



HAL
open science

Operator approach to values of stochastic games with varying stage duration

Sylvain Sorin, Guillaume Vigeral

► **To cite this version:**

Sylvain Sorin, Guillaume Vigeral. Operator approach to values of stochastic games with varying stage duration. *International Journal of Game Theory*, 2016, 10.1007/s00182-015-0512-8 . hal-01116459v1

HAL Id: hal-01116459

<https://hal.science/hal-01116459v1>

Submitted on 13 Feb 2015 (v1), last revised 8 Jan 2016 (v2)

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.

GENERALIZED ITERATIONS OF NON EXPANSIVE MAPS, EVOLUTION EQUATIONS AND VALUES OF ZERO-SUM STOCHASTIC GAMES WITH VARYING STAGE DURATION

SYLVAIN SORIN AND GUILLAUME VIGERAL

ABSTRACT. We study the links between the values of stochastic games with varying stage duration h , the corresponding Shapley operators \mathbf{T} and \mathbf{T}_h and the solution of $\dot{f}_t = (\mathbf{T} - Id)f_t$. For a large class of stochastic games we establish two kinds of results, under both the discounted or the finite duration framework. On the one hand, for a fixed horizon or discount factor, the value converges as the stage duration goes to 0. On the other hand, the asymptotic behavior of the normalized value as the horizon goes to infinity, or as the discount factor goes to 0, does not depend on the stage duration.

1. INTRODUCTION

The operator introduced by Shapley [14] to study zero-sum stochastic games is a non expansive map \mathbf{T} from a Banach space to itself. Several results have been obtained by using a similar “operator approach” to repeated games, e.g. [13], [7], [15], [11].

We link here the study of fractional iterations of a non expansive map, the solution of the evolution equation $\dot{f}_t = (\mathbf{T} - Id)f_t$ and the values of stochastic games with varying stage duration, as defined in [9].

The structure of the paper is as follows. We first recall results on non expansive maps (Section 2) and the definition of Shapley operators (Section 3). In Section 4 we explain the link between games with varying stage duration and fractional Shapley operators T_h . Section 5 and 6 are respectively devoted to the study of games with finite horizon and with discount factor. In both frameworks we establish results of two different kinds. Firstly, for a fixed evaluation (horizon or discount factor), the value of a game with varying stage duration converges as the stage duration goes to 0. This is done for general stage duration in the finite horizon case, and for a uniform stage duration in the discounted case. Secondly, the asymptotic behavior of the normalized value does not depend on the stage duration. Once again, in the finite horizon case the duration may not be uniform. In Section 7 we study the discretization of the continuous stage game. The last section provide concluding comments.

2. ITERATIONS OF NON EXPANSIVE MAPS AND EVOLUTION EQUATIONS

Consider a non expansive map \mathbf{T} from a Banach space X to itself. In this section we recall basic results concerning its iterations and the corresponding discrete and continuous dynamics.

2.1. Finite iteration.

The n -stage iteration starting from $x \in X$ is $W_n = \mathbf{T}^n(x)$ hence satisfies

$$W_n - W_{n-1} = -(Id - \mathbf{T})(W_{n-1})$$

which can be considered as a discretization of the differential equation

$$(1) \quad \dot{f}_t = -Af_t, \quad f_0 = x,$$

where the accretive (maximal monotone) operator A is $Id - \mathbf{T}$.

Date: February 2015.

This was co-funded by PGMO 2014-LMG. The second author was partially supported by the French Agence Nationale de la Recherche (ANR) "ANR GAGA: ANR-13-JS01-0004-01".

The comparison between the iterates of \mathbf{T} and the solution $f_t(x)$ of (1) is given by the generalized Chernoff's formula [6], [2], see e.g. Brézis [1], p.16:

Proposition 2.1.

$$(2) \quad \|f_t(x) - \mathbf{T}^n(x)\| \leq \|x - \mathbf{T}(x)\| \sqrt{t + (n-t)^2}.$$

In particular with $x = 0$ and $t = n$, one obtains

$$(3) \quad \left\| \frac{f_n(0)}{n} - v_n \right\| \leq \frac{\|\mathbf{T}(0)\|}{\sqrt{n}}$$

where $\mathbf{T}^n(0) = V_n = nv_n$.

A change of time shows that a solution of

$$(4) \quad \dot{g}_t = -\frac{(Id - \mathbf{T})g_t}{h}, \quad g_0 = x,$$

satisfies

$$\|g_t(x) - \mathbf{T}^n(x)\| \leq \|x - \mathbf{T}x\| \sqrt{\frac{t}{h} + (n - \frac{t}{h})^2}.$$

2.2. Interpolation.

Let $h \in [0, 1]$.

Definition:

$$(5) \quad \mathbf{T}_h = (1-h)Id + h\mathbf{T}.$$

Thus \mathbf{T}_h appears as the operator \mathbf{T} applied with rate h . In particular \mathbf{T} corresponds to $h = 1$ and Id to $h = 0$.

Then using (4) one has

$$(6) \quad \|f_t(x) - \mathbf{T}_h^n(x)\| \leq \|x - \mathbf{T}x\| \sqrt{th + (nh-t)^2},$$

hence in particular with $h = \frac{t}{n}$

$$(7) \quad \|f_t(x) - \mathbf{T}_{t/n}^n(x)\| \leq \|x - \mathbf{T}x\| \frac{t}{\sqrt{n}},$$

or

$$(8) \quad \|f_{nh}(x) - \mathbf{T}_h^n(x)\| \leq \|x - \mathbf{T}x\| h\sqrt{n}.$$

2.3. Eulerian schemes.

More generally for two sequences $\{h_k\}, \{\hat{h}_\ell\}$ in $[0, 1]$, with associated Eulerian schemes

$$x_{k+1} = \mathbf{T}_{h_{k+1}}x_k,$$

$$\hat{x}_{\ell+1} = \mathbf{T}_{\hat{h}_{\ell+1}}\hat{x}_\ell,$$

Vigeral [19] obtains:

Proposition 2.2.

$$(9) \quad \|\hat{x}_\ell - x_k\| \leq \|\hat{x}_0 - z\| + \|x_0 - z\| + \|z - \mathbf{T}z\| \sqrt{(\sigma_k - \hat{\sigma}_\ell)^2 + \tau_k + \hat{\tau}_\ell}, \quad \forall z \in X,$$

$$(10) \quad \|f_t(x) - x_k\| \leq \|x - \mathbf{T}x\| \sqrt{(\sigma_k - t)^2 + \tau_k},$$

with $x_0 = x$, $\sigma_k = \sum_{i=1}^k h_i$, $\tau_k = \sum_{i=1}^k h_i^2$, $\hat{\sigma}_\ell = \sum_{j=1}^\ell \hat{h}_j$, $\hat{\tau}_\ell = \sum_{j=1}^\ell \hat{h}_j^2$.

This gives, in the uniform case $h_i = h$

$$(11) \quad \|f_t(x) - \mathbf{T}_h^n(x)\| \leq \|x - \mathbf{T}x\| \sqrt{nh^2 + (nh-t)^2}$$

and coincides with (7) and (8) at $t = nh$.

2.4. Two approximations.

Equation (6) or more generally (10) corresponds to two approximations:

i) Comparison on a compact interval $[0, T]$ of f_t to the linear interpolation $\hat{\mathbf{T}}_{T/n}^s$ of $\{\mathbf{T}_{T/n}^m\}$, $m = 0, \dots, n$, which is, using (7) :

$$\|f_t(0) - \hat{\mathbf{T}}_{T/n}^{mt/T}(0)\| \leq K \frac{T}{\sqrt{n}}, \forall t \in [0, T],$$

or more generally

$$\|f_T(0) - \Pi \mathbf{T}_{h_i}(0)\| \leq K' \sqrt{h},$$

with $\sum h_i = T$, $h^i \leq h$.

ii) Comparison of the asymptotic behavior of f_t , solution of (1) and iterations of terms of the form $\Pi_i \mathbf{T}_{h_i}$ with stepsize $h_i \leq 1$ and total length $\sigma_k = t$:

$$\|f_t(0) - \Pi_i \mathbf{T}_{h_i}(0)\| \leq \|\mathbf{T}(0)\| \sqrt{t}.$$

2.5. General properties.

Define V_λ as the unique fixed point $V_\lambda = \mathbf{T}((1 - \lambda)V_\lambda)$ and $v_\lambda = \lambda V_\lambda$.

We recall some basic evaluations, see e.g. [18].

Proposition 2.3.

$$\begin{aligned} \|v_n\| &\leq \|\mathbf{T}(0)\|, \\ \|v_\lambda\| &\leq \|\mathbf{T}(0)\|, \\ \|v_\lambda - v_\mu\| &\leq 2\left|1 - \frac{\lambda}{\mu}\right| \|\mathbf{T}(0)\|, \\ \|V_\lambda - V_\mu\| &\leq \frac{|\lambda - \mu|}{\lambda\mu} \|\mathbf{T}(0)\|. \end{aligned}$$

Proof. By induction $\|\mathbf{T}^n(0)\| \leq n\|\mathbf{T}(0)\|$ since:

$$\|\mathbf{T}^n(0)\| \leq \|\mathbf{T}^n(0) - \mathbf{T}^{n-1}(0)\| + \|\mathbf{T}^{n-1}(0)\| \leq \|\mathbf{T}(0)\| + \|\mathbf{T}^{n-1}(0)\|.$$

Similarly:

$$\|V_\lambda\| - \|\mathbf{T}(0)\| \leq \|V_\lambda - \mathbf{T}(0)\| = \|\mathbf{T}([1 - \lambda]V_\lambda) - \mathbf{T}(0)\| \leq (1 - \lambda)\|V_\lambda\|$$

implies $\lambda\|V_\lambda\| \leq \|\mathbf{T}(0)\|$.

Finally:

$$\begin{aligned} \|V_\lambda - V_\mu\| &= \|\mathbf{T}([1 - \lambda]V_\lambda) - \mathbf{T}([1 - \mu]V_\mu)\| \\ &\leq \|[1 - \lambda]V_\lambda - [1 - \mu]V_\mu\| \\ &\leq (1 - \lambda)\|V_\lambda - V_\mu\| + |\lambda - \mu|\|V_\mu\| \end{aligned}$$

thus

$$\lambda\|V_\lambda - V_\mu\| \leq |\lambda - \mu|\|V_\mu\| \leq \frac{|\lambda - \mu|}{\mu} \|\mathbf{T}(0)\|.$$

On the other hand:

$$\|v_\lambda - v_\mu\| \leq \lambda\|V_\lambda - V_\mu\| + |\lambda - \mu|\|V_\mu\|$$

hence

$$\|v_\lambda - v_\mu\| \leq 2|\lambda - \mu|\|V_\mu\|$$

and the result follows from $\|V_\mu\| \leq \|\mathbf{T}(0)\|/\mu$. \square

3. STOCHASTIC GAMES AND SHAPLEY OPERATOR

Consider a two person zero-sum stochastic game G with finite state space Ω . I and J are compact action spaces, X and Y are the sets of regular probabilities on the corresponding Borel σ -algebra. g is a bounded payoff function from $\Omega \times I \times J$ to \mathbb{R} (with multilinear extension to $X \times Y$) and for each $(i, j) \in I \times J$, $P(i, j)$ is a transition probability from Ω to $\Delta(\Omega)$. g and P are separately continuous on I and J .

The game is played in stages. At stage n , knowing the state ω_n , player 1 (resp. 2) chooses $i_n \in I$ (resp. $j_n \in J$), the stage payoff is $g_n = g(\omega_n, i_n, j_n)$ and the next state ω_{n+1} is selected according to $P(i_n, j_n)[\omega_n]$.

One associates to G a Shapley operator, see [14], which is a non expansive map \mathbf{T} defined on the set $\{f \in F = \mathbb{R}^\Omega\}$ by:

$$(12) \quad \mathbf{T}(f)(\omega) = \mathbf{val}_{(x,y) \in X \times Y} \{g(\omega; x, y) + P(x, y)[\omega] \circ f\}$$

where $\mathbf{val}_{X \times Y}$ is the max min = min max = value operator on $X \times Y$,

$$P(x, y)[\omega](\omega') = \int_{I \times J} P(i, j)[\omega](\omega') x(di) y(dj) \text{ and } R \circ f = \sum_{\zeta} R(\zeta) f(\zeta).$$

Recall that $V_n = \mathbf{T}^n(0)$ is the value of the n -stage game with total evaluation $\sum_{m=1}^n g_m$ and that V_λ is the unnormalized value of the discounted game with total evaluation $\sum_{m=1}^{\infty} g_m (1 - \lambda)^{m-1}$.

Let us introduce Q such that, for each (i, j) , $P(i, j) = Id + Q(i, j)$.

Then $\mathbf{T} = Id + \mathbf{S}$ with

$$(13) \quad \mathbf{S}(f)(\omega) = \mathbf{val}\{g(\omega; \cdot) + Q(\cdot)[\omega] \circ f\}.$$

It follows that, for $h \in [0, 1]$

$$\mathbf{T}_h = Id + h \mathbf{S} = (1 - h)Id + h \mathbf{T}$$

is also

$$\mathbf{T}_h(f) = \mathbf{val}\{h g + [Id + h Q] \circ f\}$$

which can be written as

$$(14) \quad \mathbf{T}_h(f) = \mathbf{val}\{h g + P_h \circ f\}$$

with: $P_h = Id + hQ$.

4. STOCHASTIC GAMES WITH VARYING STAGE DURATION

Let $G = (g, Q)$ be a stochastic game as above. We consider here a continuous time Markov chain associated to the kernel Q .

Explicitly, define $P^t(i, j)$ as the continuous time Markov chain on Ω , indexed by \mathbb{R}^+ , with generator $Q(i, j)$:

$$(15) \quad \dot{P}^t(i, j) = P^t(i, j)Q(i, j).$$

One introduces two families of varying duration games, see Neyman [9], associated to $G = (g, Q)$.

4.1. Discretization.

Given a stepsize $h \in (0, 1]$, \overline{G}^h has to be considered as the discretization with mesh h of the game in continuous time \overline{G} where the state variable follows P^t and is controlled by both players, see [21], [17], [3], [8].

More precisely the players act at time $s = kh$ by choosing actions (i, j) (at random according to x , resp. y), knowing the current state. Between time s and $s + h$, the state z_t evolves with law P^t following (15) with $Q(i, j)$ and $P^s = Id$.

The associated Shapley operator is

$$\overline{\mathbf{T}}_h(f) = \mathbf{val}_{X \times Y} \{g^h + P^h \circ f\}.$$

where $g^h(\omega_0, x, y)$ stands for $\mathbb{E}[\int_0^h g(\omega_t; x, y) dt]$ and $P^h(x, y) = \int_{I \times J} P^h(i, j) x(di) y(dj)$.

4.2. Exact sequence.

One can also define an “exact” game G^h , see Neyman [9], with stage duration h and stage transition $P_h = Id + h Q$. Hence $G^h = (h g, h Q)$.

Then one has:

Proposition 4.1.

If \mathbf{T} is the Shapley operator of G , \mathbf{T}_h is the Shapley operator of the game G^h .

Proof. The one stage operator associated to the game G^h , say B_h , satisfies

$$B_h(f)(\omega) = \text{val}_{X \times Y} \{h g(\omega; x, y) + P_h(x, y)[\omega] \circ f\}$$

hence (14) gives the result. \square

5. FINITELY REPEATED GAME

We consider the exact game G^h .

5.1. Approximation.

The recursive equation for the unnormalized value V_n^h of the n -stage game G^h is

$$V_n^h(\omega) = \text{val}[h g(\omega; \cdot) + P_h(\cdot)[\omega] \circ V_{n-1}^h]$$

so that

$$V_n^h = \mathbf{T}_h V_{n-1}^h = \mathbf{T}_h^n(0)$$

Let f be the solution of (1) with \mathbf{T} satisfying (12). Using the results of Section 1 we obtain:

Proposition 5.1.

$$\|V_n^h - f_{nh}(0)\| \leq Lh\sqrt{n}.$$

5.2. Vanishing step sizes.

The normalized value of the game G^h with length $N = nh$, $\frac{V_n^h}{N}$ satisfies:

$$\left\| \frac{V_n^h}{N} - \frac{f_N(0)}{N} \right\| \leq L\sqrt{\frac{h}{N}}.$$

In fact Proposition 2.2 gives a more precise result, that we now describe.

We say that \bar{V}_t is the limit value of the continuous time game on $[0, t]$ if for any sequence of partitions of $[0, t]$ with vanishing mesh, the sequence of values of the corresponding games converges to \bar{V}_t .

Proposition 5.2.

For any h_1, \dots, h_n less than h , the unnormalized value U of the game with n stages of duration h_1, \dots, h_n satisfies

$$\|U - f_{(h_1+\dots+h_n)}(0)\| \leq L'\sqrt{h(h_1 + \dots + h_n)}.$$

The (unnormalized) limit value \bar{V}_t exists and satisfies: $\bar{V}_t = f_t(0)$.

Proof. The inequality is obtained from equation (10).

The existence of \bar{V}_t follows. \square

Note that \bar{V}_t equals also the value of the continuous time game of duration t introduced in Neyman [8].

5.3. Asymptotic analysis.

An alternative consequence of Proposition 5.2 is that for any t and any n -stage game with total duration t and normalized value u_t one has:

Proposition 5.3.

$$\|u_t - \frac{f_t(0)}{t}\| \leq \frac{L'}{\sqrt{t}}$$

In particular the asymptotic average behavior of the value of the game depends only on its duration t up to a term $O(\frac{1}{\sqrt{t}})$.

6. DISCOUNTED GAME

6.1. Discounted exact game.

For any $l \in \ell^1(\mathbb{R}^+, \mathbb{R}^+)$, and any sequence h_1, \dots, h_n, \dots in $[0, 1]$ with $\sum h_n = +\infty$, one defines the game with payoff evaluation l and stage duration sequence $(h_n)_{n \in \mathbb{N}}$ where the players move at time $t_0 = 0, t_1 = h_1, \dots, t_n = h_1 + \dots + h_n, \dots$. Between t_{n-1} and t_n the payoff has a weight $\int_{t_{n-1}}^{t_n} l(t) dt$ and transition occurs with rate h_n . The value of such a game thus satisfies a recursive equation

$$(16) \quad W(t_{n-1})(\omega) = \text{val}_{X \times Y} \left[\left(\int_{t_{n-1}}^{t_n} l(t) dt \right) g(\omega, x, y) + P_{h_n}(x, y)[\omega] \circ W(t_n) \right].$$

In the specific case of a discounted evaluation $l(t) = e^{-kt}$, $k > 0$, and a uniform stage length $h \in [0, 1]$, the stationarity of the model yields a fixed point equation

$$W_k^h(\omega) = \text{val}_{X \times Y} \left[\frac{1 - e^{-kh}}{k} g(\omega, x, y) + e^{-kh} P_h(x, y)[\omega] \circ W_k^h \right].$$

Remark 6.1. For any k and h define the stage discount factor λ by $1 - \lambda h = e^{-kh}$. Then by uniqueness of the fixed point of the above equation, $kW_k^h = \lambda V_\lambda^h$ where

$$V_\lambda^h(\omega) = \text{val}_{X \times Y} [hg(\omega, x, y) + (1 - h\lambda)P_h(x, y)[\omega] \circ V_\lambda^h].$$

In particular, for $h = 1$ we recover V_λ (see 2.5) associated to \mathbf{T} defined in (12).

Hence our definition coincides (up to normalization) with the one of Neyman (eq. (3) p. 254 in [9]). Observe that $\lambda \sim k$ as h goes to 0.

Proposition 6.1.

$$kW_k^h = \lambda V_\lambda^h = \mu V_\mu, \quad \text{with} \quad \mu = \frac{1 - e^{-kh}}{1 - (1 - h)e^{-kh}} = \frac{\lambda}{1 + \lambda - \lambda h}.$$

Proof. By definition,

$$\begin{aligned} W_k^h &= \frac{1 - e^{-kh}}{hk} \mathbf{T}_h \left(\frac{khe^{-kh}}{1 - e^{-kh}} W_k^h \right) \\ &= \frac{1 - e^{-kh}}{hk} \left[(1 - h) \frac{khe^{-kh}}{1 - e^{-kh}} W_k^h + h \mathbf{T} \left(\frac{khe^{-kh}}{1 - e^{-kh}} W_k^h \right) \right]. \end{aligned}$$

Hence

$$\begin{aligned} \frac{k(1 - (1 - h)e^{-kh})}{1 - e^{-kh}} W_k^h &= \mathbf{T} \left(\frac{khe^{-kh}}{1 - e^{-kh}} W_k^h \right) \\ &= \mathbf{T} \left((1 - \mu) \frac{k(1 - (1 - h)e^{-kh})}{1 - e^{-kh}} W_k^h \right). \end{aligned}$$

Uniqueness of the fixed point of $\mathbf{T}(1 - \mu) \cdot$ yields the result. \square

Similar covariance properties were obtained in [9].

6.2. Vanishing duration.

We now recover the convergence property in [9].

Corollary 6.1.

For a fixed k , W_k^h converges as h goes to 0. The limit, denoted \overline{W}_k , equals $\frac{1}{1+k}V_{\frac{k}{1+k}}$, hence is the only solution of:

$$(17) \quad kF = \text{val}[g + Q \circ F]$$

Proof. Immediate consequence of Proposition 6.1 since $e^{-kh} = 1 - kh + o(h)$ as h goes to 0. \square

Once again \overline{W}_k has to be interpreted as the k -discounted value of the continuous time game \overline{G} , see [4], [8].

6.3. Asymptotic behavior.

Finally we show for any h the normalized value $w_k^h = kW_k^h$ has the same asymptotic behavior (as k tends to 0) as v_k .

Proposition 6.2.

$$\|w_k^h - v_k\| \leq Ck\left(1 + \frac{k}{2}\right).$$

Proof. Let $\mu = \frac{1-e^{-kh}}{1-(1-h)e^{-kh}}$. By Proposition 2.3 and Proposition 6.1,

$$\begin{aligned} \left\|w_k^h - v_k\right\| &= \|v_\mu - v_k\| \\ &\leq C\left\|1 - \frac{\mu}{k}\right\| \\ &= C\frac{k-1 + (1-k+kh)e^{-kh}}{k - k(1-h)e^{-kh}} \\ &\leq C\frac{k-1 + (1-k+kh)(1-kh+k^2h^2/2)}{kh} \\ &= C\left(k - \frac{kh}{2} - \frac{k^2h}{2} + \frac{h^2k^2}{2}\right) \\ &\leq C\left(k + \frac{k^2}{2}\right). \end{aligned}$$

\square

In particular, letting h go to 0, we get that $\overline{w}_k = k\overline{W}_k$ has the same asymptotic behavior as v_k .

7. DISCRETIZATION OF THE CONTINUOUS TIME GAME

We consider now the game \overline{G}^h (with either a fixed finite horizon or a fixed discount factor) as h goes to 0.

7.1. Finite horizon.

The value \overline{V}_n^h of the n -stage game with stage duration h satisfies $\overline{V}_n^h = (\overline{\mathbf{T}}_h)^n(0)$. Similarly for varying stage duration, corresponding to a partition Π , one gets a recursive equation $\overline{V}_\Pi(t_n) = \overline{\mathbf{T}}_{h_n}[\overline{V}_\Pi(t_{n+1})]$.

Lemma 7.1.

$$\|\mathbf{T}_h(f) - \overline{\mathbf{T}}_h(f)\| \leq (1 + \|f\|)h O(h).$$

Proof. By nonexpansiveness of the value operator,

$$\begin{aligned} \|\mathbf{T}_h(f) - \overline{\mathbf{T}}_h(f)\| &\leq \|hg(\cdot) - \int_0^h g_t(\cdot)dt\| + \|f\|\|P_h(\cdot) - P^h(\cdot)\| \\ &= h O(h) + \|f\|h O(h) \end{aligned}$$

since $P_h = Id + hQ$ and $\mathbf{P}^h = e^{hQ} = Id + hQ + hO(h)$. \square

Proposition 7.1.

There exists C depending only of the game G such that for any finite sequence $(h_i)_{i \leq n}$ in $[0, h]$ with sum t and corresponding partition Π :

$$\|\bar{V}_\Pi(t) - \bar{V}_t\| \leq C(\sqrt{ht} + ht + ht^2).$$

In particular for a given t , $\bar{V}_\Pi(t)$ tends to \bar{V}_t as h goes to 0.

Proof. The value of any game with total length less than t is bounded by Ct , independently of h . Hence nonexpansiveness of the operators as well as the previous Lemma gives

$$\begin{aligned} \|\bar{V}_\Pi(t) - \Pi_i \mathbf{T}_{h_i}(0)\| &= \|\Pi_i \bar{\mathbf{T}}_{h_i}(0) - \Pi_i \mathbf{T}_{h_i}(0)\| \\ &\leq \left\| \bar{\mathbf{T}}_{h_1} \left(\Pi_{i \geq 2} \bar{\mathbf{T}}_{h_i}(0) \right) - \bar{\mathbf{T}}_{h_1} \left(\Pi_{i \geq 2} \mathbf{T}_{h_i}(0) \right) \right\| + \left\| \bar{\mathbf{T}}_{h_1} \left(\Pi_{i \geq 2} \mathbf{T}_{h_i}(0) \right) - \mathbf{T}_{h_1} \left(\Pi_{i \geq 2} \mathbf{T}_{h_i}(0) \right) \right\| \\ &\leq \left\| \Pi_{i \geq 2} \bar{\mathbf{T}}_{h_i}(0) - \Pi_{i \geq 2} \mathbf{T}_{h_i}(0) \right\| + Ch_1^2 \left(1 + \sum_{i=2}^n h_i \right). \end{aligned}$$

Hence by summation,

$$\begin{aligned} \|\bar{V}_\Pi(t) - \Pi_i \mathbf{T}_{h_i}(0)\| &\leq C(1+t) \sum_{i=1}^n h_i^2 \\ &\leq Ch(1+t) \sum_{i=1}^n h_i \\ &= Ch t(1+t). \end{aligned}$$

Then Proposition 5.2 yields the result. \square

Remark 7.1. For a given h , the right hand term is quadratic in t , hence we do not link the asymptotic behavior of the normalized quantity $\bar{v}_n^h = \frac{\bar{V}_n^h}{nh}$ and of $\bar{v}_t = \frac{\bar{V}_t}{t}$. However if n is a function of h converging slowly enough to infinity, the previous proposition can be used. For example for $n(h) = \frac{1}{h\sqrt{h}}$ (so that $t(h) = \frac{1}{\sqrt{h}}$), one has

$$\|\bar{v}_{n(h)}^h - \bar{v}_{t(h)}\| = O(\sqrt{h}).$$

7.2. Discounted case.

The evaluation is $l(t) = e^{-kt}$ and we consider uniform stage duration h . The value \bar{W}_k^h of the discretization of the discounted continuous game satisfies the fixed point equation

$$\bar{W}_k^h(\omega) = \mathbf{val}_{X \times Y} \left[\int_0^h e^{-kt} g(\omega_t, x, y) + e^{-kh} \mathbf{P}^h(x, y)[\omega] \circ \bar{W}_k^h \right].$$

Proposition 7.2.

For a given k , \bar{W}_k^h tends to \bar{W}_k as h goes to 0.

Proof. The equations for \bar{W}_k^h and W_k^h , as well as the nonexpansiveness of the value operator, give:

$$\begin{aligned} \|\bar{W}_k^h - W_k^h\| &\leq h \left\| \frac{1 - e^{-kh}}{kh} g(\cdot) - \frac{1}{h} \int_0^h e^{-kt} g_t(\cdot) \right\| + e^{-kh} \|\bar{W}_k^h - W_k^h\| + \|W_k^h\| \|P_h - \mathbf{P}^h\| \\ &\leq h O(h) + e^{-kh} \|\bar{W}_k^h - W_k^h\| + \frac{1}{k} h O(h) \end{aligned}$$

hence $(1 - e^{-kh}) \|\bar{W}_k^h - W_k^h\| = h O(h)$ and the result follows from Corollary 6.1. \square

Similar properties were obtained in [9] using the fact that \overline{W}_k satisfies (17).

For an alternative approach to the limit behavior of the discretization of the continuous model, relying on viscosity solution tools and extending to various information structures on the state, see [16].

8. EXTENSIONS AND CONCLUDING COMMENTS

8.1. Extensions.

Leaving the hypothesis Ω finite, one can consider two frameworks (see [5], Prop VII.1.4 and VII.1.5) where \mathbf{T} is well defined and X is the space of bounded measurable (resp. continuous) functions on Ω .

All the results concerning exact games G^h (sections 5 and 6) go through but the study of the Markov process \mathbf{P}^t (section 7) deserves further study.

8.2. Stochastic games: no signals on the state.

Consider now a finite stochastic game where the players knows only the initial distribution $m \in \Delta(\Omega)$ and the actions at each stage.

The basic equation for the the game with duration h is then

$$\hat{\mathbf{T}}^h f(m) = \text{val}[hg(m; x, y) + \sum_{ij} x^i y^j f(m * P_h(i, j))].$$

with $[m * P_h(i, j)](\omega) = \sum_z m(z) P_h(i, j)[z](\omega)$.

The equation

$$\hat{\mathbf{T}}^h = h\hat{\mathbf{T}} + (1 - h)Id$$

does not hold anymore and $\mathbf{T} - Id = -A$ has to be replaced by $\lim_{h \rightarrow 0} \frac{\hat{\mathbf{T}}^h - Id}{h}$ in (1). The study of such games with varying duration thus seems more involved.

8.3. Link with games with uncertain duration.

Equation (16) can be written as $W(t_{n-1}) = \frac{\alpha_n}{h_n} \mathbf{T}_h \left(\frac{h_n W(t_n)}{\alpha_n} \right)$ with $\alpha_n := \int_{t_{n-1}}^{t_n} l(t) dt$. When l has a compact support or the sequence h_n is finite, this yields $W(0) = \Psi(0)$, where Ψ is some generalized iterate [7, 11] of \mathbf{T} . Hence W can also be seen as the value of some game with uncertain duration. See [19] for specific remarks in the particular case of V_n^h .

8.4. Oscillations.

Several examples of stochastic games (either with a finite set of states and compact sets of actions [20], or compact set of states and finite set of actions [22]) were recently constructed for which the values v_n and v_λ do not converge. Hence the values of the corresponding continuous games do not converge when t goes to infinity or k to 0.

8.5. Comparison to the literature.

The approach here is different from the one of Neyman [9] : the proofs are based on properties of operators and not on strategies. We recover some of the results in [9] and extend them to more general frameworks (compact action or states spaces, varying duration).

8.6. Main results.

The main results can be summarized in two parts:

- for a given finite horizon (or discounted evaluation) the value of the game with vanishing stage duration converges thus defining a limit value for the associated continuous time game. Moreover the limit is described explicitly.
- as the horizon goes to ∞ the impact of the stage duration goes to 0 and the asymptotic behavior of the normalized value function is independent of the discretization. In the case of a uniform stage duration the same holds for discounted games as the discount factor goes to 0. Moreover there is an explicit formula.

REFERENCES

- [1] Brézis H. (1973) Opérateurs maximaux monotones, North-Holland.
- [2] Brézis H. and A. Pazy (1970) Accretive sets and differential equations in Banach spaces, *Israel Journal of Mathematics*, **8**, 367-383.
- [3] Guo X. and O. Hernandez-Lerma (2003) Zero-sum games for continuous-time Markov chains with unbounded transition and average payoff rates, *Journal of Applied Probability*, **40**, 327-345.
- [4] Guo X. and O. Hernandez-Lerma (2005) Zero-sum continuous-time Markov games with unbounded transition and discounted payoff rates, *Bernoulli*, **11**, 1009-1029.
- [5] Mertens J.-F., S. Sorin and S. Zamir (2015) *Repeated Games*, Cambridge University Press.
- [6] Miyadera I. and S. Oharu (1970) Approximation of semi-groups of nonlinear operators, *Tôhoku Math. Journal*, **22**, 24-47.
- [7] Neyman A. (2003) Stochastic games and nonexpansive maps, *Stochastic Games and Applications*, Neyman A. and S. Sorin (eds.), NATO Science Series C 570, Kluwer Academic Publishers, 397-415.
- [8] Neyman A. (2012) Continuous-time stochastic games, DP 616, CSR Jerusalem.
- [9] Neyman A. (2013) Stochastic games with short-stage duration, *Dynamic Games and Applications*, **3**, 236-278.
- [10] Neyman A. and S. Sorin (eds.) (2003) *Stochastic Games and Applications*, NATO Science Series, **C 570**, Kluwer Academic Publishers.
- [11] Neyman A. and S. Sorin (2010) Repeated games with public uncertain duration process, *International Journal of Game Theory*, **39**, 29-52.
- [12] Prieto-Rumeau T., Hernandez-Lerma O. (2012) *Selected Topics on Continuous-Time Controlled Markov Chains and Markov Games*, Imperial College Press.
- [13] Rosenberg D. and S. Sorin (2001) An operator approach to zero-sum repeated games, *Israel Journal of Mathematics*, **121**, 221- 246.
- [14] Shapley L. S. (1953) Stochastic games, *Proceedings of the National Academy of Sciences of the U.S.A*, **39**, 1095-1100.
- [15] Sorin S. (2003) The operator approach to zero-sum stochastic games, *Stochastic Games and Applications*, Neyman A. and S. Sorin (eds.), NATO Science Series C 570, Kluwer Academic Publishers, 375-395.
- [16] Sorin S. (2015) Limit value of dynamic zero-sum games with vanishing stage duration, preprint.
- [17] Tanaka K. and K. Wakuta (1977) On continuous Markov games with the expected average reward criterion, *Sci. Rep. Niigata Univ. Ser. A*, **14**, 15-24.
- [18] Vigeral G. (2009) Propriétés asymptotiques des jeux répétés à somme nulle, PhD Thesis, UPMC-Paris 6.
- [19] Vigeral G. (2010a) Evolution equations in discrete and continuous time for non expansive operators in Banach spaces, *ESAIM COCV*, **16**, 809-832.
- [20] Vigeral G. (2013) A zero-sum stochastic game with compact action sets and no asymptotic value, *Dynamic Games and Applications*, **3**, 172-186.
- [21] Zachrisson L.E. (1964) Markov Games, *Advances in Game Theory*, Dresher M., L. S. Shapley and A.W. Tucker (eds), Annals of Mathematical Studies, **52**, Princeton University Press, 210-253.
- [22] Ziliotto B. (2013) Zero-sum repeated games: counterexamples to the existence of the asymptotic value and the conjecture $\max\min = \limv_n$, preprint hal-00824039, to appear in *Annals of Probability*.

Sylvain Sorin, SORBONNE UNIVERSITÉS, UPMC UNIV PARIS 06, INSTITUT DE MATHÉMATIQUES DE JUSSIEU-PARIS RIVE GAUCHE, UMR 7586, CNRS, UNIV PARIS DIDEROT, SORBONNE PARIS CITÉ, F-75005, PARIS, FRANCE

E-mail address: `sylvain.sorin@imj-prg.fr`

Guillaume Vigeral, UNIVERSITÉ PARIS-DAUPHINE, CEREMADE, PLACE DU MARÉCHAL DE LATTRE DE TASSIGNY, 75775 PARIS CEDEX 16, FRANCE

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

<http://www.ceremade.dauphine.fr/vigeral/indexenglish.html>