{array}$$

For every $\lambda \in [0,1]$, let $T^i_\beta(\lambda)$ be the set of all the best replies of player i in the β-discounted game if the other three players play the stationary strategy $(\lambda, 1-\lambda)$. By the symmetries of the game it follows that $T^i_\beta(\lambda) = T^j_\beta(\lambda)$ for each pair of players i and j.

Every fixed point λ of the correspondence T^i_β is a stationary equilibrium for the β-discounted game, where all the players play the same mixed action $(\lambda, 1-\lambda)$.

Note that $\lambda = 0$ is a fixed point of T^i_β, and $\lambda = 1$ is not a fixed point. We shall see that there is a fixed point $\lambda(\beta) \in (0,1)$ such that $\lim_{\beta\to 1} \lambda(\beta) = 1$.

Fix $\beta \in (0,1)$, and assume that players 2,3 and 4 play $(\lambda, 1-\lambda)$, where $\lambda \in (0,1)$. If player 1 plays the bottom row then his expected payoff is 1, while if he plays the top row then his expected payoff is $g^1 = 4\lambda^2(1-\lambda) + \beta\lambda^3 g^1$. Player 1 is indifferent between his two actions if $g^1 = 1$, hence if

$$f_\beta(\lambda) \overset{\text{def}}{=} (4-\beta)\lambda^3 - 4\lambda^2 + 1 = 0. \qquad (62)$$

We claim that (62) has a solution $\lambda(\beta) \in (0,1)$ such that $\lim_{\beta \to 1} \lambda(\beta) = 1$. Indeed, $f_\beta(1) > 0$ for every $\beta \in (0,1)$, and $f_\beta(\beta^2) < 0$ for β sufficiently close to 1. The claim follows since f_β is continuous and $\lim_{\beta \to 1} \beta^2 = 1$.

However, the conditions of Lemmas 4.1, 4.2, 4.4 and 4.8 are not satisfied for the non-absorbing cell. Moreover, the conditions of Lemma 4.9 are not satisfied as well. It is clear that the condition of Lemma 4.1 is not satisfied for the non-absorbing cell, and Lemma 4.2 is irrelevant for this cell.

It is easily checked that the only convex combinations z of the four vectors $a_1 = (1,4,0,0), a_2 = (4,1,0,0), a_3 = (0,0,1,4)$ and $a_4 = (0,0,4,1)$ such that $z^i = 1$ if a_i has a positive weight in the combination are the four combinations that include only a single vector a_i. Therefore the conditions of Lemma 4.4 are not satisfied for the non-absorbing cell. Moreover, if the conditions of Lemma 4.9 hold, then at every phase only a single player perturbs (rather than a subset of the players).

We now prove that the conditions of Lemma 4.9 do not hold when a single player perturbs in every phase.

Assume to the contrary that $g = (g^1, g^2, g^3, g^4) \geq (1,1,1,1)$ is such an equilibrium payoff, that player 1 is the first player to perturb, $\gamma > 0$ is the overall probability in which player 1 perturbs in the first phase, and $f = (f^1, f^2, f^3, f^4)$ is the expected payoff for the players given player 1 did not perturb in the first phase.

Clearly $g = \gamma(1,4,0,0) + (1-\gamma)f$, hence $f^3, f^4 > 1$ and $f^1 = 1$. Therefore the only player that can perturb after player 1 has finished his perturbations is player 2. Let $y = (y^1, y^2, y^3, y^4)$ be the expected payoff for the players after player 2 has finished his phase of perturbations. Since $f^1 = 1$ it follows that $y^1 < 1$, a contradiction to the individual rationality of the strategy.

It is a little more technical to show that such a construction is not possible if the future expected payoff of the players should be at least $1 - \epsilon$ at each stage (given absorption has not occurred) instead of 1.

Remark: Recently Solan and Vieille [29] found an example of a four-player repeated game with absorbing states that has no perturbed equilibrium payoff.

4.9 Repeated Team Games with Absorbing States

DEFINITION 4.25 *An $n+m$-player repeated game with absorbing states is a* repeated team game with absorbing states *if $h^i = h^j$ and $u^i = u^j$ whenever*

$1 \leq i, j \leq n$ *or* $n + 1 \leq i, j \leq n + m$.

Let G be a repeated team game with absorbing states. We denote by N and M the two sets of players in each team, hence $n = |N|$ and $m = |M|$.

Let

$$d^i = \begin{cases} \max_{j \in N} c^j & i \in N \\ \max_{j \in M} c^j & i \in M \end{cases}$$

It is clear that for any equilibrium payoff $g = (g^i)_{i \in N \cup M}$, $g^i = g^j$ whenever $i, j \in N$ or $i, j \in M$. Moreover, $g^i \geq c^i$ for every $i \in N \cup M$, and therefore $g^i \geq d^i$ for every $i \in N \cup M$.

The main result of this section is:

THEOREM 4.26 *Every repeated team game with absorbing states has a perturbed equilibrium payoff.*

Proof: For every $i \in N$, define $\tilde{h}^i : X \to \mathbf{R}$ by

$$\tilde{h}^i(x) = \min\{h^i(x), d^i\}.$$

By the remark following the proof of Lemma 4.13, it follows that Lemma 4.22 holds for G. Let (x, g, μ, d) satisfy the conclusion of Lemma 4.22. We prove that x is a base of a perturbed equilibrium.

Note that since $\tilde{h}^i(x) = \tilde{h}^j(x)$ whenever $1 \leq i, j \leq n$ or $n + 1 \leq i, j \leq n + m$, it follows that $g^i = g^j$ whenever $1 \leq i, j \leq n$ or $n + 1 \leq i, j \leq n + m$.

We have 4 cases:

1) x is absorbing.

2) x is non-absorbing and $h(x) \geq g$.

3) x is non-absorbing, and there exists a single absorbing neighbor $b^{i_0} \in \mathcal{B}(x)$ such that $u(x^{-i_0}, b^{i_0}) \geq g$.

4) Non of the first three cases hold.

In case 1, condition 4.22(b).ii holds, and, together with condition 4.22(a) implies that the conditions of Lemma 4.2 hold.

In case 2, condition 4.22(b).i holds, and, together with condition 4.22(a) implies that the conditions of Lemma 4.1 hold.

In case 3, the assumption and condition 4.22(a) imply that the conditions of Lemma 4.4 hold w.r.t. the probability distribution that gives probability 1 to b^{i_0}.

Assume now that neither case 1 nor cases 2 or 3 hold. We prove that in this case the conditions of Lemma 4.4 hold.

Since cases 1 and 2 do not hold it follows that condition 4.22(b).iii holds.

Let $b^{i_0} \in \text{supp}(\mu)$ be any single absorbing neighbor. By condition 4.22(b).iii(3), $u^{i_0}(x^{-i_0}, b^{i_0}) = g^{i_0}$, and since the game is a team game, $u^i(x^{-i_0}, b^{i_0}) = g^i$ for every player i of the same team as i_0. Since case 3 does not hold, there exists a player i of the opposing team such that $u^i(x^{-i_0}, b^{i_0}) < g^i$, and therefore $u^i(x^{-i_0}, b^{i_0}) < g^i$ for every player i in the opposing team. In particular, if $\cap_i \mathcal{B}_i(x) \cap \text{supp}(\mu) \neq \emptyset$ then

$$\sum_{\cap_i \mathcal{B}_i(x) \cap \text{supp}(\mu)} \mu(b^L) u^i(x^{-L}, b^L) \leq g^i \text{ for every } i \in N \cup M \tag{63}$$

and

$$\sum_{\cap_i \mathcal{B}_i(x) \cap \text{supp}(\mu)} \mu(b^L) u(x^{-L}, b^L) \neq g. \tag{64}$$

Let $\mathcal{C} = \{b^L \in \text{supp}(\mu) \mid |L| \geq 2\}$. By condition 4.22(b).iii(2), (63) and (64), it follows that $\mathcal{C} \neq \emptyset$ and $\sum_{b^L \in \mathcal{C}} \mu(b^L) > 0$.

Let ν be the induced probability distribution of μ over $\mathcal{C}$. By condition 4.22(b).iii(2), (63) and (64) we have:

$$\sum_{b^L \in \mathcal{C}} \nu(b^L) u^i(x^{-L}, b^L) \geq g^i \text{ for every } i \in N \cup M. \tag{65}$$

Since $|L| \geq 2$ for every $b^L \in \text{supp}(\nu)$, condition 4.22(a) and (65) imply that the conditions of Lemma 4.4 hold w.r.t. $\mathcal{C}$ and ν. ■

4.10 Proof of Lemma 4.24

In this section we prove Lemma 4.24. The proof is technical, and its general idea is as follows. Let r be the number of players that have single absorbing neighbors of x with positive weight according to μ. We divide the proof into 4 cases, according to the value of r. If $r = 0$ then 4.24(a) holds trivially. If r is positive and 4.24(b) does not hold, we need to construct the probability distribution ν. The method is to different when $r = 1$ and when $r > 1$.

When $r = 2$, by changing the weights of the single absorbing neighbors of the various players in a certain way, we can obtain a probability distribution ν over $\mathcal{B}(x)$ that satisfies 4.24(a).

When $r = 3$, there can be two sub-cases. Either 4.24(b) holds, or, as when $r = 2$, by changing the weights of the single absorbing neighbors of the various players in a certain way, we can obtain a probability distribution ν over $\mathcal{B}(x)$ that satisfies 4.24(a).

When $r = 1$ one constructs the desired probability distribution ν in a different manner, by changing the weights of *all* the absorbing neighbors of x.

The rest of the section is arranged as follows. We first deal with the case $r = 1$. Then we prove two lemmas, that show how one can change the weights of single absorbing neighbors in order to get a new probability distribution for which there is an equality in (61) for one more player. These two lemmas are then applied to prove the case $r = 2$, and finally, the case $r = 3$.

For the rest of the section we fix $x \in X$ and $g \in \mathbf{R}^3$ for which there exist $\mu \in \Delta(\mathcal{B}(x))$ and $d : \cup A^i \to [0, \infty)$ such that 4.22(b).iii holds for (x, g, μ, d).

For every vector $0 < \lambda \in \mathbf{R}^3_+$, the *normalization* of λ, denoted by $\bar{\lambda}$, is defined by:

$$\bar{\lambda}_i = \frac{\lambda_i}{\sum_{j=1}^3 \lambda_j}.$$

Let $z \in \mathbf{R}^3$. The *signed form* of z is the vector of the signs of the values of $(z^1 - g^1, z^2 - g^2, z^3 - g^3)$. If $z^i - g^i = 0$ then the ith coordinate of the sign vector is 0.

Let $\mathcal{MD} = \{(\mu, d) \mid \mu \in \mathcal{B}(x), d : \cup_i A^i \to [0, \infty) \text{ and 4.22(b).iii holds for } (x, g, \mu, d)\}$. We denote by $\mathcal{M}$ the projection of $\mathcal{MD}$ over the first coordinate. By the assumption, $\mathcal{MD} \neq \emptyset$. From now on we fix a pair $(\mu, d) \in \mathcal{MD}$ such that μ has a minimal support among the elements in $\mathcal{M}$.

Let $R = \{i \in N \mid \mathcal{B}_i(x) \cap \text{supp}(\mu) \neq \emptyset\}$. R is the set of players that have single absorbing neighbors in $\text{supp}(\mu)$. We denote $r = |R|$. The proof of Lemma 4.24 is divided into several cases, according to the value of r.

If $r = 0$ or there is an equality in 4.22(b).iii(2) then 4.24(a) holds trivially. From now on we assume that $\sum_{i \in L} d(b^i) = 1$ for every $b^L \in \text{supp}(\mu)$. In particular, $d(a^i) = 1$ for every player $i \in N$ and single absorbing neighbor $a^i \in \mathcal{B}_i(x) \cap \text{supp}(\mu)$.

For every $i \in R$ we define $y^i \in X^i$ as follows:

$$y^i_{a^i} = \begin{cases} 0 & d(a^i) \neq 1 \\ \frac{\mu(a^i)}{\sum_{b^i \in \mathcal{B}_i(x)} \mu(b^i)} & d(a^i) = 1 \end{cases}$$

Note that for every $a^i \in \text{supp}(y^i)$, $w(x^{-i}, a^i) > 0$ and, by 4.22(b).iii(3), $u^i(x^{-i}, a^i) = g^i$.

If $1 \in R$ and the signed form of $u(x^{-1}, y^1)$ is $(0, +, +)$ then y^1, viewed as a probability distribution over $\mathcal{B}(x)$, satisfies 4.24(a). By the minimality of μ, it follows that $u(x^{-1}, y^1)$ does not have the signed form $(0, -, -)$. Hence we assume that if $1 \in R$ then $u(x^{-1}, y^1)$ has the signed form $(0, +, -)$ or $(0, -, +)$. Symmetric assumptions are made for players 2 and 3.

For every function $\rho : \mathcal{B}(x) \to [0, \infty)$ and player $j \in N$ we define $\text{supp}(\rho) \stackrel{\text{def}}{=} \{b^L \in \mathcal{B}(x) \mid \rho(b^L) \neq 0\}$ and

$$\langle \rho, u^j \rangle = \sum_{b^L \in \mathcal{B}(x)} \rho(b^L) u^j(x^{-L}, b^L).$$

Recall that $\mathcal{B}(\rho, a^i) = \{b^L \in \mathcal{B}(x) \cap \text{supp}(\rho) \mid i \in L, b^i = a^i\}$ for every $a^i \in A^i$. Denote

$$\langle \rho, u^j, a^i \rangle = \sum_{b^L \in \mathcal{B}(\rho, a^i)} \rho(b^L) u^j(x^{-L}, b^L).$$

Note that for every $i \in N$, $\langle \rho, u^j \rangle = \sum_{a^i \in A^i} \langle \rho, u^j, a^i \rangle$.

LEMMA 4.27 *If $r = 1$ then assertion 4.24(a) holds.*

Proof: Assume that $R = \{1\}$. Clearly if there exists $\nu \in \Delta(\mathcal{B}(x))$ such that $\text{supp}(\nu) \subset \text{supp}(\mu)$, $\text{supp}(\nu) \cap \mathcal{B}_i(x) = \emptyset$ and $\langle \nu, u^j \rangle \geq g^j$ for $j = 2, 3$ then the lemma holds. Indeed, if such ν exists then by the minimality of μ it follows that $\langle \nu, u^1 \rangle < g^1$, and therefore there exists a convex combination of μ and ν that satisfies 4.24(a).

By 4.22(b).iii(3),

$$\langle \mu, u^2, a^2 \rangle = g^2 \sum_{b^L \in \mathcal{B}(\mu, a^2)} \mu(b^L) \qquad \forall a^2 \notin \text{supp}(x^2). \tag{66}$$

By the above discussion it follows that for every $a^2 \notin \text{supp}(x^2)$ with $\mathcal{B}(\mu, a^2) \neq \emptyset$

$$\langle \mu, u^3, a^2 \rangle < g^3 \sum_{b^L \in \mathcal{B}(\mu, a^2)} \mu(b^L). \tag{67}$$

Similarly,

$$\langle \mu, u^3, a^3 \rangle = g^3 \sum_{b^L \in \mathcal{B}(\mu,a^3)} \mu(b^L) \qquad \forall a^3 \notin \operatorname{supp}(x^3) \tag{68}$$

and if $\mathcal{B}(\mu, a^3) \neq \emptyset$ we have

$$\langle \mu, u^2, a^3 \rangle < g^2 \sum_{b^L \in \mathcal{B}(\mu,a^3)} \mu(b^L). \tag{69}$$

Define

$$\nu(b^L) = \begin{cases} d(b^1)\mu(b^L) & b^L \in \mathcal{B}(x) \text{ and } 1 \in L \\ 0 & \text{Otherwise.} \end{cases}$$

Since $\operatorname{supp}(\mu) \cap \mathcal{B}_1(x) \neq \emptyset$ it follows that $\sum_{b^L \in \mathcal{B}(x)} \nu(b^L) > 0$. By 4.22(b).iii(2) and 4.22(b).iii(4)

$$\begin{aligned} g^i &\leq \sum_{b^L \in \mathcal{B}(x)} \mu(b^L) u^i(x^{-L}, b^L) \\ &= \sum_{b^L \in \mathcal{B}(x)} \mu(b^L) u^i(x^{-L}, b^L) \sum_{j \in L} d(b^j) \\ &= \sum_{a^1 \in A^1} \langle \nu, u^i \rangle + \sum_{a^2 \in A^2} d(a^2) \langle \mu, u^i, a^2 \rangle + \sum_{a^3 \in A^3} d(a^3) \langle \mu, u^i, a^3 \rangle \\ &= \sum_{a^1 \in A^1} \langle \nu, u^i \rangle + \sum_{a^2 \notin \operatorname{supp}(x^2)} d(a^2) \langle \mu, u^i, a^2 \rangle + \sum_{a^3 \notin \operatorname{supp}(x^3)} d(a^3) \langle \mu, u^i, a^3 \rangle \end{aligned}$$

where the last equality holds since $d(a^i) = 0$ whenever $a^i \in \operatorname{supp}(x^i)$. By (66), (67), (68) and (69) it follows that $\langle \bar{\nu}, u^i \rangle \geq g^i$ for $i = 2, 3$, where $\bar{\nu}$ is the normalization of ν.

Moreover,

$$\langle \nu, u^1 \rangle = \sum_{a^1 \in A^1} d(a^1) \langle \mu, u^1, a^1 \rangle.$$

By condition 4.22(b).iii(3), $\langle \mu, u^1, a^1 \rangle = g^1$ whenever $d(a^1) > 0$. Therefore $\langle \bar{\nu}, u^1 \rangle = g^1$, and $\bar{\nu}$ is the desired probability distribution. ■

The following two lemmas assert that if $2 \in R$ then we can assume w.l.o.g. that $\langle \mu, u^3 \rangle = g^3$.

LEMMA 4.28 *If* $2 \in R$ *and* $u(x^{-2}, y^2)$ *has the signed form* $(-, 0, +)$ *then there exists* $\nu \in \mathcal{M}$ *such that* $\mathrm{supp}(\mu) = \mathrm{supp}(\nu)$ *and* $\langle \nu, u^3 \rangle = g^3$. *Furthermore, if* $\langle \mu, u^2 \rangle = g^2$ *then* $\langle \nu, u^2 \rangle = g^2$.

Proof: For every $t \in [0, 1]$ define:

$$\nu_t(b^L) = \begin{cases} t\mu(b^L) & b^L \in \mathrm{supp}(\mathcal{B}_2(x)) \\ \mu(b^L) & \text{Otherwise} \end{cases}$$

i.e. as t decreases we decrease the weight of the single absorbing neighbors of player 2. Clearly $\nu_1 = \mu$, for every $t \in [0, 1]$ conditions 4.22(b).iii(3) and 4.22(b).iii(4) hold w.r.t. $\bar{\nu}_t$ and 4.22(b).iii(2) holds w.r.t. $\bar{\nu}_t$ for $i = 1, 2$. For every t such that $\langle \bar{\nu}_t, u^3 \rangle \geq g^3$ condition 4.22(b).iii(2) holds w.r.t. $\bar{\nu}_t$ for $i = 3$ too. Since μ has a minimal support, there is $t_0 \in (0, 1]$ such that $\langle \bar{\nu}_{t_0}, u^3 \rangle = g^3$. For every t, $\langle \mu, u^2 \rangle = \langle \bar{\nu}_t, u^2 \rangle$ and therefore $\bar{\nu}_{t_0}$ is the desired vector. ■

LEMMA 4.29 *If* $2 \in R$ *and* $u(x^{-2}, y^2)$ *has the signed form* $(+, 0, -)$ *Then there exists a vector* $\nu \in \mathcal{M}$ *such that* $\mathrm{supp}(\mu) = \mathrm{supp}(\nu)$ *and* $\langle \nu, u^3 \rangle = g^3$. *Furthermore, if* $\langle \mu, u^2 \rangle = g^2$ *then* $\langle \nu, u^2 \rangle = g^2$.

Proof: For every $t \in [1, \infty)$ define:

$$\nu_t(b^L) = \begin{cases} t\mu(b^L) & b^L \in \mathrm{supp}(\mathcal{B}_2(x)) \\ \mu(b^L) & \text{Otherwise} \end{cases}$$

i.e. as t increases we increase the weight of the single absorbing neighbors of player 2. Clearly $\nu_1 = \mu$, for every $t \in [1, \infty)$ conditions 4.22(b).iii(3) and 4.22(b).iii(4) hold w.r.t. $\bar{\nu}_t$ and 4.22(b).iii(2) holds w.r.t. $\bar{\nu}_t$ for $i = 1, 2$. For every t such that $\langle \bar{\nu}_t, u^3 \rangle \geq g^3$ condition 4.22(b).iii(2) holds w.r.t. $\bar{\nu}_t$ for $i = 3$ too. Since $u(x^{-2}, y^2)$ has the signed form $(+, 0, -)$, for t sufficiently large $\langle \bar{\nu}_t, u^3 \rangle < g^3$. Therefore there is $t_0 \in [1, \infty)$ such that $\langle \bar{\nu}_{t_0}, u^3 \rangle = g^3$. For every t, $\langle \mu, u^2 \rangle = \langle \bar{\nu}_t, u^2 \rangle$ and therefore $\bar{\nu}_{t_0}$ is the desired vector. ■

LEMMA 4.30 *If* $r = 2$ *then there exists* $\nu \in \mathcal{M}$ *such that 4.24(a) holds.*

Proof: Assume w.l.o.g. that $R = \{2, 3\}$. By either Lemma 4.28 or Lemma 4.29, according to the signed form of $u(x^{-3}, y^3)$, we can assume that $\langle \mu, u^2 \rangle = g^2$. By a second use of either Lemma 4.28 or Lemma 4.29, according to the signed form of $u(x^{-2}, y^2)$, we can assume w.l.o.g. that $\langle \mu, u^i \rangle = g^i$ for $i = 2, 3$, as desired. ■

LEMMA 4.31 *If* $r = 3$ *and* $(u(x^{-1}, y^1), u(x^{-2}, y^2), u(x^{-3}, y^3))$ *are not positive cyclic vectors, then there exists* $\nu \in \mathcal{M}$ *such that 4.24(a) holds.*

Proof: Up to symmetries we can assume that $u(x^{-1}, y^1)$ has the signed form $(0, +, -)$ and $u(x^{-2}, y^2)$ has the signed form $(-, 0, +)$.

Using Lemma 4.28 twice, once with $u(x^{-1}, y^1)$ and then with $u(x^{-2}, y^2)$, we can assume that $\langle \mu, u^i \rangle = g^i$ for $i = 2, 3$. If $\langle \mu, u^1 \rangle = g^1$ we are done. Hence assume $\langle \mu, u^1 \rangle > g^1$.

Case 1: $\langle \mu, u^3 \rangle$ has the signed form $(+, -, 0)$.

Let $\alpha, \beta > 0$ solve the equations

$$\begin{aligned} u^3(x^{-1}, y^1) + \alpha u^3(x^{-2}, y^2) &= (1 + \alpha) g^3 \\ u^2(x^{-1}, y^1) + \beta u^2(x^{-3}, y^3) &= (1 + \beta) g^2. \end{aligned} \tag{70}$$

Since $(u(x^{-1}, y^1), u(x^{-2}, y^2), u(x^{-3}, y^3))$ are not positive cyclic it follows that

$$\alpha u^1(x^{-2}, y^2) + \beta u^1(x^{-3}, y^3) \leq (\alpha + \beta) g^1. \tag{71}$$

Let $\nu : \mathcal{B}(x) \to [0, \infty)$ be defined as follows:

$$\nu(b^L) = \begin{cases} \mu(b^L) & b^L \in \mathcal{B}_1(x) \cap \mathrm{supp}(\mu) \\ \alpha\mu(b^L) & b^L \in \mathcal{B}_2(x) \cap \mathrm{supp}(\mu) \\ \beta\mu(b^L) & b^L \in \mathcal{B}_3(x) \cap \mathrm{supp}(\mu) \\ 0 & \text{Othewise.} \end{cases}$$

By (70), $\langle \bar{\nu}, u^i \rangle = g^i$ for $i = 2, 3$. If there is an equality in (71) then $\langle \bar{\nu}, u^1 \rangle = g^1$, and $\bar{\nu}$ is the desired vector. Otherwise $\langle \bar{\nu}, u^1 \rangle < g^1$. Let ρ be the convex combination of μ and ν that satisfies $\langle \rho, u \rangle = g$. Then ρ satisfies 4.24(a), as desired.

Case 2: $u(x^{-3}, y^3)$ has the signed form $(-, +, 0)$.

Let $\alpha, \beta > 0$ solve the equations

$$\begin{aligned} &u^3(x^{-1}, y^1) + \alpha u^3(x^{-2}, y^2) = g^3 \\ &u^2(x^{-1}, y^1) = \beta u^2(x^{-3}, y^3). \end{aligned} \tag{72}$$

For every $t \geq 0$ define

$$\nu_t(b^L) = \begin{cases} \mu(b^L) + t & b^L \in \mathcal{B}_1(x) \cap \operatorname{supp}(\mu) \\ \mu(b^L) + t\alpha & b^L \in \mathcal{B}_2(x) \cap \operatorname{supp}(\mu) \\ \mu(b^L) + t\beta & b^L \in \mathcal{B}_3(x) \cap \operatorname{supp}(\mu) \\ \mu(b^L) & \text{Otherwise.} \end{cases}$$

Note that $\sum_{b^L \in \mathcal{B}(x)} \nu_t(b^L) = 1 + tq$, where

$$q = \sum_{b^L \in \mathcal{B}_1(x)} \mu(b^L) + \alpha \sum_{b^L \in \mathcal{B}_2(x)} \mu(b^L) + \beta \sum_{b^L \in \mathcal{B}_3(x)} \mu(b^L).$$

For every $t \geq 0$ we have by (72):

$$\begin{aligned} \langle \nu_t, u^3 \rangle &= \langle \mu, u^3 \rangle + t(u^3(x^{-1}, y^1) + \alpha u^3(x^{-2}, y^2)) + t\beta g^3 \sum_{b^L \in \mathcal{B}_3(x)} \mu(b^L) = \\ &\quad g^3(1 + tq) \\ \langle \nu_t, u^2 \rangle &= \langle \mu, u^2 \rangle + t(u^2(x^{-1}, y^1) + \beta u^2(x^{-3}, y^3)) + t\alpha g^2 \sum_{b^L \in \mathcal{B}_2(x)} \mu(b^L) = \\ &\quad g^2(1 + tq) \\ \langle \nu_t, u^1 \rangle &= \langle \mu, u^1 \rangle + t(\alpha u^1(x^{-2}, y^2) + \beta u^1(x^{-3}, y^3)) + t \sum_{b^L \in \mathcal{B}_1(x)} \mu(b^L). \end{aligned}$$

Clearly $\langle \nu_0, u^1 \rangle = \langle \mu, u^1 \rangle > g^1$. Since $u^1(x^{-i}, y^i) < g^i$ for $i = 1, 2$, there exists $t_0 > 0$ such that $\langle \bar{\nu}_{t_0}, u^1 \rangle = g^1$. Then $\bar{\nu}_{t_0}$ is the desired probability distribution. ∎

5 Recursive Games with the Absorbing Property

In this section we prove that every two-player stochastic game with two non-absorbing states has an equilibrium payoff. We use Vieille's reduction [33, 34] that claims that it is sufficient to prove the existence for positive recursive stochastic games with the absorbing property.

We begin by an equivalent formulation of recursive games (section 5.1), provide an example of such a game, and show some of the equilibrium payoffs and the corresponding ϵ-equilibrium profiles (section 5.2). We then introduce two sufficient conditions for existence of an equilibrium payoff in recursive games (section 5.3), and after proving some preliminary results (section 5.4), we introduce the approximating games and derive several results on these games (section 5.5). We end by proving the main result (section 5.6) and by explaining why our approach cannot be generalized for games with more than two non-absorbing states (section 5.7).

5.1 An Equivalent Formulation

Let $G = (N, S, (A^i)_{i \in N}, h, w)$ be a two-player stochastic game, that is, $|N| = 2$. For simplicity we denote by A and B the sets of actions available for the two players, and by X and Y the spaces of stationary strategies. Strategies of the two players are denoted by σ and τ.

Let $T \subseteq S$ be the set of all the absorbing states.

DEFINITION 5.1 *The game is* positive *if* $h^2(s,a,b) > 0$ *for every absorbing state* $s \in T$ *and every pair of actions* $(a,b) \in A \times B$. *It is* recursive *if* $h^i(s,a,b) = 0$ *for every non-absorbing state* $s \notin T$, *every pair of actions* $(a,b) \in A \times B$ *and every player* $i = 1, 2$. *It satisfies* the absorbing property *if for every fully mixed stationary strategy* y *of player 2, every strategy* σ *of player 1 and every initial state* $s \in S$,

$$\Pr_{s,\sigma,y} \left(\exists t \in \mathbf{N} \text{ s.t. } s_t \in T \right) = 1$$

where s_t *is the state of the game at stage* t.

In [34] Vieille proves that in order to prove existence of an equilibrium payoff in every two-player stochastic game, it is sufficient to prove the existence for every positive recursive game with the absorbing property. Following closely the proof of Vieille reveals that he proves even more:

THEOREM 5.2 *If every positive recursive game with the absorbing property and at most n non-absorbing states has an equilibrium payoff, then every stochastic game with at most n non-absorbing states has an equilibrium payoff.*

From now on we fix a positive recursive game that satisfies the absorbing property.

Note that since the game is positive and recursive, $\lim_t (r^i_1 + \cdots + r^i_t)/t$ exists. Since the game satisfies the absorbing property, the min-max value of player 2 is positive, hence any ϵ-equilibrium profile, for ϵ sufficiently small, must be absorbing with high probability.

As in section 4.1, we assume w.l.o.g. that $h(s,\cdot,\cdot)$ is constant over each absorbing state $s \in T$, and denote this constant value by $u(s)$.

For every subset $C \subset S$ and every triplet $(s,a,b) \in S \times A \times B$ we denote $w_C(s,a,b) = \sum_{s' \in C} w_{s'}(s,a,b)$.

For every $(a,b) \in A \times B$ and every function $g : S \to \mathbf{R}^2$ we define

$$\psi_g(s,a,b) = \sum_{s' \in S} w_{s'}(s,a,b) g(s'). \tag{73}$$

$\psi_g(s,a,b)$ is the expected payoff for the players if the game is in state s, they play the pure actions (a,b), and the continuation payoff is given by g. Note that for every fixed triplet (s,a,b), the function $\psi_g(s,a,b)$ is linear in g, and therefore continuous. The multi-linear extension of $\psi_g(s,\cdot,\cdot)$ over $\Delta(A) \times \Delta(B)$ is denoted also by $\psi_g(s,\cdot,\cdot)$.

5.2 An Example

Consider the following positive recursive game:

state 1

	L	C	R
T	**1**	**1**	1/2 **2** 1/2 $3,1$ *
B	**1**	2/3 **2** 1/3 $4,2$ *	**1**

state 2

	L	R
T	$0,1$ *	**1**
B	**1**	$0,1$ *

A boldfaced letter means a transition to a state with no daily payoff. As before, an asterisk means transition to an absorbing state with the indicated absorbing payoff. If the transition is probabilistic, the probability appears to the left.

Note that if player 2 plays a fully mixed stationary strategy then the game is bound to be eventually absorbed, whatever player 1 plays. Hence the game satisfies the absorbing property.

One equilibrium payoff is $((2,1),(1,1))$. An ϵ-equilibrium strategy profile (for every $\epsilon > 0$) is:

- In state **1** the players play the mixed actions $(T, (1-\epsilon)L + \epsilon R)$.
- In state **2** both players play the mixed actions $(\frac{1}{2}, \frac{1}{2})$.
- If any player plays an action which has probability 0 to be played, then both players play the pure actions (T, L) in both states forever (this part of the strategy serves as a punishment strategy).

It is easy to verify that no player can profit more than ϵ by any deviation, and that this strategy profile yields the players the desired payoff.

Another equilibrium payoff is $((2, 21/17), (1, 19/17)$. An ϵ-equilibrium strategy profile for this payoff is more complex. Let $n_1 \in \mathbf{N}$ and $\epsilon_1 < \epsilon$ such that $(1-\epsilon_1)^{n_1} = 1/2$. Define the following strategy profile:

- In state **2**, the players play the mixed actions $(\frac{1}{2}, \frac{1}{2})$.

- Assume the game moves to state **1** at stage t_0. The players play as follows:
  - The players play the mixed actions $((1-\epsilon_1)T+\epsilon_1 B, (1-\epsilon_1)L+\epsilon_1 C)$. The players play these mixed actions until player 2 played the action C for n_1 times since stage t_0, or until both players played (B, C) at the same stage (and the game leaves state **1**).
  - If player 2 played the action C for n_1 times since stage t_0, then the players play the mixed actions $(T, (1-\epsilon)L+\epsilon R)$ until player 2 plays the action R (and the game leaves state **1**).
  - If any player plays an action which has probability 0 to be played, the players play the pure actions (T, L) in both states forever.

Note that if the players follow this profile, then the probability that the game leaves state **1** through (B, C) is 1/2, as is the probability that the game leaves state **1** through (T, R). In particular the game is bound to be eventually absorbed.

Assume that the players follow the above profile, and let $g = (g_s^i)$ be the payoff that the players receive. Clearly no player can deviate and gain in state **2**, and $g_2^i = \frac{1}{2}(0,1) + \frac{1}{2}g_1^i$ for $i = 1, 2$.

Moreover, we have:

$$g_1 = \frac{1}{2}\left(\frac{1}{3}(4,2) + \frac{2}{3}g_2\right) + \frac{1}{2}\left(\frac{1}{2}(3,1) + \frac{1}{2}g_2\right)$$

and therefore $g_1 = (2, 21/17)$.

We shall now verify that no player can profit more than ϵ by deviating in state **1**. Indeed, if player 2 deviates in state **1**, then his expected payoff is bounded by 1 (since after the punishment begins, his expected payoff is at most 1). Given the players follow the above profile, the initial state is **1**, and the first time the game leaves state **1** is through (B, C), the expected payoff for player 1 is $\frac{2}{3} + \frac{1}{3} \cdot 4 = 2$, where if the first time the game leaves state **1** is through (T, R), the expected payoff of player 1 is $\frac{1}{2} + \frac{1}{2} \cdot 3 = 2$ as well. Hence player 1 cannot profit by any deviation whenever the game is in state **1**. Therefore this strategy profile is an ϵ-equilibrium, as desired.

5.3 Sufficient Conditions for Existence of an Equilibrium Payoff

In this section we give two sets of sufficient conditions for existence of an equilibrium payoff in positive recursive games with the absorbing property.

Definition 5.3 *Let (x, y) be a stationary profile. A set $C \subseteq S$ is* stable under (x, y) *if $w_C(s, x, y) = 1$ for every $s \in C$.*

In the example presented in section 5.2, $\{s_1\}$ is stable, for example, under $((T, L), (\cdot, \cdot))$, and $\{s_1, s_2\}$ is stable under $((T, L), (T, R))$ and $((T, L), (B, L))$, as well as under $((\frac{1}{2}T + \frac{1}{2}B, L), (T, R))$.

Definition 5.4 *Let (x, y) be a stationary profile and $\epsilon > 0$. An* ϵ-perturbation *is a stationary profile (x', y') such that $\| x - x' \| < \epsilon$, $\| y - y' \| < \epsilon$, $\text{supp}(x'_s) \supseteq \text{supp}(x_s)$ and $\text{supp}(y'_s) \supseteq \text{supp}(y_s)$.*

Note that if C is stable under an ϵ-perturbation (x', y') of (x, y), then in particular C is stable under (x, y).

Definition 5.5 *Let (x, y) be a stationary profile. A set $C \subseteq S$ is* communicating *w.r.t. (x, y) if for every $s \in C$ there exists an ϵ-perturbation (x', y') of (x, y) such that C is stable under (x', y') and*

$$\Pr_{s',x',y'}(\exists t \in \mathbf{N} \text{ s.t. } s_t = s) = 1 \qquad \forall s' \in C.$$

In the example presented in section 5.2, $\{s_1\}$ is communicating w.r.t. $((T, L), (\cdot, \cdot))$, $((T, C), (\cdot, \cdot))$, $((B, L), (\cdot, \cdot))$ and $((B, R), (\cdot, \cdot))$.

A set C is communicating if the players, by changing their stationary strategies a little, can reach from any state in C any other state in C, without leaving the set. Note that if there exists an ϵ-perturbation (x', y') that satisfies definition 5.5, then for every $\epsilon' > 0$ there exists an ϵ'-perturbation that satisfies it.

Clearly every absorbing state defines a communicating set, that includes only this state. We denote by $\mathcal{C}(x, y)$ the collection of all communicating sets w.r.t. (x, y). Define for every communicating set $C \in \mathcal{C}(x, y)$ and $s \in C$

$$\begin{aligned} A^1_s(C, y) &= \{a \in A \mid w_C(s, a, y_s) < 1\} \qquad \text{and} \qquad (74)\\ B^1_s(C, x) &= \{b \in B \mid w_C(s, x_s, b) < 1\}. \end{aligned}$$

In the example presented in section 5.2, denote by $(x^\star, y^\star) = ((T, L), (T, L))$. Then $A^1_{s_1}(s_1, y^\star) = \emptyset$ and $B^1_{s_1}(s_1, x^\star) = \{L\}$.

DEFINITION 5.6 *Let* (x, y) *be a stationary profile and* $C \in \mathcal{C}(x, y)$. *Every triplet* (s, x'_s, y_s), *where* $s \in C$ *and* $x'_s \in \Delta(A^1_s(C, y))$ *is an* exit of player 1 from C. *Every triplet* (s, x_s, y'_s), *where* $s \in C$ *and* $y'_s \in \Delta(B^1_s(C, x))$ *is an* exit of player 2 from C. *Every triplet* $(s, x'_s, y'_s) \in C \times \Delta(A) \times \Delta(B)$ *such that* $\text{supp}(x_s) \cap \text{supp}(x'_s) = \text{supp}(y_s) \cap \text{supp}(y'_s) = \emptyset$ *is a* joint exit from C *if* $w_C(s, x'_s, y'_s) < 1$ *while* $w_C(s, x_s, y'_s) = w_C(s, x'_s, y_s) = 1$.

In the example presented in section 5.2, $\{s_1\} \in \mathcal{C}(x^\star, y^\star)$, (s_1, T, R) is an exit of player 2 from $\{s_1\}$ and (s_1, B, C) is a joint exit from $\{s_1\}$.

A joint exit (s, x'_s, y'_s) is *pure* if $|\text{supp}(x'_s)| = |\text{supp}(y'_s)| = 1$. An exit (s, x'_s, y_s) of player 1 is *pure* if $|\text{supp}(x'_s)| = 1$. Pure exits of player 2 are defined analogously.

We denote by $E^1_C(x, y), E^2_C(x, y)$ and $E^{1,2}_C(x, y)$ the sets of exits of player 1, player 2 and the joint exits from C respectively. Let

$$E_C(x, y) = E^1_C(x, y) \cup E^2_C(x, y) \cup E^{1,2}_C(x, y)$$

be the set of all exits from C and $E^0_C(x, y)$ be the set of all pure exits from C. We denote by $s(e), x(e)$ and $y(e)$ the three coordinates of any exit e.

For any exit e, we define the support of e by:

$$\text{supp}(e) = \text{supp}(x(e)) \times \text{supp}(y(e)).$$

Note that any two exits $e_1, e_2 \in E^0_C(x, y)$ have disjoint supports.

For simplicity we write $s \in \mathcal{C}(x, y)$ whenever $\{s\} \in \mathcal{C}(x, y)$. In this case we write $E_s(x, y)$ instead of $E_{\{s\}}(x, y)$.

Let $R = S \setminus T$ be the set of non-absorbing states, and recall that $c = (c^i_s)$ is the min-max value of the players.

LEMMA 5.7 *Let* (x, y) *be a stationary profile such that* $R \in \mathcal{C}(x, y)$. *Assume that there exists an exit* $e \in E_R(x, y)$ *and* $g = (g_s)_{s \in S} \in \mathbf{R}^{2|S|}$ *such that*

1) g_s *is constant over* R, $g_s = u(s)$ *for every absorbing state* $s \in T$, *and* $g_s = \psi_g(e)$ *for every non-absorbing state* $s \in R$.

2) $g^1_s \geq \psi^1_c(s, a, y_s)$ *for every* $s \in R$ *and* $a \in A$.

3) $g^2_s \geq \psi^2_c(s, x_s, b)$ for every $s \in R$ and $b \in B$.

4) If $e \in E^1_R(x, y)$ then $g^1 = \psi^1_g(s(e), a, y_{s(e)})$ for every $a \in \mathrm{supp}(x(e))$.

5) If $e \in E^2_R(x, y)$ then $g^2 = \psi^2_g(s(e), x_{s(e)}, b)$ for every $b \in \mathrm{supp}(y(e))$.

Then g is an equilibrium payoff.

Note the similarity between the conditions of Lemma 5.7 and Lemma 4.4.

Proof: Fix $\epsilon > 0$. As in section 4.2, the ϵ-equilibrium profile that we construct (both here and in the proof of the next sufficient condition) is defined by a profile σ that satisfies $\| \sigma(h_0) - (x, y) \| \leq \epsilon$ for every finite history $h_0 \in H_0$, supplemented with a statistical test.

Let $\delta \in (0, \epsilon)$ be sufficiently small, and (x', y') be an ϵ-perturbation of (x, y) such that

$$\mathrm{Pr}_{s',x',y'}(\exists t \in \mathbf{N} \text{ s.t. } s_t = s(e)) = 1 \qquad \forall s' \in R.$$

Since $R \in \mathcal{C}(x, y)$, such a perturbation exists.

Define a profile σ as follows:

- Whenever the game is in state $s(e)$ the players play the mixed action combination $((1 - \delta)x_{s(e)} + \delta x(e), (1 - \delta)y_{s(e)} + \delta y(e))$.
- Whenever the game is in a state $s \neq s(e)$ the players play the mixed action combination (x'_s, y'_s).

If the players follow σ then the game is bound to exit R through e, and to be absorbed. Hence, by condition 1, the expected payoff for the players is g_s, where s is the initial state.

In order to prevent the players from deviating, we choose $t_1 \in \mathbf{N}$ sufficiently large and define the following statistical test. At each stage t the players check the following.

1) Both players check whether the realized action of their opponent is compatible with σ.

2) If $e \in E^2_R(x, y)$ and the game visited the state $s(e)$ at least t_1 times, then player 2 checks whether the distribution of the realized actions of player 1, whenever the game is in $s(e)$, is δ-close to $x_{s(e)}$. If $e \in E^1_R(x, y)$, then player 1 employs a symmetric test.

3) If $e \in E_R^{1,2}(x, y)$ and the game visited the state $s(e)$ at least t_1 times, then player 1 checks whether the realized actions of player 2, whenever the game is in the state $s(e)$, restricted to $\mathrm{supp}(y(e))$, is δ-close to $y(e)$. Player 2 employs a symmetric test.

If a player fails one of these tests, this player is punished by his opponent with an ϵ-min-max strategy forever.

Since player 1 may profit by causing the game never to be absorbed, we add one more test. Let $t_2 \in \mathbf{N}$ be sufficiently large such that if no deviation is detected then absorption occurs before stage t_2 with probability greater than $1 - \epsilon$. We add the following test to σ:

4) At stage t_2 both players switch to an ϵ-min-max strategy.

The constants δ and t_1 are chosen as in the proof of Lemma 4.4, in such a way that the probability of false detection of deviation is bounded by ϵ, and no player can profit more than $2\epsilon R$ by any kind of deviation. Thus g is an (x, y)-perturbed equilibrium payoff. ■

LEMMA 5.8 *Let (x, y) be a stationary profile, and $g = (g_s)_{s \in S} \in \mathbf{R}^{2|S|}$. Assume that the following conditions hold:*

1) $g_s = u(s)$ for every absorbing state $s \in T$.

2) $g_s^1 \geq \psi_c^1(s, a, y_s)$ for every $s \in R$ and $a \in A$.

3) $g_s^2 \geq \psi_c^2(s, x_s, b)$ for every $s \in R$ and $b \in B$.

4) For every $s \in R \setminus \mathcal{C}(x, y)$ the following hold:

a) $g_s^1 = \psi_g^1(s, a, y_s)$ for every $a \in \mathrm{supp}(x_s)$.
b) $g_s^2 = \psi_g^2(s, x_s, b)$ for every $b \in \mathrm{supp}(y_s)$.

5) For every $s \in R \cap \mathcal{C}(x, y)$ there exist two exits $e_1 = e_1(s)$ and $e_2 = e_2(s)$ in $E_s(x, y)$ and $\alpha = \alpha(s) \in [0, 1]$ that satisfy the following:

a) $\psi_g^1(e_j) = g_s^1$ for each $j = 1, 2$, and $g_s^2 = \alpha\psi_g^2(e_1) + (1 - \alpha)\psi_g^2(e_2)$.
b) If $e_j \in E_s^1(x, y)$ then $g_s^1 = \psi_g^1(s, a, y_s)$ for every $a \in \mathrm{supp}(x(e_j))$.

c) If $e_j \in E_s^2(x,y)$ then $\psi_g^2(s, x_s, b_1) = \psi_g^2(s, x_s, b_2) \geq \psi_c^2(s, x_s, b_3)$ for every $b_1, b_2 \in \operatorname{supp}(y(e_j))$ and $b_3 \in B$.

d) At most one of e_1 and e_2 is an exit of player 2.

6) The Markov chain over S whose transition law is induced by (x_s, y_s) for every $s \notin \mathcal{C}(x,y)$ and by $\alpha(s)e_1(s)+(1-\alpha(s))e_2(s)$ for every $s \in \mathcal{C}(x,y)$ is absorbing (i.e. for every initial state, an absorbing state is reached with probability 1).

Then g is an equilibrium payoff.

Proof: Let $\epsilon > 0$ be given. Let $t_1, t_2 \in \mathbf{N}$ be sufficiently large. These two constants will be used in the statistical tests, but the profile σ that we construct depend on them as well.

Define a profile σ as follows:

- Whenever the game is in a state $s \in R \backslash \mathcal{C}(x,y)$ the players play (x_s, y_s).

In order to define σ at states $s \in R \cap \mathcal{C}(x,y)$, we recall that to each such state there exist two exits $e_1 = e_1(s)$ and $e_2 = e_2(s)$, and $\alpha = \alpha(s)$ that satisfy condition 5. If one of these exits is an exit of player 2, we assume it is e_2. Otherwise, we assume that $\psi_g^2(e_1) \leq \psi_g^2(e_2)$. We fix $\delta = \delta(s) \in (0, 1/t_1)$ and $n = n(s) \in \mathbf{N}$ such that $(1-\delta)^n = 1-\alpha$ if e_1 is a unilateral exit and $(1-\delta^2)^n = 1-\alpha$ if e_1 is a joint exit.

The profile σ is defined at states $s \in R \cap \mathcal{C}(x,y)$ as follows. Assume the game moves to this state at stage t_0.

- The players play the mixed action combination $((1-\delta)x_s + \delta x(e_1), (1-\delta)y_s + \delta y(e_1))$ for n stages or until an action combination in the support of e_1 is played.
- If no action combination in the support of e_1 was played in the first n stages, the players play $((1-\delta)x_s + \delta x(e_2), (1-\delta)y_s + \delta y(e_2))$ until an action combination in the support of e_2 is played.
- If an action combination in the support of either e_1 or e_2 is played, but the game *remains* in the same state, then the players act as if the game has left the state s, and immediately returned to it.

Fix an initial state s and assume that the players follow σ. By condition 6, the game is bound to be eventually absorbed. Moreover, by the definition of $\delta(s)$ and $n(s)$, the probability to leave a state $s \in R \cap \mathcal{C}(x, y)$ through an action combination in the support of e_1 is $\alpha(s)$, and through an action combination in the support of e_2 is $1 - \alpha(s)$. By conditions 1, 4 and 5(a) the expected payoff for the players is g_s.

We supplement σ with a statistical test. In order to prevent the players from playing actions which are not compatible with σ, the players check, as in lemma 5.7, that the realized action combination that is played is compatible with σ.

Let $k \in \mathbf{N}$ be sufficiently large such that the number of times that the game leaves a communicating state $s \in R \cap \mathcal{C}(x, y)$ until absorption occurs is smaller than k with probability of at least $1 - \epsilon$.

Assume that the game moves to a state $s \in \mathcal{C}(x, y)$ at stage t_0. Denote $e_1 = e_1(s)$, $e_2 = e_2(s)$ and $n = n(s)$. Each player checks his opponent's behavior as follows. At every stage t such that $t_0 < t < t_0 + n$, or until the game leaves s:

1) If $e_1 \in E^2_s(x, y)$ and $t \geq t_0 + t_1$, then player 2 checks whether the distribution of the realized actions of player 1 at stages $t_0, t_0 + 1, \ldots, t - 1$, restricted to $\text{supp}(x_s)$, is ϵ-close to x_s. If $e_1 \in E^1_s(x, y)$, then player 1 employs a symmetric test.

2) If $e_1 \in E^{1,2}_s(x, y)$ and $t \geq t_0 + t_2$, then both players check whether the realized actions of their opponent at stages $t_0, t_0 + 1, \ldots, t - 1$, restricted to $\text{supp}(x(e_1))$ and $\text{supp}(y(e_1))$, is ϵ-close to $x(e_1)$ and $y(e_1)$ respectively.

If a player fails one of these tests at a stage $t_0 \leq t < t_0 + n$, this player is punished with an ϵ-min-max strategy forever.

If no deviation is detected before stage $t_0 + n$, then each player begins to check, in a similar way, if his opponent continues to follow σ, until the game leaves the state s (i.e. replace e_1 by e_2 in the statistical tests).

Let t_3 be sufficiently large such that if no deviation is detected then leaving s occurs in t_3 stages with probability greater than $1 - \epsilon/k$. As in the proof of Lemma 5.7, at stage $t_0 + t_3$ both players switch to an ϵ-min-max strategy.

By conditions 2, 3 and 4, no player can deviate in any state $s \notin \mathcal{C}(x, y)$ and profit more than ϵR.

The constants t_1 and $t_2 = t_2(\delta)$ are chosen, as in the proof of Lemma 4.4, in such a way that no player can profit more than $2\epsilon R$ by any deviation, and the probability of false detection of deviation is bounded by ϵ. Thus g is an (x, y)-perturbed equilibrium payoff. ■

5.4 Preliminary Results

A stationary profile (x, y) is *absorbing* if by playing it, the game is bound to reach an absorbing state with probability 1.

For every state $s \in S$, let

$$v_s^i(x, y) = \mathbf{E}_{s,x,y} \left(\lim_{t \to \infty} \sum_{t=1}^{\infty} r_t^i / t \right)$$

be the expected undiscounted payoff for player i if the initial state is s and the players follow the stationary profile (x, y). The function $v(x, y) = (v_s(x, y))_{s \in S} \in \mathbf{R}^{2|S|}$ is harmonic over S w.r.t. the transition $p_{s,s'} = w_{s'}(s, x_s, y_s)$. If (x, y) is absorbing then $v(x, y)$ is the unique solution of the following system of linear equations:

$$\begin{aligned} \xi_s &= u(s) && \forall s \in T \\ \xi_s &= \psi_\xi(s, x_s, y_s) && \forall s \in R. \end{aligned} \tag{75}$$

Lemma 5.9 *Let (x, y) be an absorbing stationary profile. Let $g : S \to \mathbf{R}^2$ be such that $\psi_g^2(s, x_s, y_s) \leq g_s^2$ for every $s \in R$ and $g_s^2 = u^2(s)$ for every $s \in T$. Then $v_s^2(x, y) \leq g_s^2$ for every $s \in S$*

Proof: $v^2(x, y)$ is an harmonic function and g^2 is a sub-harmonic function over S that have the same values over T. Hence $v^2(x, y) - g^2$ is a super-harmonic function that vanishes over T. Since (x, y) is absorbing, the result follows. ■

Corollary 5.10 *Let x be a stationary strategy of player 1. Let $g : S \to \mathbf{R}^2$ satisfy for every stationary strategy y of player 2, (i) $g_s^2 \geq \psi_g^2(s, x_s, y_s)$ for every $s \in R$ and (ii) $g_s^2 = u^2(s)$ for every $s \in T$. Then $c_s^2 \leq g_s^2$ for every $s \in S$.*

Proof: Since the game is a positive recursive game with the absorbing property, the best reply of player 2 against the stationary strategy x is a stationary strategy y such that (x, y) is absorbing. By Lemma 5.9, $v^2_s(x, y) \leq g^2_s$ for every stationary strategy y such that (x, y) is absorbing and $s \in S$. Hence $c^2_s \leq g^2_s$. ■

A symmetric proof proves the following lemma:

LEMMA 5.11 *Let y be a fully mixed stationary strategy of player 2. Let $g : S \to \mathbf{R}^2$ satisfy for every stationary strategy x of player 1, (i) $\psi^1_g(s, x_s, y_s) \leq g^1_s$ for every $s \in R$ and (ii) $g^1_s = u^1(s)$ for every $s \in T$. Then $c^1_s \leq g^1_s$ for every $s \in S$.*

5.5 The ϵ-Approximating Game

5.5.1 The Game

Let $\epsilon^\star = \frac{1}{|B|}$. For every $\epsilon \in (0, \epsilon^\star)$ define the set

$$Y_s(\epsilon) = \left\{ y_s \in \Delta(B) \mid \sum_{b \in J} y^b_s \geq \epsilon^{|B|-|J|} \qquad \forall J \subseteq B \right\}. \tag{76}$$

Let $Y(\epsilon) = \times_{s \in S} Y_s(\epsilon)$. Every stationary strategy $y \in Y(\epsilon)$ is fully mixed. Since the game satisfies the absorbing property, the payoff function $v(x, y)$ is continuous over $X \times Y(\epsilon)$.

Define the ϵ-approximating game $G'(\epsilon)$ as a positive recursive game with the absorbing property (S, A, B, w, u), where player 2 is restricted to strategies in which the mixed action he plays whenever the game is in state s must be in $Y_s(\epsilon)$ (rather than in $\Delta(B)$).

5.5.2 Existence of a Stationary Equilibrium

Note that X and $Y(\epsilon)$ (for every $\epsilon \in (0, \epsilon^\star)$) are non-empty, convex and compact sets. Define the correspondence $\phi^1_{s,\epsilon} : X \times Y(\epsilon) \to \Delta(A)$ by:

$$\phi^1_{s,\epsilon}(x, y) = \text{argmax}_{x'_s \in \Delta(A)} \psi^1_{v(x,y)}(s, x'_s, y_s) \tag{77}$$

that is, player 1 maximizes his payoff locally — in every state s he chooses a mixed action that maximizes his expected payoff if the initial state is s, player 2 plays the mixed action y_s, and the continuation payoff is given by $v(x, y)$. Let $\phi^1_\epsilon = \times_{s \in S} \phi^1_{s,\epsilon}$.

Lemma 5.12 *The correspondence ϕ^1_ϵ has non-empty convex values and it is upper semi-continuous.*

Proof: Since $\psi^1_{v(x,y)}(s, x'_s, y_s)$ is linear in x'_s for every fixed (s, x, y) and X is convex and compact, ϕ^1_ϵ has non-empty and convex values. By the continuity of v^1 over the compact set $X \times Y(\epsilon)$, it follows that ϕ^1_ϵ is upper semi-continuous. ■

Define the correspondence $\phi^2_{s,\epsilon} : X \times Y(\epsilon) \to Y_s(\epsilon)$ by:

$$\phi^2_{s,\epsilon}(x, y) = \text{argmax}_{y'_s \in Y_s(\epsilon)} \psi^2_{v(x,y)}(s, x_s, y'_s). \tag{78}$$

Let $\phi^2_\epsilon = \times_{s \in S} \phi^2_{s,\epsilon}$. As in Lemma 5.12, since $Y_s(\epsilon)$ is not empty, convex and compact whenever $\epsilon \in (0, \epsilon^\star)$, and v^2 is continuous over $X \times Y(\epsilon)$, we have:

Lemma 5.13 *The correspondence ϕ^2_ϵ has non-empty convex values and it is upper semi-continuous.*

Define the correspondence $\phi_\epsilon : X \times Y(\epsilon) \to X \times Y(\epsilon)$ by

$$\phi_\epsilon(x, y) = \phi^1_\epsilon(x, y) \times \phi^2_\epsilon(x, y).$$

By Lemmas 5.12, 5.13 and by Kakutani's fixed point Theorem we get:

Lemma 5.14 *For every $\epsilon \in (0, \epsilon^\star)$ there exists $(x(\epsilon), y(\epsilon)) \in X \times Y(\epsilon)$ that is a fixed point of the correspondence ϕ_ϵ.*

5.5.3 Properties as $n \to 0$

Since the state and action spaces are finite, there exists sequences $\{\epsilon_n\}_{n \in \mathbf{N}}$ of real numbers and $\{(x(n), y(n))\}_{n \in \mathbf{N}}$ of stationary profiles such that $\epsilon_n \to 0$, $(x(n), y(n))$ is a fixed point of ϕ_{ϵ_n} for every $n \in \mathbf{N}$, and $(x(n), y(n))$ converge to a limit $(x(\infty), y(\infty))$. Moreover, it can be assumed that $\text{supp}(x_s(n))$ and $\text{supp}(y_s(n))$ are independent of n. Thus, if for some fixed pair (s, a), $x^a_s(n) > 0$ for one n, this inequality holds for every $n \in \mathbf{N}$.

In the sequel, we need that various sequences that depend on $\{x(n)\}$ and $\{y(n)\}$ have a limit. The number of those sequences is finite, hence, by taking a subsequence we will assume that the limits exist.
Remark: Using Theorem 2.5 it can be proven that we can choose for every $\epsilon > 0$ a fixed point $(x(\epsilon), y(\epsilon))$ of ϕ_ϵ such that x and y, as functions of ϵ, are Puiseux functions, hence all the limits that we take exist.

We denote for every $n \in \mathbf{N}$ and $s \in S$, $g_s(n) = v_s(x(n), y(n))$. Denote $g_s(\infty) = \lim_{n\to\infty} g_s(n)$. Define for every $x \in X$, every $y \in Y$ and every $(s, b) \in S \times B$

$$h^b_s(x, y) = \psi^2_{v(x,y)}(s, x_s, b).$$

$h^b_s(x, y)$ is the expected payoff for player 2 if the initial state is s, the players play the mixed action combination (x_s, b) and the continuation payoff is given by $v(x, y)$.

Lemma 5.15 *Let* $s_1 \in S$ *and* $b_1, b_2 \in B$. *If* $\lim_{n\to\infty} \frac{y^{b_1}_{s_1}(n)}{y^{b_2}_{s_1}(n)} < \infty$ *then for every* n *sufficiently large*

$$h^{b_1}_{s_1}(x(n), y(n)) \leq h^{b_2}_{s_1}(x(n), y(n)).$$

Proof: By definition, $y(n)$ is a best reply in $Y(\epsilon_n)$ against $x(n)$. If the lemma is not true, then for every n one can find a strategy $y'(n)$ for player 2, that is in $Y(\epsilon_n)$, and, for n sufficiently large, yields player 2 a better payoff than does $y(n)$ — a contradiction.

Assume that the lemma is not true. Then, by taking a subsequence, $h^{b_1}_{s_1}(x(n), y(n)) > h^{b_2}_{s_1}(x(n), y(n))$ for every $n \in \mathbf{N}$. Define for every n the stationary strategy $y'(n)$ for player 2 as follows:

$$y'^b_s(n) = \begin{cases} y^{b_2}_s(n)/2 & (s, b) = (s_1, b_2) \\ y^{b_1}_s(n) + y^{b_2}_s(n)/2 & (s, b) = (s_1, b_1) \\ y^b_s(n) & \text{Otherwise.} \end{cases}$$

Let us verify that $y'_s(n) \in Y_s(\epsilon_n)$ for every n sufficiently large. Otherwise, by taking a subsequence, there exists a set $J \subseteq B$ such that $b_2 \in J$, $b_1 \notin J$ and

$$\sum_{b\in J} y'^b_s(n) < \epsilon_n^{|B|-|J|} \qquad \forall n \in \mathbf{N}. \tag{79}$$

In particular, $\lim \frac{y_s^{b_2}(n)}{\epsilon_n^{|B|-|J|}} < \infty$. By the assumption, $\lim \frac{y_s^{b_1}(n)}{\epsilon_n^{|B|-|J|}} < \infty$ as well. Since $\sum_{b \in J} y_s^b(n) + y_s^{b_1}(n) \geq \epsilon_n^{|B|-|J|-1}$ it follows that there exists $b \in J \setminus \{b_2\}$ such that $\lim \frac{y_s^b(n)}{\epsilon_n^{|B|-|J|-1}} > 0$ — a contradiction to (79).

However

$$\begin{aligned}\psi^2_{g(n)}(s, x_s(n), y'_s(n)) - \psi^2_{g(n)}(s, x_s(n), y_s(n))&\\ &= \left(\psi^2_{g(n)}(s, x_s(n), b_1) - \psi^2_{g(n)}(s, x_s(n), b_2)\right) y_{s_1}^{b_2}(n)/2\\ &= \left(h_{s_1}^{b_1}(x(n), y(n)) - h_{s_1}^{b_2}(x(n), y(n))\right) y_{s_1}^{b_2}(n)/2 > 0,\end{aligned}$$

which contradicts the optimality of $y(n)$ against $x(n)$. ■

By applying Lemma 5.15 in both directions, and taking the limit as $n \to \infty$ we conclude that if player 2 plays two actions with the same order of magnitude, then the corresponding limits of his continuation payoffs are equal. We need a weaker result:

COROLLARY 5.16 *Let* $b_1, b_2 \in B$ *and* $s \in S$. *If* $\lim_{n\to\infty} \frac{y_s^{b_1}(n)}{y_s^{b_2}(n)} \in (0, \infty)$ *then*

$$\psi^2_{g(\infty)}(s, x_s(\infty), b_1) = \psi^2_{g(\infty)}(s, x_s(\infty), b_2).$$

Since all the actions in the support of $y_s(\infty)$ are played with the same order of magnitude, we have:

COROLLARY 5.17 *For every* $b \in \text{supp}(y_s(\infty))$

$$\psi^2_{g(\infty)}(s, x_s(\infty), b) = g_s^2(\infty).$$

Proof:

$$\begin{aligned} g_s^2(\infty) &= \lim_{n\to\infty} g_s^2(n)\\ &= \lim_{n\to\infty} \psi^2_{g(n)}(s, x_s(n), y_s(n))\\ &= \psi^2_{g(\infty)}(s, x_s(\infty), y_s(\infty))\\ &= \sum_{b \in \text{supp}(y_s(\infty))} y_s^b(\infty) \psi^2_{g(\infty)}(s, x_s(\infty), b).\end{aligned}$$

The result now follows from Corollary 5.16. ■

By Lemma 5.15, Corollary 5.17 and the continuity of ψ it follows that

$$\psi^2_{g(\infty)}(s, x_s(\infty), b) \leq g^2_s(\infty) \qquad \forall (s, b) \in S \times B. \tag{80}$$

By Corollary 5.10 and (80) we have

$$g^2_s(\infty) \geq c^2_s \qquad \forall s \in S. \tag{81}$$

Since $x(n)$ is a best reply against $y(n)$,

$$\psi^1_{g(n)}(s, a, y_s(n)) \leq g^1_s(n) \qquad \forall (s, a) \in S \times A \tag{82}$$

and equality holds whenever $a \in \operatorname{supp}(x_s(n))$. Taking a limit in (82) as $n \to \infty$ we get

$$\psi^1_{g(\infty)}(s, a, y_s(\infty)) \leq g^1_s(\infty) \qquad \forall (s, a) \in S \times A \tag{83}$$

and equality holds whenever $x^a_s(n) > 0$ for every n.

By Lemma 5.11 and (82), $c^1_s \leq g^1_s(n)$ for every $s \in S$ and $n \in \mathbf{N}$, and by taking the limit as $n \to \infty$, $c^1_s \leq g^1_s(\infty)$ for every $s \in S$. Therefore,

$$\psi^1_c(s, a, y_s(\infty)) \leq g^1_s(\infty) \qquad \forall (s, a) \in S \times A. \tag{84}$$

To summarize, we have asserted that $g^i_s(\infty)$ is greater than the min-max value of player i, and that no player can receive more than $g_s(\infty)$ by playing any action in any state s and then be punished with his min-max value.

5.6 Existence of an Equilibrium Payoff

In this section we prove that if there are two non-absorbing states then either the conditions of Lemma 5.7 hold or the conditions of Lemma 5.8 hold. Denote the two non-absorbing states by $R = \{s_1, s_2\}$.

In our proof we could have used Puiseux functions, as was done in Section 4. However, whereas there the degree of the function was used extensively, here it is used only to simplify arguments. Therefore we decided to prove the results using a convergent sub-sequence, as is done in the literature, including the classical paper of Vrieze and Thuijsman [36].

Before going into the details of the proof, we explain the way we are heading. The idea is, as in Section 4, to consider a sequence of stationary equilibria in the ϵ-approximating games that converges to a limit x. We denote by g the limit of the equilibrium payoff in these ϵ-approximating games.

If there is no communicating set w.r.t. x, it is easy to see that x is a stationary absorbing equilibrium profile.

Assume now that there is a communicating set w.r.t. x that contains a single state s. The idea is to find a pair of exits from the communicating set $\{s\}$ such that the continuation payoff of player 1 by both is equal to g_s^1 (that is, the limit of the expected payoff for player 1 in the ϵ-approximating game if the play leaves s through each of these exits), and the expected payoff of player 2 if the first exit is used is more than g_s^2, while if the second exit is used is less than g_s^2. Since there are only two non-absorbing states, the other state either forms another communicating set, and we apply the same argument, or it is transient w.r.t. x. Thus, we can apply Lemma 5.8, and construct an equilibrium payoff. The problem arises when both states form a single communicating set. In this case, there need not exist a pair of exits as above. However, there always exists an exit that yields both player an expected continuation payoff of *at least* g_s (for any non-absorbing state s). Since there is a unique communicating set, one can apply Lemma 5.7, and construct an equilibrium payoff.

5.6.1 Exits from a State

Fix a state $s \in S$ such that $w_s(s, x_s(\infty), y_s(\infty)) = 1$. In particular, $s \in \mathcal{C}(x(\infty), y(\infty))$. Since the game satisfies the absorbing property, $E_s(x(\infty),y(\infty)) \neq \emptyset$.

Let

$$U = \left\{(a,b) \in A \times B \mid (a,b) \in \text{supp}(e) \text{ for some } e \in E_s^0(x(\infty), y(\infty))\right\}.$$

U is the set of all pairs of actions that take part in one of the absorbing neighbors of $(x(\infty), y(\infty))$. Note that not necessarily $w_s(s,a,b) < 1$ for every $(a,b) \in U$, and that $E_s^0(x(\infty), y(\infty))$ induces naturally a partition on U.

For every $n \in \mathbf{N}$, the mixed action $(x_s(n), y_s(n))$ induces a probability distribution over $A \times B$. Since the game satisfies the absorbing property, this probability distribution gives positive probability to the set U. Let ρ_n be the conditional probability induced by $(x_s(n), y_s(n))$ over $E_s^0(x(\infty), y(\infty))$. Define $\rho_\infty \overset{\text{def}}{=} \lim_n \rho_n$.

We define two additional subsets of $A \times B$:

$$U_1 = \{(a,b) \in (A \times B) \setminus U \mid w_s(s,a,b) < 1\}$$

and

$$U_2 = \{(a,b) \in (A \times B) \setminus U \mid w_s(s,a,b) = 1\}.$$

U_1 is the set of pairs where exiting $\{s\}$ occurs with positive probability, while U_2 is the set of pairs that cause the play to stay in s with probability 1.

Any pair in U_1 is dominated by a pair of actions in U: for every $(a,b) \in U_1$ there exists a pair $(a_0, b_0) \in U$ such that $\lim_n \frac{x_s^a(n) y_s^b(n)}{x_s^{a_0}(n) y_s^{b_0}(n)} = 0$. Clearly, if $(a,b) \in U_2$ then $g_s(n) = \psi_{g(n)}(s,a,b)$. Therefore

$$g_s(n) = \sum_{(a,b)\in A\times B} x_s^a(n) y_s^b(n) \psi_{g(n)}(s,a,b) = \sum_{e \in E_s^0(x(\infty),y(\infty))} \rho_n(e)\psi_{g(n)}(e) + o(1),$$

where $o(1)$ means a function that converges to 0 as $n \to \infty$. The elements in U contribute the first term on the right hand side, and the elements in U_1 contribute the second term. The elements in U_2 are absorbed on the left hand side.

By taking the limit as $n \to \infty$ we have:

$$g_s(\infty) = \sum_{e \in E_s^0(x(\infty),y(\infty))} \rho_\infty(e)\psi_{g(\infty)}(e). \tag{85}$$

This equation is the equivalent of Lemma 4.21 for communicating sets.

For every $a \in \text{supp}(x_s(n)) \setminus \text{supp}(x_s(\infty)) \setminus A_s^1(s, y(\infty))$ we define

$$B(a) = \{b \in B \mid (s,a,b) \in E_s^0(x(\infty), y(\infty))\}.$$

$B(a)$ is the collection of the elements in U whose first coordinate is a. Note that if $b \in B(a)$ then (s,a,b) is a joint exit from s.

If $\sum_{b\in B(a)} \rho_\infty(s,a,b) > 0$ then we define an exit $e_a \in E_s(x(\infty), y(\infty))$ by:

$$e_a = \left(s, a, \frac{\sum_{b\in B(a)} \rho_\infty(s,a,b) \cdot b}{\sum_{b\in B(a)} \rho_\infty(s,a,b)}\right).$$

e_a is not necessarily a pure exit. It is, in a sense, an exit that includes all the exits from s where player 1 plays a.

Since for every $n \in \mathbf{N}$ and action $a \in \text{supp}(x_s(n))$, $\psi^1_{g(n)}(s, a, y_s(n)) = g^1_s(n)$, one can show as above that

$$\psi^1_{g(\infty)}(e_a) = g^1_s(\infty). \tag{86}$$

Recall that $B^1_s(s, x(\infty))$ is the set of all actions of player 2 that cause the game to leave s with positive probability, given player 1 plays $x_s(\infty)$. If $\sum_{b \in B^1_s(s,x(\infty))} \rho_\infty(s, x_s(\infty), b) > 0$, we define an exit $e_0 \in E_s(x(\infty), y(\infty))$ by:

$$e_0 = \left(s, x_s(\infty), \frac{\sum_{b \in B^1_s(s,x(\infty))} \rho_\infty(s, x_s(\infty), b) \cdot b}{\sum_{b \in B^1_s(s,x(\infty))} \rho_\infty(s, x_s(\infty), b)} \right).$$

As above, since $\psi^1_{g(n)}(s, x_s(\infty), y_s(n)) = g^1_s(n)$ for every $n \in \mathbf{N}$, one can show that

$$\psi^1_{g(\infty)}(e_0) = g^1_s(\infty). \tag{87}$$

For every $a \in A^1_s(s, x(\infty))$ for which $\rho_\infty(s, a, y_s(\infty)) > 0$ define an exit $e_a \in E_s(x(\infty), y(\infty))$ by:

$$e_a = (s, a, y_s(\infty)).$$

By (83) we have

$$\psi^1_{g(\infty)}(e_a) = g^1_s(\infty). \tag{88}$$

The following lemma states that there always exists a pair of exits that satisfy various properties, as described in the beginning of this section.

LEMMA 5.18 *There exist two exits* $e_1, e_2 \in E_s(x(\infty), y(\infty))$ *and* $\alpha \in [0, 1]$ *such that*

1. $g^1_s(\infty) = g^1_s(e_j)$ *for* $j = 1, 2$.
2. $g^2_s(\infty) = \alpha \psi^2_{g(\infty)}(e_1) + (1 - \alpha) \psi^2_{g(\infty)}(e_2)$.
3. *At most one of* e_1, e_2 *is a unilateral exit of player 2.*
4. *If there exists* $e \in \text{supp}(\rho_\infty)$ *such that* $w_T(e) > 0$, *then* $w_T(e_1) > 0$.

Proof: Consider the set E' that contains the exits e_0 and $\{e_a\}$ that were defined above. Note that the supports of these exits form a partition of U, which is coarser than the partition $E^0_s(x(\infty), y(\infty))$. Thus, for every $e' \in E'$,

$$\rho_\infty(e') = \sum\{\rho_\infty(e) \mid e \in E^0_s(x(\infty), y(\infty)) \text{ and } \mathrm{supp}(e) \subseteq \mathrm{supp}(e')\}.$$

By 85,

$$g_s(\infty) = \sum_{e' \in E'} \rho_\infty(e')\psi_{g(\infty)}(e'). \tag{89}$$

Note that by (86), (87) and (88), $g^1_s(\infty) = \psi^1_{g(\infty)}(e')$ for every $e' \in E'$.

We will now distinguish between two cases.

Case 1: If there exists $e' \in E'$ with $w_T(e') > 0$, we choose e_1 to satisfy $w_T(e_1) > 0$. If $\psi_{g(\infty)}(e_1) = g_s(\infty)$ we are done (choose $\alpha = 1$ and $e_2 = e_1$). Otherwise, assume w.l.o.g. that $\psi^2_{g(\infty)}(e_1) > g^2_s(\infty)$. by (89) there exists $e_2 \in E'$ with $\psi^2_{g(\infty)}(e_2) < g^2_s(\infty)$. α in, then, the unique number that satisfies condition (2).

Case 3: Otherwise, by (89) it follows that there exist two exits $e_1, e_2 \in E'$ and $\alpha \in [0, 1]$ such that condition (2) holds. By (86), (87) and (88) condition (1) holds, by assumption condition (4) holds, and by the construction of E' condition (3) holds. ∎

5.6.2 The Main Result

The proof of the main result is divided into two cases: whether R is communicating under $(x(\infty), y(\infty))$ or not.

Consider the following two conditions:

A.1. R is communicating under $(x(\infty), y(\infty))$.

A.2. For every $s \in R$ such that $s \in \mathcal{C}(x(\infty), y(\infty))$, and every $e \in E_s(x(\infty), y(\infty))$ such that $\rho_\infty(e) > 0$ we have $w_T(e) = 0$.

We shall prove that if conditions **A** hold then the conditions of Lemma 5.7 hold, while if they do not hold then the conditions of Lemma 5.8 hold.

Lemma 5.19 *If conditions* **A** *hold then the conditions of Lemma 5.7 hold w.r.t.* $(x(\infty), y(\infty))$.

Proof: By taking a subsequence and exchanging the names of s_1 and s_2 if necessary, we can assume that exactly one of the following holds:

B.1. Either $g^2_{s_1}(n) > g^2_{s_2}(n)$ for every $n \in \mathbf{N}$.

B.2. Or, $g^2_{s_1}(n) = g^2_{s_2}(n)$ and $g^1_{s_1}(n) > g^1_{s_2}(n)$ for every $n \in \mathbf{N}$.

B.3. Or, $g^2_{s_1}(n) = g^2_{s_2}(n)$, $g^1_{s_1}(n) = g^1_{s_2}(n)$ and $w_T(s_1, x_{s_1}(n), y_{s_1}(n)) > 0$ for every $n \in \mathbf{N}$.

Step 1: Construction of an exit.
Note that $w_T(s_1, x_{s_1}(n), y_{s_1}(n)) > 0$ for every $n \in \mathbf{N}$. Otherwise, it follows that $g_{s_1}(n) = g_{s_2}(n)$ for every $n \in \mathbf{N}$, which contradicts all of the assumptions **B**.

Let $A^\star$ be the set of all actions $a \in A$ such that

i) $w_T(s_1, a, y_{s_1}(n)) > 0$ for every n.

ii) $\lim_{n\to\infty} \frac{w_T(s_1,a,y_{s_1}(n))}{w_T(s_1,a',y_{s_1}(n))} > 0$ for every a' such that $w_T(s_1, a', y_{s_1}(n)) > 0$ for every n.

i.e. the actions of player 1 that are absorbing with highest order of magnitude against $y_{s_1}(n)$. Let $B^\star$ be the set of all actions $b \in B$ such that

- $w_T(s_1, x_{s_1}(n), b) > 0$ for every n.
- $\lim \frac{y^b_{s_1}(n)}{y^{b'}_{s_1}(n)} > 0$ for every b' such that $w_T(s_1, x_{s_1}(n), b') > 0$ for every n.

i.e. the actions of player 2 that are absorbing against $x_{s_1}(n)$ and player 2 plays with highest order of magnitude.

Since $w_T(s_1, x_{s_1}(n), y_{s_1}(n)) > 0$ for every n it follows that $A^\star, B^\star \neq \emptyset$.

Let $x^\star_{s_1}(n)$ and $y^\star_{s_1}(n)$ be the conditional probability distribution that is induced by $x_{s_1}(n)$ and $y_{s_1}(n)$ over $A^\star$ and $B^\star$ respectively. Denote $(x^\star_{s_1}(\infty), y^\star_{s_1}(\infty)) = \lim_{n\to\infty}(x^\star_{s_1}(n), y^\star_{s_1}(n))$.

Let $e = (s_1, x^\star_{s_1}(\infty), y^\star_{s_1}(\infty))$. By the definition of $A^\star$ and $B^\star$ it follows that $\mathrm{supp}(x^\star_{s_1}(\infty)) = \mathrm{supp}(x^\star_{s_1}(n))$ and $\mathrm{supp}(y^\star_{s_1}(\infty)) = \mathrm{supp}(y^\star_{s_1}(n))$ for every $n \in \mathbf{N}$. Therefore e is an exit from R.

Step 2: $g^2_{s_1}(\infty) \leq \psi^2_{g(\infty)}(e)$.

Assume to the contrary that $g^2_{s_1}(\infty) > \psi^2_{g(\infty)}(e)$. In particular, for n sufficiently large, $g^2_{s_1}(n) > \psi^2_{g(n)}(e)$. By the definition of $A^\star$ and since $g^2_{s_1}(n) \geq g^2_{s_2}(n)$ it follows that $g^2_{s_1}(n) > \psi^2_{g(n)}(s_1, x_{s_1}(n), y^\star_{s_1}(n))$ for n sufficiently large. Since $g^2_{s_1}(n) = \psi^2_{g(n)}(s_1, x_{s_1}(n), y_{s_1}(n))$, it follows that there exists an action $b_0 \in B$ such that $g^2_{s_1}(n) < \psi^2_{g(n)}(s_1, x_{s_1}(n), b_0)$ for n sufficiently large. By Lemma 5.15, $\lim_n \frac{y^{b_0}_{s_1}(n)}{y^b_{s_1}(n)} = \infty$ for every action $b \in B^\star$, and by the definition of $B^\star$, $b_0 \notin B^\star$. In particular, $w_T(s_1, x_{s_1}(n), b_0) = 0$, which implies that for every n, $g^2_{s_1}(n) \geq \psi^2_{g(n)}(s_1, x_{s_1}(n), b_0)$ — a contradiction.

Step 3: $g^1_{s_1}(\infty) \leq \psi^1_{g(\infty)}(e)$.
Assume to the contrary that $g^1_{s_1}(\infty) > \psi^1_{g(\infty)}(e)$. In particular, $g^1_{s_1}(n) > \psi^1_{g(n)}(e)$ for n sufficiently large. We shall now prove that this implies that $g^1_{s_1}(n) < g^1_{s_2}(n)$ *and* there exists $b_0 \in B$ such that $w_{s_2}(s_1, x^\star(\infty), b_0) > 0$ and $\lim_n \frac{y^{b_0}_{s_1}(n)}{y^b_{s_1}(n)} = \infty$ for every $b \in B^\star$. Indeed, otherwise it follows by the definition of $B^\star$ that for every n sufficiently large, $g^1_{s_1}(n) > \psi^1_{g(n)}(s_1, x^\star(n), y(n))$, which contradicts assumption **C**.1.

Hence **B**.1 holds, and therefore $g^2_{s_1}(n) > g^2_{s_2}(n)$ for every $n \in \mathbf{N}$. By the definition of $B^\star$, $w_T(s_1, x_{s_1}(n), b_0) = 0$ for every n, and therefore $g^2_{s_1}(n) > \psi^2_{g(n)}(s_1, x_{s_1}(n), b_0)$ for every n sufficiently large, and we get the same contradiction as in step 2.

Step 4: Definition of the equilibrium payoff.
Define $z = (z_s)_{s \in S} \in \mathbf{R}^{2|S|}$ by:

$$z_s = \begin{cases} u(s) & s \in T \\ \frac{\sum_{s' \in T} w_{s'}(e) u(s')}{w_T(e)} & s \in R \end{cases}$$

Step 5: The conditions of Lemma 5.7 hold w.r.t. $(x(\infty), y(\infty))$ and z.
Condition 1 of Lemma 5.7 follows from the definition of z. Condition 2 follows from Step 3 and (84) while condition 3 follows from Step 2, (80) and (81). Condition 4 follows from (83) and condition 5 follows from Corollary 5.16. ■

Lemma 5.20 *If conditions* **A** *do not hold then the conditions of Lemma 5.8 hold w.r.t.* $(x(\infty), y(\infty))$.

Proof: Define $z = g(\infty)$. We prove that the conditions of Lemma 5.8 hold w.r.t. $(x(\infty), y(\infty))$ and z. Condition 1 holds by the definition of z. Condition 2 follows from (84) while condition 3 follow from (80) and (81). Condition 4 follows from (83) and Corollary 5.17. Conditions 5(a) and 5(d) follow from Lemma 5.18, whereas condition 5(b) follows from (83) and condition 5(c) follows from corollary 5.17, (80) and (81).

Since conditions **A** do not hold, it follows that condition 6 of Lemma 5.8 holds. ■

5.7 More Than Two Non-Absorbing States

Why does our approach fail for games with more than two non-absorbing states ? The reason is that if conditions **A** hold then the equilibrium payoff that we construct need not be equal to $g(\infty)$, and we run into a similar problem as discussed in section 1.5.2.

As an example, consider the following positive recursive game with the absorbing property and four non-absorbing states:

state 1

$$\begin{array}{c|c|c|} & 1-\frac{1}{n} & \frac{1}{n} \\ \hline 1-\frac{1}{n} & \mathbf{2} & 1,1 \quad ^{*} \\ \hline \frac{1}{n} & 0,6 \quad ^{*} & \mathbf{3} \\ \hline \end{array}$$

state 2

$$\begin{array}{c|c|c|} & 1-\frac{1}{n} & \frac{1}{n} \\ \hline 1-\frac{1}{n} & \mathbf{1} & -1,0 \quad ^{*} \\ \hline \frac{1}{n} & 0,5 \quad ^{*} & \mathbf{4} \\ \hline \end{array}$$

state 3

$$\begin{array}{c|c|c|} & 1/2 & 1/2 \\ \hline 1/2 & \mathbf{1} & 1,3 \quad ^{*} \\ \hline 1/2 & 1,3 \quad ^{*} & \mathbf{1} \\ \hline \end{array}$$

state 4

$$\begin{array}{c|c|c|} & 1/2 & 1/2 \\ \hline 1/2 & \mathbf{2} & -1,3 \quad ^{*} \\ \hline 1/2 & -1,3 \quad ^{*} & \mathbf{2} \\ \hline \end{array}$$

Figure 1

Let $(x(n), y(n))$ be the stationary profile indicated in Figure 1. First let us verify that $(x(n), y(n))$ is a fixed point of the correspondence $\phi_{1/n}$ defined in section 5.5. Indeed, both players are indifferent between their actions in states 3 and 4, and in states 1 and 2 player 2 must play each action with a probability of at least $1/n$. It can be checked that player 2 prefers to play in these two states L, hence his strategy is a best reply in $Y(1/n)$. The expected payoff for player 1 by this stationary strategy profile is $g^1(n) = (\frac{1/n}{2-1/n}, \frac{-1/n}{2-1/n}, \frac{1}{2-1/n}, \frac{-1}{2-1/n})$. One can check that player 1 is indifferent between his actions in states 1 and 2, and therefore his strategy is a best reply against $y(n)$.

Let $C = \{s_1, s_2\}$. Note that C is communicating w.r.t. $(x(\infty), y(\infty))$, and that the exits from C w.r.t. $(x(\infty), y(\infty))$ are $(s_1, B, L), (s_1, T, R), (s_2, B, L)$ and (s_2, T, R). It turns out that ρ_∞ is the uniform distribution over these four exits and therefore

$$g^1(\infty) = g^2(\infty) = \frac{1}{4}(0, 6) + \frac{1}{4}(1, 1) + \frac{1}{4}(0, 5) + \frac{1}{4}(-1, 0) = (0, 3).$$

However, there does not exist any way to exit from C in such a way that is individually rational for both players and yields the players an expected payoff $(0, 3)$.

The proof given by Vieille [35] for the existence of an equilibrium payoff in positive recursive games with the absorbing property and arbitrary number of non-absorbing states, uses different approximating games and best reply correspondence. Moreover, after proving that the approximating game has a stationary equilibrium, a lot of work is needed to construct an ϵ-equilibrium profile in the original game. We hope that one can prove the existence of an equilibrium payoff for an arbitrary number of non-absorbing states by taking our approach, and finding another payoff function for player 2 or other constraints on his strategy space (or both).

References

[1] Amir R. (1987) Sequential Games of Resource Extraction: Existence of Nash Equilibrium, Cowles Foundation D.P. #825

[2] Aumann R.J. and Maschler B.M. (1995) Repeated Games with Incomplete Information, The MIT Press

[3] Benedetti R. and Risler J.J. (1990) Real Algebraic and Semi-Algebraic Sets, Hermann

[4] Bewley T. and Kohlberg E. (1976) The Asymptotic Theory of Stochastic Games, *Math. Oper. Res.*, **1**, 197-208

[5] Blackwell D. and Ferguson T. S. (1968) The Big Match, *The Annals of Math. Stat.*, **39**, 159-163

[6] Bornstein G., Budescu D. and Zamir S. (1996) Cooperation in Intergroup, N-Person and Two-Person Games of Chicken, D.P. #96, Center for Rationality and Interactive Decision Theory, Hebrew University, Jerusalem

[7] Bornstein G., Erev I. and Goren H. (1994) The Effect of Repeated Play in the IPG and IPD Team Games, *J. of Conflict Resolution*, **38**, 690-707

[8] Bornstein G., Goren H. and Winter E. (1996) Experimental Study of Repeated Team-Games, *European J. Political Economy*, **12**, 629-639

[9] Dutta P. and Sundaram R. (1993) How Different can Strategic Models Be?, *J. Econ. Theory*, **60**, 42-61

[10] Everett H. (1957) Recursive Games, Contributions to the theory of Games, **3**, 47-78, Princeton, N.J. Annals of Mathematical Studies, **39**, Princeton University Press.

[11] Filar J.A. (1985) Player Aggregation in the Traveling Inspector Model, *IEEE Trans. Automated Control*, AC-30, 723-729

[12] Fink A.M. (1964) Equilibrium in a Stochastic n-Person Game, *J. Sci. Hiroshima Univ.*,**28**, 89-93

[13] Flesch J., Thuijsman F. and Vrieze K. (1997) Cyclic Markov Equilibria in Stochastic Games, *Int. J. Game Th.*, **26**, 303-314

[14] Forster O. (1981) Lectures on Riemann Surfaces, Graduate Texts in Mathematics, Springer Verlag

[15] Gillette D. (1957) Stochastic Games with Zero Stop Probabilities, *Contributions to the Theory of Games*, **3**, Princeton University Press

[16] Kohlberg E. (1974) Repeated Games with Absorbing States, *Ann. Stat.*, **2**, 724-738

[17] Kohlberg E. and Zamir S. (1974) Repeated Games of Incomplete Information: The Symmetric Case, *Ann. Stat.*, **2**, 1040-1041

[18] Lamperti J.W. (1996) Probability, John Wiley and sons

[19] Levhari D. and Mirman L. (1980) The Great Fish War: An Example Using a Dynamic Cournot-Nash Solution, *Bell J. Econ.*, **11**(1), 322-334

[20] Mertens J.F. and Neyman A. (1981) Stochastic Games, *Int. J. Game Th.*, **10**, 53-66

[21] Mertens J.F., Sorin S. and Zamir S. (1994) Repeated Games, CORE Discussion Paper 9421

[22] Neyman A. (1996) Lecture Notes of a course in “Stochastic Games”, The Hebrew University, Jerusalem.

[23] Neyman A. and Sorin S. (1995) Equilibria in Repeated Games of Incomplete Information: The Deterministic Symmetric Case, D.P. #82, Center for Rationality and Interactive Decision Theory, Hebrew University, Jerusalem

[24] Neyman A. and Sorin S. (1995) Equilibria in Repeated Games of Incomplete Information: The General Symmetric Case, D.P. #86, Center for Rationality and Interactive Decision Theory, Hebrew University, Jerusalem

[25] Rogers P.D. (1969) Non-Zerosum Stochastic Games, Ph.D. Thesis, report ORC 69-8, Operations Research Center, University of California, Berkeley

[26] Shapley L.S. (1953) Stochastic Games, *Proc. Nat. Acad. Sci. U.S.A.*, **39**, 1095-1100

[27] Shubik M. and Whitt W. (1973) Fiat Money in an Economy with one Nondurable Good and No Credit: a Noncooperative Sequential Game. In *Topics in Differential Games*, Blaquiere A. (ed), 401-448

[28] Sobel M.J. (1971) Noncooperative Stochastic Games, *Ann. MAth. Stat.*, **42**, No. 6, 1930-1935

[29] Solan E. and Vieille N. (1998), Quitting Games. *In preparation*

[30] Sorin S. (1986) Asymptotic Properties of a Non-zerosum Stochastic Games, *Int. J. Game Theory*, **15**, 101-107

[31] Thuijsman F. and Raghavan T.E.S. (1997) Perfect Information Stochastic Games and Related Classes, *Int. J. Game Theory*, (forthcoming)

[32] Vieille N. (1993) Solvable States in Stochastic Games, *Int. J. Game Theory*, **21**, 395-404

[33] Vieille N. (1994) On Equilibria in Undiscounted Stochastic Games, *D.P. 9446, CEREMADE*

[34] Vieille N. (1997) Equilibrium in 2-Person Stochastic Games I: A Reduction, *preprint*

[35] Vieille N. (1997) Equilibrium in 2-Person Stochastic Games II: The Case of Recursive Games, *preprint*

[36] Vrieze O.J. and Thuijsman F. (1989) On Equilibria in Repeated Games With Absorbing States, *Int. J. Game Th.*, **18**, 293-310

[37] Walker R.J. (1950) Algebraic Curves, Princeton University Press

[38] Winston W. (1978) A Stochastic Game Model of a Weapons Development Competition, *Siam J. Control and Optimization*, **16**, 411-419