



HAL
open science

Explicit Formulas for Repeated Games with Absorbing States

Rida Laraki

► **To cite this version:**

| Rida Laraki. Explicit Formulas for Repeated Games with Absorbing States. 2009. hal-00362421

HAL Id: hal-00362421

<https://hal.science/hal-00362421>

Preprint submitted on 27 Feb 2009

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.



ÉCOLE POLYTECHNIQUE



CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

**EXPLICIT FORMULAS FOR REPEATED GAMES WITH
ABSORBING STATES**

Rida LARAKI

Février 2009

Cahier n° 2009-06

DEPARTEMENT D'ECONOMIE

Route de Saclay

91128 PALAISEAU CEDEX

(33) 1 69333033

<http://www.enseignement.polytechnique.fr/economie/>

<mailto:chantal.poujouly@polytechnique.edu>

Explicit Formulas for Repeated Games with Absorbing States*

Rida LARAKI†

January 26, 2009

Abstract

Explicit formulas for the asymptotic value $\lim_{\lambda \rightarrow 0} v(\lambda)$ and the asymptotic minmax $\lim_{\lambda \rightarrow 0} w(\lambda)$ of finite λ -discounted absorbing games are provided. New simple proofs for the existence of the limits as λ goes zero are given. Similar characterizations for stationary Nash equilibrium payoffs are obtained. The results may be extended to absorbing games with compact action sets and jointly continuous payoff functions.

1 Introduction

Aumann and Maschler [1] introduced and studied long interaction two player zero-sum repeated games with incomplete information. They introduced the asymptotic and uniform approaches and obtained an explicit formulas for the asymptotic value when a player is fully informed (the famous *Cav(u)* theorem). Mertens and Zamir [8] completed this work and obtained their elegant system of functional equations [8] that characterizes the asymptotic value of a repeated game with incomplete information on both sides. Unfortunately, very few repeated games have an explicit characterization of the asymptotic value .

Stochastic games are repeated games in which a state variable follows a markov chain controlled by the actions of the players. Shapley [10] introduced the two player zero-sum model with finitely many states and actions (*i.e.* the finite model). He proved the existence of the value of the λ -discounted game $v(\lambda)$ by introducing a dynamic programming principle (the Shapley operator).

Kohlberg [5] proved the existence of the asymptotic value $v = \lim_{\lambda \rightarrow 0} v(\lambda)$ in the subclass of finite *absorbing games* (*i.e.*, stochastic games in which only one state is non-absorbing). His operator approach uses the additional information obtained from the derivative of the Shapley operator at $\lambda = 0$ to deduce the existence of $\lim_{\lambda \rightarrow 0} v(\lambda)$ and its characterization via variational inequalities.

Laraki ([6], [7]) used a variational approach to study games in which each player controls a martingale (including Aumann and Maschler repeated games). As in differential games with fixed duration, the approach starts with the dynamic programming principle, follows with an optimal strategy for some player for a fixed discounted factor λ , relax the constraint of the other player and lets the discount factor λ carefully tend to zero. This allows to show the existence of $\lim_{\lambda \rightarrow 0} v(\lambda)$ and its variational characterization.

The same approach gives, for absorbing games, a new proof for the existence of $\lim_{\lambda \rightarrow 0} v(\lambda)$ and its characterization as the value of a one-shot game. When the probability of absorption is controlled by only one player (as in the big match), the formula can be simplified to the value of

*I would like to thank Michel Balinski, Eilon Solan, Sylvain Sorin and Xavier Venel for their very useful comments.

†Main position: CNRS, Laboratoire d'Econométrie, Département d'Economie, Ecole Polytechnique, Palaiseau, France. Associated to: Équipe Combinatoire, Université Paris 6. Email: rida.laraki@polytechnique.edu

an underlying finite game. Coulomb [3] provided another explicit formula for $\lim_{\lambda \rightarrow 0} v(\lambda)$ using an algebraic approach.

Rosenberg and Sorin [9] extended the Kohlberg result to absorbing games when action sets are infinite but compact and payoffs and transitions are separately continuous. However, no explicit formula is known for v .

The minmax $w(\lambda)$ of a multi-player λ -discounted absorbing game is the level at which a team of players could punish another player. From Bewley and Kohlberg [2] one deduces the existence of $\lim w(\lambda)$ for any finite absorbing game. However, no explicit formula exists and it is not known if the limit exists in infinite absorbing games.

The variational approach allows (1) to prove the existence of the asymptotic minmax $w = \lim_{\lambda \rightarrow 0} w(\lambda)$ of any multi-player jointly continuous compact absorbing game and (2) provides an explicit formula for w . Similar formulas could be obtained for stationary Nash equilibrium payoffs.

2 The value

Consider two finite sets I and J , two (payoff) functions f, g from $I \times J$ to $[-1, 1]$ and a (probability transition) function p from $I \times J$ to $[0, 1]$.

The game is played in discrete time. At stage $t = 1, 2, \dots$ player I chooses at random $i_t \in I$ (according to some mixed action $x_t \in \Delta(I)$ ¹) and, simultaneously, player J chooses at random $j_t \in J$ (according to some mixed action $y_t \in \Delta(J)$ ²):

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

(ii) with probability $1 - p(i_t, j_t)$ the game is absorbed and the payoff is $g(i_t, j_t)$ in all future stages;

and

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

If the stream of payoffs is $r(t)$, $t = 1, 2, \dots$, the λ -discounted-payoff of the game is $\sum_{t=1}^{\infty} \lambda(1 - \lambda)^{t-1} r(t)$. Player I maximizes the expected discounted-payoff and player J minimizes that payoff.

$M_+(I) = \{\alpha = (\alpha^i)_{i \in I} : \alpha^i \in [0, +\infty)\}$ is the set of positive measures on I (the I -dimensional positive orthant). For any i and j , let $p^*(i, j) = 1 - p(i, j)$ and $f^*(i, j) = [1 - p(i, j)] \times g(i, j)$. For any $(\alpha, j) \in M_+(I) \times J$ and $\varphi : I \times J \rightarrow [-1, 1]$, φ is extended linearly as follows $\varphi(\alpha, j) = \sum_{i \in I} \alpha^i \varphi(i, j)$. Note that $\Delta(I) \subset M_+(I)$.

Proposition 1 (Shapley 1953) G_λ has a value, $v(\lambda)$. It is the unique real in $[0, 1]$ satisfying,

$$v(\lambda) = \max_{x \in \Delta(I)} \min_{j \in J} [\lambda f(x, j) + (1 - \lambda) p(x, j) v(\lambda) + (1 - \lambda) f^*(x, j)]. \quad (1)$$

Equation (1) implies that player I has an optimal stationary strategy (that plays the same mixed action x at each period). This implies in particular that the proposition holds even if the players have no memory or do not observe past actions.

Example: a quitting-game

| | | |
|---|----|----|
| | C | Q |
| C | 0 | 1* |
| Q | 1* | 0* |

The game stops with probability 1 if one of the players plays Q. There are two absorbing payoffs 1 and 0 (they are marked, as usual, with a *). The absorbing payoff 1 is achieved at some period

¹ $\Delta(I) = \{(x^i)_{i \in I} : x^i \in [0, 1], \sum_{i \in I} x^i = 1\}$ is the set probabilities over I .

² $\Delta(J) = \{(y^j)_{j \in J} : y^j \in [0, 1], \sum_{j \in J} y^j = 1\}$ is the set of probabilities over J .

if (C,Q) or (Q,C) is played. The absorbing payoff 0^* is achieved if (Q,Q) is played. The game is non-absorbed if both players decide to continue and play (C,C).

Consider the following strategy stationary profile in which player I plays at each period $(xC, (1-x)Q)$ and player 2 $(yC, (1-y)Q)$. The corresponding discounted payoff $r_\lambda(x, y)$ satisfies

$$r_\lambda(x, y) = xy(\lambda \times 0 + (1-\lambda)r_\lambda(x, y)) + ((1-x)y + (1-y)x),$$

so that

$$r_\lambda(x, y) = \frac{x + y - 2xy}{1 - xy(1 - \lambda)}.$$

The value $v_\lambda \in [0, 1]$ satisfies:

$$\begin{aligned} v_\lambda &= \text{value} \left(\begin{array}{c|cc} & \text{C} & \text{Q} \\ \hline \text{C} & (1-\lambda)v_\lambda & 1 \\ \hline \text{Q} & 1 & 0 \end{array} \right) \\ &= \max_{x \in [0,1]} \min_{y \in [0,1]} [xy(1-\lambda)v_\lambda + x(1-y) + y(1-x)] \\ &= \min_{y \in [0,1]} \max_{x \in [0,1]} [xy(1-\lambda)v_\lambda + x(1-y) + y(1-x)]. \end{aligned}$$

It may be checked that

$$v_\lambda = x_\lambda = y_\lambda = \frac{1 - \sqrt{\lambda}}{1 - \lambda}.$$

The value and the optimal strategies are not rational fractions of λ (but admit a puiseux series in power of λ). Bewley and Kohlberg [2] show this to hold for all finite stochastic games and deduce from it the existence of $\lim v(\lambda)$.

Lemma 2 $v(\lambda)$ satisfies

$$v(\lambda) = \max_{x \in \Delta(I)} \min_{j \in J} \frac{\lambda f(x, j) + (1-\lambda) f^*(x, j)}{\lambda p(x, j) + p^*(x, j)}.$$

Proof. If in the λ -discounted game, player I plays the stationary strategy x and player J plays a pure stationary strategy $j \in J$, the λ -discounted reward $r(\lambda, x, j)$ satisfies:

$$r(\lambda, x, j) = \lambda f(x, j) + (1-\lambda)p(x, j)r(\lambda, x, j) + (1-\lambda)f^*(x, j).$$

Since $p^* = (1-p)$,

$$r(\lambda, x, j) = \frac{\lambda f(x, j) + (1-\lambda) f^*(x, j)}{\lambda p(x, j) + p^*(x, j)}.$$

The maximizer has a stationary optimal strategy and the minimizer has a pure stationary best reply: this proves the lemma. ■

In the following, $\alpha \perp x$ means that for every $i \in I$, $x^i > 0 \Rightarrow \alpha^i = 0$.

Theorem 3 As λ goes to zero $v(\lambda)$ converges to

$$v = \sup_{x \in \Delta(I)} \sup_{\alpha \perp x \in M_+(I)} \min_{j \in J} \left(\frac{f^*(x, j)}{p^*(x, j)} \mathbf{1}_{\{p^*(x, j) > 0\}} + \frac{f(x, j) + f^*(\alpha, j)}{p(x, j) + p^*(\alpha, j)} \mathbf{1}_{\{p^*(x, j) = 0\}} \right).$$

Proof. Let $w = \lim_{n \rightarrow \infty} v(\lambda_n)$ be an accumulation point of $v(\lambda)$.

Step 1: Consider an optimal stationary strategy $x(\lambda_n)$ for player I and go to the limit using Shapley's dynamic programming principle. From the formula of $v(\lambda_n)$, there exists $x(\lambda_n) \in \Delta(I)$ such that for every $j \in J$,

$$v(\lambda_n) \leq \frac{\lambda_n f(x(\lambda_n), j) + (1-\lambda_n) f^*(x(\lambda_n), j)}{\lambda_n p(x(\lambda_n), j) + p^*(x(\lambda_n), j)}. \quad (2)$$

By the compactness of $\Delta(I)$ it may be supposed that $x(\lambda_n) \rightarrow x$.

Case 1: $p^*(x, j) > 0$. Letting λ_n go to zero implies $w \leq \frac{f^*(x, j)}{p^*(x, j)}$.

Case 2: $p^*(x, j) = \sum_{i \in I} x^i p^*(i, j) = 0$. Thus, $\sum_{i \in S(x)} p^*(i, j) = 0$ where $S(x) = \{i \in I : x^i > 0\}$ is the support of x . Let $\alpha(\lambda_n) = \left(\frac{x^i(\lambda_n)}{\lambda_n} 1_{\{x^i=0\}} \right)_{i \in I} \in M_+(I)$ so that $\alpha(\lambda_n) \perp x$. Consequently,

$$\begin{aligned} \sum_{i \in I} \frac{x^i(\lambda_n)}{\lambda_n} p^*(i, j) &= \sum_{i \notin S(x)} \frac{x^i(\lambda_n)}{\lambda_n} p^*(i, j) \\ &= \sum_{i \in I} \alpha^i(\lambda_n) p^*(i, j) \\ &= p^*(\alpha(\lambda_n), j), \end{aligned}$$

and

$$\sum_{i \in I} \frac{x^i(\lambda_n)}{\lambda_n} f^*(i, j) = \sum_{i \in I} \alpha^i(\lambda_n) f^*(i, j) = f^*(\alpha(\lambda_n), j),$$

so, from equation (2), and because $p(x, j) = 1$,

$$w \leq \liminf_{n \rightarrow \infty} \frac{f(x, j) + (1 - \lambda_n) f^*(\alpha(\lambda_n), j)}{p(x, j) + p^*(\alpha(\lambda_n), j)}. \quad (3)$$

Since J is finite, for any $\varepsilon > 0$, there is $N(\varepsilon)$ such that, for every $j \in J$, $w \leq \frac{f(x, j) + f^*(\alpha(\lambda_{N(\varepsilon)}), j)}{p(x, j) + p^*(\alpha(\lambda_{N(\varepsilon)}), j)} + \varepsilon$.

Consequently, $w \leq v$.

Step 2: Construct a strategy for player I in the λ_n -discounted game that guarantees v as $\lambda_n \rightarrow 0$. Let $(\alpha_\varepsilon, x_\varepsilon) \in M_+(I) \times \Delta(I)$ be ε -optimal for the maximizer in the formula of v . For λ_n small enough, let $x_\varepsilon(\lambda_n)$ be proportional to $x_\varepsilon + \lambda_n \alpha_\varepsilon$ ($x_\varepsilon(\lambda_n) = \mu_n(x_\varepsilon + \lambda_n \alpha_\varepsilon)$ for some $\mu_n > 0$). Let $r(\lambda_n)$ be the unique real in the interval $[0, 1]$ that satisfies,

$$r(\lambda_n) = \min_{j \in J} \left[\frac{\lambda_n [f(x_\varepsilon(\lambda_n), j)] + (1 - \lambda_n) (p(x_\varepsilon(\lambda_n), j)) r(\lambda_n)}{\lambda_n p(x_\varepsilon(\lambda_n), j) + p^*(x_\varepsilon(\lambda_n), j)} \right]. \quad (4)$$

By the linearity of f , p , f^* and p^* on x ,

$$\begin{aligned} r(\lambda_n) &= \min_j \frac{\lambda_n f(x_\varepsilon + \lambda_n \alpha_\varepsilon, j) + (1 - \lambda_n) f^*(x_\varepsilon + \lambda_n \alpha_\varepsilon, j)}{\lambda_n p(x_\varepsilon + \lambda_n \alpha_\varepsilon, j) + p^*(x_\varepsilon + \lambda_n \alpha_\varepsilon, j)} \\ &= \min_j \frac{\lambda_n f(x_\varepsilon, j) + \lambda_n^2 f(\alpha_\varepsilon, j) + (1 - \lambda_n) f^*(x_\varepsilon, j) + (1 - \lambda_n) \lambda_n f^*(\alpha_\varepsilon, j)}{\lambda_n p(x_\varepsilon, j) + \lambda_n^2 p(\alpha_\varepsilon, j) + p^*(x_\varepsilon, j) + \lambda_n p^*(\alpha_\varepsilon, j)}. \end{aligned}$$

Also, $v(\lambda_n) \geq r(\lambda_n)$ since $r(\lambda_n)$ is the payoff of player I if he plays the stationary strategy $x_\varepsilon(\lambda_n)$. Let $j_{\lambda_n} \in J$ be an optimal stationary pure best response for player J against $x_\varepsilon(\lambda_n)$ (an element of the arg min in (4)). Since J is finite and $r(\lambda_n)$ bounded, one can switch to a subsequence and suppose that j_{λ_n} is constant ($= j$) and that $r(\lambda_n) \rightarrow r$. If $p^*(x_\varepsilon, j) > 0$ then $r = \frac{f^*(x_\varepsilon, j)}{p^*(x_\varepsilon, j)}$. If $p^*(x_\varepsilon, j) = 0$, clearly $r = \frac{f(x_\varepsilon, j) + f^*(\alpha_\varepsilon, j)}{p(x_\varepsilon, j) + p^*(\alpha_\varepsilon, j)}$. Consequently, $w \geq v$. ■

This proof shows that for each $\varepsilon > 0$, a player always admits an ε -optimal strategy in the λ -discounted game proportional to $x_\varepsilon + \lambda \alpha_\varepsilon$ for all λ small enough. The quitting game example shows that a 0-optimal strategy of the λ -discounted game is not always of that form. This identifies the asymptotic value as the value of what may be called the asymptotic game.

For any $(\alpha, \beta) \in M_+(I) \times M_+(J)$ and $\varphi : I \times J \rightarrow [-1, 1]$, φ is extended linearly as follows $\varphi(\alpha, \beta) = \sum_{i \in I, j \in J} \alpha^i \beta^j \varphi(i, j)$. For player I let

$$\Lambda(I) = \{(x, \alpha) \in \Delta(I) \times M_+(I) : \alpha \perp x\}$$

and similarly for player J.

Corollary 4 v satisfies the following equations:

$$\begin{aligned}
v &= \sup_{(x,\alpha) \in \Lambda(I)} \inf_{(y,\beta) \in \Lambda(J)} \left(\begin{array}{l} \frac{f^*(x,y)}{p^*(x,y)} \mathbf{1}_{\{p^*(x,y) > 0\}} \\ + \frac{f(x,y) + f^*(\alpha,y) + f^*(x,\beta)}{p(x,y) + p^*(\alpha,y) + p^*(x,\beta)} \mathbf{1}_{\{p^*(x,y) = 0\}} \end{array} \right) \\
&= \inf_{(y,\beta) \in \Lambda(I)} \sup_{(x,\alpha) \in \Lambda(I)} \left(\begin{array}{l} \frac{f^*(x,y)}{p^*(x,y)} \mathbf{1}_{\{p^*(x,y) > 0\}} \\ + \frac{f(x,y) + f^*(\alpha,y) + f^*(x,\beta)}{p(x,y) + p^*(\alpha,y) + p^*(x,\beta)} \mathbf{1}_{\{p^*(x,y) = 0\}} \end{array} \right) \\
&= \sup_{(x,\alpha) \in \Lambda(I)} \inf_{y \in \Delta(J)} \left(\begin{array}{l} \frac{f^*(x,y)}{p^*(x,y)} \mathbf{1}_{\{p^*(x,y) > 0\}} \\ + \frac{f(x,y) + f^*(\alpha,y)}{p(x,y) + p^*(\alpha,y)} \mathbf{1}_{\{p^*(x,y) = 0\}} \end{array} \right).
\end{aligned}$$

Proof. Consider an ε -optimal strategy $x_\varepsilon(\lambda)$ proportional to $x_\varepsilon + \lambda\alpha_\varepsilon$ in the λ -discounted game. Taking any strategy of Player J proportional to $y(\lambda) = y + \lambda\beta$ yields

$$v(\lambda) - \varepsilon \leq \frac{\lambda f(x_\varepsilon + \lambda\alpha_\varepsilon, y + \lambda\beta) + (1 - \lambda) f^*(x_\varepsilon + \lambda\alpha_\varepsilon, y + \lambda\beta)}{\lambda p(x_\varepsilon + \lambda\alpha_\varepsilon, y + \lambda\beta) + p^*(x_\varepsilon + \lambda\alpha_\varepsilon, y + \lambda\beta)}.$$

$p^*(x_\varepsilon, y) > 0$ implies $v = \lim v(\lambda) \leq \frac{f^*(x_\varepsilon, y)}{p^*(x_\varepsilon, y)}$. If $p^*(x_\varepsilon, y) = 0$ then $f^*(x_\varepsilon, y) = 0$. Using the multi-linearity of f , f^* , p and p^* and dividing by λ imply:

$$v(\lambda) - \varepsilon \leq \frac{f(x_\varepsilon + \lambda\alpha_\varepsilon, y + \lambda\beta) + (1 - \lambda) f^*(\alpha_\varepsilon, y) + (1 - \lambda) f^*(x_\varepsilon, \beta) + (1 - \lambda) \lambda f^*(\alpha_\varepsilon, \beta)}{p(x_\varepsilon + \lambda\alpha_\varepsilon, y + \lambda\beta) + p^*(\alpha_\varepsilon, y) + f^*(x_\varepsilon, \beta) + \lambda p^*(\alpha_\varepsilon, \beta)}.$$

Going to the limit,

$$v \leq \frac{f(x_\varepsilon, y) + f^*(\alpha_\varepsilon, y) + f^*(x_\varepsilon, \beta)}{p(x_\varepsilon, y) + p^*(\alpha_\varepsilon, y) + f^*(x_\varepsilon, \beta)},$$

which holds for all (y, β) . Thus,

$$v \leq \sup_{(x,\alpha) \in \Lambda(I)} \inf_{(y,\beta) \in \Lambda(J)} \left(\begin{array}{l} \frac{f^*(x,y)}{p^*(x,y)} \mathbf{1}_{\{p^*(x,y) > 0\}} \\ + \frac{f(x,y) + f^*(\alpha,y) + f^*(x,\beta)}{p(x,y) + p^*(\alpha,y) + p^*(x,\beta)} \mathbf{1}_{\{p^*(x,y) = 0\}} \end{array} \right).$$

And similarly for the other inequality. Since the inf sup is always higher than the sup inf the first two equalities follow.

Taking $\beta = 0$ in the last inequality implies:

$$v \leq \sup_{(x,\alpha) \in \Lambda(I)} \inf_{y \in \Delta(J)} \left(\begin{array}{l} \frac{f^*(x,y)}{p^*(x,y)} \mathbf{1}_{\{p^*(x,y) > 0\}} \\ + \frac{f(x,y) + f^*(\alpha,y)}{p(x,y) + p^*(\alpha,y)} \mathbf{1}_{\{p^*(x,y) = 0\}} \end{array} \right),$$

and from the formula of v in theorem 3, one obtains the last equality of the corollary. ■

3 Absorption controlled by one player

Consider the following zero-sum absorbing game (the big-match).

| | | |
|-----|-------|-------|
| | L | R |
| T | 1^* | 0^* |
| B | 0 | 1 |

It is easy to show that $v(\lambda) = \frac{1}{2}$ and that the unique optimal strategy for player I is to play T with probability $\frac{\lambda}{1+\lambda}$. Consequently, $v = \frac{1}{2}$ which also happens to be the value of the underlying one-shot game $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. On the other hand, the asymptotic value of the quitting-game is 1,

which is not the value of the underlying one-shot game $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. A natural question arises: what are the absorbing games when v is the value of an underlying one-shot game?

A game is partially-controlled by player I if the transition function $p(i, j)$ depends only on i (but not the associated payoffs).

Proposition 5 *If a zero-sum absorbing game is partially-controlled by player I, the asymptotic value equals the value of the underlying game, namely:*

$$v = \max_{x \in \Delta(I)} \min_{j \in J} \sum_{i \notin I^*} x^i f(i, j) + \sum_{i \in I^*} x^i g(i, j)$$

where $I^* = \{i : p^*(i) > 0\}$ is the set of absorbing actions of player I.

Proof. Step 1 $v \leq u$.

Let $x_\varepsilon \in \Delta(I)$ and $\alpha_\varepsilon \perp x_\varepsilon \in M_+(I)$ be ε -optimal in the formula of v . If $p^*(x_\varepsilon) > 0$ then,

$$v - \varepsilon \leq \min_{j \in J} \left(\frac{f^*(x_\varepsilon, j)}{p^*(x_\varepsilon)} \right) = \min_{j \in J} \left(\sum_{i \in I} \frac{x^i p^*(i)}{p^*(x_\varepsilon)} g(i, j) \right) = \min_{j \in J} \left(\sum_{i \in I^*} \frac{x^i p^*(i)}{p^*(x_\varepsilon)} g(i, j) \right) \leq \max_{z \in \Delta(I^*)} \min_{j \in J} \sum_{i \in I^*} z^i g(i, j) \leq u.$$

If $p^*(x_\varepsilon) = 0$ then $x_\varepsilon^i = 0$ for $i \in I^*$ and $p^*(i) = 0$ when $i \notin I^*$ so that

$$\begin{aligned} v - \varepsilon &\leq \min_{j \in J} \left(\frac{f(x_\varepsilon, j) + f^*(\alpha_\varepsilon, j)}{p(x_\varepsilon) + p^*(\alpha_\varepsilon)} \right) \\ &= \min_{j \in J} \left(\sum_{i \notin I^*} \frac{x_\varepsilon^i}{p(x_\varepsilon) + p^*(\alpha_\varepsilon)} f(i, j) + \sum_{i \in I^*} \frac{\alpha_\varepsilon^i p^*(i)}{p(x_\varepsilon) + p^*(\alpha_\varepsilon)} g(i, j) \right). \end{aligned}$$

But when $i \notin I^*$, $p(i) = 1$, thus:

$$\begin{aligned} v - \varepsilon &\leq \min_{j \in J} \left(\sum_{i \notin I^*} \frac{x_\varepsilon^i p(i)}{p(x_\varepsilon) + p^*(\alpha_\varepsilon)} f(i, j) + \sum_{i \in I^*} \frac{\alpha_\varepsilon^i p^*(i)}{p(x_\varepsilon) + p^*(\alpha_\varepsilon)} g(i, j) \right) \\ &\leq u. \end{aligned}$$

Step 2 $v \geq u$.

Let x_0 be optimal for player I in the one shot matrix game associated with u . Define (x_1, α_1) as follows. If $p^*(x_0) = 0$, let $(x_1, \alpha_1) = (x_0, 0)$. This clearly implies that $v \geq u$. If $p^*(x_0) > 0$ then for all $i \in I^*$, let $x_1^i = 0$ (so that x_1 is non-absorbing) and for $i \notin I^*$ let $\alpha_1^i = 0$. This will imply that,

$$\begin{aligned} v &\geq \min_{j \in J} \left(\frac{f(x_1, j) + f^*(\alpha_1, j)}{p(x_1) + p^*(\alpha_1)} \right) \\ &= \min_{j \in J} \left(\sum_{i \notin I^*} \frac{x_1^i}{p(x_1) + p^*(\alpha_1)} f(i, j) + \sum_{i \in I^*} \frac{\alpha_1^i p^*(i)}{p(x_1) + p^*(\alpha_1)} g(i, j) \right) \\ &= \min_{j \in J} \left(\sum_{i \notin I^*} \frac{x_1^i p(i)}{p(x_1) + p^*(\alpha_1)} f(i, j) + \sum_{i \in I^*} \frac{\alpha_1^i p^*(i)}{p(x_1) + p^*(\alpha_1)} g(i, j) \right). \end{aligned}$$

Complete the definition of (x_1, α_1) as follows. For $i \notin I^*$, let $x_0^i = \frac{x_1^i p(i)}{p(x_1) + p^*(\alpha_1)}$ (x_1 is proportional to x_0 on the $I \setminus I^*$) and for $i \in I^*$ let $x_0^i = \frac{\alpha_1^i p^*(i)}{p(x_1) + p^*(\alpha_1)}$ (α_1 is proportional to x_0 on I^*). Consequently,

$$v \geq \min_{j \in J} \left[\sum_{i \notin I} x_0^i g(i, j) + \sum_{i \in I} x_0^i g(i, j) \right] = u.$$

■

4 The minmax

A team of N players (named I) play against player (J). Assume the finiteness of all the strategy sets. Each player k in team I has a finite set of actions I_k . Player J has a finite set of actions J . Let $I = I_1 \times \dots \times I_N$ and f, g from $I \times J \rightarrow [-1, 1]$ and $p : I \times J \rightarrow [0, 1]$. The game is played as above, except that (1) at each period, players in team I randomize independently (they are not allowed to correlate their random moves); and (2) team I minimizes the expected λ -discounted-payoff and player J maximizes the payoff (players in I try to punish player J).

Let $\Delta = \Delta(I_1) \times \dots \times \Delta(I_N)$, $p^*(\cdot) = 1 - p(\cdot)$, $f^*(\cdot) = p^*(\cdot) \times g(\cdot)$ and $M_+ = M_+(I_1) \times \dots \times M_+(I_N)$. For $x \in X$, $j \in J$, $k \in N$ and $\alpha \in M_+$, a function $\varphi : I \times J \rightarrow [-1, 1]$ is extended multi-linearly as follows:

$$\begin{aligned}\varphi(x, j) &= \sum_{i=(i_1, \dots, i_N) \in I} x_1^{i_1} \times \dots \times x_N^{i_N} \varphi(i, j) \\ \varphi(\alpha_k, x_{-k}, j) &= \sum_{i=(i_1, \dots, i_N) \in I} x_1^{i_1} \times \dots \times x_{k-1}^{i_{k-1}} \times \alpha_k^{i_k} \times x_{k+1}^{i_{k+1}} \dots \times x_N^{i_N} \varphi(i, j).\end{aligned}$$

Let $w(\lambda)$ denote the minimum payoff that team I can guarantee against player J. From Bewley and Kohlberg [2] one can deduce the existence of $w(\lambda)$. However, no explicit formula exists.

Theorem 6 $w(\lambda) = \min_{x \in \Delta} \max_{j \in J} \frac{\lambda f(x, j) + (1-\lambda) f^*(x, j)}{\lambda p(x, j) + p^*(x, j)}$ and, as $\lambda \rightarrow 0$, converges to

$$\begin{aligned}w &= \inf_{(x, \alpha) \in \Delta \times M_+ : \forall k, \alpha_k \perp x_k} \max_{j \in J} \left(\begin{array}{l} \frac{f^*(x, j)}{p^*(x, j)} \mathbf{1}_{\{p^*(x, j) > 0\}} \\ + \frac{f(x, j) + \sum_{k=1}^N f^*(\alpha_k, x_{-k}, j)}{p(x, j) + \sum_{k=1}^N p^*(\alpha_k, x_{-k}, j)} \mathbf{1}_{\{p^*(x, j) = 0\}} \end{array} \right) \\ &= \inf_{(x, \alpha) \in \Delta \times M_+ : \forall k, \alpha_k \perp x_k} \max_{y \in \Delta(J)} \left(\begin{array}{l} \frac{f^*(x, y)}{p^*(x, y)} \mathbf{1}_{\{p^*(x, y) > 0\}} \\ + \frac{f(x, y) + \sum_{k=1}^N f^*(\alpha_k, x_{-k}, y)}{p(x, y) + \sum_{k=1}^N p^*(\alpha_k, x_{-k}, y)} \mathbf{1}_{\{p^*(x, y) = 0\}} \end{array} \right).\end{aligned}$$

Proof. For the first formula, follow the ideas in the proof of theorem 3 and corollary 4. Let $v = \lim_{n \rightarrow \infty} w(\lambda_n)$ where $\lambda_n \rightarrow 0$.

Modifications in step 1 in theorem 3: let $x(\lambda_n) \rightarrow x$ be such that for every $j \in J$,

$$w(\lambda_n) \geq \frac{\lambda_n f(x(\lambda_n), j) + (1 - \lambda_n) f^*(x(\lambda_n), j)}{\lambda_n p(x(\lambda_n), j) + p^*(x(\lambda_n), j)}.$$

Let $y(\lambda_n) = x(\lambda_n) - x \rightarrow 0$ so that:

$$\begin{aligned}p^*(x(\lambda_n), j) &= \sum_{i=(i_1, \dots, i_N) \in I} x_1^{i_1}(\lambda_n) \times \dots \times x_N^{i_N}(\lambda_n) p(i, j) \\ &= \sum_{i=(i_1, \dots, i_N) \in I} (y_1^{i_1}(\lambda_n) + x_1^{i_1}) \times \dots \times (y_N^{i_N}(\lambda_n) + x_N^{i_N}) p(i, j) \\ &= p^*(x, j) + \sum_{k=1}^N p^*(y_k(\lambda_n), x_{-k}, j) + o\left(\sum_{k=1}^N p^*(y_k(\lambda_n), x_{-k}, j)\right)\end{aligned}$$

If $p^*(x, j) > 0$ then $w \geq \frac{f^*(x, j)}{p^*(x, j)}$. If $p^*(x, j) = 0$ and if $\alpha_k(\lambda_n) = \left(\frac{x_k^{i_k}(\lambda_n)}{\lambda_n} \mathbf{1}_{\{x_k^{i_k} = 0\}} \right)_{i_k \in I_k} \in M_+(I_k)$

then $\alpha_k(\lambda_n) \perp x_k$ and

$$p^*\left(\frac{x(\lambda_n)}{\lambda_n}, j\right) = \sum_{k=1}^N p^*(\alpha_k(\lambda_n), x_{-k}, j) + o\left(\sum_{k=1}^N p^*(\alpha_k(\lambda_n), x_{-k}, j)\right)$$

and the same is true for f^* so that

$$w \geq \limsup_{n \rightarrow \infty} \frac{f(x, j) + \sum_{k=1}^N f^*(\alpha_k(\lambda_n), x_{-k}, j)}{p(x, j) + \sum_{k=1}^N f^*(\alpha_k(\lambda_n), x_{-k}, j)}$$

which implies that $w \geq v$.

Modifications in step 2 in theorem 3: take $(\alpha^\varepsilon, x^\varepsilon)$ to be ε -optimal for the minimizer in the formula of w and define $x_k^\varepsilon(\lambda_n)$ to be proportional to $x_k^\varepsilon + \lambda_n \alpha_k^\varepsilon$ and deduce that $w \leq v$.

For the second formula, follow corollary 4. For each $\varepsilon > 0$, the proof above implies that players in I have an ε -optimal strategy $(x_k^\varepsilon(\lambda))_{k \in I}$ where $x_k^\varepsilon(\lambda)$ is proportional to $x_k^\varepsilon + \lambda \alpha_k^\varepsilon$ in the λ -discounted game for all λ small enough. This implies that for any $y \in \Delta(J)$,

$$w(\lambda) + \varepsilon \geq \frac{\lambda f(x_k^\varepsilon(\lambda), y) + (1 - \lambda) f^*(x_k^\varepsilon(\lambda), y)}{\lambda p(x_k^\varepsilon(\lambda), y) + p^*(x_k^\varepsilon(\lambda), y)}.$$

where the right hand is a fractional function of λ . Consequently, it admits a limit which may be computed as in step 1 (using the multi-linearity of payoffs and transitions). This will imply that

$$w \geq \inf_{x \in \Delta} \inf_{\alpha \in M_+ : \forall k, \alpha_k \perp x_k} \max_{y \in \Delta(J)} \left(\frac{\frac{f^*(x, y)}{p^*(x, y)} \mathbf{1}_{\{p^*(x, y) > 0\}}}{\frac{f(x, y) + \sum_{k=1}^N f^*(\alpha_k, x_{-k}, y)}{p(x, y) + \sum_{k=1}^N p^*(\alpha_k, x_{-k}, y)} \mathbf{1}_{\{p^*(x, y) = 0\}}} \right).$$

The first formula of w and the fact that $J \subset \Delta(J)$ imply the other inequality. ■

5 Stationary Nash Equilibria

Consider a N player absorbing game where each player $k \in N$ has a finite set of actions I^k . Define the payoff functions $f^k : I \rightarrow [-1, 1]$ and $g^k : I \rightarrow [-1, 1]$, $k \in \{1, \dots, N\}$ and a probability transition $p : I \rightarrow [0, 1]$ where $I = I^1 \times \dots \times I^N$. The game is played as above except that if at stage t player $k = 1, \dots, N$ chooses the action $i_t^k \in I^k$ then player k receives $f^k(i_t^1, \dots, i_t^N)$ and if the game is absorbed he receives $g^k(i_t^1, \dots, i_t^N)$. From Fink [4], the λ -discounted game admits a stationary Nash equilibrium. Notations and calculus as above allow to establish that $x_\lambda \in \Delta$ with the corresponding payoff $u_\lambda = (u_\lambda^1, \dots, u_\lambda^N) \in \mathbf{R}^N$ is a stationary equilibrium iff:

$$\begin{aligned} x_\lambda^k &\in \arg \max_{x^k \in \Delta(I^k)} \frac{\lambda f^k(x^k, x_\lambda^{-k}) + (1 - \lambda) f^{*k}(x^k, x_\lambda^{-k})}{\lambda p(x^k, x_\lambda^{-k}) + p^*(x^k, x_\lambda^{-k})} \\ u_\lambda &= \max_{x^k \in \Delta(I^k)} \frac{\lambda f^k(x^k, x_\lambda^{-k}) + (1 - \lambda) f^{*k}(x^k, x_\lambda^{-k})}{\lambda p(x^k, x_\lambda^{-k}) + p^*(x^k, x_\lambda^{-k})}, \end{aligned}$$

Arguments as above prove the following.

Theorem 7 *Let $u = (u^1, \dots, u^N) \in [-1, 1]^N$ be an accumulation point of u_λ . Then u should be a limit equilibrium payoff of the game where the set of strategies of player k is:*

$$\Lambda(I^k) = \left\{ (x^k, \alpha^k) \in \Delta(I^k) \times M_+(I^k) : \alpha^k \perp x^k \right\}$$

and where the payoff function of player k is

$$\frac{f^{*k}(x)}{p^*(x)} \mathbf{1}_{\{p^*(x) > 0\}} + \frac{f^k(x) + \sum_{j=1}^N f^{*k}(\alpha^j, x^{-j})}{p(x) + \sum_{j=1}^N p^*(\alpha^j, x^{-j})} \mathbf{1}_{\{p^*(x) = 0\}}$$

More precisely, for any $\varepsilon > 0$, there exists $(x_\varepsilon, \alpha_\varepsilon)$ such that:

$$\begin{aligned} u^k &= \lim_{\varepsilon \rightarrow 0} \frac{f^{*k}(x_\varepsilon)}{p^*(x_\varepsilon)} \mathbf{1}_{\{p^*(x_\varepsilon) > 0\}} + \frac{f^k(x_\varepsilon) + \sum_{j=1}^N f^{*k}(\alpha_\varepsilon^j, x_\varepsilon^{-j})}{p(x_\varepsilon) + \sum_{j=1}^N p^*(\alpha_\varepsilon^j, x_\varepsilon^{-j})} \mathbf{1}_{\{p^*(x_\varepsilon) = 0\}} \\ &\geq \sup_{(x^k, \alpha^k) \in \Lambda(I^k)} \left(\frac{\frac{f^{*k}(x^k, x_\varepsilon^{-k})}{p^*(x^k, x_\varepsilon^{-k})} \mathbf{1}_{\{p^*(x^k, x_\varepsilon^{-k}) > 0\}}}{\frac{f^k(x^k, x_\varepsilon^{-k}) + f^{*k}(\alpha^k, x_\varepsilon^{-k}) + \sum_{j \neq k}^N f^{*j}(\alpha_\varepsilon^j, x^k, x_\varepsilon^{-\{k, j\}})}{p(x^k, x_\varepsilon^{-k}) + p^*(\alpha^k, x_\varepsilon^{-k}) + \sum_{j \neq k}^N p^*(\alpha_\varepsilon^j, x^k, x_\varepsilon^{-\{k, j\}})}} \mathbf{1}_{\{p^*(x^k, x_\varepsilon^{-k}) = 0\}} \right) - \varepsilon \end{aligned}$$

6 Compact continuous games

Let us extend the model of zero-sum game. I and J are now assumed to be compact and metric sets. The game is separately (resp. jointly) continuous if f , g and p are separately (resp. jointly) continuous functions on $I \times J$. $\Delta(K)$, $K = I, J$, is the set of Borel probability measures on K and $M_+(K)$ is the set of Borel positive measure on K . They are endowed with the weak* topology. For $(\alpha, \beta) \in M_+(I) \times M_+(J)$ and $\varphi : I \times J \rightarrow [-1, 1]$ measurable, $\varphi(\alpha, \beta) = \int_{I \times J} \varphi(i, j) d\alpha(i) d\beta(j)$.

This framework was introduced in Rosenberg and Sorin [9]. Following the approach of Kohlberg [5], Rosenberg and Sorin considered the Shapley operator $r \rightarrow \Phi(\lambda, r)$ where

$$\begin{aligned} \Phi(\lambda, r) &= \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} [\lambda f(x, y) + (1 - \lambda) p(x, y) r + (1 - \lambda) f^*(x, y)] \\ &= \min_{y \in \Delta(J)} \max_{x \in \Delta(I)} [\lambda f(x, y) + (1 - \lambda) p(x, y) r + (1 - \lambda) f^*(x, y)]. \end{aligned}$$

The operator is well defined and the existence of the value is guaranteed via Sion's minmax theorem. As Shapley did, the operator is $(1 - \lambda)$ -contracting so that the value of the λ -discounted game $v(\lambda)$ is the unique fixed point. Kohlberg [5], in finite absorbing games and Rosenberg and Sorin [9] in separately continuous absorbing games proved the existence of $v = \lim v(\lambda)$ and provided a variational characterization of v using the information obtained from the derivative of $\Phi(\lambda, r)$ around $\lambda \approx 0$. Notations for a multi-player absorbing game are introduced similarly.

Theorem 8 *If the game is jointly continuous, all the results proved above for finite games still hold (for $\lim v(\lambda)$, $\lim w(\lambda)$ and Nash equilibria).*

Proof. Let us show how the first part of theorem 3 is modified. Let $w = \lim_{n \rightarrow \infty} v(\lambda_n)$ where $\lambda_n \rightarrow 0$. Take an optimal strategy $x(\lambda_n)$ of player I in the λ_n -discounted game and suppose w.l.o.g. that it converges to some x . Consider any strategy j of Player J so that:

$$v(\lambda_n) \leq \frac{\lambda_n f(x(\lambda_n), j) + (1 - \lambda_n) f^*(x(\lambda_n), j)}{\lambda_n p(x(\lambda_n), j) + p^*(x(\lambda_n), j)}$$

If $p^*(x, j) > 0$ then $v \leq \frac{f^*(x, j)}{p^*(x, j)}$. If $p^*(x, j) = 0$ then $p^*(i, j) = 0$ on $i \in S(x)$ the support of x . Define $\alpha(\lambda_n) \in M_+(I)$ to be $d\alpha(\lambda_n)(i) = \frac{dx(\lambda_n)(i)}{\lambda_n} \mathbf{1}_{\{i \notin S(x)\}}$. Let $s_n \geq 0$ to be such that $\alpha(\lambda_n) = s_n \sigma(\lambda_n)$ and $\sigma(\lambda_n) \in \Delta(I)$ and assume w.l.o.g. that $\sigma(\lambda_n) \rightarrow \sigma$ and $s_n \rightarrow t \in [0, +\infty]$ (by compactness of $\Delta(I)$). Using joint continuity, the fact that $p(x, j) = 1$ and that payoffs are uniformly bounded by 1 imply that for any $\varepsilon > 0$, there is $N(\varepsilon)$ such that for all $n \geq N(\varepsilon)$ and all $j \in J$

$$\begin{aligned} \frac{f(x(\lambda_n), j) + (1 - \lambda_n) f^*(\alpha(\lambda_n), j)}{p(x(\lambda_n), j) + p^*(\alpha(\lambda_n), j)} &\leq \frac{f(x, j) + \varepsilon + f^*(\alpha(\lambda_n), j) - \lambda_n f^*(\alpha(\lambda_n), j)}{p(x, j) - \varepsilon + p^*(\alpha(\lambda_n), j)} \\ &\leq \frac{f(x, j) + f^*(\alpha(\lambda_n), j)}{p(x, j) + p^*(\alpha(\lambda_n), j)} + \frac{2\varepsilon}{1 - \varepsilon} + \lambda_n \end{aligned}$$

Consequently,

$$w \leq \sup_{x \in \Delta(I)} \sup_{\alpha \perp x \in M_+(I)} \min_{j \in J} \left(\frac{f^*(x, j)}{p^*(x, j)} \mathbf{1}_{\{p^*(x, j) > 0\}} + \frac{f(x, j) + f^*(\alpha, j)}{p(x, j) + p^*(\alpha, j)} \mathbf{1}_{\{p^*(x, j) = 0\}} \right).$$

Step 2 of theorem 1 needs no modification. The other proofs are adapted in a similar way. ■

References

- [1] Aumann R.J. and M. Maschler (1995). *Repeated Games with Incomplete Information*, M.I.T. Press.
- [2] Bewley, T. and Kohlberg E. (1976). The Asymptotic Theory of Stochastic Games. *Mathematics of Operation Research*, **1**, 197-208.
- [3] Coulomb, J. M. (2001). Repeated Games with Absorbing States and Signaling Structure. *Mathematics of Operation Research*, **26**, 286-303.
- [4] Fink, A. M. (1964). Equilibrium in a Stochastic N -Person Game. *J. Sci. Hiroshima Univ.* **28**, 89-93.
- [5] Kohlberg, E. (1974). Repeated Games with Absorbing States. *Annals of Statistics*, **2**, 724-738.
- [6] Laraki, R. (2001a). The Splitting Game and Applications. *International Journal of Game Theory*, **30**, 359-376.
- [7] Laraki, R. (2001b). Variational Inequalities, System of Functional Equations, and Incomplete Information Repeated Games. *SIAM Journal of Control and Optimization*, **40(2)**, 516-524.
- [8] Mertens J.-F. and S. Zamir (1971). The Value of Two-Person Zero-Sum Repeated Games with Lack of Information on Both Sides, *International Journal of Game Theory*, **1**, 39-64.
- [9] Rosenberg, D. and S. Sorin (2001). An Operator Approach to Zero-sum Repeated Games. *Israel Journal of Mathematics*, **121**, 221-246.
- [10] Shapley, L. S. (1953) Stochastic Games, *Proceedings of the National Academy of Sciences of the U. S. A.*, **39**, 1095-1100.