HAL
open science

Existence of the limit value of two person zero-sum discounted repeated games via comparison theorems

Sylvain Sorin, Guillaume Vigeral

To cite this version:

Sylvain Sorin, Guillaume Vigeral. Existence of the limit value of two person zero-sum discounted repeated games via comparison theorems. Journal of Optimization Theory and Applications, 2013, 157 (2), pp.564-576. 10.1007/s10957-012-0193-4 . hal-00661844

HAL Id: hal-00661844

https://hal.science/hal-00661844v1

Submitted on 20 Jan 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

EXISTENCE OF THE LIMIT VALUE OF TWO PERSON ZERO-SUM DISCOUNTED REPEATED GAMES VIA COMPARISON THEOREMS

SYLVAIN SORIN AND GUILLAUME VIGERAL

Abstract. We give new proofs of existence of the limit of the discounted values for two person zero-sum games in the following frameworks: incomplete information, absorbing, recursive. The idea of these new proofs is to use some comparison criteria.

AMS Classification: 91A15, 91A20, 49J40, 47J20

Keywords: stochastic games, repeated games, incomplete information, asymptotic value, comparison principle, variational inequalities

1. Introduction

The purpose of this article is to present a unified approach to the existence of the limit value for two person zero-sum discounted games. The main tools used in the proofs are

- the fact that the discounted value v_λ satisfies the Shapley equation [1],
- properties of accumulation points of the discounted values, and of the corresponding optimal strategies,
- comparison of two accumulation points leading to uniqueness and characterization.

We apply this program for three well known classes of games, each time covering the case where action spaces are compact.

The first class corresponds to games with incomplete information, where the results were initially obtained in Aumann and Maschler [2] and Mertens and Zamir [3] (including also the asymptotic study of the finitely repeated games). In that case we follow a quite similar approach to Laraki [4]. For absorbing games the results are due to Kohlberg [5] for finitely many actions, later extended in Rosenberg and Sorin [6] for the compact case. An explicit formula for the limit was recently obtained in Laraki [7] and we obtain a related one. The case of recursive games was first handled in Everett [8], with a different notion of limit value involving asymptotic payoff on plays. It was later shown in Sorin [9] that these results implied also the existence of the limit value for two person zero-sum discounted games.

Date: June 2011.
The research of the first author was supported by grant ANR-08-BLAN- 0294-01 (France).
The research of the second author was supported by grant ANR-10-BLAN 0112 (France).

2. Model, notations and basic Lemmas

Let G be a two person zero-sum stochastic game defined by a finite state space Ω, compact metric action spaces I and J for player 1 and 2 (with mixed extensions $X = \Delta(I)$, set of Borel probabilities on I and $Y = \Delta(J)$ respectively), a separately continuous real bounded payoff g on $I \times J \times \Omega$ and a separately continuous transition ρ from $I \times J \times \Omega$ to $\Delta(\Omega)$.

The game is played in discrete time. At stage t, given the state ω_t the payers choose moves $i_t \in I, j_t \in J$, the stage payoff is $g_t = g(i_t, j_t, \omega_t)$ and the new state ω_{t+1} is selected according to $\rho(i_t, j_t, \omega_t)$ and is announced to the players. Given $\lambda \in]0,1]$ the total evaluation in the λ-discounted game is $\sum_{t=1}^{\infty} \lambda(1-\lambda)^{t-1} g_t$.

The Shapley operator $\Phi(\lambda, f)$ [1] is then defined, for $\lambda \in [0,1]$ and f in some closed subset $\mathcal{F}$ of the set of bounded functions from Ω to $I\!R$, by the formula

$$\begin{aligned}\Phi(\lambda, f)(\omega) &= \min_Y \max_X \left\{\lambda g(x,y,\omega) + (1-\lambda)\mathsf{E}_{\rho(x,y,\omega)} f(\cdot)\right\} \\ &= \max_X \min_Y \left\{\lambda g(x,y,\omega) + (1-\lambda)\mathsf{E}_{\rho(x,y,\omega)} f(\cdot)\right\}\end{aligned}$$

where g and ρ are bilinearly extended to $X \times Y$. For $\lambda > 0$, the only fixed point of $\Phi(\lambda, f)$ is the value v_λ of the discounted game.

The sets of optimal actions of each player in the above formula are denoted by $X_\lambda(f)(\omega)$ and $Y_\lambda(f)(\omega)$. Let $\mathbf{X} = X^\Omega$ and similarly $\mathbf{Y} = Y^\Omega$. For simplicity for any $(\mathbf{x}, \mathbf{y}) \in \mathbf{X} \times \mathbf{Y}$ we denote $\rho(\mathbf{x}, \mathbf{y}, \omega) = \rho(\mathbf{x}(\omega), \mathbf{y}(\omega), \omega)$. Also define $\mathbf{X}_\lambda(f) = \Pi_\Omega X_\lambda(f)(\omega)$ and $\mathbf{Y}_\lambda(f) = \Pi_\Omega Y_\lambda(f)(\omega)$.

$\mathcal{S}$ denotes the set of fixed points of the projective operator $\Phi(0, .)$, and $\mathcal{S}_0$ is the set of accumulation points of the family v_λ.

The following lemmas are easy to establish in this finite state framework:

Lemma 1. $\mathcal{S}_0 \subset \mathcal{S}$.

Lemma 2. *Assume that v_{λ_n} converges to $v \in \mathcal{S}_0$ and that a sequence of optimal actions $\mathbf{x}_{\lambda_n} \in \mathbf{X}_{\lambda_n}(v_{\lambda_n})$ converges to $\mathbf{x}$. Then $\mathbf{x} \in \mathbf{X}_0(v)$.*

Lemma 3. *Let v and v' be in $\mathcal{S}$ and $\Omega_1 = \mathrm{Argmax}(v - v')$. For any $x \in \mathbf{X}_0(v)$, $y \in \mathbf{Y}_0(v')$, and $\omega \in \Omega_1$, the probability $\rho(\mathbf{x}, \mathbf{y}, \omega)$ is supported by Ω_1.*

Proof. Since $v \in \mathcal{S}$ and $x \in \mathbf{X}_0(v)$:

$$v(\omega) = \Phi(0, v)(\omega) \leq \mathsf{E}_{\rho(\mathbf{x},\mathbf{y},\omega)} v(\cdot).$$

Using a dual inequality as well:

$$v(\omega) - v'(\omega) \leq \mathsf{E}_{\rho(\mathbf{x},\mathbf{y},\omega)}(v - v')(\cdot)$$

and the result follows. □

3. Games with incomplete information

We consider here two person zero-sum games with incomplete information (independent case and standard signalling). π is a product probability $p \otimes q$ on a finite product space $K \times L$, with $p \in P = \Delta(K)$, $q \in Q = \Delta(L)$. g is a payoff function from $I \times J \times K \times L$ to $I\!R$ where I and J are finite action sets. Given the parameter (k, ℓ) selected according to π, each player knows his own component (k for player 1, ℓ for player 2) and holds a prior on the other player's component. From stage 1 on, the parameter is fixed, the repeated game with payoff $g(.,.,k,\ell)$ is played. The moves of the players at stage t are $\{i_t, j_t\}$, the payoff is $g_t = g(i_t, j_t, k, \ell)$ and the information of the players after stage t is $\{i_t, j_t\}$. $\overline{X} = \Delta(I)^K$ and $\overline{Y} = \Delta(J)^L$ are the type-dependent mixed action sets of the players; g is extended on $\overline{X} \times \overline{Y} \times M$ by $g(x, y, p, q) = \sum_{k,\ell} p^k q^\ell g(x^k, y^\ell, k, \ell)$.

Given (x, y, p, q), let $x(i) = \sum_k x_i^k p^k$ be the probability of action i and $p(i)$ be the conditional probability on K given the action i, explicitly $p^k(i) = \frac{p^k x_i^k}{x(i)}$ (and similarly for y and q).

While this framework is not a particular case of section 2 since the set $P \times Q$ that will play the rôle of the state space is not finite, it is still possible to introduce a Shapley operator for this game. This operator is defined on the set $\mathcal{F}$ of continuous concave-convex fonctions on $P \times Q$ by:

$$\begin{aligned} \Phi(\lambda, f)(p, q) &= \min_{y \in \overline{Y}} \max_{x \in \overline{X}} \left\{ \lambda g(p, q, x, y) + (1 - \lambda) \sum_{i,j} x(i) y(j) f(p(i), q(j)) \right\} \qquad (1) \\ &= \max_{x \in \overline{X}} \min_{y \in \overline{Y}} \left\{ \lambda g(p, q, x, y) + (1 - \lambda) \sum_{i,j} x(i) y(j) f(p(i), q(j)) \right\} \end{aligned}$$

and the value $v_\lambda(p, q)$ of the λ-discounted game is the unique fixed point of $\Phi(\lambda, .)$ on $\mathcal{F}$. These relations are due to Aumann and Maschler (1966) [2] and Mertens and Zamir (1971) [3].

$\overline{X}_\lambda(f)(p, q)$ denotes the set of optimal strategies of player 1 in $\Phi(\lambda, f)(p, q)$.

In this framework any $f \in \mathcal{F}$ is a fixed point of the projective operator $\Phi(0, .)$, that is $\mathcal{F} = \mathcal{S}$.

Notice that if C is a bound for the payoff function g, then any v_λ is bounded by C as well, and is moreover C-Lipschitz. The family $\{v_\lambda\}$ is thus relatively compact for the topology of the uniform convergence, hence $\mathcal{S}_0$, the set of accumulation points of the family $\{v_\lambda\}$, is nonempty.

To ease the notations we will denote the sum $\sum_{i,j} x(i) y(j) f(p(i), q(j))$ by $\mathsf{E}_{\rho(x,p) \times \rho'(y,q)} f(\tilde{p}, \tilde{q})$. Notice that $\tilde{p}$ only depends on x and p, and that $\tilde{q}$ only depends on y and q.

For any $f \in \mathcal{F}$, Jensen's inequality ensures that $\mathsf{E}_{\rho(x,p)} f(\tilde{p}, q) \leq f(p, q)$. The strategies of player 1 for

which the equality holds for all $f \in \mathcal{F}$ are called non revealing and denoted $NR(p) = \{x \in X; \forall f \in \mathcal{F}, \mathsf{E}_{\rho(x,p)} f(\tilde{p}) = f(p)\} = \{x \in X; \tilde{p} = p,\ \rho(x,p)\ a.s.\}$. The set $NR(q)$ of non revealing strategies of player 2 is defined similarly.

Finally, the non revealing value u is

$$u(p,q) \quad = \quad \min_{y \in NR(q)} \max_{x \in NR(p)} g(x,y,p,q) = \max_{x \in NR(p)} \min_{y \in NR(q)} g(x,y,p,q).$$

The existence of $\lim v_\lambda$ was first proved in [2] for games with incomplete information on one side. It was then generalized in [3] for games with incomplete information on both sides, with a characterization of the limit v being the only solution of the system

$$\begin{cases} v = & \mathtt{Cav}_p \min(u,v) \\ v = & \mathtt{Vex}_q \max(u,v) \end{cases}$$

where $\mathtt{Cav}(f)$ (resp. $\mathtt{Vex}$(f)) denotes the smallest concave function in the first variable which is larger than f (resp. the largest convex function in the second variable which is less than f).

A shorter proof of this result (including characterization) was established in [4]. The tools used in the following proof are quite similar to the one used in [4], but the structure differs.

Lemmas 1 and 2 still hold in this framework; we now prove a more precise version of Lemma 3 using the geometry of $P \times Q$. Let $\mathscr{C}(P \times Q)$ be the set of real continuous functions on $P \times Q$.

Lemma 4. *Let $v \in \mathcal{S}$ and let $f \in \mathscr{C}(P \times Q)$ be concave with respect to the first variable. If (p,q) is an extreme point of* $\mathrm{Argmax}(v - f)$ *then: $\overline{X}_0(v)(p,q) \subset NR(p)$.*

Proof. Let $x \in \overline{X}_0(v)(p,q)$ and $y \in NR(q)$, then

$$v(p,q) \leq \mathsf{E}_{\rho(x,p) \times \rho'(y,q)} v(\tilde{p}, \tilde{q}) = \mathsf{E}_{\rho(x,p)} v(\tilde{p}, q)$$

while by Jensen's inequality

$$f(p,q) \geq \mathsf{E}_{\rho(x,p)} f(\tilde{p}, q)$$

so

$$\mathsf{E}_{\rho(x,p)} (v - f)(\tilde{p}, q) \geq (v - f)(p,q)$$

which implies that $x \in NR(p)$ since (p,q) is an extreme point of $\mathrm{Argmax}(v - f)$.

□

Remark that $v \in \mathcal{F}$ implies that $NR(p) \subset \overline{X}_0(v)(p,q)$ since v is a saddle function, hence in fact $NR(p) = \overline{X}_0(v)(p,q)$ in the previous lemma.

Lemma 5. *Let $x_\lambda \in \overline{X}_\lambda(v_\lambda)(p,q)$ and $y \in NR(q)$, then*

$$v_\lambda(p,q) \leq g(x_\lambda, y, p, q).$$

Proof. By definition of v_λ and x_λ,

$$\begin{aligned} v_\lambda(p,q) &\leq \lambda g(x_\lambda, y, p, q) + (1-\lambda)\mathsf{E}_{\rho(x_\lambda,p)\times\rho'(y,q)} v_\lambda(\tilde{p},\tilde{q}) \\ &\leq \lambda g(x_\lambda, y, p, q) + (1-\lambda) v_\lambda(p,q) \end{aligned}$$

using Jensen's inequality and the fact that $y \in NR(q)$. Hence

$$v_\lambda(p,q) \leq g(x_\lambda, y, p, q).$$

□

Recall that $\mathcal{S}_0 \subset \mathcal{S}$ is the set of accumulation points of $\{v_\lambda\}$ for the uniform norm.

Proposition 6. *Let $v \in \mathcal{S}_0$.*
i) Let $f \in \mathscr{C}(P \times Q)$ be concave with respect to the first variable. Then at any extreme point (p,q) of $\mathrm{Argmax}(v-f)$

$$v(p,q) \leq u(p,q).$$

ii) Let $f' \in \mathscr{C}(P \times Q)$ be convex with respect to the second variable. Then at any extreme point (p,q) of $\mathrm{Argmin}(v'-f')$

$$v'(p,q) \geq u(p,q).$$

Proof. We prove *i*). Apply Lemma 5 to any sequence $\{v_{\lambda_n}\}$ converging to v. By Lemma 2, their exists $x \in \overline{X}_0(v)(p,q)$ such that

$$v(p,q) \leq \inf_{y \in NR(q)} g(x, y, p, q).$$

Lemma 4 implies that $x \in NR(p)$ (since $v \in \mathcal{S}$), and the result follows by definition of u.
ii) is established in a dual way. □

Proposition 6 implies the following corollaries of existence and caracterization of $\lim v_\lambda$:

Corollary 7. *v_λ converges uniformly as λ tend to 0.*

Proof. Let v and v' in $\mathcal{S}_0$ and let (p,q) be any extreme point of $\mathrm{Argmax}(v-v')$. Since v' is concave in its first variable, Proposition 6 *i*) with $f = v'(\cdot, q)$ implies that $v(p,q) \leq u(p,q)$. Apply now Proposition 6 *ii*) to $f' = v(p,\cdot)$ to get $v'(p,q) \geq u(p,q)$. This yields $v(p,q) \leq v'(p,q)$, hence $v \leq v'$, thus uniqueness. □

Corollary 8. *Any accumulation point v of v_λ satisfies the Mertens-Zamir system:*

$$\begin{cases} v = & \mathtt{Cav}_p \min(u,v) \\ v = & \mathtt{Vex}_q \max(u,v) \end{cases}$$

Proof. Let v be an accumulation point of the family $\{v_\lambda\}$. It is enough to prove that $v \leq \mathtt{Cav}_p \min(u,v)$ since, v being concave in p, the other inequality is trivial and a dual argument gives the dual equality. Denote $f = \mathtt{Cav}_p \min(u,v)$, and let (p,q) be any extreme point of $\mathrm{Argmax}(v-f)$. Since f is concave in p, Proposition 6 implies that $v(p,q) \leq u(p,q)$. Hence,

$$v(p,q) \leq \min(u,v)(p,q) \leq f(p,q)$$

and thus $v \leq f$. □

Remarks

The proof above also shows that v is the smallest among the functions satisfying $w = \mathtt{Cav}_p \min(u,w)$.

A similar approach applies word for word to the dependent case, as defined in Mertens and Zamir [3]. The case where the action sets I and J are compact metric can also be handled in the same way, using the martingales $(\tilde{p},\tilde{q})$ of regular conditional probabilities.

4. Absorbing games

We consider here a special class of stochastic games as defined in Section 2. We are given two separately continuous (payoff) functions g, g^* from $I \times J$ to $[-1,1]$ and a separately continuous (probability of absorption) function p from $I \times J$ to $[0,1]$.

The repeated game with absorbing states is played in discrete time as follows. At stage $t = 1,2,...$ (if absorption has not yet occurred) player 1 chooses $i_t \in I$ and, simultaneously, player 2 chooses $j_t \in J$:

(i) the payoff at stage t is $g(i_t, j_t)$;

(ii) with probability $p^*(i_t, j_t) := 1 - p(i_t, j_t)$ absorption is reached and the payoff in all future stages $s > t$ is $g^*(i_t, j_t)$;

(iii) with probability $p(i_t, j_t)$ the situation is repeated at stage $t+1$.

Recall that the asymptotic analysis for these games is due to Kohlberg [5] in the case where I and J are finite.

As usual denote $X = \Delta(I)$ and $Y = \Delta(J)$; g, p and p^* are bilinearly extended to $X \times Y$. Let $p^*(x,y)\overline{g}^*(x,y) = \int_{I\times J} p^*(i,j)g^*(i,j)x(di)y(dj)$. $\overline{g}^*(x,y)$ is thus the expected absorbing payoff conditionally to absorption.

The Shapley operator of the game is then defined on $I\!R$ by

$$\begin{aligned}\Phi(\lambda, f) &= \min_{y\in Y}\max_{x\in X}\left\{\lambda g(x,y) + (1-\lambda)(p(x,y)f + p^*(x,y)\overline{g}^*(x,y)\right\}\\ &= \max_{x\in X}\min_{y\in Y}\left\{\lambda g(x,y) + (1-\lambda)(p(x,y)f + p^*(x,y)\overline{g}^*(x,y)\right\}.\end{aligned}$$

The following lemma is analogous to Lemma 4:

Lemma 9.

i) Let $f \in I\!R$ such that $f \geq \Phi(0,f)$ and $y \in Y_0(f)$. Then for any $x \in X$,

$$p^*(x,y) > 0 \implies f \geq \overline{g}^*(x,y).$$

ii) Let $f \in I\!R$ such that $f \leq \Phi(0,f)$ and $x \in X_0(f)$. Then for any $y \in Y$,

$$p^*(x,y) > 0 \implies f \leq \overline{g}^*(x,y).$$

Proof. We prove $i)$. Given $x \in X$ and $y \in Y_0(f)$,

$$\begin{aligned} f &\geq \Phi(0,f)\\ &\geq p(x,y)f + p^*(x,y)\overline{g}^*(x,y)\end{aligned}$$

and $p(x,y) = 1 - p^*(x,y)$, hence the result. □

Given $\lambda \in]0,1[$, $x \in X$ and $y \in Y$, let $r_\lambda(x,y)$ be the induced payoff in the discounted game by the corresponding stationary strategies: $r_\lambda(x,y) = \mathsf{E}_{x,y}\sum \lambda(1-\lambda)^{t-1}g_t$.

Lemma 10.

$$r_\lambda(x,y) \leq \begin{cases} g(x,y) & \text{if } p^*(x,y) = 0\\ \max(g(x,y), \overline{g}^*(x,y)) & \text{if } p^*(x,y) > 0.\end{cases}$$

Proof.

$$r_\lambda(x,y) = \lambda g(x,y) + (1-\lambda)\left[p(x,y)r_\lambda(x,y) + p^*(x,y)\overline{g}^*(x,y)\right]$$

hence

$$r_\lambda(x,y) = \frac{\lambda g(x,y) + (1-\lambda)p^*(x,y)\overline{g}^*(x,y)}{\lambda + (1-\lambda)p^*(x,y)}$$

□

Lemma 5 also has an analogue in this setup:

Lemma 11. *Let $\lambda \in]0,1[$, $x_\lambda \in X_\lambda(v_\lambda)$ and $y \in Y$, then*

$$v_\lambda \leq \begin{cases} g(x_\lambda, y) & \text{if } p^*(x_\lambda, y) = 0 \\ \max(g(x_\lambda, y), \overline{g}^*(x_\lambda, y)) & \text{if } p^*(x_\lambda, y) > 0. \end{cases}$$

Proof. Since x_λ is optimal in the discounted game, for any $y \in Y$,

$$v_\lambda \leq r_\lambda(x_\lambda, y)$$

and the assertion follows from Lemma 10. □

As in the previous part, combining the two lemmas yields:

Proposition 12. *Assume that $v_{\lambda_n} \to v$ and $x_{\lambda_n} \to x$ with $x_{\lambda_n} \in X_{\lambda_n}(v_{\lambda_n})$. Let v' such that $v' \geq \Phi(0, v')$ and $y \in Y_0(v')$, then*

$$v \leq \max(g(x, y), v').$$

Proof. For any n and any $y \in Y$, Lemma 11 implies that either $v_{\lambda_n} \leq g(x_{\lambda_n}, y)$ or that $p^*(x_{\lambda_n}, y) > 0$ and $v_{\lambda_n} \leq \max(g(x_{\lambda_n}, y), \overline{g}^*(x_{\lambda_n}, y))$. In that second case, since $y \in Y_0(v')$, the first assertion in Lemma 9 ensures that $\overline{g}^*(x_{\lambda_n}, y) \leq v'$, so in both cases we have $v_{\lambda_n} \leq \max(g(x_{\lambda_n}, y), v')$. Passing to the limit yields the result. □

Corollary 13. *v_λ converges as λ goes to 0.*

Proof. Suppose on the contrary that there are two sequences $v_{\lambda_n} \to v$ and $v_{\lambda'_n} \to v'$ with $v > v'$. Up to an extraction, one can assume that $x_{\lambda_n} \in X_{\lambda_n}(v_{\lambda_n})$ converges to x and similarly $y_{\lambda'_n} \in Y_{\lambda'_n}(v_{\lambda'_n})$ converges to y. By Lemma 2, $v' = \Phi(0, v')$ and $y \in Y_0(v')$ so applying Proposition 12 we get $v \leq \max(g(x, y), v')$, hence $v \leq g(x, y)$. A dual reasoning yields $v' \geq g(x, y)$, a contradiction. □

We now identify the limit v of the absorbing game.

Definition 14. Define the function $W : X \times Y \to I\!R$ by

$$W(x, y) = \text{med}\left(g(x, y), \sup_{x'; p^*(x', y) > 0} \overline{g}^*(x', y), \inf_{y'; p^*(x, y') > 0} \overline{g}^*(x, y')\right)$$

where $\text{med}(\cdot, \cdot, \cdot)$ denotes the median of three numbers, with the usual convention that a supremum (resp. an infimum) over an empty set equals $-\infty$ (resp. $+\infty$).

Corollary 15. *The limit v is the value of the zero-sum game, denoted by Υ, with action spaces X and Y and payoff W.*

Proof. It is enough to show that $v \leq w := \sup_x \inf_y W(x,y)$ as a dual argument yields the conclusion. Assume by contradiction that $w < v$.

Let $\varepsilon > 0$ with $w + 2\varepsilon < v$. Consider $x \in X_0(v)$ an accumulation point of $x_\lambda \in X_\lambda(v_\lambda)$ and let y be an ε-best response to x in the game Υ. Lemma 9 *ii*) implies that

$$\inf_{y';p^*(x,y')>0} \overline{g}^*(x,y') \geq v > w + \varepsilon \geq W(x,y),$$

so that

$$W(x,y) = \max\left(g(x,y), \sup_{x';p^*(x',y)>0} \overline{g}^*(x',y)\right).$$

Thus, $\sup_{x';p^*(x',y)>0} \overline{g}^*(x',y) \leq w + \varepsilon < v - \varepsilon$ and similarly $g(x,y) < v - \varepsilon$. The corresponding inequalities hold with x_λ, for λ small enough:

$$p^*(x_\lambda, y)[\overline{g}^*(x_\lambda, y) - (v - \varepsilon)] \leq 0, \qquad g(x_\lambda, y) \leq v - \varepsilon$$

leading by Lemma 10 to $v_\lambda \leq v - \varepsilon$, a contradiction.

□

Remarks

The proof of Corollary 15 establishes in itself the existence of the limit v (by doing the same reasoning with any accumulation point of v_λ).

Also notice that this proves that the game Υ has a value, which is not obvious a priori.

5. Recursive games

Recursive games are another special class of stochastic games as defined in Section 2. We are given a finite set $\Omega = \Omega_0 \cup \Omega^*$, two compact metric sets I and J, a payoff function g^* from Ω^* to $\mathbb{R}$ and a separately continuous function ρ from $I \times J \times \Omega_0$ to $\Delta(\Omega)$. Ω^* is the set of absorbing states while Ω_0 is the set of recursive states.

The repeated recursive game is played in discrete time as follows. At stage $t = 1, 2, ...$, if absorbtion has not yet occurred and current state is $\omega_t \in \Omega_0$, player 1 chooses $i_t \in I$ and, simultaneously, player 2 chooses $j_t \in J$:

(i) the payoff at stage t is 0;

(ii) the state ω_{t+1} is chosen with probability distribution $\rho(\omega_{t+1}|i_t, j_t, \omega_t)$;

(iii) if $\omega_{t+1} \in \Omega^*$ absorbtion is reached and the payoff in all future stages $s > t$ is $g^*(\omega_{t+1})$;

(iv) if $\omega_{t+1} \in \Omega_0$ absorbtion is not reached and the game goes on.

The study of those recursive games was first done by Everett [8], who proved that the game has a value when considering the asymptotic payoff on plays.

As before denote $X = \Delta(I)$ and $Y = \Delta(J)$, $\mathbf{X} = X^\Omega$ and similarly $\mathbf{Y} = Y^\Omega$; ρ is bilinearly extended to $\mathbf{X} \times \mathbf{Y}$. Recall that in this framework the Shapley operator is defined from $\mathbb{R}^{|\Omega_0|}$ to itself by

$$\begin{aligned}\Phi(\lambda, f)(\omega) &= \min_{y\in Y}\max_{x\in X}\left\{(1-\lambda)\sum_{\omega'\in\Omega}\rho(\omega'|x,y,\omega)f(\omega')\right\}\\ &= \max_{x\in X}\min_{y\in Y}\left\{(1-\lambda)\sum_{\omega'\in\Omega}\rho(\omega'|x,y,\omega)f(\omega')\right\}\end{aligned}$$

where by convention $f(\omega') = g^*(\omega')$ whenever $\omega' \in \Omega^*$.

Proposition 16. *Let $v \in \mathcal{S}_0$, and v' such that $\max_\Omega v(\omega) - v'(\omega) > 0$. Assume that $v'(\omega) \geq \Phi(0, v')(\omega)$ for all $\omega \in \Omega_1 := \mathrm{Argmax}_\Omega(v - v')$. Then $v(\cdot) \leq 0$ on Ω_1.*

Proof. Denote by Ω_2 the Argmax of v on the set Ω_1, it is enough to prove that $v(\cdot) \leq 0$ on Ω_2, so we assume the contrary. Up to extraction, $v_{\lambda_n} \to v$, $\mathbf{x}_{\lambda_n} \in \mathbf{X}_{\lambda_n}(v_{\lambda_n}) \to \mathbf{x}$ and there exists $\omega_0 \in \Omega_2$ which realizes the maximum of v_{λ_n} on Ω_2 for every n. In particular $v(\omega_0) > 0$. Since $\mathbf{x}_{\lambda_n}$ is optimal, we get for any $\mathbf{y} \in \mathbf{Y}$:

$$v_{\lambda_n}(\omega_0) \leq (1-\lambda_n)\left[\sum_{\omega'\in\Omega_2}\rho(\omega'|\mathbf{x}_{\lambda_n},\mathbf{y},\omega_0)v_{\lambda_n}(\omega') + \sum_{\omega'\in\Omega\setminus\Omega_2}\rho(\omega'|\mathbf{x}_{\lambda_n},\mathbf{y},\omega_0)v_{\lambda_n}(\omega')\right]$$

so by definition of ω_0,

$$(1-(1-\lambda_n)\rho(\Omega_2|\mathbf{x}_{\lambda_n},\mathbf{y},\omega_0))\, v_{\lambda_n}(\omega_0) \leq (1-\lambda_n)\sum_{\omega'\in\Omega\setminus\Omega_2}\rho(\omega'|\mathbf{x}_{\lambda_n},\mathbf{y},\omega_0)v_{\lambda_n}(\omega').$$

For simplicity denote $\rho_n = \rho(\Omega_2|\mathbf{x}_{\lambda_n},\mathbf{y},\omega_0)$. If $\rho_n = 1$ for infinitely many n, we immediately get $v(\omega_0) \leq 0$ and the requested contradiction, hence we assume that it is not the case. Hence up to an extraction, μ_n defined by $\mu_n(w') = \dfrac{\rho(\omega'|\mathbf{x}_{\lambda_n},\mathbf{y},\omega_0)}{1-\rho_n}$ is a probability measure on $\Omega\setminus\Omega_2$. Then for n large enough, we get an analogue of inequalities in Lemmas 5 and 11:

$$\begin{aligned}v_{\lambda_n}(\omega_0) &\leq \frac{1-\lambda_n}{1-(1-\lambda_n)\rho_n}\sum_{\omega'\in\Omega\setminus\Omega_2}\rho(\omega'|\mathbf{x}_{\lambda_n},\mathbf{y},\omega_0)v_{\lambda_n}(\omega')\\ &= \frac{(1-\lambda_n)(1-\rho_n)}{\lambda_n+(1-\lambda_n)(1-\rho_n)}\sum_{\omega'\in\Omega\setminus\Omega_2}\frac{\rho(\omega'|\mathbf{x}_{\lambda_n},\mathbf{y},\omega_0)}{1-\rho_n}v_{\lambda_n}(\omega')\\ &\leq \max\left(0, \sum_{\omega'\in\Omega\setminus\Omega_2}\mu_n(\omega')v_{\lambda_n}(\omega')\right). \qquad (2)\end{aligned}$$

On the other hand, choose now $\mathbf{y} \in \mathbf{Y}_0(v')$. Since $\omega_0 \in \Omega_2$,

$$\begin{aligned}v'(\omega_0) &\geq \Phi(0,v')(\omega_0)\\ &\geq \left[\sum_{\omega'\in\Omega_2}\rho(\omega'|\mathbf{x}_{\lambda_n},\mathbf{y},\omega_0)v'(\omega') + \sum_{\omega'\in\Omega\setminus\Omega_2}\rho(\omega'|\mathbf{x}_{\lambda_n},\mathbf{y},\omega_0)v'(\omega')\right]\end{aligned}$$

so using the fact that v' is constant on Ω_2, we get an analogue to Lemmas 4 and 9:

$$v'(\omega_0) \geq \sum_{\omega' \in \Omega \setminus \Omega_2} \mu_n(\omega') v'(\omega'). \tag{3}$$

Letting n go to infinity in inequalities (2) and (3), and using $v(\omega_0) > 0$, we obtain by compactness the existence of $\mu \in \Delta(\Omega \setminus \Omega_2)$ such that

$$v(\omega_0) \leq \sum_{\omega' \in \Omega \setminus \Omega_2} \mu(\omega') v(\omega') \tag{4}$$

$$v'(\omega_0) \geq \sum_{\omega' \in \Omega \setminus \Omega_2} \mu(\omega') v'(\omega'). \tag{5}$$

Substracting (5) from (4) yields

$$(v - v')(\omega_0) \leq \sum_{\omega' \in \Omega \setminus \Omega_2} \mu(\omega')(v - v')(\omega')$$

and since $\omega_0 \in \Omega_1 = \mathrm{Argmax}_\Omega(v - v')$, this implies that the support of μ is included in Ω_1 and that the inequality in (4) is an equality. This in turn forces the support of μ to be included in $\Omega_2 = \mathrm{Argmax}_{\Omega_1} v$, a contradiction to the construction of μ. □

Corollary 17. *v_λ converges as λ goes to 0.*

Proof. Assume that there are two accumulation points v and v' with $\max_\Omega\{v - v'\} > 0$, and denote $\Omega_1 = \mathrm{Argmax}_\Omega(v - v')$. Then Proposition 16 implies that $v(\cdot) \leq 0$ on Ω_1. A dual argument yields that $v'(\cdot) \geq 0$ on Ω_1, a contradiction. □

We now recover a characterization of the limit due to Everett [8]:

Corollary 18. *$\mathcal{S}_0 \subset \overline{\mathscr{L}^+} \cap \overline{\mathscr{L}^-}$, where $\overline{A}$ is the closure of A and*

$$\mathscr{L}^+ := \left\{ f \in \mathbb{R}^\Omega, \begin{array}{ll} \Phi(0,f)(\omega) \leq f(\omega) & \forall \omega \in \Omega_0 \\ \Phi(0,f)(\omega) = f(\omega) & \implies f(\omega) \geq 0 \\ f(\omega) \geq g^*(\omega) & \forall \omega \in \Omega^* \end{array} \right\} \tag{6}$$

and symmetrically

$$\mathscr{L}^- := \left\{ f \in \mathbb{R}^\Omega, \begin{array}{ll} \Phi(0,f)(\omega) \geq f(\omega) & \forall \omega \in \Omega_0 \\ \Phi(0,f)(\omega) = f(\omega) & \implies f(\omega) \leq 0 \\ f(\omega) \leq g^*(\omega) & \forall \omega \in \Omega^* \end{array} \right\}. \tag{7}$$

We will need the following lemma:

Lemma 19. *For any $\varepsilon \geq 0$, there exist $\Omega' \subset \Omega_0$ and $v' \in \mathscr{F}$ such that the couple (Ω', v') satisfies*

a) $v'(\omega) = g^*(\omega)$ *for all* $\omega \in \Omega^*$.

b) $v'(\omega) = v(\omega) - \varepsilon$ *on* Ω'.

c) $v(\omega) \geq v'(\omega) > v(\omega) - \varepsilon$ *on* $\Omega_0 \backslash \Omega'$.

d) *For any* $\omega \in \Omega_0 \backslash \Omega'$, $\Phi(0, v')(\omega) > v'(\omega)$.

e) *For any* $\omega \in \Omega'$, $\Phi(0, v')(\omega) = v'(\omega)$.

Proof. This was proved in [10] but we recall the proof for the sake of completeness.

Let $\mathscr{E}$ be the set of couples $(\Omega'', v'') \in \mathscr{P}(\Omega_0) \times \mathscr{F}$ that satisfies properties a) to d). This set is nonempty since $(\Omega_0, v - \varepsilon \mathbb{1}_{\omega \in \Omega_0}) \in \mathscr{E}$. Since Ω_0 is finite, we can choose a couple (Ω', v') in $\mathscr{E}$ such that there is no (Ω'', v'') in $\mathscr{E}$ with $\Omega'' \subsetneq \Omega'$. Let $\widetilde{\Omega}$ be the set on which $\Phi(0, v')(\omega) = v'(\omega)$; we now prove that $\widetilde{\Omega} = \Omega'$, hence that (Ω', v') also satisfies property e).

By contradiction assume that $\widetilde{\Omega} \subsetneq \Omega'$ and consider, for small $\alpha > 0$, $v_\alpha := v' + \alpha \mathbb{1}_{\omega \in \Omega' \backslash \widetilde{\Omega}}$. The couple $(\widetilde{\Omega}, v_\alpha)$ clearly satisfies property a) to c) for $\alpha < \varepsilon$. It also satisfies property d) for α small enough by continuity of $\Phi(0, \cdot)$. So for small α the couple $(\widetilde{\Omega}, v_\alpha)$ is in $\mathscr{E}$, contradicting the minimality of Ω'.

□

We can now prove Corollary 18:

Proof of Corollary 18. Let $v \in \mathcal{S}_0$, let $\varepsilon > 0$ and define (v', Ω') as in Lemma 19. By properties a) to c) , $\|v - v'\|_\infty \leq \varepsilon$. If $\Omega' = \emptyset$, then property d) implies that $v' \in \mathscr{L}^-$. If Ω' is nonempty, then by properties b) c) and e) $\Omega' = \mathrm{Argmax}(v - v')$ and $\Phi(0, v')(\cdot) = v'(\cdot)$ on Ω'. Hence, Proposition 16 yields that $v(\cdot) \leq 0$ on Ω'. So $v'(\cdot) \leq 0$ on Ω' and $v' \in \mathscr{L}^-$ as well.

This implies that $v \in \overline{\mathscr{L}^-}$, and by duality $v \in \overline{\mathscr{L}^+}$. □

Remark 20. This corollary implies in itself that v_λ converges, as there is at most one element in the intersection, see [6] and Proposition 9 in [9].

6. Conclusion

In this paper we proved the convergence of the discounted values of two person zero-sum discounted games in several frameworks, using a unified approach relying on a comparaison principle beetween the accumulation points of the discounted values. We also obtained some characterization of the limit. We believe that this approach can be extended in several directions :

- For other classes of stochastic games, where the existence of a limit value is still to establish : stochastic games with a finite number of states and compact sets of actions, absorbing games with incomplete information, and more generally stochastic games with incomplete information.

- For other evaluations of the payoff beyond the discounted case, for example to prove the convergence of the value of the n-stage game when n tends to infinity.

References

[1] Shapley L. S.: Stochastic Games. Proceedings of the National Academy of Sciences of the U.S.A., **39**, 1095-1100 (1953)

[2] Aumann R.J. and M. Maschler: Repeated Games with Incomplete Information, M.I.T. Press (1995)

[3] Mertens J.-F. and S. Zamir: The Value of Two-Person Zero-Sum Repeated Games with Lack of Information on Both Sides. International Journal of Game Theory, **1**, 39-64 (1971)

[4] Laraki R.: Variational Inequalities, System of Functional Equations, and Incomplete Information Repeated Games. SIAM J. Control and Optimization, **40**, 516-524 (2001)

[5] Kohlberg E.: Repeated Games with Absorbing States. Annals of Statistics, **2**, 724-738 (1974)

[6] Rosenberg D. and S. Sorin: An Operator Approach to Zero-Sum Repeated Games. Israel Journal of Mathematics, **121**, 221-246 (2001)

[7] Laraki R.: Explicit Formulas for Repeated Games with Absorbing States. International Journal of Game Theory, **39**, 53-69 (2010)

[8] Everett H.: Recursive Games. Contributions to the Theory of Games, III (H. W. Kuhn and A.W. Tucker, eds.), Annals of Mathematical Studies 39, Princeton University Press 47-78 (1957)

[9] Sorin S.: The Operator Approach to Zero-Sum Stochastic Games. Chapter 27 in A. Neyman and S. Sorin (eds.), Stochastic Games and Applications, Kluwer Academic Publishers (2003)

[10] Vigeral G.: Propriétés Asymptotiques des Jeux Répétés à Somme Nulle. Thèse de doctorat, Université Pierre et Marie Curie (2009)

Sylvain Sorin, Combinatoire et Optimisation, IMJ, CNRS UMR 7586, Faculté de Mathématiques, Université P. et M. Curie - Paris 6, Tour 15-16, 1 étage, 4 Place Jussieu, 75005 Paris and Laboratoire d'Econométrie, Ecole Polytechnique, France

E-mail address: `sorin@math.jussieu.fr`

`http://www.math.jussieu.fr/ sorin/`

Guillaume Vigeral (corresponding author), Université Paris-Dauphine, CEREMADE, Place du Maréchal De Lattre de Tassigny. 75775 Paris cedex 16, France

E-mail address: `vigeral@ceremade.dauphine.fr`

`http://www.ceremade.dauphine.fr/ vigeral/indexenglish.html`