The asymptotic value in finite stochastic games

Miquel Oliu-Barton

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
ISSN 0364-765X | EISSN 1526-5471 | 00 | 0000 | 0001

INFORMS
DOI 10.1287/xxxx.0000.0000

The asymptotic value in finite stochastic games

Miquel Oliu-Barton
Combinatoire et Optimisation, Institut de Mathématiques de Jussieu (CNRS, UMR 7586)
Université Pierre et Marie Curie (Paris 6), 75005 Paris, France
miquel.oliu.barton@normalesup.org
https://sites.google.com/site/oliubarton/

We provide a direct, self-contained proof for the existence of $\lim_{\lambda\to 0} v_\lambda$, where v_λ is the value of λ-discounted finite two-person zero-sum stochastic game.

Key words: Stochastic game; Asymptotic value; Shapley operator
MSC2000 subject classification: 91A06, 91A15

1. Introduction. Two-person zero-sum stochastic games were introduced by Shapley [4]. They are described by a 5-tuple $(\Omega, \mathcal{I}, \mathcal{J}, q, g)$, where Ω is a finite set of states, $\mathcal{I}$ and $\mathcal{J}$ are finite sets of actions, $g : \Omega \times \mathcal{I} \times \mathcal{J} \to [0,1]$ is the payoff, $q : \Omega \times \mathcal{I} \times \mathcal{J} \to \Delta(\Omega)$ the transition and, for any finite set X, $\Delta(X)$ denotes the set of probability distributions over X. The functions g and q are bilinearly extended to $\Omega \times \Delta(\mathcal{I}) \times \Delta(\mathcal{J})$. The stochastic game with initial state $\omega \in \Omega$ and discount factor $\lambda \in (0,1]$ is denoted by $\Gamma_\lambda(\omega)$ and is played as follows: at stage $m \geq 1$, knowing the current state ω_m, the players choose actions $(i_m, j_m) \in \mathcal{I} \times \mathcal{J}$; their choice produces a stage payoff $g(\omega_m, i_m, j_m)$ and influences the transition: a new state ω_{m+1} is chosen according to the probability distribution $q(\cdot|\omega_m, i_m, j_m)$. At the end of the game, player 1 receives $\sum_{m\geq 1} \lambda(1-\lambda)^{m-1} g(\omega_m, i_m, j_m)$ from player 2. The game $\Gamma_\lambda(\omega)$ has a value $v_\lambda(\omega)$, and $v_\lambda = (v_\lambda(\omega))_{\omega\in\Omega}$ is the unique fixed point of the so-called Shapley operator [4], i.e. $v_\lambda = \Phi(\lambda, v_\lambda)$, where for all $f \in \mathbb{R}^\Omega$:

$$\Phi(\lambda, f)(\omega) = \mathrm{val}_{(s,t)\in\Delta(\mathcal{I})\times\Delta(\mathcal{J})}\{\lambda g(\omega, s, t) + (1-\lambda)\mathbb{E}_{q(\cdot|\omega,s,t)}[f(\widetilde{\omega})]\}. \tag{1}$$

The Shapley operator provides optimal stationary strategies for both players. In particular, the result holds for any signalling structure on past actions. The existence of $\lim_{\lambda\to 0} v_\lambda$ was established by Bewley and Kohlberg [1], using Tarski-Seidenberg elimination theorem.

The purpose of this note is to provide a direct, self-contained proof for the existence of $\lim_{\lambda\to 0} v_\lambda$.

1.1. The payoff induced by a couple of stationary strategies. Let $(x,y) \in \Delta(\mathcal{I})^\Omega \times \Delta(\mathcal{J})^\Omega$ be a pair of stationary strategies. Every time the state $\omega \in \Omega$ is reached the next state is distributed according to $q(\cdot|\omega, x(\omega), y(\omega))$ and the stage payoff is $g(\omega, x(\omega), y(\omega))$. Thus, the sequence of states $(\omega_m)_m$ is a Markov chain with transition $Q = (q(\omega'|\omega, x(\omega), y(\omega))_{(\omega,\omega')\in\Omega^2}$ and the stage payoffs can be described by a vector $g = (g(\omega, x(\omega), y(\omega))_{\omega\in\Omega}$. Let $\gamma_\lambda(\omega, x, y)$ be the expected payoff induced by (x,y) in $\Gamma_\lambda(\omega)$. Then $\gamma_\lambda = (\gamma_\lambda(\omega, x, y))_{\omega\in\Omega}$ is the unique solution of $\gamma_\lambda = \lambda g + (1-\lambda)Q\gamma_\lambda$. From Cramer's rule – the matrix $\mathrm{Id} - (1-\lambda)Q$ is invertible – one deduces that

$\gamma_\lambda(\omega,x,y)$ is a rational fraction in λ and in the variables $(Q(\omega',\omega))_{(\omega,\omega')\in\Omega^2}$ and $(g(\omega))_{\omega\in\Omega}$. Suppose now that $y\in\Delta(\mathcal{J})^\Omega$ is fixed. Then, for any pair $(\omega,\omega')\in\Omega^2$, $Q(\omega,\omega')=\sum_{i\in I}x^i(\omega)q(\omega'|\omega,i,y)$ and $g(\omega)=\sum_{i\in I}x^i(\omega)g(\omega,i,y)$. Thus, $\gamma_\lambda(\omega,x,y)$ is a rational fraction in λ and in the variables $(x^i(\omega))_{(\omega,i)\in\Omega\times\mathcal{I}}$. Moreover, the monomials both in the numerator and in the denominator can all be written in the following form:

$$c\lambda^a\prod_{(\omega,i)\in\Omega\times\mathcal{I}}(x^i(\omega))^{A(\omega,i)},\tag{2}$$

where $c\in\mathbb{R}$ is some constant depending on the pair (ω,y), $a\in\{0,\dots,|\Omega|\}$ and $A(\omega,i)\in\{0,1\}$ for all $(\omega,i)\in\Omega\times\mathcal{I}$. Hence, by setting

$$\mathcal{M}^+=\{(A,a)\mid A\in\{0,1\}^{\Omega\times\mathcal{I}},\ a\in\{0,\dots,|\Omega|\}\},$$

one can express the expected payoff as:

$$\gamma_\lambda(\omega,x,y)=\frac{\displaystyle\sum_{(A,a)\in\mathcal{M}^+}c(A,a)\lambda^a\prod_{(\omega,i)\in\Omega\times\mathcal{I}}(x^i(\omega))^{A(\omega,i)}}{\displaystyle\sum_{(A,a)\in\mathcal{M}^+}c'(A,a)\lambda^a\prod_{(\omega,i)\in\Omega\times\mathcal{I}}(x^i(\omega))^{A(\omega,i)}},\tag{3}$$

where the real constants $(c(A,a),c'(A,a))_{(A,a)\in\mathcal{M}^+}$ depend on (ω,y) but not on (λ,x).

1.2. The asymptotic payoff induced by a sequence of stationary strategies. Consider now a sequence $(\lambda_n,x_n)_n$, where $\lambda_n\in(0,1]$ is a discount factor and $x_n\in\Delta(\mathcal{I})^\Omega$ is a stationary strategy, for all $n\in\mathbb{N}$. The aim of this section is to construct a vector which summarizes all relevant information for computing the limit of $\gamma_{\lambda_n}(\omega,x_n,y)$, as n tends to infinity, for a fixed stationary strategy $y\in\Delta(\mathcal{J})^\Omega$.

Definition 1. A sequence $(\lambda_n,x_n)_n$ in $(0,1]\times\Delta(\mathcal{I})^\Omega$ is *regular* if $\lim_{n\to\infty}\lambda_n=0$ and if for any two monomials of the form (2) their ratio converges in $[0,+\infty]$ as n tends to infinity. [1]
Regular sequences can be characterized by a vector. Indeed, introduce the following set:

$$\mathcal{M}:=\mathcal{M}^+-\mathcal{M}^+=\{(A,a)\mid A\in\{-1,0,1\}^{\Omega\times\mathcal{I}},\ a\in\{-|\Omega|,\dots,0,\dots,|\Omega|\}\}.$$

The sequence $(\lambda_n,x_n)_n$ is regular if for all $(A,a)\in\mathcal{M}$ the following limit exists in $[0,+\infty]$:

$$L[(\lambda_n,x_n)_n](A,a):=\lim_{n\to\infty}\lambda_n^a\prod_{(\omega,i)\in\Omega\times\mathcal{I}}(x_n^i(\omega))^{A(\omega,i)}.\tag{4}$$

The regularity of a sequence depends on the existence of finitely many limits. Thus, for any family $(x_\lambda)_{\lambda\in(0,1]}$ of stationary strategies there exists $(\lambda_n)_n$ such that $(\lambda_n,x_{\lambda_n})_n$ is regular.

Proposition 1. *Let $y\in\Delta(\mathcal{J})^\Omega$ and $\omega\in\Omega$ be fixed. For any regular sequence $(\lambda_n,x_n)_n$, $\lim_{n\to\infty}\gamma_{\lambda_n}(\omega,x_n,y)$ exists and depends only on the vector $L[(\lambda_n,x_n)_n]$.*

Proof. Let $(\lambda_n,x_n)_n$ be regular and let $L=L[(\lambda_n,x_n)_n]$. We have already seen in Section 1.1 that the expected payoff induced by (x_n,y) in $\Gamma_{\lambda_n}(\omega)$ can be written as a rational fraction whose monomials are all of the form (2). A sharper result can be derived from [5, Proposition 3]. The key observation is that, for any ω', the mean λ-discounted time $t_\lambda(\omega,\omega')=\sum_{m\geq 1}\lambda(1-\lambda)^{m-1}Q^{m-1}(\omega,\omega')$ can be written has a hitting time of an auxiliary Markov chain whose transitions are in the set $\{0,\lambda,((1-\lambda)Q(\omega,\omega'))_{(\omega,\omega')\in\Omega^2}\}$. Thus, using a classical result from Friedlin and Wentzell for finite

[1] We use here the natural convention that $\frac{0}{0}=0^0=1$ and $0^\beta=0$, $0^{-\beta}=\frac{\beta}{0}=+\infty$, for all $\beta>0$.

Markov chains, one deduces that the monomials in $\gamma_{\lambda_n}(\omega, x_n, y)$ have *nonnegative* coefficients. The monomials are no longer as in (2), but rather of the form:

$$m_n := c(1-\lambda_n)^b \lambda_n^a \prod_{(\omega,i)\in\Omega\times\mathcal{I}} (x_n^i(\omega))^{A(\omega,i)}, \tag{5}$$

where $c > 0$ (and does not depend on (λ_n, x_n)), $b \in \{0, \dots, |\Omega|\}$ and $(A, a) \in \mathcal{M}^+$. Note that the regularity of $(\lambda_n, x_n)_n$ ensures that, for any two monomials m_n and m'_n of the form (5), their ratio converges and is determined by L, and the constants $c, c' > 0$. Use the vector L to define an order relation in the set of the monomials in $\gamma_{\lambda_n}(\omega, x_n, y)$: $m_n \preceq m'_n$ if and only if $\lim_{n\to\infty} m_n/m'_n \in [0, +\infty)$. The set is totally ordered. Dividing numerator and denominator by some maximal element m^*_n, and taking $n \to \infty$ we obtain that:

$$\lim_{n\to\infty} \gamma_{\lambda_n}(\omega, x_n, y) = \frac{\sum_{(A,a)\in\mathcal{M}^+} c(A,a) L(A - A^*, a - a^*)}{\sum_{(A,a)\in\mathcal{M}^+} c'(A,a) L(A - A^*, a - a^*)}, \tag{6}$$

for some *nonnegative* constants $(c(A,a), c'(A,a))_{(A,a)\in\mathcal{M}^+}$ The maximality of m^* ensures that $L(A - A^*, a - a^*) \in [0, +\infty)$, for all $(A, a) \in \mathcal{M}^+$ and that not all are 0. The result follows. □

1.3. Canonical strategies. For any $\mathbf{c} = (\mathbf{c}(\omega, i))$ and $\mathbf{e} = (\mathbf{e}(\omega, i))$ in $\mathbb{R}_+^{\Omega\times I}$, we define a family of stationary strategies $(\mathbf{x}_\lambda)_\lambda$ as follows:

$$\mathbf{x}^i_\lambda(\omega) := \frac{\mathbf{c}(\omega, i)\lambda^{\mathbf{e}(\omega,i)}}{\sum_{i'\in\mathcal{I}} \mathbf{c}(\omega, i')\lambda^{\mathbf{e}(\omega,i')}}, \quad \forall(\omega, i) \in \Omega \times \mathcal{I},\ \forall\lambda \in (0, 1]. \tag{7}$$

Assume, in addition, that $\sum_{i\in\mathcal{I},\ \mathbf{e}(\omega,i)=0} \mathbf{c}(\omega, i) = 1$ for all ω, so that

$$\mathbf{x}^i_\lambda(\omega) \sim_{\lambda\to 0} \mathbf{c}(\omega, i)\lambda^{\mathbf{e}(\omega,i)}, \quad \forall(\omega, i) \in \Omega \times \mathcal{I}. \tag{8}$$

The exponent determines the order of magnitude of the probability of playing the action i at state ω asymptotically; the coefficient $\mathbf{c}(\omega, i)$ its intensity.

Definition 2. A family of strategies $(\mathbf{x}_\lambda)_{\lambda\in(0,1]}$ is *canonical* if it is induced by some $\mathbf{x} = (\mathbf{c}, \mathbf{e})$ in the following set:

$$\mathbf{X} = \{(\mathbf{c}, \mathbf{e}) \in (\mathbb{R}^*_+ \times \mathbb{R}_+)^{\Omega\times\mathcal{I}} \mid \forall\omega \in \Omega,\ \sum_{i\in\mathcal{I},\ \mathbf{e}(\omega,i)=0} \mathbf{c}(\omega, i) = 1\}.$$

Note that the coefficients are taken strictly positive.
For all $(A, a) \in \mathcal{M}$ and $\mathbf{x} = (\mathbf{c}, \mathbf{e}) \in \mathbf{X}$ the following limit exists:

$$L_{\mathbf{x}}(A, a) := \lim_{\lambda\to 0} \lambda^a \prod_{(\omega,i)} (\mathbf{x}^i_\lambda(\omega))^{A(\omega,i)}. \tag{9}$$

Indeed, a direct consequence of (8) is that:

$$L_{\mathbf{x}}(A, a) = \lim_{\lambda\to 0} \lambda^{a+\sum_{(\omega,i)} A(\omega,i)\mathbf{e}(\omega,i)} \prod_{(\omega,i)} \mathbf{c}(\omega, i)^{A(\omega,i)},$$

where $\prod_{(\omega,i)} \mathbf{c}(\omega, i)^{A(\omega,i)} > 0$. Thus:

$$L_{\mathbf{x}}(A, a) \in \begin{cases} \{0\}, & \text{iff } a + \sum_{(\omega,i)} A(\omega, i)\mathbf{e}(\omega, i) > 0, \\ \{+\infty\}, & \text{iff } a + \sum_{(\omega,i)} A(\omega, i)\mathbf{e}(\omega, i) < 0, \\ (0, +\infty), & \text{iff } a + \sum_{(\omega,i)} A(\omega, i)\mathbf{e}(\omega, i) = 0. \end{cases} \tag{10}$$

Thus, for any $\mathbf{x} \in \mathbf{X}$ and any vanishing sequence $(\lambda_n)_n$ of discount factors, the sequence $(\lambda_n, \mathbf{x}_{\lambda_n})_n$ is regular. Moreover, $L_{\mathbf{x}} = L[(\lambda_n, \mathbf{x}_{\lambda_n})_n]$ for any such sequence.

2. Main results.

2.1. Representation of a regular $(\lambda_n, x_n)_n$ by a canonical strategy. Fix some regular sequence $(\lambda_n, x_n)_n$ throughout this section and let $L = L[(\lambda_n, x_n)_n] \in [0, +\infty]^{\mathcal{M}}$ the vector defined in (4). Notice that L has many elementary properties:

$(P1)$ $L(0,0) = 1$ and, for all $(A,a) \neq 0$, $L(A,a) = +\infty$ if and only if $L(-A,-a) = 0$;

$(P2)$ For all $\mu \in \mathbb{R}$, $L(0,\mu) := \lim_{n\to\infty} \lambda_n^{\mu} = 0 \Leftrightarrow \mu > 0$ and $L(0,\mu) \in (0,+\infty) \Leftrightarrow \mu = 0$. In particular, $L(0,\mu) \in \{0, 1, +\infty\}$ for all $\mu \in \mathbb{R}$;

$(P3)$ For all $(\omega, i) \in \Omega \times \mathcal{I}$, $L((0,\ldots,1^{(\omega,i)},\ldots,0),0) := \lim_{n\to\infty} x_n^i(\omega) \in [0,1]$;

$(P4)$ If $L(A,a) < +\infty$, then $L(\mu A, \mu a) := \lim_{n\to\infty} \lambda_n^{\mu a} \prod_{(\omega,i)} (x_n^i(\omega))^{\mu A(\omega,i)} = L(A,a)^{\mu}$;

$(P5)$ If $L(A,a) < +\infty$ and $L(B,b) < +\infty$, then $L(A+B, a+b) = L(A,a)L(B,b)$.

Proposition 2. *There exists $\mathbf{x} \in \mathbf{X}$ such that $L_{\mathbf{x}} = L$.*

Proof. Note that $\prod_{(\omega,i)} \mathbf{c}(\omega,i)^{A(\omega,i)} > 0$ for any $A \in \{-1,0,1\}^{\Omega\times I}$. Thus, from (10) and $(P1)$ one deduces the following necessary and sufficient conditions on the coefficients and the exponents $(\mathbf{c}, \mathbf{e})$ of $\mathbf{x}$ for having $L_{\mathbf{x}} = L$:

$$\sum\nolimits_{(\omega,i)} A(\omega,i)\mathbf{e}(\omega,i) + a > 0, \qquad \forall (A,a) \in \mathcal{M} \text{ s.t. } L(A,a) = 0, \tag{11}$$

$$\sum\nolimits_{(\omega,i)} A(\omega,i)\mathbf{e}(\omega,i) + a = 0, \ \forall (A,a) \in \mathcal{M} \text{ s.t. } L(A,a) \in (0,+\infty), \tag{12}$$

$$\prod\nolimits_{(\omega,i)} \mathbf{c}(\omega,i)^{A(\omega,i)} = L(A,a), \ \forall (A,a) \in \mathcal{M} \text{ s.t. } L(A,a) \in (0,+\infty). \tag{13}$$

Notation: Let $\mathcal{L}_0 := \{(A,a) \in \mathcal{M} \mid L(A,a) = 0\}$ and $\mathcal{L}_+ := \{(A,a) \in \mathcal{M} \mid L(A,a) \in (0,+\infty)\}$. Put $\mathcal{L} := \mathcal{L}_0 \cup \mathcal{L}_+$.

Solving for the exponents. Let us prove that the system (11)-(12) has a solution. One and only one of the systems (11)-(12) and (14)-(15)-(16) is consistent (see Mertens, Sorin and Zamir [3], part A, page 28):

$$\sum\nolimits_{(A,a)\in\mathcal{L}} \mu(A,a)A = 0, \quad \mu_{|\mathcal{L}_0} \geq 0, \tag{14}$$

$$-\sum\nolimits_{(A,a)\in\mathcal{L}} \mu(A,a)a \geq 0, \tag{15}$$

$$-\sum\nolimits_{(A,a)\in\mathcal{L}} \mu(A,a)a + \sum\nolimits_{(A,a)\in\mathcal{L}_0} \mu(A,a) > 0, \tag{16}$$

Let us prove that the system (14)-(15)-(16), with unknowns $\mu = (\mu(A,a))_{(A,a)} \in \mathbb{R}^{\mathcal{L}}$, is inconsistent. In (14), $\mu_{|\mathcal{L}_0} := (\mu(A,a))_{(A,a)\in\mathcal{L}_0}$ denotes the restriction of μ to $\mathcal{L}_0$. Assume (14). On the one hand, by $(P4)$-$(P5)$, for all $\mu \in \mathbb{R}^{\mathcal{L}}$:

$$\prod_{(A,a)\in\mathcal{L}_+} L(A,a)^{\mu(A,a)} = L\left(\sum\nolimits_{(A,a)\in\mathcal{L}_+} \mu(A,a)A, \sum\nolimits_{(A,a)\in\mathcal{L}_+} \mu(A,a)a\right) \in (0,+\infty). \tag{17}$$

On the other hand, by $(P4)$-$(P5)$, for all $\mu \in \mathbb{R}^{\mathcal{L}}$ such that $\mu_{|\mathcal{L}_0} \geq 0$ one has:

$$\prod_{(A,a)\in\mathcal{L}_0} L(A,a)^{\mu(A,a)} = L\Bigg(\sum_{(A,a)\in\mathcal{L}_0} \mu(A,a)A, \sum_{(A,a)\in\mathcal{L}_0} \mu(A,a)a\Bigg) = \begin{cases} 1 & \text{if } \mu_{|\mathcal{L}_0} = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

Multiplying (17) and (18) yields, by assumption (14) :

$$L\left(0, \sum\nolimits_{(A,a)\in\mathcal{L}} \mu(A,a)a\right) \in \begin{cases} (0,+\infty) & \text{if } \mu_{|\mathcal{L}_0} = 0, \\ \{0\} & \text{otherwise.} \end{cases} \tag{19}$$

By $(P2)$, the first case implies $\sum_{(A,a)\in\mathcal{L}}\mu(A,a)a=0$, which contradicts (16), and the second case implies $\sum_{(A,a)\in\mathcal{L}}\mu(A,a)a=0$, which contradicts (15). The system (14)-(15)-(16) being inconsistent, the existence of a solution to (11)-(12) in $\mathbb{R}^{\Omega\times\mathcal{I}}$ follows. By $(P3)$, $((0,\dots,1^{(\omega,i)},\dots,0),0)\in\mathcal{L}_0\cup\mathcal{L}_+$. Thus, $\mathbf{e}(\omega,i)=(0,\dots,1^{(\omega,i)},\dots,0)\mathbf{e}\geq 0$, by (11) and (12).

Solving for the coefficients. Taking the logarithm in (13) yields:

$$\sum\nolimits_{(\omega,i)}A(\omega,i)\ln\mathbf{c}(\omega,i)=\ln(L(A,a)),\quad\forall(A,a)\in\mathcal{L}_+,\tag{20}$$

which is a linear system in $\mathbf{d}=(\ln\mathbf{c}(\omega,i))_{(\omega,i)}\in\mathbb{R}^{\Omega\times\mathcal{I}}$. As before, one and only one of the systems (20) and (21) is consistent:

$$\sum\nolimits_{(A,a)\in\mathcal{L}_+}\mu(A,a)A=0,\quad\sum\nolimits_{(A,a)\in\mathcal{L}_+}\mu(A,a)\ln(L(A,a))>0.\tag{21}$$

Let us prove that the system (21), with unkowns $\mu=(\mu(A,a))_{(A,a)}\in\mathbb{R}^{\mathcal{L}_+}$, is inconsistent. Suppose that $\sum_{(A,a)\in\mathcal{L}_+}\mu(A,a)A=0$. Then, by $(P4)$-$(P5)$:

$$\prod\nolimits_{(A,a)\in\mathcal{L}_+}L(A,a)^{\mu(A,a)}=L\left(0,\sum\nolimits_{(A,a)\in\mathcal{L}_+}\mu(A,a)a\right)\in(0,+\infty).$$

By $(P2)$, this implies $\sum_{(A,a)\in\mathcal{L}_+}\mu(A,a)a=0$ and, a fortiori, $\prod_{(A,a)\in\mathcal{L}_+}L(A,a)^{\mu(A,a)}=1$, so that (21) fails. Consequently, there exists $\mathbf{c}=(\exp(\mathbf{d}(\omega,i))\in(\mathbb{R}_+)^{\Omega\times I}$ satisfying (13). □

2.2. Existence of $\lim_{\lambda\to 0}v_\lambda$.

THEOREM 1. *The limit of $(v_\lambda)_\lambda$, as λ tends to 0, exists. Moreover, there exists $\mathbf{x}\in\mathbf{X}$ such that $(\mathbf{x}_\lambda)_\lambda$ is* asymptotically optimal, *i.e. for all $\varepsilon>0$, there exists $\lambda_0\in(0,1]$ such that:*

$$\gamma_\lambda(\omega,\mathbf{x}_\lambda,y)\geq\lim\nolimits_{\lambda\to 0}v_\lambda(\omega)-\varepsilon,\quad\forall\omega\in\Omega,\ \forall y\in\Delta(\mathcal{J})^\Omega,\ \forall\lambda\in(0,\lambda_0).$$

Proof. Let $\omega\in\Omega$ be fixed. Let $(x_\lambda)_{\lambda>0}$ be a family of optimal stationary strategies in $(\Gamma_\lambda(\omega))_{\lambda>0}$ and let $(\lambda_n)_n$ be a sequence of discount factors such that $\lim_{n\to\infty}v_{\lambda_n}(\omega)=\limsup_{\lambda\to 0}v_\lambda(\omega)$. The optimality of x_{λ_n} implies that $\gamma_{\lambda_n}(\omega,x_{\lambda_n},j)\geq v_{\lambda_n}(\omega)$, for all $j\in\mathcal{J}^\Omega$. Indeed, against a stationary strategy of player 1, player 2 faces a Markov decision process. Thus, player 2 has a pure stationary best reply. Up to some subsequence, $(\lambda_n,x_{\lambda_n})_n$ is regular. By Proposition 2, there exists $\mathbf{x}\in\mathbf{X}$ such that $L_{\mathbf{x}}=L[(\lambda_n,x_{\lambda_n})_n]$. Thus, by Proposition 1,

$$\lim\nolimits_{n\to\infty}\gamma_{\lambda_n}(\omega,x_{\lambda_n},j)=\lim\nolimits_{n\to\infty}\gamma_{\lambda_n}(\omega,\mathbf{x}_{\lambda_n},j),\quad\forall j\in\mathcal{J}^\Omega.$$

On the other hand, the limit $\lim_{\lambda\to 0}\gamma_\lambda(\omega,\mathbf{x}_\lambda,j)$ exists. Consequently:

$$\lim\nolimits_{\lambda\to 0}\gamma_\lambda(\omega,\mathbf{x}_\lambda,j)=\lim\nolimits_{n\to\infty}\gamma_{\lambda_n}(\omega,x_{\lambda_n},j)\geq\limsup\nolimits_{\lambda\to 0}v_\lambda(\omega),\quad\forall j\in\mathcal{J}^\Omega.\tag{22}$$

It follows that for all $\varepsilon>0$ there exists $\lambda_0\in(0,1]$ such that:

$$\min_{j\in\mathcal{J}^\Omega}\gamma_\lambda(\omega,\mathbf{x}_\lambda,j)\geq\limsup\nolimits_{\lambda\to 0}v_\lambda(\omega)-\varepsilon,\quad\forall\lambda\in(0,\lambda_0).\tag{23}$$

The latter implies that $v_\lambda(\omega)\geq\limsup_{\lambda\to 0}v_\lambda(\omega)-\varepsilon$, for all $\lambda\in(0,\lambda_0)$, and the existence of $\lim_{\lambda\to 0}v_\lambda$ follows by taking the lim inf. The canonical strategy $\mathbf{x}$ has the desired property. □

2.3. Concluding remarks.

(1) Consider an infinitely repeated stochastic game where the past actions are observed. The existence of the uniform value is due to Mertens and Neyman [2] and relies on the following result:

THEOREM 2. *Let $f:(0,1)\to\mathbb{R}^\Omega$ be a function such that:*
(*a*) $\|f_\lambda - f_{\lambda'}\| \le \int_\lambda^{\lambda'}\varphi(x)dx$, *for all* $0<\lambda<\lambda'<1$ *and for some* $\varphi\in L^1\big((0,1],\mathbb{R}_+\big)$*;*
(*b*) *There exists* $\lambda_0>0$ *such that* $\Phi(\lambda,f_\lambda)\ge f_\lambda$, *for all* $\lambda\in(0,\lambda_0)$.[2]
Then, player 1 *can guarantee* $\lim_{\lambda\to 0} f_\lambda$ *in* Γ_∞.

One can use Theorem 1 to prove the existence of the uniform value. Indeed, for any $x\in\Delta(\mathcal{I})^\Omega$, $\omega\in\Omega$ and $\lambda\in(0,1]$, let $w_\lambda^x(\omega):=\min_{j\in\mathcal{J}^\Omega}\gamma_\lambda(\omega,x,j)$ be the payoff guaranteed by x in $\Gamma_\lambda(\omega)$. One can check that $w_\lambda^x\le\Phi(\lambda,w_\lambda^x)$, for all $\lambda\in(0,1]$. Besides, for any $\mathbf{x}\in\mathbf{X}$, the functions $(\lambda\mapsto w_\lambda^{\mathbf{x}_\lambda}(\omega))_{\omega\in\Omega}$ are of bounded variation, so that player 1 can guarantee $\lim_{\lambda\to 0} w_\lambda^{\mathbf{x}_\lambda}$, for any $\mathbf{x}\in\mathbf{X}$. In particular, if $(\mathbf{x}_\lambda)_\lambda$ is asymptotically optimal, player 1 can guarantee $\lim_{\lambda\to 0} v_\lambda$.

(2) The existence of an $\mathbf{x}\in\mathbf{X}$ such that $(\mathbf{x}_\lambda)_\lambda$ is asymptotically optimal was already noticed by Solan and Vieille [6]. The result was deduced from the semi-algebraicity of $\lambda\mapsto v_\lambda$, obtained in [1] using Tarski-Seidenberg elimination theorem.

(3) In the system (11)-(12) for the exponents (first part of the proof of Proposition 2) note that all the entries of A are in $\{-1,0,1\}$. This implies the existence of a solution having all its coordinates in $\{0,1/N,2/N,\dots\}$, for some $N\le|\Omega||\mathcal{I}|^{\sqrt{|\Omega||\mathcal{I}|}}$.

(4) Our approach fails without the finiteness assumption on $\mathcal{I}$, $\mathcal{J}$ and Ω. An example where $\mathcal{I}$ and $\mathcal{J}$ are compact, q is continuous, g is independent of the actions and the family $(v_\lambda)_\lambda$ does not converge is due to Vigeral [7].

Acknowledgments. I am particularly indebted to Sylvain Sorin for his careful reading and comments on earlier versions and to Nicolas Vieille for his valuable remarks. I would like to thank Eilon Solan for his accurate comments and remarks, and also Fabien Gensbittel, Mario Bravo and Guillaume Vigeral for the discussions at an early stage of this work.

References

[1] Bewley, T., E. Kohlberg. 1976. The asymptotic theory of stochastic games. *Math. Oper. Res.* 197–208.

[2] Mertens, J.F., A. Neyman. 1981. Stochastic games. *Int. J. Game Theory* **10** 53–66.

[3] Mertens, J.F., S. Sorin, S. Zamir. 1994. *Repeated Games*. CORE DP 9420-21-22.

[4] Shapley, L.S. 1953. Stochastic games. *Proc. Nat. Acad. Sci.* **39** 1095–1100.

[5] Solan, E. 2003. Continuity of the value of competitive markov decision processes. *Journal of Theoretical Probability* **16**(4) 1–15.

[6] Solan, E., N. Vieille. 2010. Computing uniformly optimal strategies in two-player stochastic games. *Econ. Theory* **42** 237–253.

[7] Vigeral, G. 2012. A zero-sum stochastic game with compact action sets and no asymptotic value (preprint).

[2] Φ is the Shapley operator, defined in (1).