

ON A CLASS OF RECURSIVE GAMES

NICOLAS VIEILLE
HEC School of Management
Jouy-en-Josas, France

1. Introduction

This chapter complements [9], and completes the proof of existence of uniform equilibrium payoffs in two-player non-zero-sum stochastic games with finite state and action sets. It is devoted to the analysis of games that are recursive in the sense of [6], and that have some further properties.

We shall follow the notations in use in earlier chapters. In particular, we let S be the state space, and A and B be the action sets of players 1 and 2 respectively. All three sets S, A and B are finite. Generic elements of S, A and B will be denoted by z, a and b . We let $p(\cdot|z, a, b)$ be the transition function of the game and $r : S \times A \times B \rightarrow \mathbb{R}^2$ be the (stage) payoff function of the game. Generic stationary strategies for the two players will be denoted by α and β . The sets of stationary strategies of the two players are respectively denoted by $\Sigma_s = \Delta(A)^S$ and $T_s = \Delta(B)^S$. The subset of S consisting of absorbing states is denoted by S^* and we set $S_0 = S \setminus S^*$. For each $z \in S^*$, we may assume w.l.o.g., for the purpose of this chapter, that $r(z, \cdot, \cdot)$ is constant, and we write $r(z)$. For each $C \subset S$, $\theta_C := \inf \{n \geq 1 : z_n \notin C\}$ is the first exit time from C .

All games considered here satisfy the following three assumptions:

- **Recursive** For each $z \notin S^*$, $r(z, \cdot, \cdot) = 0$.
- **Positive** For each $z \in S^*$, $r^2(z) > 0$.
- **Absorbing** For every initial state z , and each stationary profile (α, β) such that $\beta_{z'}(b) > 0$ for every $(z', b) \in S \times B$, one has $\mathbf{P}_{z, \alpha, \beta}(\theta_{S_0} < +\infty) = 1$.

Assumptions **Recursive** and **Positive** together ensure that player 2 would rather reach S^* than remain forever within S_0 . Assumption **Absorbing** ensures that S^* is a.s. reached in finite time, provided that player 2 assigns positive probability to each action in each state. Note that, for

each initial state z and each pair (σ, τ) of strategies, one has

$$\gamma(z, \sigma, \tau) := \mathbf{E}_{z, \sigma, \tau} \left[r(z_{\theta_{S_0}}) \mathbf{1}_{\theta_{S_0} < +\infty} \right] = \lim_{n \rightarrow +\infty} \gamma_n(z, \sigma, \tau).$$

We present the proof of the following result.

Theorem 1 *Every finite stochastic game that satisfies **Recursive**, **Positive** and **Absorbing** has a uniform equilibrium payoff.*

Together with the results of [9], Theorem 1 implies the existence of uniform equilibrium payoffs in every finite two-player stochastic game. In [9], it was shown that one could further assume $r^1(z) < 0$ for each $z \in S^*$. It is not clear whether adding this assumption would simplify the proof of Theorem 1.

Let Γ be a stochastic game that satisfies **Recursive**, **Positive** and **Absorbing**. We shall define a family $(\Gamma_\varepsilon)_{\varepsilon > 0}$ of auxiliary games, in which player 2's strategy choice is constrained. For each $\varepsilon > 0$, we define a modified best-reply map in the space of stationary profiles of Γ_ε , with fixed point $(\alpha^\varepsilon, \beta^\varepsilon)$. We will prove that, for each z , $\lim_{\varepsilon \rightarrow 0} \gamma(z, \alpha^\varepsilon, \beta^\varepsilon)$ exists and is a uniform equilibrium payoff of Γ . We use extensively the tools introduced in [5].

Section 2 gives an example of a game with no stationary ε -equilibrium. This contrasts with zero-sum recursive games, where stationary ε -optimal strategies do exist (see [1]). In Section 3, we define the constrained games $(\Gamma_\varepsilon)_{\varepsilon > 0}$, and the modified best-reply map. Section 4 discusses the asymptotics as ε goes to zero, in a non-rigorous way. We limit ourselves to games with two non-absorbing states, and we will add further assumptions. This case contains already most of the features of the general proof, with the benefit of a simple setup.

2. Example

We consider the recursive game Γ described in Figure 1. It is a variant of the example in Flesch et al. [2].

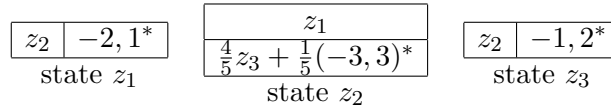


Figure 1

The game has three non-absorbing states, labelled z_1, z_2, z_3 and three absorbing states with respective payoffs $(-2, 1)$, $(-3, 3)$ and $(-1, 2)$. In each non-absorbing state, one of the two players is a dummy, while the other

player may choose between two actions. Since current payoffs are zero until an absorbing state is hit, only the transitions are indicated.

In both states z_1 and z_3 , player 2 chooses one of two columns. In state z_1 (resp. z_3), the **Left** column leads to state z_2 (resp. to z_2), while the **Right** column leads to the absorbing state with payoff $(-2, 1)$ (resp. with payoff $(-1, 2)$). In state z_2 , player 1 has to choose one of two rows. The **Top** row leads to state z_1 , while the **Bottom** row results in a non-deterministic transition: with probability $\frac{4}{5}$, the play moves to state z_3 ; it otherwise moves to the absorbing state with payoff vector $(-3, 3)$. Plainly, the game satisfies both the **Absorbing** and the **Positive** conditions.

Let $\varepsilon \in (0, \frac{1}{5})$ be given. We claim that the game Γ has no stationary ε -equilibrium, in the sense that there is no stationary profile (α, β) that would be an ε -equilibrium of the game with payoff $\gamma(z, \cdot, \cdot)$, for each $z \in S$. Indeed, argue by contradiction and let (α, β) be such a stationary profile. If α assigns positive probability to the **Bottom** row (in state z_2), player 2 may obtain a payoff of 3, whatever the initial state. It must therefore be the case that $\gamma^2(z_2, \alpha, \beta) \geq 3 - \varepsilon$, which implies $\gamma^1(z_2, \alpha, \beta) \leq -3 + \varepsilon$ —a contradiction, since player 1 can guarantee -2 by always choosing the **Top** row.

Assume now that α assigns probability one to the **Top** row. Plainly, it must be that β assigns a positive probability to the **Right** column in state z_1 (otherwise, $\gamma^2(z_1, \alpha, \beta) = 0$). Therefore, $\gamma^1(z_2, \alpha, \beta) = -2$ and $\gamma^2(z_2, \alpha, \beta) = 1$. Starting from z_3 , player 2 may obtain a payoff of 2, using the **Right** column. Thus, it must be that β assigns a probability of at least $1 - \varepsilon$ to the **Right** column in state z_3 . Given any such β , $\gamma^1(z_2, \text{Bottom}, \beta) \geq -\frac{3}{2}$, a contradiction.

In this example, the following is true. For each $z \in S$, the game with payoff function $\gamma(z, \cdot, \cdot)$ has a stationary 0-equilibrium. Whether this always holds is an open problem.

3. Constrained Games

3.1. INTRODUCTION

Given $\varepsilon > 0$, we let

$$T_s(\varepsilon) = \{\beta \in T_s \text{ such that } \beta_z(b) \geq \varepsilon \text{ for every } z \in S_0, b \in B\}$$

denote the set of stationary strategies that assign a probability of at least ε to each action in each state. By the **Absorbing** property, the function $(\alpha, \beta) \mapsto \gamma(z, \alpha, \beta)$ is continuous over $\Sigma_s \times T_s(\varepsilon)$, for each $z \in S$. We let Γ_ε be the game obtained from Γ in which the strategy spaces of the two players are restricted to Σ_s and $T_s(\varepsilon)$ respectively.

It is natural to look for a stationary equilibrium $(\alpha^\varepsilon, \beta^\varepsilon)$ of the game Γ_ε . The existence of such an equilibrium follows by standard arguments. One may then analyze the asymptotic properties of $(\alpha^\varepsilon, \beta^\varepsilon)$, as ε goes to zero. Up to a subsequence, both limits $(\alpha, \beta) := \lim_{\varepsilon \rightarrow 0} (\alpha^\varepsilon, \beta^\varepsilon)$ and $\gamma := \lim_{\varepsilon \rightarrow 0} \gamma(\alpha^\varepsilon, \beta^\varepsilon)$ exist.

This approach was used by Vieille [7] and more generally by Flesch et al. [2] in the analysis of recursive games with at most one non-absorbing state. It was also used by Solan [4] for games that satisfy **Recursive**, **Absorbing**, **Positive**, and with at most two non-absorbing states. Solan managed to construct a (non-stationary) ε -equilibrium of Γ , by perturbing the limit profile (α, β) in an appropriate, history-dependent way. However, the limit payoff γ need not be an equilibrium payoff of Γ and this approach does not extend to larger games.

In the proof sketched below, we adopt a slightly different approach. The profile $(\alpha^\varepsilon, \beta^\varepsilon)$ is not defined to be an arbitrary stationary equilibrium of Γ_ε , but rather to be a fixed point of a suitably modified best-reply map for Γ_ε . The definition of this best-reply map permits a much refined asymptotic analysis.

3.2. THE MODIFIED BEST REPLIES

Choose integers $n_0 = 0, n_1, \dots, n_{|B| \times |S|}$ such that $n_p > |S|(n_{p-1} + 1)$, for each $p \in \{1, \dots, |B| \times |S|\}$. We abbreviate $n_{|B| \times |S|}$ to N . For each $\varepsilon > 0$, we define a set-valued map $\Phi^\varepsilon = \Phi_1 \times \Phi_2^\varepsilon$ on the convex compact set $\Sigma_s \times T_s(\varepsilon^N)$. In a first approximation, Φ^ε may be interpreted as a selection of the best-reply map for the game Γ_{ε^N} . The map Φ_1 depends only on the variable β , while Φ_2^ε depends on both variables.

For later use, observe that for every stationary profile (α, β) and every initial state \bar{z} , the probability $\mathbf{P}_{\bar{z}, \alpha, \beta}(z_{\theta_{S_0}} = z^*)$ of hitting the absorbing state $z^* \in S^*$ is a rational function of the probabilities $\alpha_z(a), \beta_z(b)$, $((z, a, b) \in S_0 \times A \times B)$ assigned to the different actions in the different states. Therefore, $\gamma(\bar{z}, \alpha, \beta)$ is also a rational function of the same variables.

3.2.1. Definition of Φ_1

For $\beta \in T_s(\varepsilon^N)$ and $z \in S$, we let $\gamma_M^1(z, \beta) = \sup_\sigma \gamma^1(z, \sigma, \beta)$ be the best possible payoff for player 1 against β , when starting from z . We define

$$\Phi_1(\beta) := \left\{ \alpha \in \Sigma_s : \mathbf{E} [\gamma_M^1(\cdot, \beta) | z, \alpha_z, \beta_z] = \max_{a \in A} \mathbf{E} [\gamma_M^1(\beta, \cdot) | z, a, \beta_z] \forall z \in S_0 \right\}.$$

Note that $\Phi_1(\beta)$ is a face of the polytope Σ_s of stationary strategies of player 1. It is clear that Φ_1 is upper hemicontinuous, and has non-empty values. Using the **Absorbing** property, it can be shown that $\Phi_1(\beta)$ coincides with the set

$$\{\alpha \in \Sigma_s : \gamma^1(z, \alpha, \beta) = \gamma_M^1(z, \beta) \text{ for every } z \in S\}$$

of stationary best replies to β .

3.2.2. Definition of Φ_2^ε

We now describe Φ_2^ε . Unlike Φ_1 , it depends on ε . Let $\varepsilon > 0$ be given, and let (α, β) be a stationary pair. The definition of $\Phi_2^\varepsilon(\alpha, \beta)$ hinges on a criterion that measures how optimal it is to play once action $b \in B$ in state $z \in S$, if future payoffs are given by $\gamma^2(\alpha, \beta)$. The most obvious such measure is the expectation of future payoffs, given by

$$\mathbf{E} [\gamma^2(\cdot, \alpha, \beta) | z, \alpha_z, b].$$

For reasons that will become evident later, we need to compare actions in different states, which the above criterion fails to do, since it gets intertwined with the comparison of the two payoffs $\gamma^2(z, \alpha, \beta)$ and $\gamma^2(z', \alpha, \beta)$.

To disentangle the two comparisons, we define the *cost* of action b in state z against α as

$$c(b; z, \alpha, \beta) := \max_B \mathbf{E} [\gamma^2(\cdot, \alpha, \beta) | z, \alpha_z, \cdot] - \mathbf{E} [\gamma^2(\cdot, \alpha, \beta) | z, \alpha_z, b].$$

Thus, the cost of b at z is the expected continuation payoff by playing b , *relative* to the highest expected continuation payoff at state z . The following properties clearly hold:

- P.1** For every $z \in S, b \in B, \alpha \in \Sigma_s$ and $\beta \in T_s(\varepsilon^N)$, one has $\min_B c(\cdot; z, \alpha, \beta) = 0$;
- P.2** For fixed $b \in B$ and $z \in S$, the function $(\alpha, \beta) \mapsto c(b; z, \alpha, \beta)$ is semi-algebraic (see [3]).

Given (α, β) , we denote by $C_0(\alpha, \beta), \dots, C_{L(\alpha, \beta)}(\alpha, \beta)$ the level sets for the function $(b, z) \mapsto c(b; z, \alpha, \beta)$, ranked by increasing cost. Note that $C_0(\alpha, \beta)$ is the set of pairs (z, b) such that $c(b; z, \alpha, \beta) = 0$.

Define $p_0 = 0$, and $p_l = \sum_{i=0}^{l-1} |C_i(\alpha, \beta)|$, for $0 < l \leq L(\alpha, \beta)$. Thus, for $(z, b) \in C_l(\alpha, \beta)$, p_l is the number of state-action pairs (z', b') that are strictly better than (z, b) , i.e., such that $c(b'; z', \alpha, \beta) < c(b; z, \alpha, \beta)$.

We define $\Phi_2^\varepsilon(\alpha, \beta)$ as the set of stationary strategies $\tilde{\beta} \in T_s(\varepsilon^N)$ such that for every $l \in \{0, \dots, L(\alpha, \beta)\}$ and $(z, b) \in C_l(\alpha, \beta)$, one has

$$\varepsilon^{n_{p_l+1}-1} \leq \tilde{\beta}_z(b) \leq \varepsilon^{n_{p_l}}. \quad (1)$$

By **P.1**, for every $z \in S$, there is at least one $b \in B$, such that $(z, b) \in C_0(\alpha, \beta)$. It easily follows that $\Phi_2^\varepsilon(\alpha, \beta)$ is non-empty, provided ε is small enough. For the asymptotic analysis below, the truly important consequence of (1) is the following observation: for every $z, z' \in S, b, b' \in B$, and $\tilde{\beta} \in \Phi_2^\varepsilon(\alpha, \beta)$

$$c(b'; z', \alpha, \beta) > c(b; z, \alpha, \beta) \Rightarrow \tilde{\beta}_{z'}(b') \leq \varepsilon \left[\tilde{\beta}_z(b) \right]^{|S|}. \quad (2)$$

The specific form of $\Phi_2^\varepsilon(\alpha, \beta)$ is in other respects somewhat irrelevant.

3.2.3. Existence of a Fixed Point and First Properties

By Kakutani's theorem, the map

$$\Phi^\varepsilon : (\alpha, \beta) \in \Sigma_s \times T_s(\varepsilon^N) \mapsto \Phi_1(\alpha) \times \Phi_2^\varepsilon(\alpha, \beta)$$

has a fixed point in $\Sigma_s \times T_s(\varepsilon^N)$, for ε small enough. Define the fixed point correspondence F as

$$\varepsilon \in (0, 1) \mapsto F(\varepsilon) := \{(\alpha, \beta) \in (0, 1) \times \Sigma_s \times T_s(\varepsilon^N), \alpha \in \Phi_1(\beta), \beta \in \Phi_2^\varepsilon(\alpha, \beta)\}.$$

By **P.2**, and the definitions of Φ_1 and Φ_2^ε , the graph of F is a semialgebraic set. Therefore, the function F has a semialgebraic selection (see [3]): there exists $\varepsilon_0 > 0$ and a semialgebraic map $\varepsilon \in (0, \varepsilon_0) \mapsto (\alpha^\varepsilon, \beta^\varepsilon)$ such that $(\alpha^\varepsilon, \beta^\varepsilon) \in F(\varepsilon)$ for each $\varepsilon \in (0, \varepsilon_0)$. It follows also that, for every $(z, a, b) \in S \times A \times B$, the maps $\varepsilon \in (0, \varepsilon_0) \mapsto \alpha_z^\varepsilon(a)$ and $\varepsilon \in (0, \varepsilon_0) \mapsto \beta_z^\varepsilon(b)$ have expansions in Puiseux series in ε that converge for ε small enough. We may further assume that, for each $z \in S$, the supports of α_z^ε and of β_z^ε are independent of $\varepsilon \in (0, \varepsilon_0)$.

4. Asymptotic Analysis

The purpose of this section is to sketch the asymptotic analysis. Throughout, we let $(\alpha^\varepsilon, \beta^\varepsilon)$ be a semialgebraic profile such that $(\alpha^\varepsilon, \beta^\varepsilon)$ is a fixed point of Φ^ε for each $\varepsilon > 0$ small enough.

We first describe the general idea of the analysis. The complete analysis is involved. Thus, for simplicity, we shall sketch it only in a specific case, which contains most of the features of the general case.

4.1. GENERAL COMMENTS

We shall rely on the tools introduced in [5]. It is proven there that both limits $\gamma = \lim_\varepsilon \gamma(\alpha^\varepsilon, \beta^\varepsilon)$ and $(\alpha^0, \beta^0) = \lim_\varepsilon (\alpha^\varepsilon, \beta^\varepsilon) \in \Sigma_s \times T_s$ do exist. Our chief goal is to show that γ is a uniform equilibrium payoff of Γ .¹

¹I.e., for each $z \in S$, $\gamma(z)$ is a uniform equilibrium payoff of the game, provided the initial state is z .

Recall also from [5] that the family of stationary profiles $(\alpha^\varepsilon, \beta^\varepsilon)_{\varepsilon \in (0, \varepsilon_0)}$ induces a hierarchical decomposition of S_0 into (a forest of) *communicating* sets. This decomposition reflects how the behavior of the Markov chain induced by $(\alpha^\varepsilon, \beta^\varepsilon)$ depends on ε .

A communicating set C is defined by the property that, given any $z, z' \in C$, the probability $\mathbf{P}_{z, \alpha^\varepsilon, \beta^\varepsilon}(\theta_{C \setminus z'} < \theta_C)$ that the play will hit z' before leaving C goes to one as ε goes to zero.²

The leaves of the forest, i.e., the smallest communicating sets, coincide with the subsets of S_0 which are recurrent for the Markov chain induced by (α^0, β^0) . The roots D^1, \dots, D^H are the largest communicating sets.

We let $T := S_0 \setminus (D^1 \dots \cup D^H)$ be the set of states which belong to no communicating set.

For $1 \leq h \leq H$, we denote by $Q^h(\cdot) := \lim_{\varepsilon \rightarrow 0} \mathbf{P}_{z, \alpha^\varepsilon, \beta^\varepsilon}(z_{\theta_{D^h}} = \cdot)$ the (asymptotic) law of the exit state from D^h , as defined by $(\alpha^\varepsilon, \beta^\varepsilon)_{\varepsilon \in (0, \varepsilon_0)}$. The distribution Q^h is independent of $z \in D^h$. The limit payoff $\gamma(\cdot)$ is constant on D^h and equal to $\mathbf{E}_{Q^h}[\gamma(\cdot)]$.

Lemma 2 *Let $\Omega := \{\{D^1\}, \dots, \{D^H\}\} \cup T \cup S^*$. Let \tilde{p} be the transition function over Ω defined by: (i) states in S^* are absorbing, (ii) $\tilde{p}(\cdot|z) := p(\cdot|z, \alpha_z^0, \beta_z^0)$ for $z \in T$, (iii) $\tilde{p}(\cdot|D^h) = Q^h$. Then the Markov chain defined by \tilde{p} is absorbing.*

Proof. There would otherwise be a communicating set included in T , or a communicating set which strictly contains some D^h . In either case, this would contradict the fact that D^1, \dots, D^H are the roots of the forest. ■

The next proposition presents no difficulty. It uses the previous lemma.

Proposition 3 *Let γ be a payoff vector. Assume that: (i) for each $z \in T$, the pair of mixed actions (α_z^0, β_z^0) is a Nash equilibrium of the matrix game with payoff matrix $(\mathbf{E}[\gamma(\cdot)|z, a, b])_{a \in A, b \in B}$; (ii) for each $1 \leq h \leq H$, the distribution Q^h is a controllable exit distribution for γ . Then γ is a uniform equilibrium payoff of Γ .*

Proof. Let us briefly describe a corresponding ε -equilibrium profile (σ, τ) . Whenever the current state belongs to T , the profile (σ, τ) plays (α^0, β^0) , irrespective of past play. Whenever the game enters some set D^h , the players switch to a profile (σ^h, τ^h) associated with the controllable exit distribution Q^h . Finally, the players switch to punishment strategies if the game has not reached S^* by stage N_0 , where N_0 is large enough. ■

We shall prove that both assumptions of Proposition 3 are satisfied. We start with the first one.

²Recall that a communicating set is a weak communicating set w.r.t. (α^0, β^0) . The converse need not hold.

Lemma 4 *For each $z \in S_0$, the pair of mixed actions (α_z^0, β_z^0) is a Nash equilibrium of the matrix game $(\mathbf{E}[\gamma(\cdot)|z, a, b])_{a \in A, b \in B}$.*

Proof. Let $z \in S_0$ be given. We first prove that α_z^0 is a best reply to β_z^0 in that matrix game. By assumption, for each $\varepsilon > 0$ close enough to zero, α_z^ε maximizes $\mathbf{E}[\gamma^1(\cdot, \alpha^\varepsilon, \beta^\varepsilon)|z, \cdot, \beta_z^\varepsilon]$ over $\Delta(A)$. Letting ε go to zero, this implies that α_z^0 maximizes $\mathbf{E}[\gamma^1(\cdot)|z, \cdot, \beta_z^0]$ over $\Delta(A)$.

We next prove that β_z^0 is a best reply to α_z^0 . Let $b \in B$. Since the map $\varepsilon \mapsto c(b; z, \alpha^\varepsilon, \beta^\varepsilon)$ is semialgebraic, it has a constant sign in a neighborhood of zero. If $c(b; z, \alpha^\varepsilon, \beta^\varepsilon) > 0$, for ε small enough, one has $\beta_z^\varepsilon(b) \leq \varepsilon$, by definition of Φ_z^ε . Therefore, $\beta_z^0(b) > 0$ implies $c(b; z, \alpha^\varepsilon, \beta^\varepsilon) = 0$, hence b maximizes $\mathbf{E}[\gamma^2(\cdot, \alpha^\varepsilon, \beta^\varepsilon)|z, \alpha_z^\varepsilon, \cdot]$, for each $\varepsilon > 0$ small enough. Again, the result follows by letting ε go to zero. ■

It is much more difficult to prove that assumption (ii) in Proposition 3 is satisfied. We here describe the main ideas. We let h be given. We abbreviate D^h and Q^h to D and Q .

For expository purposes, we assume that:

- A1** For every $(z, a, b) \in D \times A \times B$, one has $p(D|z, a, b) = 0$ whenever $p(D|z, a, b) < 1$.
- A2** For any two distinct triples $(z_1, a_1, b_1), (z_2, a_2, b_2) \in D \times A \times B$ such that $p(D|z_1, a_1, b_1) = p(D|z_2, a_2, b_2) = 0$, the distributions $p(\cdot|z_1, a_1, b_1)$ and $p(\cdot|z_2, a_2, b_2)$ have disjoint support.

From [5], we know that Q is a convex combination of unilateral exits and joint exits:

$$Q = \sum_{l \in L_1} \mu_l Q_l + \sum_{l \in L_2} \mu_l Q_l + \sum_{l \in L_3} \mu_l Q_l \quad (3)$$

where:

- for $l \in L_1$: $Q_l = p(\cdot|z^l, a^l, \beta_{z^l}^0)$ for some $(z^l, a^l) \in D \times A$;
- for $l \in L_2$: $Q_l = p(\cdot|z^l, \alpha_{z^l}^0, b^l)$ for some $(z^l, b^l) \in D \times B$;
- for $l \in L_3$: $Q_l = p(\cdot|z^l, a^l, b^l)$ for some $(z^l, a^l, b^l) \in D \times A \times B$ such that $p(D|z^l, \alpha_{z^l}^0, b^l) = 1 = p(D|z^l, a^l, \beta_{z^l}^0)$.

We refer to elements of L_1, L_2, L_3 as unilateral exits of player 1, unilateral exits of player 2, and joint exits respectively. W.l.o.g., we assume that $\mu_l > 0$, for every $l \in L_1 \cup L_2 \cup L_3$. By Assumption **A2**, the decomposition (3) is unique. For interpretation, one has $\mu_l = \lim_{\varepsilon \rightarrow 0} \mathbf{P}_{z, \alpha^\varepsilon, \beta^\varepsilon}((z_{\theta_D-1}, a_{\theta_D-1}) = (z^l, a^l))$ for each $l \in L_1$. Similar equalities hold for $l \in L_2 \cup L_3$.

Assumption (2) in Proposition 3 is satisfied if $\mathbf{E}_{Q_l}[\gamma^1(\cdot)] = \gamma^1(z)$ ($z \in D$) for every $l \in L_1$, and if $\mathbf{E}_{Q_l}[\gamma^2(\cdot)] = \gamma^2(z)$ ($z \in D$) for every $l \in L_2$. Indeed, each player is then indifferent between exiting on his own and waiting for some other type of exit to occur.

While it easily follows from the fixed-point property that $\mathbf{E}_{Q_l} [\gamma^1(\cdot)] = \gamma^1(z)$ ($z \in D$) for each $l \in L_1$, there is no reason why $\mathbf{E}_{Q_l} [\gamma^2(\cdot)]$ should be independent of $l \in L_2$. If not, player 2 would favor the unilateral exits $l \in L_2$ for which $\mathbf{E}_{Q_l} [\gamma^2(\cdot)]$ is highest. It is also clear that no statistical test can be designed that would force player 2 to choose his various unilateral exits according to the weights μ_l , $l \in L_2$.

In the next section, we show in a simple setting how the definition of the modified best reply Φ^ε allows us to recover some properties of the quantities $\mathbf{E}_{Q_l} [\gamma^1(\cdot)]$, for $l \in L_2$ (expected exit payoffs of player 1, associated to unilateral exits of player 2). We later sketch how to deal with the general case.

4.2. A SIMPLE SETTING

We will consider a game with two non-absorbing states, labeled \bar{z} and \tilde{z} . We will not define the game completely, but rather will assume that the Puiseux profile $(\alpha^\varepsilon, \beta^\varepsilon)$ has the following properties:

1. The unique maximal communicating set is $D = S_0 = \{\bar{z}, \tilde{z}\}$. In particular, the limit payoff $\gamma(z) = \lim_{\varepsilon \rightarrow 0} \gamma(z, \alpha^\varepsilon, \beta^\varepsilon)$ is independent of the initial state $z \in \{\bar{z}, \tilde{z}\}$;
2. There exist $m, m' \in L_2$, such that $\mathbf{E}_{Q_m} [\gamma^2(\cdot)] < \mathbf{E}_{Q_{m'}} [\gamma^2(\cdot)] < \gamma^2(\bar{z})$.

4.2.1. First Remarks

We first state without proof a few facts, which either follow directly from Lemma 4 or can be derived by a minor modification of the proof:

- For each $l \in L_1$, $\mathbf{E}_{Q_l} [\gamma^1(\cdot)] = \gamma^1(\bar{z}) (= \gamma^1(\tilde{z}))$;
- For each $l \in L_2$, $\mathbf{E}_{Q_l} [\gamma^2(\cdot)] \leq \gamma^2(\bar{z})$;
- Since $\gamma^2(z) = \max_B \mathbf{E} [\gamma^2(\cdot) | z, \alpha_z^0, \cdot]$ for $z \in \{\bar{z}, \tilde{z}\}$, and since $\gamma^2(\bar{z}) = \gamma^2(\tilde{z})$, comparing the (limit) costs of two actions $b, b' \in B$ in the two states $z, z' \in \{\bar{z}, \tilde{z}\}$ amounts to comparing expected continuation payoffs:

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} c(b; z, \alpha^\varepsilon, \beta^\varepsilon) &> \lim_{\varepsilon \rightarrow 0} c(b'; z', \alpha^\varepsilon, \beta^\varepsilon) \\ &\Updownarrow \\ \mathbf{E} [\gamma^2(\cdot) | z, \alpha_z^0, b] &< \mathbf{E} [\gamma^2(\cdot) | z', \alpha_{z'}^0, b'] \end{aligned} \tag{4}$$

Since β^ε is a Puiseux strategy, there exist, for each $(z, b) \in \{\bar{z}, \tilde{z}\} \times B$, numbers $p_z(b) > 0$ and $d_z(b) \geq 0$ such that $\lim_{\varepsilon \rightarrow 0} \frac{\beta_z^\varepsilon(b)}{p_z(b)\varepsilon^{d_z(b)}} = 1$. Similarly, for each $(z, a) \in \{\bar{z}, \tilde{z}\} \times A$ such that $\alpha_z^\varepsilon(a) > 0$ for each $\varepsilon > 0$ small enough, there exist $p_z(a) > 0$, and $d_z(a) \geq 0$ such that $\lim_{\varepsilon \rightarrow 0} \frac{\alpha_z^\varepsilon(a)}{p_z(a)\varepsilon^{d_z(a)}} = 1$. By definition of Φ_1 , $p_z(a) > 0$ only if a maximizes $\mathbf{E} [\gamma^1(\cdot, \alpha^\varepsilon, \beta^\varepsilon) | z, \cdot, \beta_z^\varepsilon]$.

We conclude this section with a crucial observation. From (2) and the definition of Φ_2^ε , it follows that, for any two pairs (z, b) and (z', b') in $\{\bar{z}, \tilde{z}\} \times B$, one has

$$\mathbf{E} [\gamma^2(\cdot) | z, \alpha_z^0, b] < \mathbf{E} [\gamma^2(\cdot) | z', \alpha_{z'}^0, b'] \Rightarrow d_z(b) > 2d_{z'}(b'). \quad (5)$$

4.2.2. $\{\bar{z}, \tilde{z}\}$ -Graphs and Degrees of Transitions

The exit distribution Q from D can be expressed in terms of $\{\bar{z}, \tilde{z}\}$ -graphs (see [5]). We shall here have a closer look. Since $\{\bar{z}, \tilde{z}\}$ is a communicating set, there exists an action pair $(a, b) \in A \times B$ such that $p(\{\bar{z}, \tilde{z}\} | \bar{z}, a, b) = 1$ and $p(\tilde{z} | \bar{z}, a, b) > 0$. Define the degree $\deg(\bar{z} \rightarrow \tilde{z})$ of the transition from \bar{z} to \tilde{z} as the minimum of $d_{\bar{z}}(a) + d_{\tilde{z}}(b)$ over such pairs $(a, b) \in A \times B$. Define the degree $\deg(\tilde{z} \rightarrow \bar{z})$ of the transition from \tilde{z} to \bar{z} similarly.

Given $l \in L_1 \cup L_2 \cup L_3$, we define the degree $\deg(l)$ of the exit labeled l as follows. If $l \in L_3$ with $Q_l = p(\cdot | \tilde{z}, a^l, b^l)$, we set

$$\deg(l) = d(\bar{z} \rightarrow \tilde{z}) + d_{\tilde{z}}(a^l) + d_{\tilde{z}}(b^l).$$

If $l \in L_1$ with $Q_l = p(\cdot | \tilde{z}, a^l, \beta_{\tilde{z}})$, we set

$$\deg(l) = d(\bar{z} \rightarrow \tilde{z}) + d_{\tilde{z}}(a^l).$$

The degree of other types of exits is defined accordingly. The following observation is an immediate consequence of Freidlin and Wentzell's formula (see [5], Lemma 1).

Lemma 5 $\deg(l)$ is independent of $l \in L_1 \cup L_2 \cup L_3$.

4.2.3. Exits of Player 2 and Continuation Payoffs of Player 1

We derive some implications of (5) and of Lemma 5. Let $m, m' \in L_2$ be such that $\mathbf{E} [\gamma^2(\cdot) | z^m, \alpha_{z^m}, b^m] < \mathbf{E} [\gamma^2(\cdot) | z^{m'}, \alpha_{z^{m'}}, b^{m'}]$. By definition of Φ^ε and using (5), one has

$$d_{z_m}(b_m) > d_{z_{m'}}(b_{m'}). \quad (6)$$

Since $\deg(m) = \deg(m')$, it must be that $z_m \neq z_{m'}$. To fix the ideas, we assume $z_m = \bar{z}$, and $z_{m'} = \tilde{z}$. For similar reasons, for each $b \in B$,

$$p(\{\bar{z}, \tilde{z}\} | \bar{z}, \alpha_{\bar{z}}^0, b) < 1 \Rightarrow \mathbf{E} [\gamma^2(\cdot) | \bar{z}, \alpha_{\bar{z}}^0, b] \leq \mathbf{E}_{Q_m} [\gamma^2(\cdot)], \quad (7)$$

and a similar implication holds in state \tilde{z} .

We partition the set L_2 of unilateral exits into $\bar{L}_2 := \{(\bar{z}, b) \in L_2\}$ and $\tilde{L}_2 := L_2 \setminus \bar{L}_2 = \{(\tilde{z}, b) \in L_2\}$. By assumption, both sets \bar{L}_2 and \tilde{L}_2

are nonempty. Using (7), $\mathbf{E}_{Q_l} [\gamma^2(\cdot)] = \mathbf{E}_{Q_m} [\gamma^2(\cdot)]$ for each $l \in \bar{L}_2$ and $\mathbf{E}_{Q_l} [\gamma^2(\cdot)] = \mathbf{E}_{Q_{m'}} [\gamma^2(\cdot)]$ for each $l \in \tilde{L}_2$.

Let \bar{Q}_2 (resp. \tilde{Q}_2) denote the renormalizations of Q over exits in \bar{L}_2 (resp. exits in \tilde{L}_2):

$$\bar{Q}_2 = \frac{\sum_{l \in \bar{L}_2} \mu_l Q_l}{\sum_{l \in \bar{L}_2} \mu_l} \text{ and } \tilde{Q}_2 = \frac{\sum_{l \in \tilde{L}_2} \mu_l Q_l}{\sum_{l \in \tilde{L}_2} \mu_l}.$$

We prove in the next two lemmas that player 1 is indifferent between the two classes \bar{L}_2 and \tilde{L}_2 of unilateral exits.

Lemma 6 $\mathbf{E}_{\bar{Q}_2} [\gamma^1(\cdot)] = \gamma^1(D)$.

Proof. We first argue that player 1, using α^0 , *prevents* the transition from \bar{z} to \tilde{z} : it is impossible for player 2 to leave \bar{z} with positive probability without leaving $\{\bar{z}, \tilde{z}\}$ with positive probability. Indeed, let us proceed by contradiction and assume that, for some $b \in B$, both $p(\{\bar{z}, \tilde{z}\} | \bar{z}, \alpha_{\bar{z}}^0, b) = 1$ and $p(\tilde{z} | \bar{z}, \alpha_{\bar{z}}^0, b) > 0$ hold. In particular,

$$\mathbf{E} [\gamma^2 | \bar{z}, \alpha_{\bar{z}}^0, b] = \gamma^2(\bar{z}) > \mathbf{E} [\gamma^2(\cdot) | \bar{z}, \alpha_{\bar{z}}^0, b_m].$$

Therefore,

$$d_{\bar{z}}(b_m) > 2d_{\bar{z}}(b).$$

Since $d_{\bar{z}}(b_m) > 2d_{\bar{z}}(b_{m'})$, since $\deg(\bar{z} \rightarrow \tilde{z}) \leq d_{\bar{z}}(b)$, and $\deg(\tilde{z} \rightarrow \bar{z}) \geq 0$, this yields

$$\deg(\tilde{z} \rightarrow \bar{z}) + d_{\bar{z}}(b_m) > \deg(\bar{z} \rightarrow \tilde{z}) + d_{\bar{z}}(b_{m'}),$$

a contradiction to $\deg(m) = \deg(m')$.

Next, observe that α^0 is a best reply to β^ε , for each ε small enough. In particular, $\lim_{\varepsilon \rightarrow 0} \gamma^1(\bar{z}, \alpha^0, \beta^\varepsilon) = \gamma^1(\bar{z})$. On the other hand, by the optional stopping theorem, one has $\gamma^1(\bar{z}, \alpha^0, \beta^\varepsilon) = \mathbf{E}_{\bar{z}, \alpha^0, \beta^\varepsilon} [\mathbf{E} [\gamma^1(\cdot) | \bar{z}, \alpha_{\bar{z}}^0, b_\nu]]$, where $\nu := \inf \{n : p(\bar{z} | \bar{z}, \alpha_{\bar{z}}^0, b_n) < 1\}$ is the first stage in which the probability of leaving \bar{z} is strictly positive. Using this remark, it is not difficult to verify that $\lim_{\varepsilon \rightarrow 0} \gamma^1(\bar{z}, \alpha^0, \beta^\varepsilon) = \mathbf{E}_{\bar{Q}_2} [\gamma^1(\cdot)]$. ■

Lemma 7 $\mathbf{E}_{\tilde{Q}_2} [\gamma^1(\cdot)] = \gamma^1(D)$.

Proof. If player 1 prevents the transition from \tilde{z} from \bar{z} , in the sense of (the proof of) Lemma 6, the result follows by the same proof, permuting the roles of \bar{z} and \tilde{z} . However, this need not hold. It may for instance be the case that \tilde{z} is transient under (α^0, β^0) , in which case $\lim_{\varepsilon \rightarrow 0} \gamma^1(\tilde{z}, \alpha^0, \beta^\varepsilon) = \mathbf{E}_{\tilde{Q}_2} [\gamma^1(\cdot)]$. Therefore, in general, the present result cannot follow from a mere adaptation of the proof of Lemma 6.

We let $b^* \in B$ be any action such that both $p(\{\bar{z}, \tilde{z}\}|\tilde{z}, \alpha_{\tilde{z}}^0, b^*) = 1$ and $p(\bar{z}|\tilde{z}, \alpha_{\tilde{z}}^0, b^*) > 0$ hold. Such an action exists if player 1 does not prevent the transition from \tilde{z} to \bar{z} .

We shall infer additional properties by modifying the *degrees* of the actions of player 1. For $z \in \{\bar{z}, \tilde{z}\}$, let $A_z := \{a \in A : \alpha_z^\varepsilon(a) > 0 \text{ for } \varepsilon \in (0, \varepsilon_0)\}$ denote the support of α_z^ε . For each $z \in \{\bar{z}, \tilde{z}\}$ and $a \in A_z$, we let $\bar{d}_z(a) = d_z(a)$ unless $(z, a) = (z^l, a^l)$ for some $l \in L_1 \cup L_3$. In the latter case, we choose $\bar{d}_z(a) > d_z(a)$. Define the stationary strategy $\bar{\alpha}^\varepsilon \in \Sigma_s$ by

$$\bar{\alpha}_z^\varepsilon(a) = \frac{p_z(a) \varepsilon^{\bar{d}_z(a)}}{\sum_{A_z} p_z(a') \varepsilon^{\bar{d}_z(a')}} :$$

the strategy $\bar{\alpha}^\varepsilon$ is obtained from α^ε by modifying the degrees of the different actions. Plainly, $\bar{\alpha}^\varepsilon \in \Phi_1(\beta^\varepsilon)$, hence

$$\lim_{\varepsilon \rightarrow 0} \gamma^1(\bar{\alpha}^\varepsilon, \beta^\varepsilon) = \gamma^1.$$

Plainly also, the limit payoff $\lim_\varepsilon \gamma^1(\bar{\alpha}^\varepsilon, \beta^\varepsilon)$ can be written $\sum_{l \in L_1} \bar{\mu}_l Q_l + \sum_{l \in L_2} \bar{\mu}_l Q_l + \sum_{l \in L_3} \bar{\mu}_l Q_l$ for some nonnegative weights $(\bar{\mu}_l)_l$.³

We argue below that

C.1 $\bar{\mu}_l = 0$, for each $l \in L_1 \cup L_3$.

C.2 $\bar{\mu}_l = \frac{\mu_l}{\sum_{L_2} \mu_k}$, for $l \in L_2$.

For $l \in L$, we write $\deg(l)$ or $\overline{\deg}(l)$ depending on whether it is computed with $(d_z(a))$ or with $(\bar{d}_z(a))$. We adopt the same convention for the degrees of the transitions $\bar{z} \rightarrow \tilde{z}$ and $\tilde{z} \rightarrow \bar{z}$.

Clearly, $\overline{\deg}(l) > \deg(l)$ for each $l \in L_1 \cup L_3$. We shall prove the following claim.

Claim: Assume $a \in A$ satisfies

$$p(\{\bar{z}, \tilde{z}\}|\bar{z}, a, \beta_{\bar{z}}^0) = 1, \quad (8)$$

and

$$p(\tilde{z}|\bar{z}, a, b) > 0 \quad d_{\bar{z}}(a) + d_{\bar{z}}(b) = \deg(\bar{z} \rightarrow \tilde{z}) \text{ for some } b \in B. \quad (9)$$

Then $\bar{d}_{\bar{z}}(a) = d_{\bar{z}}(a)$.

In a sense, this claim states that any action that is involved in the transition from \bar{z} to \tilde{z} is unaffected by the change in the degrees. The same result also holds when permuting the states \bar{z} and \tilde{z} .

³It can be checked that the decomposition of the new exit distribution involves only the exits listed in L .

The above claim (and the symmetric one) implies in particular that

$$\overline{\deg}(\bar{z} \rightarrow \tilde{z}) = \deg(\bar{z} \rightarrow \tilde{z})$$

$$\overline{\deg}(\tilde{z} \rightarrow \bar{z}) = \deg(\tilde{z} \rightarrow \bar{z}).$$

Since the strategy of player 2 has not been changed, this readily implies that $\overline{\deg}(l) = \deg(l)$, for each $l \in L_2$. This yields **C.1**.

The claim also shows that the relative weight of the different exits in L_2 is not changed. Assertion **C2** can then be derived.

We now turn to the proof of the above claim. Let $\bar{a} \in A$ be an action for which (8) and (9) hold, and let $\bar{b} \in B$ be the corresponding action of player 2. By way of contradiction, assume that $\bar{d}_{\bar{z}}(\bar{a}) > d_{\bar{z}}(\bar{a})$. Therefore, $\bar{a} = a^l$, for some $l \in L_1 \cup L_3$. Since $p(\{\bar{z}, \tilde{z}\} | \bar{z}, \bar{a}, \beta_{\bar{z}}^0) = 1$, $l \notin L_1$. Thus, $l \in L_3$: one has $Q_l = p(\cdot | \bar{z}, \bar{a}, b^l)$, for some $b^l \in B$ such that $p(\{\bar{z}, \tilde{z}\} | \bar{z}, \alpha_{\bar{z}}^0, b^l) = 1$.

By Lemma 5, one has $\deg(m') = \deg(l)$, which reads

$$\deg(\bar{z} \rightarrow \tilde{z}) + d_{\bar{z}}(b_{m'}) = \deg(\tilde{z} \rightarrow \bar{z}) + d_{\bar{z}}(\bar{a}) + d_{\bar{z}}(b^l). \quad (10)$$

Recall now that $\deg(\bar{z} \rightarrow \tilde{z}) = d_{\bar{z}}(\bar{a}) + d_{\bar{z}}(\bar{b})$. On the other hand, $\deg(\tilde{z} \rightarrow \bar{z}) \leq d_{\bar{z}}(b^*)$ by the choice of b^* . Substituting in (10) yields

$$d_{\bar{z}}(b_{m'}) \leq d_{\bar{z}}(b^l) + d_{\bar{z}}(b^*) - d_{\bar{z}}(\bar{b}) \leq d_{\bar{z}}(b^l) + d_{\bar{z}}(b^*). \quad (11)$$

Observe now that neither b^l in state \bar{z} , nor b^* in state \tilde{z} , is a unilateral exit of player 2: $p(\{\bar{z}, \tilde{z}\} | \bar{z}, \alpha_{\bar{z}}^0, b^l) = 1 = p(\{\bar{z}, \tilde{z}\} | \tilde{z}, \alpha_{\tilde{z}}^0, b^*)$. Therefore, one has $\mathbf{E}[\gamma^2(\cdot) | \bar{z}, \alpha_{\bar{z}}^0, b^l] = \gamma^2(\bar{z}) = \mathbf{E}[\gamma^2(\cdot) | \tilde{z}, \alpha_{\tilde{z}}^0, b^*]$. By (5), this implies

$$d_{\bar{z}}(b_{m'}) > 2d_{\bar{z}}(b^l) \text{ and } d_{\bar{z}}(b_{m'}) > 2d_{\bar{z}}(b^*),$$

a contradiction to (11). ■

Let us rephrase the above results in a form that is better suited to generalization.

1. The expectation $\mathbf{E}_{Q_l}[\gamma^2(\cdot)]$ may take only two possible values for $l \in L_2$. Denote by L_2^i , $i = 1, 2$ the two level sets of $l \in L_2 \mapsto \mathbf{E}_{Q_l}[\gamma^2(\cdot)]$ and by $Q_{L_2^i} := \frac{\sum_{l \in L_2^i} \mu_l Q_l}{\sum_{l \in L_2^i} \mu_l}$ the renormalization of Q over exits in L_2^i .
2. For each $i = 1, 2$, there is a communicating set $F^i \subseteq \{\bar{z}, \tilde{z}\}$ such that the following conditions are met:
 - (a) exits in L_2^i are available in F^i : for each $l \in L_2^i$, $z^l \in F^i$;
 - (b) by playing α^0 , player 1 prevents transitions from F^i to $\{\bar{z}, \tilde{z}\} \setminus F^i$ (if $F^i = \{\bar{z}, \tilde{z}\}$, this condition is empty);

(c) player 1 is indifferent between $Q_{L_2^1}$ and $Q_{L_2^2}$:

$$\mathbf{E}_{Q_{L_2^1}} [\gamma^1(\cdot)] = \mathbf{E}_{Q_{L_2^2}} [\gamma^1(\cdot)] = \gamma^1(D);$$

(d) the best unilateral exits that are available to player 2 in F^i are those in L_2^i :

$$\mathbf{E} [\gamma^2(\cdot)|z, \alpha_z^0, b] \leq \mathbf{E}_{Q_{L_2^i}} [\gamma^2(\cdot)],$$

for every $(z, b) \in F^i \times B$ such that $p(\{\bar{z}, \tilde{z}\}|z, \alpha_z^0, b) < 1$.

3. For every $l \in L_1$, $\mathbf{E}_{Q_l} [\gamma^1(\cdot)] = \gamma^1(\bar{z})$.

4. Finally, $\mathbf{E} [\gamma^1(\cdot)|z, a, \beta_z^0] \leq \gamma^1(\bar{z})$ and $\mathbf{E} [\gamma^2(\cdot)|z, \alpha_z^0, b] \leq \gamma^2(\bar{z})$, for every $(z, a, b) \in \{\bar{z}, \tilde{z}\} \times A \times B$.

As shown in [5], this implies that Q is controllable w.r.t. γ .

4.3. THE GENERAL CASE

We briefly indicate how the analysis of the simple setting can be generalized. We make no attempt at providing a proof. As above, we let D be a maximal communicating subset, and write the decomposition of the corresponding exit distribution as

$$Q = \sum_{l \in L_1} \mu_l Q_l + \sum_{l \in L_2} \mu_l Q_l + \sum_{l \in L_3} \mu_l Q_l.$$

The main issue is to find a partition (L_2^1, \dots, L_2^H) of L_2 and communicating subsets (F^1, \dots, F^H) of D that satisfy properties **1-4** of the previous section.

For $l \in L_2$, we let $D^1(l) \subset D^2(l) \subset \dots \subset D^M(l)$ be the communicating subsets of D which contain z^l . Denote by $D(l)$ the first one in this sequence that has the property that it is much more difficult to leave $D(l)$ than to reach z_l starting anywhere in $D(l)$ (set $D(l) = D$ if no such set exists). We will not define this property formally. It should be thought of as an extension of property **2.b**.

We now let (L^1, \dots, L^P) be the partition of L_2 into level sets of the map $l \in L_2 \mapsto \mathbf{E}_{Q_l} [\gamma^2(\cdot)]$. For each $1 \leq p \leq P$, define the equivalence relation \mathcal{R}_p on L^p

$$l \mathcal{R}_p l' \Leftrightarrow D(l) = D(l').$$

We finally define the sets L_2^1, \dots, L_2^H to be the equivalence classes of the relations \mathcal{R}_p , $1 \leq p \leq P$. For $1 \leq h \leq H$, we set $D^h = D(l)$, where $l \in L^h$. It can be shown that properties **1-4** of the previous section are satisfied.

References

1. Everett, H. (1957) Recursive games, in M. Dresher, A. W. Tucker and P. Wolfe (eds.), *Contributions to the Theory of Games, Vol. III*, Annals of Mathematics Studies 39, Princeton University Press, Princeton, NJ, pp. 47–78.
2. Flesch, J., Thuijsman, F. and Vrieze, O.J. (1996) Recursive repeated games with absorbing states, *Mathematics of Operations Research* **21**, 1016–1022.
3. Neyman, A. (2003) Real algebraic tools in stochastic games, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Kluwer Academic Publishers, Dordrecht, Chapter 6, pp. 57–75.
4. Solan, E. (2000) Stochastic games with two non-absorbing states, *Israel Journal of Mathematics* **119**, 29–54.
5. Solan, E. (2003) Perturbations of Markov chains with applications to stochastic games, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Kluwer Academic Publishers, Dordrecht, Chapter 17, pp. 265–280.
6. Thuijsman, F. (2003) Recursive games, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Kluwer Academic Publishers, Dordrecht, Chapter 16, pp. 253–264.
7. Vieille, N. (1992) Contributions à la théorie des jeux répétés, Thèse de doctorat, Université Paris VI.
8. Vieille, N. (2000) Two-player stochastic games II: The case of recursive games, *Israel Journal of Mathematics* **119**, 93–126.
9. Vieille, N. (2003) Two-player non-zero-sum games: A reduction, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Kluwer Academic Publishers, Dordrecht, Chapter 18, pp. 283–294.