Dynamic Games and Applications (2020) 10:555–572
https://doi.org/10.1007/s13235-019-00333-z

Limit Optimal Trajectories in Zero-Sum Stochastic Games

Sylvain Sorin[1] · **Guillaume Vigeral**[2]

Published online: 28 October 2019

Abstract
We consider zero-sum stochastic games. For every discount factor λ, a time normalization allows to represent the discounted game as being played during the interval [0, 1]. We introduce the trajectories of cumulated expected payoff and of cumulated occupation measure on the state space up to time $t \in [0, 1]$, under ε-optimal strategies. A limit optimal trajectory is defined as an accumulation point as (λ, ε) tend to 0. We study existence, uniqueness and characterization of these limit optimal trajectories for compact absorbing games.

Keywords Zero-sum · Stochastic game · Absorbing game

1 Introduction

The analysis of two-person zero-sum repeated games in discrete time may be performed along two lines: (1) *Asymptotic approach* to each probability distribution θ on the set of stages $(m = 1, 2, \ldots)$ one associates the game G_θ where the evaluation of the stream of stage payoffs $\{g_m\}$ is $\sum_{m=1}^{+\infty} \theta_m g_m$, and one denotes its value by v_θ. Given a preordered family $\{\Theta, \succ\}$ of probability distributions, one studies whether v_θ converges as $\theta \in \Theta$ "goes to ∞" according to $\succ$. Typical examples correspond to n-stage games $(\theta_m = \frac{1}{n} I_{m \le n}, n \to \infty)$, λ-discounted games $(\theta_m = \lambda(1-\lambda)^{m-1}, \lambda \to 0)$, or more generally decreasing evaluations $(\theta_m \ge \theta_{m+1})$ with $\theta_1 \to 0$. The game has an *asymptotic value* v^* when these limits exist

Some of the results of this paper were presented in "Atelier Franco-Chilien: Dynamiques, optimisation et apprentissage" Valparaiso, November 2010, and a preliminary version of this paper was given at the Game Theory Conference in Stony Brook, July 2012. This research was supported by Grant PGMO 0294-01 (France).

✉ Guillaume Vigeral
vigeral@ceremade.dauphine.fr
http://www.ceremade.dauphine.fr/vigeral/indexenglish.html

Sylvain Sorin
sylvain.sorin@imj-prg.fr
https://webusers.imj-prg.fr/sylvain.sorin

[1] Institut de Mathématiques de Jussieu-Paris Rive Gauche, UMR 7586, CNRS, Sorbonne Université, UPMC Paris 06, 75005 Paris, France

[2] CNRS, CEREMADE, Université Paris-Dauphine, PSL Research University, Place du Maréchal De Lattre de Tassigny, 75775 Paris Cedex 16, France

and coincide. (2) *Uniform approach* for each strategy of player 1, one evaluates the amount that can be obtained against any strategy of the opponent in every sufficiently long game for $\{\Theta, \succ\}$. This allows to define a min max and a max min, and the game has a *uniform value* v_∞ when min max = max min.

The second approach is stronger than the first one (existence of v_∞ implies existence of v^* and their equality), but there are games with asymptotic value and no uniform value (repeated games with incomplete information on both sides, Aumann and Maschler [1], Mertens and Zamir [7]; stochastic games with signals on the moves, see, e.g., [11]). The first approach deals only with families of values, while the second explicitly consider strategies. The main difference is that in the first case ε-optimal strategies of the players may depend on the evaluation represented by θ.

We focus here on a class of games where this dependence has a smooth representation and allows for a more precise analysis. Basically in addition to the asymptotic properties of the value, one studies the asymptotic behavior along the play induced by ε-optimal strategies.
In order to do this, we first normalize the duration of the game using the evaluation θ. We thus consider each game G_θ as being played on [0, 1], stage n lasting from time $t_{n-1} = \sum_{m<n} \theta_m$ to time $t_n = t_{n-1} + \theta_n$ (with $t_0 = 0$).
Note that here time t corresponds to the fraction t of the total duration of the game, as evaluated through θ. In particular, given ε-optimal strategies in G_θ the stream of expected stage payoffs generates a bounded measurable trajectory on [0, 1] and one will consider its asymptotic behavior.

The next section introduces the basic definitions and concepts that allow to describe our results. The main proofs are in Sect. 3. Further examples are in Sects. 4 and 5.

To end this quick overview, let us recall that there are games without asymptotic value: stochastic games with compact action spaces, Vigeral [15]; finite stochastic games with signals on the state, Ziliotto [16]; or in a more general framework, Sorin and Vigeral [13].

2 Limit Optimal Trajectories

Let Γ be a two-person zero-sum stochastic game [10] with state space Ω, action spaces I and J, stage payoff g and transition ρ from $\Omega \times I \times J$ to $\mathbb{R}$ (resp. $\Delta(\Omega)$). We assume that Ω is finite, I and J are compact metric sets, g and ρ are continuous functions. We keep the same notations for the multilinear extensions to $X = \Delta(I)$ and $Y = \Delta(J)$, where as usual $\Delta(A)$ denotes the set of probabilities on A.

For any pair of stationary strategies $(\mathbf{x}, \mathbf{y}) \in X^\Omega \times Y^\Omega = \mathbf{X} \times \mathbf{Y}$, any state $\omega \in \Omega$ and any stage n, denote by $c_n^{\omega,\mathbf{x},\mathbf{y}}$ the expected payoff at stage n under these stationary strategies, given the initial state ω, and by $q_n^{\omega,\mathbf{x},\mathbf{y}} \in \Delta(\Omega)$ the corresponding distribution of the state at stage n. Hence, $c_n^{\omega,\mathbf{x},\mathbf{y}} = \langle q_n^{\omega,\mathbf{x},\mathbf{y}}, \bar{g}(\mathbf{x}, \mathbf{y})\rangle$ where $\bar{g}(\mathbf{x}, \mathbf{y})$ stands for the vector payoff with component in state $\zeta \in \Omega$ given by $g(\zeta; \mathbf{x}(\zeta), \mathbf{y}(\zeta))$.

Definition 1 For any $(\mathbf{x}, \mathbf{y}) \in X^\Omega \times Y^\Omega$, any discount factor $\lambda \in (0, 1]$, and any starting state ω, define the function $l_\lambda^{\omega,\mathbf{x},\mathbf{y}} : [0, 1] \to \mathbb{R}$ by

$$l_\lambda^{\omega,\mathbf{x},\mathbf{y}}(t_n) = \lambda \sum_{i=1}^{n} (1-\lambda)^{i-1} c_i^{\omega,\mathbf{x},\mathbf{y}}$$

for $t_n = \lambda \sum_{i=1}^{n} (1-\lambda)^{i-1}$ and a linear interpolation between these dates $\{t_n\}$.

Thus, $l_\lambda^{\omega,\mathbf{x},\mathbf{y}}(t_n)$ corresponds to the expectation of the accumulated payoff for the n first stages, or equivalently up to time t_n, and $l_\lambda^{\omega,\mathbf{x},\mathbf{y}}(t)$ to the same at the fraction t of the game, both under $\mathbf{x}$ and $\mathbf{y}$ in the λ-discounted game starting from ω.

Let $M(\Omega)$ denote the set of positive measures on Ω. We introduce similarly the expected accumulated occupation measure at time t under $\mathbf{x}$ and $\mathbf{y}$ in the λ-discounted game starting from ω as follows:

Definition 2 For any $(\mathbf{x},\mathbf{y}) \in X^\Omega \times Y^\Omega$, any discount factor λ, and any starting state ω, define the function $Q_\lambda^{\omega,\mathbf{x},\mathbf{y}} : [0,1] \to M(\Omega)$ by

$$Q_\lambda^{\omega,\mathbf{x},\mathbf{y}}(t_n) = \lambda \sum_{i=1}^{n} (1-\lambda)^{i-1} q_i^{\omega,\mathbf{x},\mathbf{y}}$$

and by a linear interpolation between these dates $\{t_n\}$.

Note that for any $t \in [0,1]$, $Q_\lambda^{\omega,\mathbf{x},\mathbf{y}}(t) \in t\,\Delta(\Omega)$.

Denote by $l_\lambda^{\mathbf{x},\mathbf{y}}$ and $Q_\lambda^{\mathbf{x},\mathbf{y}}$ the Ω-vectors of functions $l_\lambda^{\omega,\mathbf{x},\mathbf{y}}(\cdot)$ and $Q_\lambda^{\omega,\mathbf{x},\mathbf{y}}(\cdot)$, respectively.

Limit trajectories for the payoff and occupation measures will be defined as accumulation points of $l_\lambda^{\mathbf{x}_\lambda,\mathbf{y}_\lambda}$ and $Q_\lambda^{\mathbf{x}_\lambda,\mathbf{y}_\lambda}$ under ε-optimal strategies ($\varepsilon \geq 0$) $\mathbf{x}_\lambda$ and $\mathbf{y}_\lambda$ in the λ- discounted game as (ε,λ) tend to 0.
More precisely, denote by $\mathbf{X}_\lambda^\varepsilon$ (resp. $\mathbf{Y}_\lambda^\varepsilon$) the set of ε-optimal stationary strategies in the λ-discounted game Γ_λ (with value v_λ) for player 1 (resp. for player 2).
Then we introduce:

Definition 3 $l = (l^\omega : [0,1] \to \mathbb{R})_{\omega\in\Omega}$ is a limit optimal trajectory for the expected accumulated payoff (LOTP) if :

$$\forall\varepsilon > 0,\ \exists\lambda_0 > 0,\ \forall\lambda < \lambda_0,\ \exists\mathbf{x}_\lambda \in X_\lambda^\varepsilon,\ \exists\mathbf{y}_\lambda \in Y_\lambda^\varepsilon,$$
$$\forall\omega \in \Omega,\ \forall t \in [0,1],\ \left|l^\omega(t) - l_\lambda^{\omega,\mathbf{x}_\lambda,\mathbf{y}_\lambda}(t)\right| \leq \varepsilon.$$

$Q = (Q^\omega : [0,1] \to M(\Omega))_{\omega\in\Omega}$ is a limit optimal trajectory for the expected accumulated occupation measure (LOTM) if :

$$\forall\varepsilon > 0,\ \exists\lambda_0 > 0,\ \forall\lambda < \lambda_0,\ \exists\mathbf{x}_\lambda \in X_\lambda^\varepsilon,\ \exists\mathbf{y}_\lambda \in Y_\lambda^\varepsilon,$$
$$\forall\omega \in \Omega,\ \forall t \in [0,1],\ \left\|Q^\omega(t) - Q_\lambda^{\omega,\mathbf{x}_\lambda,\mathbf{y}_\lambda}(t)\right\| \leq \varepsilon.$$

Alternative weaker and stronger definitions are as follows:

Definition 4 $l = (l^\omega : [0,1] \to \mathbb{R})_{\omega\in\Omega}$ is a weak limit optimal trajectory for the expected accumulated payoff (WLOTP) if :

$$\forall\varepsilon > 0,\ \exists\lambda \leq \varepsilon,\ \exists\mathbf{x}_\lambda \in X_\lambda^\varepsilon,\ \exists\mathbf{y}_\lambda \in Y_\lambda^\varepsilon,\ \forall\omega \in \Omega,\ \forall t \in [0,1],\ \left|l^\omega(t) - l_\lambda^{\omega,\mathbf{x}_\lambda,\mathbf{y}_\lambda}(t)\right| \leq \varepsilon.$$

$l = (l^\omega : [0,1] \to \mathbb{R})_{\omega\in\Omega}$ is a strong limit optimal trajectory for the expected accumulated payoff (SLOTP) if :

$$\forall\varepsilon > 0,\ \exists\lambda_0 > 0,\ \exists\varepsilon_0 > 0,\ \forall\lambda < \lambda_0,\ \forall\mathbf{x}_\lambda \in X_\lambda^{\varepsilon_0},$$
$$\forall\mathbf{y}_\lambda \in Y_\lambda^{\varepsilon_0},\ \forall\omega \in \Omega,\ \forall t \in [0,1],\ \left|l^\omega(t) - l_\lambda^{\omega,\mathbf{x}_\lambda,\mathbf{y}_\lambda}(t)\right| \leq \varepsilon.$$

and similar definitions for weak (resp. strong) limit optimal trajectory for the expected accumulated occupation measure, (WLOTM) (resp. (SLOTM)).

Remark 5 Thus, LOTP (resp. LOTM) basically are limit points of the functions $t \to l_\lambda^\omega(t)$ (resp. $t \to Q_\lambda^\omega(t)$) as $lambda$ goes to 0 and players play almost optimally in the λ-discounted game. The weak definitions are only accumulation points of the same trajectories, while the strong ones also ask for some robustness with respect to the almost optimal strategies. It is readily checked that

- WLOTP (and WLOTM) always exists by standard arguments of equicontinuity.
- If a LOTP l exists, v_λ converges to $l(1)$.
- If a SLOTP exists, it is unique.
- No SLOTM exists in general (just consider a game where payoff is always 0).

Remark 6 If the game has a uniform value v and both players use ε-optimal strategies, the average expected payoff is essentially constant along the play. Hence $\ell(t) = tv$ is a LOTP.

A first approach to this topic concerns one player games (or games where one player controls the transitions), where there is no finiteness assumption on Ω. In such games, if v_λ converges uniformly, then there exists a SLOTP and it is linear w.r.t. t, which means that the expected payoff is constant along the trajectory (Sorin et al. [14]). The same article provides an example of a two-player game with finite action and countable state spaces, where LOTP is not unique.

Let us also mention recent results of Oliu-Barton and Ziliotto [8] establishing the existence of linear SLOTP for finite stochastic games and optimal strategies: the class of games is larger and they allow for behavioral 0-optimal strategies (hence not only stationary). Our results deal with compact action spaces and ε-optimal stationary strategies.

More precisely, the main contributions of the current paper are:

(a) For absorbing games, existence of a linear LOTP, and existence of a "geometric" algebraic LOTM.
(b) For finite absorbing games, existence of a SLOTP.
(c) An example of a finite game where LOTM is not semialgebraic.
(d) An example of compact absorbing game with non-uniqueness of LOTP.

As a final comment, let us underline the fact that the previous concepts and definitions can be extended to any repeated game, for any evaluation and any type of strategies.

3 Absorbing Games

An absorbing game Γ is defined by two sets of actions I and J, two stage payoff functions g, g^* from $I \times J$ to $[-1, 1]$ and a probability of absorption p^* from $I \times J$ to $[0, 1]$. I and J are compact metric sets; g, g^* and p^* are (jointly) continuous functions. The repeated game is played in discrete time as follows. At stage $t = 1, 2, \ldots$ (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 $p^*(i_t, j_t)$ absorption is reached and the payoff in all future stages $s > t$ is $g^*(i_t, j_t)$;
(iii) with probability $p(i_t, j_t) := 1 - p^*(i_t, j_t)$ the game is repeated at stage $t + 1$.

Recall that the asymptotic analysis for these games is due to Kohlberg [3] in the case where I and J are finite and Rosenberg and Sorin [9] in the current framework. In either case, the value v_λ of the discounted game Γ_λ converges to some v as λ goes to 0. This does not require any assumption on the information of the players. In case of full observation of the actions—or even only of the stage payoff—a uniform value exists, see Mertens and Neyman [5] in the finite case and Mertens et al. [6] for compact actions.

Recall that $X = \Delta(I)$ and $Y = \Delta(J)$ are the sets of probabilities on I and J. The functions g, p and p^* are bilinearly extended to $X \times Y$. Let

$$G^*(x, y) := p^*(x, y)\overline{g}^*(x, y) := \int_{I\times J} p^*(i, j)g^*(i, j)x(\mathrm{d}i)y(\mathrm{d}j).$$

$\overline{g}^*(x, y)$ is thus the expected absorbing payoff conditionally to absorption (and is thus only defined for $p^*(x, y) \neq 0$).

3.1 An Auxiliary Game

Consider the two-person zero-sum (one-shot) game $\mathbf{A}$, with actions $(x, x', a) \in S = X^2 \times \mathbb{R}^+$ and $(y, y', b) \in T = Y^2 \times \mathbb{R}^+$, defined by the payoff function

$$A(x, x', a, y, y', b) = \frac{g(x, y) + a\, G^*(x', y) + b\, G^*(x, y')}{1 + a\, p^*(x', y) + b\, p^*(x, y')}. \tag{1}$$

3.1.1 General Properties

The following proposition extends to the compact case results due to Laraki [4] (Theorem 3 and Corollary 4) in the finite case (which were later on simplified by Cardaliaguet et al. [2]).

Proposition 7 (1) *The game* $\mathbf{A}$ *has a value, which is* $v = \lim v_\lambda$. *More precisely,*

$$\begin{aligned} v &= \max_{x\in X} \sup_{(x',a)\in X\times\mathbb{R}^+} \inf_{(y,y',b)\in T} A(x, x', a, y, y', b) \\ &= \min_{y\in Y} \inf_{(y',b)\in Y\times\mathbb{R}^+} \sup_{(x,x',a)\in S} A(x, x', a, y, y', b). \end{aligned}$$

(2) *Moreover, if* (x, x', a) *is* ε*-optimal in the game* $\mathbf{A}$*, then for any* λ *small enough the stationary strategy* $\hat{x}_\lambda := \frac{x+\lambda a x'}{1+\lambda a}$ *is* 2ε*-optimal in* Γ_λ.

Proof (1) Consider an accumulation point w of the family $\{v_\lambda\}$ and let $\lambda_n \to 0$ such that v_{λ_n} converges to w. We will show that

$$w \leq \sup_{(x,x',a)\in S} \inf_{(y,y',b)\in T} \frac{g(x, y) + a\, G^*(x', y) + b\, G^*(x, y')}{1 + a\, p^*(x', y) + b\, p^*(x, y')} \tag{2}$$

A dual argument proves at the same time that the family $\{v_\lambda\}$ converges and that the auxiliary game $\mathbf{A}$ has a value. Let $r_\lambda(x, y)$ be the payoff in the game Γ_λ, induced by a pair of stationary strategies $(x, y) \in X \times Y$. It satisfies

$$r_\lambda(x, y) = \lambda g(x, y) + (1 - \lambda)[(1 - p^*(x, y))r_\lambda(x, y) + G^*(x, y)] \tag{3}$$

hence

$$r_\lambda(x, y) = \frac{\lambda g(x, y) + (1 - \lambda)G^*(x, y)}{\lambda + (1 - \lambda)p^*(x, y)}. \tag{4}$$

In particular for any $x_\lambda \in X$ optimal for player 1 in Γ_λ, one obtains

$$v_\lambda \leq \frac{\lambda g(x_\lambda, y) + (1 - \lambda)G^*(x_\lambda, y)}{\lambda + (1 - \lambda)p^*(x_\lambda, y)}, \quad \forall y \in Y, \tag{5}$$

that one can write

$$v_\lambda \leq \frac{g(x_\lambda, y) + \frac{(1-\lambda)}{\lambda} G^*(x_\lambda, y)}{1 + \frac{(1-\lambda)}{\lambda} p^*(x_\lambda, y)}, \quad \forall y \in Y. \tag{6}$$

Let $\overline{x} \in X$ be an accumulation point of $\{x_{\lambda_n}\}$ and given $\varepsilon > 0$ let $\overline{\lambda}$ in the sequence $\{\lambda_n\}$ such that

$$|g(\overline{x}, y) - g(x_{\overline{\lambda}}, y)| \leq \varepsilon, \quad \forall y \in Y$$

(we use the fact that g is uniformly continuous on $X \times Y$) and

$$|v_{\overline{\lambda}} - w| \leq \varepsilon.$$

Then with $\overline{a} = \frac{(1-\overline{\lambda})}{\overline{\lambda}}$ and $\overline{x}' = x_{\overline{\lambda}}$, (6) implies

$$w - \varepsilon \leq \frac{g(\overline{x}, y) + \overline{a} G^*(\overline{x}', y)}{1 + \overline{a} p^*(\overline{x}', y)} + \varepsilon, \quad \forall y \in Y. \tag{7}$$

On the other hand, going to the limit in (5) leads to

$$w\, p^*(\overline{x}, y') \leq G^*(\overline{x}, y'), \qquad \forall y' \in Y. \tag{8}$$

We multiply (7) by the denominator $1 + \overline{a} p^*(\overline{x}', y)$, and we add to (8) multiplied by $b \in \mathbf{R}_+$ to obtain the property: $\forall \varepsilon > 0$, $\exists\, \overline{x}, \overline{x}' \in X$ and $\overline{a} \in \mathbf{R}_+$ such that

$$w \leq \frac{g(\overline{x}, y) + \overline{a}\, G^*(\overline{x}', y) + b\, G^*(\overline{x}, y')}{1 + \overline{a}\, p^*(\overline{x}', y) + b\, p^*(\overline{x}, y')} + 2\varepsilon, \qquad \forall y, y' \in Y, b \in \mathbf{R}_+ \tag{9}$$

which implies (2). Note moreover that $(*)$ $\overline{x}$ is independent of ε, which allows to replace sup by max hence the result.

(2) Let (x, x', a) be ε-optimal in the game $\mathbf{A}$ and $\hat{x}_\lambda := \dfrac{x + \lambda a x'}{1 + \lambda a}$. Using (4) one obtains

$$r_\lambda(\hat{x}_\lambda, y) = \frac{\lambda[g(x, y) + \lambda a g(x', y)] + (1-\lambda)[G^*(x, y) + \lambda a G^*(x', y)]}{\lambda(1 + \lambda a) + (1-\lambda)(p^*(x, y) + \lambda a p^*(x', y))}.$$

Note that

$$A\left(x, x', a, y, y, \frac{1-\lambda}{\lambda}\right) = \frac{\lambda g(x, y) + \lambda a\, G^*(x', y) + (1-\lambda)\, G^*(x, y)}{\lambda + \lambda a\, p^*(x', y) + (1-\lambda)\, p^*(x, y)}.$$

Thus,

$$\left| r_\lambda(\hat{x}_\lambda, y) - A\left(x, x', a, y, y, \frac{1-\lambda}{\lambda}\right)\right| \leq 4C\lambda a$$

where C is a bound on the payoffs. Hence, for any $y \in Y$

$$v - r_\lambda(\hat{x}_\lambda, y) \leq \varepsilon + 4C\lambda a \leq 2\varepsilon$$

for λ small enough.

A is an *auxiliary limit game* in the sense that:

(i) There is a map ϕ_1 from $S \times (0, 1]$ to X (that associates with a strategy of player 1 in **A** and a discount factor a stationary strategy of player 1 in Γ).
(ii) There is a map ψ_1 from $Y \times (0, 1]$ to T (that associates with a stationary strategy of player 2 in Γ and a discount factor a stationary strategy of player 2 in **A**).

(iii)

$$r_\lambda(\phi_1(\lambda,s),y) \geq A(s,\psi_1(\lambda,y)) - o(1), \quad \forall s \in S, \forall y \in Y$$

(iv) A dual property holds with some ϕ_2 and ψ_2.

These properties imply: $\lim v_\lambda$ exists and equals $v(\mathbf{A})$.

We then recover Corollary 3.2 in Sorin and Vigeral [12], with a new proof that will be useful in the sequel.

Corollary 8

$$\begin{aligned} v &= \min_{y\in Y}\max_{x\in X} \mathrm{med}\left(g(x,y);\ \sup_{x''|p^*(x'',y)>0}\left\{\overline{g}^*(x'',y)\right\};\ \inf_{y''|p^*(x,y'')>0}\left\{\overline{g}^*(x,y'')\right\}\right) \\ &= \max_{x\in X}\min_{y\in Y} \mathrm{med}\left(g(x,y);\ \sup_{x''|p^*(x'',y)>0}\left\{\overline{g}^*(x'',y)\right\};\ \inf_{y''|p^*(x,y'')>0}\left\{\overline{g}^*(x,y'')\right\}\right) \end{aligned} \tag{10}$$

where med *is the median of three numbers, and with the usual convention that* $\sup_{x''\in\emptyset} = -\infty$*;* $\inf_{y''\in\emptyset} = +\infty$*. Moreover if* (x,x',a) *(resp* (y,y',ε)*) is* ε*-optimal in* $\mathbf{A}$*, then* x *(resp.* y*) is* ε*-optimal in* (10)*.*

Proof For any $\varepsilon > 0$, fix a triplet $(y, y'_\varepsilon, b_\varepsilon) \in T$ of the second player ε-optimal in $\mathbf{A}$, where we can assume that y does not depend on ε by the previous proof $(*)$. Then for any $x'' \in X$ such that $p^*(x'',y) > 0$, one has

$$v + \varepsilon \geq \lim_{a\to+\infty} A(x,x'',a,y,y'_\varepsilon,b_\varepsilon) = \overline{g}^*(x'',y) \tag{11}$$

thus $v+\varepsilon \geq h^+(y) := \sup_{x''|p^*(x'',y)>0}\{\overline{g}^*(x'',y)\}$. Denote similarly $h^-(x) = \inf_{y''|p^*(x,y'')>0}\{\overline{g}^*(x,y'')\}$. On the other hand, for any x

$$v + \varepsilon \geq A(x,x'',0,y,y'_\varepsilon,b_\varepsilon) = \frac{g(x,y) + b_\varepsilon\, G^*(x,y'_\varepsilon)}{1 + b_\varepsilon\, p^*(x,y'_\varepsilon)}$$

Now if $p^*(x,y'_\varepsilon) > 0$, $\dfrac{g(x,y) + b_\varepsilon\, G^*(x,y'_\varepsilon)}{1 + b_\varepsilon\, p^*(x,y'_\varepsilon)} \geq \min\{g(x,y), \overline{g}^*(x,y'_\varepsilon)\}$, hence in all cases

$$v + \varepsilon \geq A(x,x'',0,y,y'_\varepsilon,b_\varepsilon) \geq \min\{g(x,y), h^-(x)\}.$$

Thus, for any $x \in X$

$$v + \varepsilon \geq \mathrm{med}\left(g(x,y); h^+(y); h^-(x)\right) \tag{12}$$

Letting ε go to 0 and using the dual inequality establish the results.

3.1.2 Further Properties of Optimal Strategies

We establish here more precise results concerning the decomposition of the payoff induced by ε-optimal strategies in the game $\mathbf{A}$.

Proposition 9 *Let* (x,x',a) *and* (y,y',b) *be* ε*-optimal in the game* $\mathbf{A}$*.*

(a) *If* $p^*(x,y) > 0$*, then* $|\overline{g}^*(x,y) - v| \leq \varepsilon$
(b) $|g(x,y) - v| \leq 2(1 + ap^*(x',y) + bp^*(x,y'))\varepsilon$

(c) *If* $ap^*(x', y) + bp^*(x, y') > 0$, *then* $\left|\frac{aG^*(x', y) + bG^*(x, y')}{ap^*(x', y) + bp^*(x, y')} - v\right| \leq 3 \frac{1 + ap^*(x', y) + bp^*(x, y')}{ap^*(x', y) + bp^*(x, y')}\varepsilon$.

Proof (a) This is exactly Eq. (11) and its dual.

(b) From $v + \varepsilon \geq A(x, x', 0, y, y', b)$, we get

$$\frac{g(x, y) + bG^*(x, y')}{1 + bp^*(x, y')} \leq v + \varepsilon.$$

On the other hand, $v - \varepsilon \leq \lim_{b\to\infty} A(x, x', a, y, y', b)$ hence $G^*(x, y') \geq (v - \varepsilon)p^*(x, y')$. Combining both inequalities yields

$$\begin{aligned} g(x, y) &\leq (v + \varepsilon)(1 + bp^*(x, y')) - b(v - \varepsilon)p^*(x, y') \\ &\leq v + \varepsilon(1 + 2bp^*(x, y')) \\ &\leq v + 2\varepsilon(1 + ap^*(x', y) + bp^*(x, y')) \end{aligned} \tag{13}$$

and the dual inequality is similar.

(c) Since $A(x, x', a, y, y', b) \geq v - \varepsilon$, one has

$$\begin{aligned} &(v - \varepsilon)(1 + ap^*(x', y) + bp^*(x, y')) \leq g(x, y) + aG^*(x', y) + bG^*(x, y') \\ &\leq v + 2\varepsilon(1 + ap^*(x', y) + bp^*(x, y')) + aG^*(x', y) + bG^*(x, y') \text{ by (13)} \end{aligned}$$

hence

$$v(ap^*(x', y) + bp^*(x, y')) - aG^*(x', y) - bG^*(x, y') \leq 3\varepsilon(1 + ap^*(x', y) + bp^*(x, y'))$$

and the dual inequality is similar.

3.2 Asymptotic Properties in Γ_λ

Since the game is absorbing, we write simply $Q_\lambda^{x,y}(t)$ for $Q_\lambda^{\omega_0,x,y}(t)(\omega_0)$, where ω_0 is the non-absorbing state.

Lemma 10 *Let x_λ and y_λ be two families of (non-necessarily optimal) stationary strategies of player 1 and player 2, respectively. Assume that $\frac{p^*(x_\lambda, y_\lambda)}{\lambda}$ converges to some γ in $[0, +\infty]$ as λ goes to 0. Then, $Q_\lambda^{x_\lambda,y_\lambda}(t)$ converges, uniformly in t, to $\frac{1 - (1 - t)^{1+\gamma}}{1 + \gamma}$, as λ goes to 0, with the natural convention that $\frac{1 - (1 - t)^{1+\gamma}}{1 + \gamma} = 0$ for $\gamma = +\infty$.*

Proof By Definition 2, for any λ and $t_n = \lambda \sum_{i=1}^{n}(1 - \lambda)^{i-1} = 1 - (1 - \lambda)^n$,

$$\begin{aligned} Q_\lambda^{x_\lambda,y_\lambda}(t_n) &= \lambda \sum_{i=1}^{n}(1 - \lambda)^{i-1}(1 - p^*(x_\lambda, y_\lambda))^{i-1} \\ &= \frac{1 - ((1 - \lambda)(1 - p^*(x_\lambda, y_\lambda)))^n}{1 + \frac{p^*(x_\lambda, y_\lambda)}{\lambda} - p^*(x_\lambda, y_\lambda)} \end{aligned}$$

with linear interpolation between these dates.

Remark first that this implies that $Q_\lambda^{x_\lambda,y_\lambda}(t) \leq [1+\frac{p^*(x_\lambda,y_\lambda)}{\lambda}]^{-1}$ for all t and λ, which gives at the limit the desired result if $\gamma=+\infty$.

Assume now that $\gamma \in [0,+\infty[$, and thus that $p^*(x_\lambda, y_\lambda)$ tends to 0 as λ goes to 0. Fix t and λ, and let n be the integer part of $\frac{\ln(1-t)}{\ln(1-\lambda)}$ so that $t_n \leq t \leq t_{n+1}$. Since $Q_\lambda^{x_\lambda,y_\lambda}(t_n)$ is decreasing in n,

$$\begin{aligned} Q_\lambda^{x_\lambda,y_\lambda}(t) &\leq Q_\lambda^{x_\lambda,y_\lambda}(t_n) \\ &= \frac{1-((1-\lambda)(1-p^*(x_\lambda,y_\lambda)))^n}{1+\frac{p^*(x_\lambda,y_\lambda)}{\lambda}-p^*(x_\lambda,y_\lambda)} \\ &\leq \frac{1-((1-\lambda)(1-p^*(x_\lambda,y_\lambda)))^{\frac{\ln(1-t)}{\ln(1-\lambda)}-1}}{1+\frac{p^*(x_\lambda,y_\lambda)}{\lambda}-p^*(x_\lambda,y_\lambda)} \\ &= \frac{1-(1-t)^{1+\frac{\ln(1-p^*(x_\lambda,y_\lambda))}{\ln(1-\lambda)}-\ln(1-\lambda)}}{1+\frac{p^*(x_\lambda,y_\lambda)}{\lambda}-p^*(x_\lambda,y_\lambda)} \end{aligned} \tag{14}$$

Similarly,

$$\begin{aligned} Q_\lambda^{x_\lambda,y_\lambda}(t) &\geq Q_\lambda^{x_\lambda,y_\lambda}(t_{n+1}) \\ &= \frac{1-((1-\lambda)(1-p^*(x_\lambda,y_\lambda)))^{n+1}}{1+\frac{p^*(x_\lambda,y_\lambda)}{\lambda}-p^*(x_\lambda,y_\lambda)} \\ &\geq \frac{1-((1-\lambda)(1-p^*(x_\lambda,y_\lambda)))^{\frac{\ln(1-t)}{\ln(1-\lambda)}+1}}{1+\frac{p^*(x_\lambda,y_\lambda)}{\lambda}-p^*(x_\lambda,y_\lambda)} \\ &= \frac{1-(1-t)^{1+\frac{\ln(1-p^*(x_\lambda,y_\lambda))}{\ln(1-\lambda)}+\ln(1-\lambda)}}{1+\frac{p^*(x_\lambda,y_\lambda)}{\lambda}-p^*(x_\lambda,y_\lambda)} \end{aligned} \tag{15}$$

Letting λ go to 0 in (14) and (15) yields the result.

For any $(x,x',a) \in X^2 \times \mathbb{R}^+$ and $(y,y',b) \in Y^2 \times \mathbb{R}^+$ define

$$\gamma(x,x',a,y,y',b) = \begin{cases} +\infty & \text{if } p^*(x,y)>0 \\ ap^*(x',y)+bp^*(x,y') & \text{if } p^*(x,y)=0 \end{cases}$$

An immediate consequence of the previous lemma is

Corollary 11 *Let $(x,x',a) \in X^2 \times \mathbb{R}^+$ and $(y,y',b) \in Y^2 \times \mathbb{R}^+$ and denote $\hat{x}_\lambda := \frac{x+\lambda a x'}{1+\lambda a}$ and $\hat{y}_\lambda := \frac{y+\lambda b y'}{1+\lambda b}$. Then, $Q_\lambda^{\hat{x}_\lambda,\hat{y}_\lambda}(t)$ converges uniformly in t to $\dfrac{1-(1-t)^{1+\gamma(x,x',a,y,y',b)}}{1+\gamma(x,x',a,y,y',b)}$ as λ goes to 0.*

Proposition 12 *Any absorbing game has a LOTM $Q(t)=\frac{1-(1-t)^{1+\gamma}}{1+\gamma}$ for some $\gamma \in [0,+\infty]$, and a LOTP $l(t)=tv$.*

Proof For every n, let (x,x'_n,a_n) and (y,y'_n,b_n) be $\frac{1}{n}$-optimal strategies for each player in $\mathbf{A}$ (recall that x and y can be chosen independently of n). Up to extraction, $\gamma_n := \gamma(x,x'_n,a_n,y,y'_n,b_n)$ converges to some γ in $[0,+\infty]$.

Fix $\varepsilon>0$ and let $n \geq \frac{2}{\varepsilon}$ such that

$$\left| \frac{1-(1-t)^{1+\gamma_n}}{1+\gamma_n} - \frac{1-(1-t)^{1+\gamma}}{1+\gamma} \right| \leq \frac{\varepsilon}{2} \tag{16}$$

on [0,1]. By Proposition 7, the strategies $\hat{x}^n_\lambda := \frac{x+\lambda a_n x'_n}{1+\lambda a_n}$ and $\hat{y}^n_\lambda := \frac{y+\lambda b_n y'_n}{1+\lambda b_n}$ are ε-optimal in Γ_λ for λ small enough. Corollary 11 and Eq. (16) imply that

$$\left| Q_\lambda^{\hat{x}^n_\lambda, \hat{y}^n_\lambda}(t) - \frac{1-(1-t)^{1+\gamma}}{1+\gamma} \right| \leq \varepsilon \tag{17}$$

for all λ small enough and $t \in [0, 1]$. This answers the first part of the proposition.

Clearly,

$$l_\lambda^{\hat{x}^n_\lambda, \hat{y}^n_\lambda}(t) = Q_\lambda^{\hat{x}^n_\lambda, \hat{y}^n}(t) g(\hat{x}^n_\lambda, \hat{y}^n_\lambda) + \left(t - Q_\lambda^{\hat{x}^n_\lambda, \hat{y}^n}(t)\right) \overline{g}^*(\hat{x}^n_\lambda, \hat{y}^n_\lambda). \tag{18}$$

Recall that the payoff function is assumed bounded by 1. Then, Eqs. (17) and (18) imply that for λ small enough and every t,

$$\left| l_\lambda^{\hat{x}^n_\lambda, \hat{y}^n_\lambda}(t) - \frac{1-(1-t)^{1+\gamma}}{1+\gamma} g(\hat{x}^n_\lambda, \hat{y}^n_\lambda) - \left(t - \frac{1-(1-t)^{1+\gamma}}{1+\gamma}\right) \overline{g}^*(\hat{x}^n_\lambda, \hat{y}^n_\lambda) \right| \leq 2\varepsilon$$

Since x^n_λ and y^n_λ converge to x and y, we then have for λ small enough

$$\left| l_\lambda^{\hat{x}^n_\lambda, \hat{y}^n_\lambda}(t) - \frac{1-(1-t)^{1+\gamma}}{1+\gamma} g(x, y) - \left(t - \frac{1-(1-t)^{1+\gamma}}{1+\gamma}\right) \overline{g}^*(\hat{x}^n_\lambda, \hat{y}^n_\lambda) \right| \leq 3\varepsilon. \tag{19}$$

We now consider four separate cases. Basically either $\gamma = 0$ or $+\infty$ and Eq. (19) implies that $l(t)$ is linear and hence equals tv since $l(1) = v$ by near optimality of the strategies $\hat{x}^n_\lambda$ and $\hat{y}^n_\lambda$ in Γ_λ ; or $\gamma \in]0, +\infty[$ and then both $g(x, y)$ and $\overline{g}^*(\hat{x}^n_\lambda, \hat{y}^n_\lambda)$ are close to v by Proposition 9, which once again implies $l(t) = tv$.

Case 1 $p^*(x, y) > 0$. Then $\overline{g}^*(\hat{x}^n_\lambda, \hat{y}^n_\lambda)$ converges to $\overline{g}^*(x, y)$ as λ go to 0, and by Proposition 9(a) $|\overline{g}^*(x, y) - v| \leq \varepsilon$. Since $\gamma = +\infty$ in that case, Eq. (19) yields $|l_\lambda^{\hat{x}^n_\lambda, \hat{y}^n_\lambda}(t) - tv| \leq 5\varepsilon$ for all λ small enough, uniformly in t.

Case 2 $p^*(x, y) = 0$ and $\gamma = 0$, hence $\gamma_n \leq 1$ (up to choosing a larger n). Then Proposition 9(b) implies $|g(x, y) - v| \leq 4\varepsilon$, and Eq. (19) yields $|l_\lambda^{\hat{x}^n_\lambda, \hat{y}^n_\lambda}(t) - tv| \leq 7\varepsilon$ for all λ small enough, uniformly in t.

Case 3 $p^*(x, y) = 0$ and $\gamma \in]0, +\infty[$, hence $\frac{\gamma}{2} \leq \gamma_n \leq 1+\gamma$ (up to choosing a larger n). Then, Proposition 9(b) implies $|g(x, y) - v| \leq 2(2+\gamma)\varepsilon$. Moreover, $p^*(x, y) = 0$ implies that $\overline{g}^*(\hat{x}^n_\lambda, \hat{y}^n_\lambda)$ converges to $\dfrac{aG^*(x'_n, y) + bG^*(x, y'_n)}{ap^*(x'_n, y) + bp^*(x, y'_n)}$ as λ go to 0, and by Proposition 9(c) $\left| \dfrac{aG^*(x'_n, y) + bG^*(x, y'_n)}{ap^*(x'_n, y) + bp^*(x, y'_n)} - v \right| \leq 3\frac{1+\gamma/2}{\gamma/2}\varepsilon$.

Hence, Eq. (19) yields $|l_\lambda^{\hat{x}^n_\lambda, \hat{y}^n_\lambda}(t) - tv| \leq (7 + 2\gamma + 3\frac{1+\gamma/2}{\gamma/2})\varepsilon$ for all λ small enough, uniformly in t.

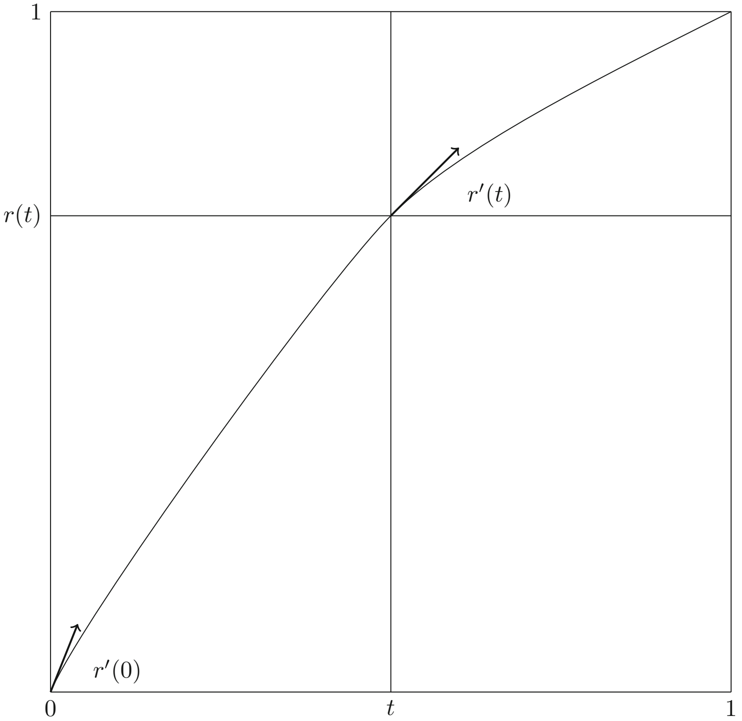
Case 4 $p^*(x, y) = 0$ and $\gamma = +\infty$, hence $\gamma_n \geq 1$ (up to choosing a larger n). Then $\overline{g}^*(\hat{x}^n_\lambda, \hat{y}^n_\lambda)$ converges to $\dfrac{aG^*(x'_n, y) + bG^*(x, y'_n)}{ap^*(x'_n, y) + bp^*(x, y'_n)}$ as λ go to 0, and by Proposition 9(c) $\left| \dfrac{aG^*(x'_n, y) + bG^*(x, y'_n)}{ap^*(x'_n, y) + bp^*(x, y'_n)} - v \right| \leq 6\varepsilon$. Thus, Eq. (19) yields $|l_\lambda^{\hat{x}^n_\lambda, \hat{y}^n_\lambda}(t) - tv| \leq 9\varepsilon$ for all λ small enough, uniformly in t.

As claimed, in every case we see that $l(t) = tv$.

Remark 13 Recall that $Q(\cdot)$ and $l(\cdot)$ represent expected *cumulated* occupation measure and payoff. By deriving these quantities with respect to t, we get that the asymptotic probability

$q(t)$ of still being in the non-absorbing state at time t is $(1-t)^{\gamma}$ and that asymptotically the current payoff is v at any time.

Remark 14 Let us give a simple heuristic behind the form $(1-t)^{\gamma}$ for $q(t)$. Assuming that this quantity is well defined and smooth, note that at time t the remaining game has a length $1-t$ and weight $q(t) = 1 - r(t)$ hence by renormalization (see figure below).



$$\frac{r'(t)}{1-r(t)}(1-t) = r'(0)$$

so that

$$-\frac{r'(t)}{1-r(t)} = -\frac{r'(0)}{1-t}$$

which leads, with $r(0) = 0$ to $r(t) = 1 - (1-t)^{\gamma}$ for some γ.

Let us illustrate now the four cases in the preceding proof by giving examples.

Example 15 Consider the absorbing game

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

(where a * denotes an absorbing payoff) with asymptotic value 1. Let $x = 1/2T + 1/2B$. Then (x, x, n) is $1/n$-optimal in $\mathbf{A}$ for player 1, while any (y, y', b) is optimal for player 2. Since $p^*(x, y) > 0$ for all y, case 1 (see proof of Proposition 12) occurs for any choice of (y, y', b); hence, the corresponding γ is $+\infty$ and $Q(t) = 0$ for all t.

Notice that in Γ_λ the only optimal stationary strategy is $(1/2, 1/2)$ for each player, leading to the same asymptotic trajectory $Q(.) = 0$.

Example 16 Consider the absorbing game

	L	R
T	1^*	0
B	0	1

with asymptotic value 1. Then (B, T, n) is $1/n$-optimal in $\mathbf{A}$, while any (y, y', b) is optimal. The associated γ is $ny(L)$, and hence, either $y(L) = 0$ and case 2 holds with $\gamma = 0$ and $Q(t) = t$, or $y(L) > 0$ and case 4 occurs with $\gamma = +\infty$ and $Q(t) = 0$.

Notice that in Γ_λ the only optimal stationary strategy is $x_\lambda = y_\lambda = (\frac{\sqrt{\lambda}}{1+\sqrt{\lambda}}, \frac{1}{1+\sqrt{\lambda}})$ for each player. Since $p^*(x_\lambda, y_\lambda) = \frac{\lambda}{(1+\sqrt{\lambda})^2} \sim \lambda$, Lemma 10 implies that the asymptotic trajectory associated with optimal strategies is $Q(t) = t - \frac{t^2}{2}$. Moreover, for any $\gamma \geq 0$ the strategy of player 2 $z_\lambda = (\frac{\gamma\sqrt{\lambda}}{1+\sqrt{\lambda}}, 1 - \frac{\gamma\sqrt{\lambda}}{1+\sqrt{\lambda}})$ is ε-optimal in Γ_λ for λ small enough, and $p^*(x_\lambda, z_\lambda) \sim \gamma\lambda$ hence any $Q(t)$ of the form $\frac{1-(1-t)^{1+\gamma}}{1+\gamma}$ is an asymptotic behavior.

Example 17 Consider the Big Match

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

with asymptotic value 1/2. Let $y = 1/2L + 1/2R$. Then, $(B, T, 1)$ and $(y, y, 0)$ are optimal in $\mathbf{A}$, with $\gamma = 1$. Hence, case 3 holds, and the corresponding $Q(t)$ is $t - \frac{t^2}{2}$. The optimal strategies in Γ_λ are $(\frac{\lambda}{1+\lambda}, \frac{1}{1+\lambda})$ and $(1/2, 1/2)$, respectively, leading to the same $Q(t) = t - \frac{t^2}{2}$.

3.3 Finite Case

We now prove that when the game is finite, the limit payoff trajectory is linear for every couple of nearly optimal stationary strategies, not only those given by Proposition 7. That is, $l(t) = tv$ is a SLOTP.

Proposition 18 *Let Γ be a finite absorbing game with asymptotic value v, x_λ and y_λ families of $\varepsilon(\lambda)$-optimal stationary strategies in Γ_λ, with $\varepsilon(\lambda)$ going to 0 as λ goes to 0. Then for every $t \in [0, 1]$, $l_\lambda^{x_\lambda, y_\lambda}(t)$ converges to tv as λ goes to 0.*

We will use in the proof of this proposition the following elementary lemma given without proof.

Lemma 19 *Let a, b, c, d be real numbers with c and d positive. Then* $\min(\frac{a}{c}, \frac{b}{d}) \leq \frac{a+b}{c+d} \leq \max(\frac{a}{c}, \frac{b}{d})$ *with equality if and only if* $\frac{a}{c} = \frac{b}{d}$

Proof of Proposition 18 The result is clear for $t = 0$ or 1, assume by contradiction that it is false for some $t \in]0, 1[$. Hence, there is a sequence λ_n going to 0 and optimal strategies x_{λ_n} and y_{λ_n} such that $l_{\lambda_n}^{x_{\lambda_n}, y_{\lambda_n}}(t)$ converges to tw with $w \neq v$. Up to extraction of subsequences, x_{λ_n} and y_{λ_n} converge to x and y, respectively. Also up to extraction, all the following limits exist in $[0, +\infty]$: $\alpha(i) = \lim_{n\to\infty} \frac{x_{\lambda_n}(i)}{\lambda_n}$, $\beta(j) = \lim_{n\to\infty} \frac{y_{\lambda_n}(j)}{\lambda_n}$, and $\gamma = \lim_{n\to\infty} \frac{p^*(x_{\lambda_n}, y_{\lambda_n})}{\lambda_n}$. If

$\gamma \neq 0$ (and hence $p^*(x_{\lambda_n}, y_{\lambda_n}) > 0$ for n large enough), denote $\overline{g}^*(x, y) := \lim \frac{G^*(x_{\lambda_n}, y_{\lambda_n})}{p^*(x_{\lambda_n}, y_{\lambda_n})}$, which also exists up to extraction.

Recall formula (4):

$$r_{\lambda_n}(x_{\lambda_n}, y_{\lambda_n}) = \frac{\lambda_n g(x_{\lambda_n}, y_{\lambda_n}) + (1-\lambda_n)G^*(x_{\lambda_n}, y_{\lambda_n})}{\lambda_n + (1-\lambda_n)p^*(x_{\lambda_n}, y_{\lambda_n})} \tag{20}$$

and since x_λ and y_λ are families of $\varepsilon(\lambda)$-optimal strategies in Γ_λ, $r_{\lambda_n}(x_{\lambda_n}, y_{\lambda_n})$ converges to v as n tends to infinity.

Recall that by Lemma 10, at the limit

$$l(t) = \frac{1-(1-t)^{1+\gamma}}{1+\gamma} g(x, y) + \left(t - \frac{1-(1-t)^{1+\gamma}}{1+\gamma}\right)\overline{g}^*(x, y).$$

We first claim that $\gamma \in]0, +\infty[$.
If $\gamma = 0$, $l(t) = tg(x, y)$, and by near optimality of x_λ and y_λ $v = l(1) = g(x, y)$, and hence, $l(t) = tv$ a contradiction.
If $\gamma = +\infty$, $l(t) = t\overline{g}^*(x, y)$, and by near optimality of x_λ and y_λ $v = l(1) = g^*(x, y)$, and hence, $l(t) = tv$ a contradiction.

Hence, $\gamma \in]0, +\infty[$, and w is a non-trivial convex combination of $g(x, y)$ and $\overline{g}^*(x, y)$. Since $v = l(1)$ is also a convex combination of $g(x, y)$ and $\overline{g}^*(x, y)$, the assumption that $v \neq w$ implies $g(x, y) \neq \overline{g}^*(x, y)$. Assume without loss of generality $g(x, y) < v < \overline{g}^*(x, y)$.

We classify the actions i of player 1 in four categories I_1 to I_4:

- $i \in I_1$ if $x(i) > 0$,
- $i \in I_2$ if $x(i) = 0$ and $\alpha(i) = +\infty$,
- $i \in I_3$ if $x(i) = 0$ and $\alpha(i) \in]0, +\infty[$,
- $i \in I_4$ if $x(i) = 0$ and $\alpha(i) = 0$.

Hence, actions of category 1 are of order 1, actions of category 3 are of order λ, actions of category 4 are of order $o(\lambda)$, and actions of category 2 are played with probability going to 0 but large with respect to λ. Define categories J_1 to J_4 of player 2 in a similar way. By definition,

$$\frac{p^*(x_{\lambda_n}, y_{\lambda_n})}{\lambda_n} = \sum_{I \times J} \frac{p^*(i, j)x_{\lambda_n}(i)y_{\lambda_n}(j)}{\lambda_n} \tag{21}$$

Recall that the left-hand side converges to $\gamma \in]0, +\infty[$, hence up to extraction $\frac{p^*(i,j)x_{\lambda_n}(i)y_{\lambda_n}(j)}{\lambda_n}$ converge in $[0, +\infty[$ for any i and j, denote by δ_{ij} the limit. If i and j are of category k and l with $k + l > 4$, then $\delta_{ij} = 0$. If $k + l < 4$, then $\frac{x_{\lambda_n}(i)y_{\lambda_n}(j)}{\lambda_n}$ diverges to $+\infty$ which implies that $p^*(i, j) = 0 = \delta_{ij}$.

Hence going to the limit in (21), we get that $\gamma = \gamma_{1,3} + \gamma_{2,2} + \gamma_{3,1}$ where $\gamma_{1,3} := \sum_{I_1 \times J_3} p^*(i, j)x(i)\beta(j)$, $\gamma_{2,2} := \sum_{I_2 \times J_2} \delta_{ij}$, and $\gamma_{3,1} := \sum_{I_3 \times J_1} p^*(i, j)\alpha(i)y(j)$. Recall that $\gamma > 0$ thus at least one of $\gamma_{1,3}$, $\gamma_{2,2}$ or $\gamma_{3,1}$ is positive as well.

Similarly, passing to the limit in the definition of $G^*(x_{\lambda_n}, y_{\lambda_n}) := \sum_{I \times J} G^*(i, j)x_{\lambda_n}(i)y_{\lambda_n}(j)$ yields $\mu := \lim \frac{G^*_{kl}(x_{\lambda_n}, y_{\lambda_n})}{\lambda_n} = \mu_{1,3} + \mu_{2,2} + \mu_{3,1}$ where $\mu_{1,3} := \sum_{I_1 \times J_3} x(i)\beta(j)G^*(i, j)$, $\mu_{2,2} := \sum_{I_2 \times J_2} \delta_{ij}\overline{g}^*(i, j)$, and $\mu_{3,1} := \sum_{I_3 \times J_1} \alpha(i)y(j)G^*(i, j)$. Note that if $\gamma_{k,4-k} = 0$ for some k, then $\mu_{k,4-k} = 0$ as well.

Finally, going to the limit in Eq. (20) yields

$$v = \frac{g(x, y) + \mu_{1,3} + \mu_{2,2} + \mu_{3,1}}{1 + \gamma_{1,3} + \gamma_{2,2} + \gamma_{3,1}}$$

and similarly going to the limit in the definition of $\overline{g}^*(x,y) := \lim \frac{G^*(x_{\lambda_n}, y_{\lambda_n})}{p^*(x_{\lambda_n}, y_{\lambda_n})}$ yields

$$\overline{g}^*(x,y) = \frac{\mu_{1,3} + \mu_{2,2} + \mu_{3,1}}{\gamma_{1,3} + \gamma_{2,2} + \gamma_{3,1}}$$

Recall that we assumed $\overline{g}^*(x,y) > v$. By Lemma 19, there exists k such that $\gamma_{k,4-k} > 0$ and $\frac{\mu_{k,4-k}}{\gamma_{k,4-k}} \geq \overline{g}^*(x,y) > v$.

Assume first that $k = 1$. Consider now the following strategy $y'_{\lambda_n} : y'_{\lambda_n}(j) = 0$ for $j \in J_3$, and $y'_{\lambda_n}(j) = y_{\lambda_n}(j)$ for all other j except for an arbitrary $j_0 \in J_1$ for which $y'_{\lambda_n}(j_0) = y_{\lambda_n}(j_0) + \sum_{j \in J_3} y_{\lambda_n}(j)$. The only effect of this deviation is that now $\gamma'_{1,3} = \mu'_{1,3} = 0$. Hence,

$$\lim_{n\to\infty} r_{\lambda_n}(x_{\lambda_n}, y'_{\lambda_n}) = \frac{g(x,y) + \mu_{2,2} + \mu_{3,1}}{1 + \gamma_{2,2} + \gamma_{3,1}}$$

which is strictly less than v by Lemma 19 since $v < \frac{\mu_{1,3}}{\gamma_{1,3}}$; this contradicts the $\epsilon(\lambda_n)$-optimality of x_{λ_n}.

Assume next that $k = 2$. Consider now the following strategy $y'_{\lambda_n} : y'_{\lambda_n}(j) = 0$ for $j \in J_2$, and $y'_{\lambda_n}(j) = y_{\lambda_n}(j)$ for all other j except for an arbitrary $j_0 \in J_1$ for which $y'_{\lambda_n}(j_0) = y_{\lambda_n}(j_0) + \sum_{j \in J_2} y_{\lambda_n}(j)$. The only effect of this deviation is that now $\gamma'_{2,2} = \mu'_{2,2} = 0$. Hence,

$$\lim_{n\to\infty} r_{\lambda_n}(x_{\lambda_n}, y'_{\lambda_n}) = \frac{g(x,y) + \mu_{1,3} + \mu_{3,1}}{1 + \gamma_{2,2} + \gamma_{3,1}}$$

which is strictly less than v by Lemma 19 since $v < \frac{\mu_{2,2}}{\gamma_{2,2}}$; this again contradicts the $\epsilon(\lambda_n)$-optimality of x_{λ_n}.

Finally assume that $k = 3$. Consider now the following strategy $x'_{\lambda_n} : x_{\lambda_n}(i) = 2x_{\lambda_n}(i)$ for $i \in I_3$ and $x'_{\lambda_n}(i) = x_{\lambda_n}(i)$ for all other i except for an arbitrary $i_0 \in I_1$ for which $x'_{\lambda_n}(i_0) = x_{\lambda_n}(i_0) - \sum_{j \in J_3} x_{\lambda_n}(j)$ (which is nonnegative for n large enough). The only effect of this deviation is that now $\gamma'_{3,1} = 2\gamma_{3,1}$ and $\mu'_{3,1} = 2\mu_{3,1}$. Hence,

$$\lim_{n\to\infty} r_{\lambda_n}(x'_{\lambda_n}, y_{\lambda_n}) = \frac{g(x,y) + \mu_{1,3} + \mu_{2,2} + 2\mu_{3,1}}{1 + \gamma_{1,3} + \gamma_{2,2} + 2\gamma_{3,1}}$$

which is strictly more than v by Lemma 19 since $v < \frac{\mu_{3,1}}{\gamma_{3,1}}$; this contradicts the $\varepsilon(\lambda_n)$-optimality of y_{λ_n}.

4 Finite Stochastic Games

4.1 Non-algebraic Limit Trajectories

Consider the following zero-sum stochastic game with two non-absorbing states and two actions for each player. In the first state s_1 (which is the starting state), the payoff and transitions are as follows:

	L	R
U	1*	0+
D	0	1

where $*$ denotes absorption and $+$ that there is a deterministic transition to state 2. Starting from the second state s_2, the game is a linear variation of the Big Match:

	L	R
U	1^*	-1^*
D	-1	1

Since $v_\lambda(s_2) = 0$ for all λ and since there is no return once the play has entered state s_2, it implies that the optimal play in state s_1 is the same as in the Big Match, in which $\gamma = 1$. So in both states the optimal strategies in Γ_λ are $D + \lambda U$ for player 1 and $1/2L + 1/2R$ for player 2. By a scaling of time, the preceding section tells us that at the limit game, the probability of being in state 2 at time t, given that there were transition from s_1 to s_2 at time z, is $\frac{1-t}{1-z}$. Since (also from the preceding section) the time of transition from s_1 to another state (which is s_2 with probability 1/2) has a uniform law on [0, 1], the probability of being in s_2 at time t is

$$p(t) = \frac{1}{2}\int_0^t \frac{1-t}{1-z}\mathrm{d}z = -\frac{(1-t)\ln(1-t)}{2}.$$

Notice that this not an algebraic function of t as it was always the case in the preceding section. Similarly, the probability of absorption before time t is $1 - (1-t)(1 - \frac{\ln(1-t)}{2})$ and is also non-algebraic.

4.2 No ε-Optimal Strategies of the Form $x + a\lambda x'$

In the following game with two non-absorbing states, the payoff is always 1 in state a and -1 in state b with the following deterministic transitions

state a

	L	R
U	a	b
D	b	1^*

state b

	L	R
U	b	a
D	a	-1^*

It is easy to see that the asymptotic value is 0 and that optimal strategies in the λ-discounted game put a weight $\sim \sqrt{\lambda}$ on D and R in both states, hence the absorbing probability is of the order of λ per stage in each state. We show that strategies of the form $(x_a + C_a\lambda x'_a, x_b + C_b\lambda x'_b)$ cannot guarantee more than -1 to player 1, as λ goes to 0.

- If $x_a(D)x_b(D) > 0$, player 2 plays L in a and R in b inducing an absorbing payoff of -1. From a, we reach eventually b where -1 has a positive probability.
- If $x_a(D) = 0, x_b(D) > 0$, player plays R in both games inducing an absorbing payoff of -1. The payoff will be absorbing with high probability in finite time and the relative probability of 1^* vanishes with λ.
- If $x_a(D) > 0, x_b(D) = 0$, player plays L in both games inducing a non- absorbing payoff of -1. For λ small, most of the time the state is b.
- If $x_a(D) = 0, x_b(D) = 0$, player plays R in game a and L in game b. The event "absorbing payoff of 1" occurs at stage n if $\omega_n = a$ and $i_n = D$. Hence, $\omega_{n-1} = b$ and $i_{n-1} = D$. Now this event "$i_n = D$ and $i_{n-1} = D$" has probability of order λ^2. Then,

the absorbing component of the λ discounted payoffs converges to 0 with λ. Moreover, the non-absorbing payoff is mainly -1.

5 An Absorbing Game with Compact Action Sets and Nonlinear LOTP

We consider the following absorbing game with compact action sets. There are three states, two absorbing 0^* and -1^*, and the non-absorbing state ω, in which the payoff is 1 whatever the actions taken. The sets of action are $X = Y = \{0\} \cup \{1/n,\ n \in \mathbf{N}^*\}$ with the usual distance. The probabilities of absorption are given by:

$$\rho(0^*|x, y) = \begin{cases} 0 & \text{if } x = y \\ \sqrt{y} & \text{if } x \neq y \end{cases}$$

and

$$\rho(-1^*|x, y) = \begin{cases} y & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases}$$

It is easily checked that both functions $\rho(0^*|\cdot\cdot)$ and $\rho(-1^*|\cdot\cdot)$ are (jointly) continuous.

Proposition 20 *For any discount factor $\lambda \in]0, 1]$, 0 (resp. 1) is optimal for player 1 (resp. player 2) in the λ-discounted game, and $v_\lambda = \lambda$.*
The corresponding payoff trajectory is: $l(t) = 0$ on $[0, 1]$.

Proof Action 0 of player 1 ensures that there will never be absorption to state -1^* and thus that the stage payoff from stage 2 on is nonnegative. Action 1 of player 2 ensures that there will be absorption with probability 1 at the end of stage 1 and thus that the stage payoff from stage 2 on is non-positive. Since the payoff in stage 1 is 1 irrespective of player's actions, the proposition is established.

Notice that the play under this couple of optimal strategies is simple: there is immediate absorption to 0^*, and in particular the limit payoff trajectory is linear and equals 0 for every time t.

We now prove that there are other ε-optimal strategies, with a different limit payoff trajectory. Denote $\{\lambda\} := \frac{1}{[1/\lambda]}$ where $[\cdot]$ is the integer part ; hence, $\lambda \leq \{\lambda\} < \frac{\lambda}{1-\lambda}$ and $1/\{\lambda\} \in \mathbb{N}^*$ for all $\lambda \in]0, 1]$.

Proposition 21 *For any discount factor $\lambda \in]0, 1]$, $\{\lambda\}$ is λ-optimal for player 1 and $\sqrt{\lambda}$-optimal for player 2 in the λ-discounted game. The corresponding payoff trajectory is: $l(t) = t - t^2$.*

Proof If both players play $\{\lambda\}$, the payoff in the λ-discounted game is, according to formula (4),

$$r_\lambda(\{\lambda\}, \{\lambda\}) = \frac{\lambda - (1 - \lambda)\{\lambda\}}{\lambda + (1 - \lambda)\{\lambda\}}$$

which is nonnegative since $\{\lambda\} < \frac{\lambda}{1-\lambda}$. On the other hand, since $\lambda \leq \{\lambda\}$, one gets $r_\lambda(\{\lambda\}, \{\lambda\}) \leq \lambda$.

If player 1 plays $\{\lambda\}$, while player 2 plays $y \neq \{\lambda\}$, there is no absorption to -1^*; hence, $r_\lambda(\{\lambda\}, y) \geq 0$. Thus, $\{\lambda\}$ is λ-optimal for player 1.

If player 2 plays $\{\lambda\}$, while player 1 plays $x \neq \{\lambda\}$, then according once again to formula (4),

$$\begin{aligned} r_\lambda(x, \{\lambda\}) &= \frac{\lambda}{\lambda + (1-\lambda)\sqrt{\{\lambda\}}} \\ &\leq \frac{\lambda}{\lambda + (1-\lambda)\sqrt{\lambda}} \\ &\leq \sqrt{\lambda}. \end{aligned}$$

Thus, $\{\lambda\}$ is $\sqrt{\lambda}$-optimal for player 2.

Notice that while the limit value is 0 and $(\{\lambda\}, \{\lambda\})$ is a couple of near-optimal strategies, along the induced play the non-absorbing payoff is 1 and the absorbing payoff is -1. One can compute that the associated γ is 1, and hence, under these strategies $Q(t) = t - \frac{t^2}{2}$. So that the accumulated limit payoff up to time t is $t - t^2$, which is nonlinear and positive for every $t \in]0, 1[$.

Basically the players use a jointly controlled procedure either to follow $(\{\lambda\}, \{\lambda\})$ or to get at most (resp. at least) 0.

6 Concluding Comments

A first series of interesting open questions is directly related to the results presented here such as:

- Extension of Proposition 18 to non-stationary strategies, or more global analysis in the framework of
- arbitrary (not discounted) evaluations where optimal strategies are not stationary,
- general stochastic games.

It is also natural to consider other families of repeated games: a first class that is of interest is games with incomplete information. The natural equivalent of LOTM in this framework is the way information is transmitted during the game.

References

1. Aumann RJ, Maschler M (1994) Repeated games with incomplete information. MIT Press, New York
2. Cardaliaguet P, Laraki R, Sorin S (2012) A continuous time approach for the asymptotic value in two-person zero-sum repeated games. SIAM J Control Optim 50:1573–1596
3. Kohlberg E (1974) Repeated games with absorbing states. Ann Stat 2:724–738
4. Laraki R (2010) Explicit formulas for repeated games with absorbing states. Int J Game Theor 39:53–69
5. Mertens J-F, Neyman A (1981) Stochastic games. Int J Game Theor 10:53–66
6. Mertens J-F, Neyman A, Rosenberg D (2009) Absorbing games with compact action spaces. Math Oper Res 34:257–262
7. Mertens J-F, Zamir S (1971) The value of two-person zero-sum repeated games with lack of information on both sides. Int J Game Theor 1:39–64
8. Oliu-Barton M, Ziliotto B (2018) Constant payoff in zero-sum stochastic games. arXiv:1811.04518v1
9. Rosenberg D, Sorin S (2001) An operator approach to zero-sum repeated games. Isr J Math 121:221–246
10. Shapley LS (1953) Stochastic games. Proc Natl Acad Sci USA 39:1095–1100
11. Sorin S (2002) A first course on zero-sum repeated games. Springer, New York
12. Sorin S, Vigeral G (2013) Existence of the limit value of two person zero-sum discounted repeated games via comparison theorems. J Optim Theor Appl 157:564–576

13. Sorin S, Vigeral G (2015) Reversibility and oscillations in zero-sum discounted stochastic games. J Dyn Games 2:103–115
14. Sorin S, Venel X, Vigeral G (2010) Asymptotic properties of optimal trajectories in dynamic programming. Sankhya A 72:237–245
15. Vigeral G (2013) A zero-sum stochastic game with compact action sets and no asymptotic value. Dyn Games Appl 3:172–186
16. Ziliotto B (2016) Zero-sum repeated games: counterexamples to the existence of the asymptotic value and the conjecture $maxmin = \lim v_n$. Ann Probab 44:1107–1133