Constant payoff in zero-sum stochastic games

Bruno Ziliotto (CNRS, Paris Dauphine University)

joint work with Miquel Oliu Barton (Paris Dauphine University)

November 14th, 2017

Stochastic game model (Shapley 1953)

Two adversary players interact repeatedly, in an environment that can evolve along the game. Each possible environment is called a *state*. At each stage :

- Players choose simultaneously an action,
- The actions and the current state determine the stage payoff,
- A new state is randomly drawn from a distribution determined by the current state and the actions,
- Players know the current state and past actions.

Stochastic game model (Shapley 1953)

Two adversary players interact repeatedly, in an environment that can evolve along the game. Each possible environment is called a *state*. At each stage :

- Players choose simultaneously an action,
- The actions and the current state determine the stage payoff,
- A new state is randomly drawn from a distribution determined by the current state and the actions,
- Players know the current state and past actions.

Stochastic game model (Shapley 1953)

Two adversary players interact repeatedly, in an environment that can evolve along the game. Each possible environment is called a *state*. At each stage :

- Players choose simultaneously an action,
- The actions and the current state determine the stage payoff,
- A new state is randomly drawn from a distribution determined by the current state and the actions,
- Players know the current state and past actions.

Stochastic game model (Shapley 1953)

Two adversary players interact repeatedly, in an environment that can evolve along the game. Each possible environment is called a *state*. At each stage :

- Players choose simultaneously an action,
- The actions and the current state determine the stage payoff,
- A new state is randomly drawn from a distribution determined by the current state and the actions,
- Players know the current state and past actions.

Two adversary players interact repeatedly, in an environment that can evolve along the game. Each possible environment is called a *state*. At each stage :

- Players choose simultaneously an action,
- The actions and the current state determine the stage payoff,
- A new state is randomly drawn from a distribution determined by the current state and the actions,
- Players know the current state and past actions.

A stochastic game is described by the following elements :

- State space K,
- Action set I (resp. J) for Player 1 (resp. 2),
- Transition function $q : K \times I \times J \to \Delta(K)$,
- Payoff function $g : K \times I \times J \to \mathbb{R}$.

The sets K, I and J are assumed to be finite.

A stochastic game is described by the following elements :

- State space K,
- Action set I (resp. J) for Player 1 (resp. 2),
- Transition function $q : K \times I \times J \to \Delta(K)$,
- Payoff function $g : K \times I \times J \to \mathbb{R}$.

The sets K, I and J are assumed to be finite.

In the game starting from k, at each stage $m \geq 1$, the play is as follows :

- Player 1 (resp. 2) chooses $i_m \in I$ (resp. $j_m \in J$)
- The stage payoff is $g_m := g(k_m, i_m, j_m)$
- Nature draws k_{m+1} from the probability distribution $q(k_m, i_m, j_m)$
- (k_{m+1}, i_m, j_m) is publicly announced to the players.

In the game starting from k, at each stage $m \geq 1$, the play is as follows :

- Player 1 (resp. 2) chooses $i_m \in I$ (resp. $j_m \in J$)
- The stage payoff is $g_m := g(k_m, i_m, j_m)$
- Nature draws k_{m+1} from the probability distribution $q(k_m, i_m, j_m)$
- (k_{m+1}, i_m, j_m) is publicly announced to the players.

In the game starting from k, at each stage $m \geq 1$, the play is as follows :

- Player 1 (resp. 2) chooses $i_m \in I$ (resp. $j_m \in J$)
- The stage payoff is $g_m := g(k_m, i_m, j_m)$
- Nature draws k_{m+1} from the probability distribution $q(k_m, i_m, j_m)$
- (k_{m+1}, i_m, j_m) is publicly announced to the players.

In the game starting from k, at each stage $m \geq 1$, the play is as follows :

- Player 1 (resp. 2) chooses $i_m \in I$ (resp. $j_m \in J$)
- The stage payoff is $g_m := g(k_m, i_m, j_m)$
- Nature draws k_{m+1} from the probability distribution $q(k_m, i_m, j_m)$
- (k_{m+1}, i_m, j_m) is publicly announced to the players.

In the game starting from k, at each stage $m \geq 1$, the play is as follows :

- Player 1 (resp. 2) chooses $i_m \in I$ (resp. $j_m \in J$)
- The stage payoff is $g_m := g(k_m, i_m, j_m)$
- Nature draws k_{m+1} from the probability distribution $q(k_m, i_m, j_m)$
- (k_{m+1}, i_m, j_m) is publicly announced to the players.

- State space $K := \{\omega, 0^*, 1^*\}$
- Action sets $I := \{Top, Bottom\}$ and $J := \{Left, Right\}$

FIGURE: Transition and payoff functions in state ω

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

Strategies

A *strategy* for Player 1 (resp. 2) assigns to each stage m and each history $(k, i_1, j_1, k_2, i_2, j_2, ..., k_{m-1}, i_{m-1}, j_{m-1}, k_m)$ an element $x \in \Delta(I)$ (resp. $y \in \Delta(J)$).

The set of strategies for Player 1 (resp. 2) is denoted by Σ (resp. $\mathcal{T}$).

Probability measure on the set of all possible histories

$(k, \sigma, \tau) \in K \times \Sigma \times \mathcal{T}$ naturally induces a probability measure $\mathbb{P}^k_{\sigma,\tau}$ on the set of all infinite histories of the game $\{(k_1, i_1, j_1, k_2, i_2, j_2, ..., k_m, i_m, j_m, ...)\}$

Strategies

A *strategy* for Player 1 (resp. 2) assigns to each stage m and each history $(k, i_1, j_1, k_2, i_2, j_2, ..., k_{m-1}, i_{m-1}, j_{m-1}, k_m)$ an element $x \in \Delta(I)$ (resp. $y \in \Delta(J)$).

The set of strategies for Player 1 (resp. 2) is denoted by Σ (resp. $\mathcal{T}$).

Probability measure on the set of all possible histories

$(k, \sigma, \tau) \in K \times \Sigma \times \mathcal{T}$ naturally induces a probability measure $\mathbb{P}^k_{\sigma,\tau}$ on the set of all infinite histories of the game $\{(k_1, i_1, j_1, k_2, i_2, j_2, ..., k_m, i_m, j_m, ...)\}$

Strategies

A *strategy* for Player 1 (resp. 2) assigns to each stage m and each history $(k, i_1, j_1, k_2, i_2, j_2, ..., k_{m-1}, i_{m-1}, j_{m-1}, k_m)$ an element $x \in \Delta(I)$ (resp. $y \in \Delta(J)$).

The set of strategies for Player 1 (resp. 2) is denoted by Σ (resp. $\mathcal{T}$).

Probability measure on the set of all possible histories

$(k, \sigma, \tau) \in K \times \Sigma \times \mathcal{T}$ naturally induces a probability measure $\mathbb{P}^k_{\sigma,\tau}$ on the set of all infinite histories of the game $\{(k_1, i_1, j_1, k_2, i_2, j_2, ..., k_m, i_m, j_m, ...)\}$

Let $\lambda \in (0,1]$. The λ-*discounted game* is the game with strategy sets Σ and $\mathcal{T}$, and payoff defined by

$$\forall(\sigma,\tau) \in \Sigma \times \mathcal{T} \quad \gamma_\lambda^k(\sigma,\tau) = \mathbb{E}_{\sigma,\tau}^k \left(\sum_{m \geq 1} \lambda(1-\lambda)^{m-1} g_m \right).$$

The λ-discounted game has a value, denoted by $v_\lambda(k)$:

$$v_\lambda(k) = \max_{\sigma \in \Sigma} \min_{\tau \in \mathcal{T}} \gamma_\lambda^k(\sigma,\tau) = \min_{\tau \in \mathcal{T}} \max_{\sigma \in \Sigma} \gamma_\lambda^k(\sigma,\tau)$$

Let $\lambda \in (0,1]$. The λ-*discounted game* is the game with strategy sets Σ and $\mathcal{T}$, and payoff defined by

$$\forall(\sigma,\tau) \in \Sigma \times \mathcal{T} \quad \gamma_\lambda^k(\sigma,\tau) = \mathbb{E}_{\sigma,\tau}^k \left(\sum_{m \geq 1} \lambda(1-\lambda)^{m-1} g_m \right).$$

The λ-discounted game has a value, denoted by $v_\lambda(k)$:

$$v_\lambda(k) = \max_{\sigma \in \Sigma} \min_{\tau \in \mathcal{T}} \gamma_\lambda^k(\sigma,\tau) = \min_{\tau \in \mathcal{T}} \max_{\sigma \in \Sigma} \gamma_\lambda^k(\sigma,\tau)$$

Shapley equation

$$v_\lambda(k) = \underset{x\in\Delta(I),\ y\in\Delta(J)}{\mathrm{Val}} \left\{ \lambda g(k,x,y) + (1-\lambda)\mathbb{E}^k_{x,y}\left[v_\lambda(\tilde{k})\right] \right\}$$

Shapley equation is notably useful to :

- Compute v_λ and optimal strategies,
- Prove theoretical properties, like the existence of optimal stationary strategies (strategies that only depend on the current state).

$$v_\lambda(k) = \operatorname*{Val}_{x\in\Delta(I),\ y\in\Delta(J)} \left\{ \lambda g(k,x,y) + (1-\lambda)\mathbb{E}^k_{x,y}\left[v_\lambda(\tilde{k})\right] \right\}$$

Shapley equation is notably useful to :

- Compute v_λ and optimal strategies,
- Prove theoretical properties, like the existence of optimal stationary strategies (strategies that only depend on the current state).

FIGURE: Transition and payoff functions in state ω

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

- $v_\lambda(\omega) = \frac{1}{2}$.

- Optimal (stationary) strategy for Player 1 and 2 in $\Gamma_\lambda(\omega)$: $\sigma_\lambda = \frac{\lambda}{1+\lambda} \cdot T + \frac{1}{1+\lambda} \cdot B$ and $\tau_\lambda = \frac{1}{2} \cdot L + \frac{1}{2} \cdot R$.

- Assume that players are very patient, and play optimal strategies... What will be the outcome of the game ?

- Patient players=λ close to 0.

- Does (v_λ) converge when λ goes to 0 ? When the answer is positive, the game is said to have a *limit value*.

Bewley-Kohlberg 1976

Any stochastic game has a limit value $v^* : K \to \mathbb{R}$.

- Assume that players are very patient, and play optimal strategies... What will be the outcome of the game ?

- Patient players=λ close to 0.

- Does (v_λ) converge when λ goes to 0 ? When the answer is positive, the game is said to have a *limit value*.

Bewley-Kohlberg 1976

Any stochastic game has a limit value $v^* : K \to \mathbb{R}$.

- Assume that players are very patient, and play optimal strategies... What will be the outcome of the game ?
- Patient players=λ close to 0.
- Does (v_λ) converge when λ goes to 0 ? When the answer is positive, the game is said to have a *limit value*.

Bewley-Kohlberg 1976

Any stochastic game has a limit value $v^* : K \to \mathbb{R}$.

- Assume that players are very patient, and play optimal strategies... What will be the outcome of the game ?
- Patient players=λ close to 0.
- Does (v_λ) converge when λ goes to 0 ? When the answer is positive, the game is said to have a *limit value*.

Bewley-Kohlberg 1976

Any stochastic game has a limit value $v^* : K \to \mathbb{R}$.

Constant payoff

- In a λ-discounted game, when players are patient and play optimally, Player 1 gets approximately payoff $v^*(k)$, and Player 2 gets payoff $-v^*(k)$.

- But maybe Player 1 gets a payoff way better than $v^*(k)$ at the beginning of the game, then much worse than $v^*(k)$ in the middle of the game, etc.

- In many situations, Player 1 would prefer not being ruined at any time of the game !

Question : do players approximately obtain the same payoff at any point in the game ?

Constant payoff

- In a λ-discounted game, when players are patient and play optimally, Player 1 gets approximately payoff $v^*(k)$, and Player 2 gets payoff $-v^*(k)$.
- But maybe Player 1 gets a payoff way better than $v^*(k)$ at the beginning of the game, then much worse than $v^*(k)$ in the middle of the game, etc.
- In many situations, Player 1 would prefer not being ruined at any time of the game !

Question : do players approximately obtain the same payoff at any point in the game ?

Constant payoff

- In a λ-discounted game, when players are patient and play optimally, Player 1 gets approximately payoff $v^*(k)$, and Player 2 gets payoff $-v^*(k)$.
- But maybe Player 1 gets a payoff way better than $v^*(k)$ at the beginning of the game, then much worse than $v^*(k)$ in the middle of the game, etc.
- In many situations, Player 1 would prefer not being ruined at any time of the game !

Question : do players approximately obtain the same payoff at any point in the game ?

- In a λ-discounted game, when players are patient and play optimally, Player 1 gets approximately payoff $v^*(k)$, and Player 2 gets payoff $-v^*(k)$.
- But maybe Player 1 gets a payoff way better than $v^*(k)$ at the beginning of the game, then much worse than $v^*(k)$ in the middle of the game, etc.
- In many situations, Player 1 would prefer not being ruined at any time of the game !

Question : do players approximately obtain the same payoff at any point in the game ?

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

- $v_\lambda(\omega) = \frac{1}{2}$.
- Optimal (stationary) strategy for Player 1 and 2 in $\Gamma_\lambda(\omega)$:
  $\sigma_\lambda = \frac{\lambda}{1+\lambda} \cdot T + \frac{1}{1+\lambda} \cdot B$ and $\tau_\lambda = \frac{1}{2} \cdot L + \frac{1}{2} \cdot R$.
- When Player 1 plays σ_λ and Player 2 plays τ_λ, the expected stage payoff is always $1/2$.
- When Player 1 plays σ_λ and Player 2 always plays L : the payoff is close to 0 at the beginning of the game, and close to 1 at the end of the game.

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

- $v_\lambda(\omega) = \dfrac{1}{2}$.

- Optimal (stationary) strategy for Player 1 and 2 in $\Gamma_\lambda(\omega)$:
  $\sigma_\lambda = \dfrac{\lambda}{1+\lambda} \cdot T + \dfrac{1}{1+\lambda} \cdot B$ and $\tau_\lambda = \dfrac{1}{2} \cdot L + \dfrac{1}{2} \cdot R$.

- When Player 1 plays σ_λ and Player 2 plays τ_λ, the expected stage payoff is always $1/2$.

- When Player 1 plays σ_λ and Player 2 always plays L : the payoff is close to 0 at the beginning of the game, and close to 1 at the end of the game.

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

- $v_\lambda(\omega) = \dfrac{1}{2}$.

- Optimal (stationary) strategy for Player 1 and 2 in $\Gamma_\lambda(\omega)$:
  $\sigma_\lambda = \dfrac{\lambda}{1+\lambda} \cdot T + \dfrac{1}{1+\lambda} \cdot B$ and $\tau_\lambda = \dfrac{1}{2} \cdot L + \dfrac{1}{2} \cdot R$.

- When Player 1 plays σ_λ and Player 2 plays τ_λ, the expected stage payoff is always $1/2$.

- When Player 1 plays σ_λ and Player 2 always plays L : the payoff is close to 0 at the beginning of the game, and close to 1 at the end of the game.

Constant payoff : formal definition

For all $t \in [0,1]$ and discount λ, let

$$m(t,\lambda) := \inf \left\{ M \geq 1, \sum_{m=1}^{M} \lambda(1-\lambda)^{m-1} \geq t \right\}$$

When λ is small, the interpretation is the following : at stage $m(t,\lambda)$, "a fraction t" of the game has been played.

Definition

The payoff is *constant* if for any $\epsilon > 0$, there exists $\lambda_0 \in (0,1]$ such that for any $\lambda \leq \lambda_0$, for any pair of strategies $(\sigma_\lambda, \tau_\lambda)$ that are optimal in the λ-discounted game, for any $t \in [0,1]$, we have

$$\left| \mathbb{E}^k_{\sigma_\lambda,\tau_\lambda} \left[\sum_{m=1}^{m(t,\lambda)} \lambda(1-\lambda)^{m-1} g_m \right] - tv^*(k) \right| \leq \epsilon.$$

Interpretation : after a fraction t of the game, the total payoff is approximately t times the optimal payoff.

Constant payoff : formal definition

For all $t \in [0,1]$ and discount λ, let

$$m(t,\lambda) := \inf \left\{ M \geq 1, \sum_{m=1}^{M} \lambda(1-\lambda)^{m-1} \geq t \right\}$$

When λ is small, the interpretation is the following : at stage $m(t,\lambda)$, "a fraction t" of the game has been played.

Definition

The payoff is *constant* if for any $\epsilon > 0$, there exists $\lambda_0 \in (0,1]$ such that for any $\lambda \leq \lambda_0$, for any pair of strategies $(\sigma_\lambda, \tau_\lambda)$ that are optimal in the λ-discounted game, for any $t \in [0,1]$, we have

$$\left| \mathbb{E}^{k}_{\sigma_\lambda,\tau_\lambda} \left[\sum_{m=1}^{m(t,\lambda)} \lambda(1-\lambda)^{m-1} g_m \right] - t v^*(k) \right| \leq \epsilon.$$

Interpretation : after a fraction t of the game, the total payoff is approximately t times the optimal payoff.

Constant payoff : formal definition

For all $t \in [0,1]$ and discount λ, let

$$m(t,\lambda) := \inf \left\{ M \geq 1,\ \sum_{m=1}^{M} \lambda(1-\lambda)^{m-1} \geq t \right\}$$

When λ is small, the interpretation is the following : at stage $m(t,\lambda)$, "a fraction t" of the game has been played.

Definition

The payoff is *constant* if for any $\epsilon > 0$, there exists $\lambda_0 \in (0,1]$ such that for any $\lambda \leq \lambda_0$, for any pair of strategies $(\sigma_\lambda, \tau_\lambda)$ that are optimal in the λ-discounted game, for any $t \in [0,1]$, we have

$$\left| \mathbb{E}^k_{\sigma_\lambda,\tau_\lambda} \left[\sum_{m=1}^{m(t,\lambda)} \lambda(1-\lambda)^{m-1} g_m \right] - t v^*(k) \right| \leq \epsilon.$$

Interpretation : after a fraction t of the game, the total payoff is approximately t times the optimal payoff.

The Sphynx

Tell me, stranger, what creature walks
with four legs in the morning
with two legs at noon
and with three legs at night ?

Oedipus

A man !

The Sphynx

Tell me, stranger, what creature walks
with four legs in the morning
with two legs at noon
and with three legs at night ?

Oedipus

A man !

The Sphynx

Tell me, game theorist, how much do you get
in (a long) stochastic game
in the morning
at noon
at night ?

The game theorist

Good question, Sphynx !
Until recently this was open...

The Sphynx

Tell me, game theorist, how much do you get
in (a long) stochastic game
in the morning
at noon
at night ?

The game theorist

Good question, Sphynx !
Until recently this was open...

The following models have constant payoff :

- Repeated games with incomplete information on one side (Aumann and Maschler 1967),
- 1-Player stochastic games=Markov Decision Processes (Sorin, Venel and Vigeral, 2010),
- Absorbing games (Oliu-Barton 2015, Sorin and Vigeral 2017).

Oliu-Barton and Z. 2017

In any stochastic game, the payoff is constant : for any $\epsilon > 0$, there exists $\lambda_0 \in (0, 1]$ such that for any $\lambda \leq \lambda_0$, for any pair of strategies $(\sigma_\lambda, \tau_\lambda)$ that are optimal in the λ-discounted game, for any $t \in [0, 1]$, we have

$$\left| \mathbb{E}^k_{\sigma_\lambda, \tau_\lambda} \left[\sum_{m=1}^{m(t,\lambda)} \lambda (1-\lambda)^{m-1} g_m \right] - t v^*(k) \right| \leq \epsilon.$$

- We show the proof for stationary strategies. We assume WLOG that the payoff depends only on the state.

- Fix two families of **stationary** strategies $(\sigma_\lambda)_{\lambda\in(0,1]}$ and $(\tau_\lambda)_{\lambda\in(0,1]}$ such that for any λ, σ_λ (resp. τ_λ) is optimal for Player 1 (resp. 2) in the λ-discounted game.

- We want to prove that for all $k \in K$,

$$\lim_{\lambda\to 0} \mathbb{E}^k_{\sigma_\lambda,\tau_\lambda}\left(\sum_{m=1}^{m(t,\lambda)} \lambda(1-\lambda)^{m-1} g_m\right) = tv^*(k)$$

Sketch of Proof

- We show the proof for stationary strategies. We assume WLOG that the payoff depends only on the state.

- Fix two families of **stationary** strategies $(\sigma_\lambda)_{\lambda \in (0,1]}$ and $(\tau_\lambda)_{\lambda \in (0,1]}$ such that for any λ, σ_λ (resp. τ_λ) is optimal for Player 1 (resp. 2) in the λ-discounted game.

- We want to prove that for all $k \in K$,

$$\lim_{\lambda \to 0} \mathbb{E}^k_{\sigma_\lambda, \tau_\lambda} \left(\sum_{m=1}^{m(t,\lambda)} \lambda (1-\lambda)^{m-1} g_m \right) = t v^*(k)$$

- We show the proof for stationary strategies. We assume WLOG that the payoff depends only on the state.

- Fix two families of **stationary** strategies $(\sigma_\lambda)_{\lambda \in (0,1]}$ and $(\tau_\lambda)_{\lambda \in (0,1]}$ such that for any λ, σ_λ (resp. τ_λ) is optimal for Player 1 (resp. 2) in the λ-discounted game.

- We want to prove that for all $k \in K$,

$$\lim_{\lambda \to 0} \mathbb{E}^k_{\sigma_\lambda, \tau_\lambda} \left(\sum_{m=1}^{m(t,\lambda)} \lambda (1-\lambda)^{m-1} g_m \right) = t v^*(k)$$

Puiseux serie

A function $f:(0,1]\to\mathbb{R}$ is a *Puiseux serie* if it has the following expansion :

$$f(\lambda)=\sum_{m\geq 0} A_m \lambda^{m/N},$$

for some positive integer N and real coefficients (A_m).

Shapley equation

$$v_\lambda(k) \quad = \quad \underset{x\in\Delta(I),\ y\in\Delta(J)}{\mathrm{Val}} \left\{\lambda g(k,x,y)+(1-\lambda)\mathbb{E}^k_{x,y}\left[v_\lambda(k_2)\right]\right\}$$

Shapley equation can be written as a finite number of polynomial inequalities in λ. Consequence :

For λ small enough, for all $i\in I$ and $j\in J$, $v_\lambda(k)$, $\sigma_\lambda(k,i)$ and $\tau_\lambda(k,j)$ can be written as Puiseux series.

Puiseux serie

A function $f : (0,1] \to \mathbb{R}$ is a *Puiseux serie* if it has the following expansion :

$$f(\lambda) = \sum_{m \geq 0} A_m \lambda^{m/N},$$

for some positive integer N and real coefficients (A_m).

Shapley equation

$$v_\lambda(k) \quad = \quad \underset{x \in \Delta(I),\ y \in \Delta(J)}{\text{Val}} \left\{ \lambda g(k,x,y) + (1-\lambda)\mathbb{E}^k_{x,y}\left[v_\lambda(k_2)\right] \right\}$$

Shapley equation can be written as a finite number of polynomial inequalities in λ. Consequence :

For λ small enough, for all $i \in I$ and $j \in J$, $v_\lambda(k)$, $\sigma_\lambda(k,i)$ and $\tau_\lambda(k,j)$ can be written as Puiseux series.

Puiseux serie

A function $f : (0,1] \to \mathbb{R}$ is a *Puiseux serie* if it has the following expansion :

$$f(\lambda) = \sum_{m \geq 0} A_m \lambda^{m/N},$$

for some positive integer N and real coefficients (A_m).

Shapley equation

$$v_\lambda(k) = \underset{x \in \Delta(I),\ y \in \Delta(J)}{\mathrm{Val}} \left\{ \lambda g(k,x,y) + (1-\lambda)\mathbb{E}^k_{x,y}\left[v_\lambda(k_2)\right] \right\}$$

Shapley equation can be written as a finite number of polynomial inequalities in λ. Consequence :

For λ small enough, for all $i \in I$ and $j \in J$, $v_\lambda(k)$, $\sigma_\lambda(k,i)$ and $\tau_\lambda(k,j)$ can be written as Puiseux series.

- Given $\lambda \in (0,1]$, the pair of optimal stationary strategies $(\sigma_\lambda, \tau_\lambda)$ induces a Markov chain $(k_m)_{m\geq 1}$ on the state space, which transition is denoted by Q_λ.

- There exists $N \geq 1$, positive integers $(a_{k,k'})$ and reals $(C_{k,k'})$ such that

$$Q_\lambda(k'|k) = C_{k,k'} \lambda^{\frac{a_{k,k'}}{N}} + o\left(\lambda^{\frac{a_{k,k'}}{N}}\right)$$

- Define

$$S(k,t) := \lim_{\lambda \to 0} \mathbb{E}^k_{\sigma_\lambda,\tau_\lambda}\left(\sum_{m=1}^{m(t,\lambda)} \lambda(1-\lambda)^{m-1} g_m\right)$$

- Given $\lambda \in (0,1]$, the pair of optimal stationary strategies $(\sigma_\lambda, \tau_\lambda)$ induces a Markov chain $(k_m)_{m\geq 1}$ on the state space, which transition is denoted by Q_λ.

- There exists $N \geq 1$, positive integers $(a_{k,k'})$ and reals $(C_{k,k'})$ such that

$$Q_\lambda(k'|k) = C_{k,k'} \lambda^{\frac{a_{k,k'}}{N}} + o\left(\lambda^{\frac{a_{k,k'}}{N}}\right)$$

- Define

$$S(k,t) := \lim_{\lambda \to 0} \mathbb{E}^k_{\sigma_\lambda, \tau_\lambda} \left(\sum_{m=1}^{m(t,\lambda)} \lambda(1-\lambda)^{m-1} g_m \right)$$

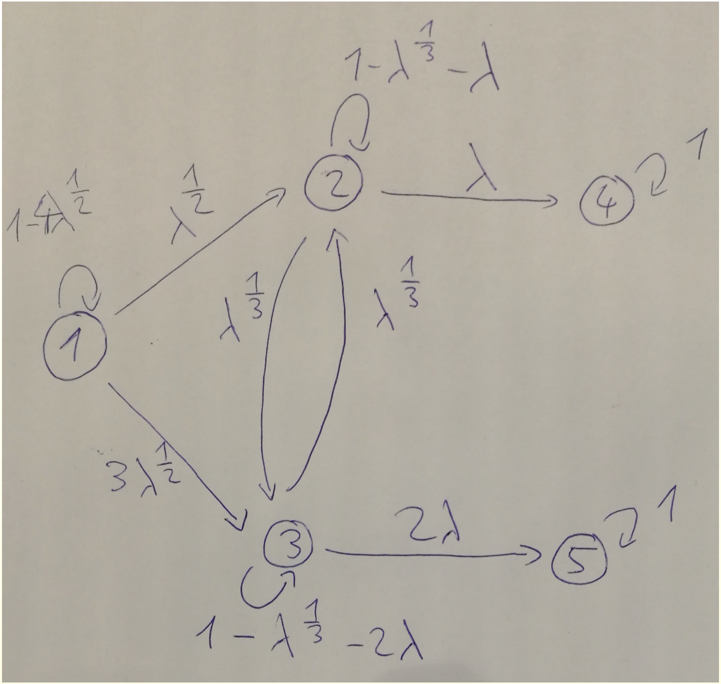
- Given $\lambda \in (0, 1]$, the pair of optimal stationary strategies $(\sigma_\lambda, \tau_\lambda)$ induces a Markov chain $(k_m)_{m \geq 1}$ on the state space, which transition is denoted by Q_λ.

- There exists $N \geq 1$, positive integers $(a_{k,k'})$ and reals $(C_{k,k'})$ such that

$$Q_\lambda(k'|k) = C_{k,k'} \lambda^{\frac{a_{k,k'}}{N}} + o\left(\lambda^{\frac{a_{k,k'}}{N}}\right)$$

- Define

$$S(k, t) := \lim_{\lambda \to 0} \mathbb{E}^k_{\sigma_\lambda, \tau_\lambda} \left(\sum_{m=1}^{m(t,\lambda)} \lambda (1-\lambda)^{m-1} g_m \right)$$


1 - λ^(1/3) - λ
λ^(1/2)
λ
1
λ^(1/3)
λ^(1/3)
3λ^(1/2)
2λ
1
1 - λ^(1/3) - 2λ

A recurrent class R is a subset of K such that for any initial state $k \in R$:

- The number of visits to any state $k' \in R$ before stage $\lambda^{\frac{1}{2N}-1}$ goes almost surely to infinity as λ goes to 0 ;
- The probability of leaving R before stage $\lambda^{\frac{1}{2N}-1}$ tends to 0 as λ goes to 0.

A state is *transient* if it does not belong to any recurrent class.

Freidlin and Koralov (2016)

- The state space K can be partitioned into recurrent classes and transient states.
- Starting from a transient state, the probability of reaching a recurrent class before stage $\lambda^{\frac{1}{2N}-1}$ tends to 1 as λ tends to 0,
- Starting in a recurrent class R, the probability of leaving R before stage $1/\lambda$ is bounded away from 1.

A recurrent class R is a subset of K such that for any initial state $k \in R$:

- The number of visits to any state $k' \in R$ before stage $\lambda^{\frac{1}{2N}-1}$ goes almost surely to infinity as λ goes to 0 ;
- The probability of leaving R before stage $\lambda^{\frac{1}{2N}-1}$ tends to 0 as λ goes to 0.

A state is *transient* if it does not belong to any recurrent class.

Freidlin and Koralov (2016)

- The state space K can be partitioned into recurrent classes and transient states.
- Starting from a transient state, the probability of reaching a recurrent class before stage $\lambda^{\frac{1}{2N}-1}$ tends to 1 as λ tends to 0,
- Starting in a recurrent class R, the probability of leaving R before stage $1/\lambda$ is bounded away from 1.

A recurrent class R is a subset of K such that for any initial state $k \in R$:

- The number of visits to any state $k' \in R$ before stage $\lambda^{\frac{1}{2N}-1}$ goes almost surely to infinity as λ goes to 0 ;
- The probability of leaving R before stage $\lambda^{\frac{1}{2N}-1}$ tends to 0 as λ goes to 0.

A state is *transient* if it does not belong to any recurrent class.

Freidlin and Koralov (2016)

- The state space K can be partitioned into recurrent classes and transient states.
- Starting from a transient state, the probability of reaching a recurrent class before stage $\lambda^{\frac{1}{2N}-1}$ tends to 1 as λ tends to 0,
- Starting in a recurrent class R, the probability of leaving R before stage $1/\lambda$ is bounded away from 1.

- Take some initial state k in a recurrent class R,
- Let $\mu \in \Delta(R)$ be the limit law of the frequency of visits of states in R before exiting R,
- Let $p \in \Delta(K \setminus R)$ be the limit law of the exit state,
- Let $\beta_t = \lim_{\lambda \to 0} \mathbb{E}^k_{x_\lambda, y_\lambda} \left[\sum_{m=1}^{T_R \wedge m(t,\lambda)} \lambda(1-\lambda)^{m-1} \right]$, where T_R is the exit time from R.

Proposition

For all $t \in [0,1]$, $\beta_t > 0$, and

$$S(k,t) = \beta_t < \mu, g > + (t - \beta_t) < p, v^* >$$

The payoff is constant if and only if $< \mu, g >=< p, v^* >$

- Take some initial state k in a recurrent class R,
- Let $\mu \in \Delta(R)$ be the limit law of the frequency of visits of states in R before exiting R,
- Let $p \in \Delta(K \setminus R)$ be the limit law of the exit state,
- Let $\beta_t = \lim_{\lambda \to 0} \mathbb{E}^k_{x_\lambda, y_\lambda} \left[\sum_{m=1}^{T_R \wedge m(t,\lambda)} \lambda (1-\lambda)^{m-1} \right]$, where T_R is the exit time from R.

Proposition

For all $t \in [0,1]$, $\beta_t > 0$, and

$$S(k,t) = \beta_t < \mu, g > + (t - \beta_t) < p, v^* >$$

The payoff is constant if and only if $< \mu, g > = < p, v^* >$

- Take some initial state k in a recurrent class R,
- Let $\mu \in \Delta(R)$ be the limit law of the frequency of visits of states in R before exiting R,
- Let $p \in \Delta(K \setminus R)$ be the limit law of the exit state,
- Let $\beta_t = \lim_{\lambda \to 0} \mathbb{E}^k_{x_\lambda, y_\lambda} \left[\sum_{m=1}^{T_R \wedge m(t,\lambda)} \lambda(1-\lambda)^{m-1} \right]$, where T_R is the exit time from R.

Proposition

For all $t \in [0,1]$, $\beta_t > 0$, and

$$S(k,t) = \beta_t < \mu, g > + (t - \beta_t) < p, v^* >$$

The payoff is constant if and only if $< \mu, g > = < p, v^* >$

Define $\pi^k \in \Delta(K)$ by

$$\forall k' \in K \quad \pi^k(k') := \lim_{\lambda \to 0} \mathbb{E}_\lambda^k \left[\sum_{m \geq 1} \lambda (1-\lambda)^{m-1} 1_{\{k_m = k'\}} \right].$$

$\pi^k = \beta_1 \mu + (1-\beta_1)\pi^p$

If for all k, $< \pi^k, v^* >= v^*(k)$, then the payoff is constant.

Define $\pi^k \in \Delta(K)$ by

$$\forall k' \in K \quad \pi^k(k') := \lim_{\lambda \to 0} \mathbb{E}_\lambda^k \left[\sum_{m \geq 1} \lambda(1-\lambda)^{m-1} 1_{\{k_m = k'\}} \right].$$

$\pi^k = \beta_1 \mu + (1-\beta_1)\pi^p$

If for all k, $< \pi^k, v^* >= v^*(k)$, then the payoff is constant.

Define $\pi^k \in \Delta(K)$ by

$$\forall k' \in K \quad \pi^k(k') := \lim_{\lambda \to 0} \mathbb{E}_\lambda^k \left[\sum_{m \geq 1} \lambda(1-\lambda)^{m-1} 1_{\{k_m = k'\}} \right].$$

$\pi^k = \beta_1 \mu + (1 - \beta_1) \pi^p$

If for all k, $< \pi^k, v^* >= v^*(k)$, then the payoff is constant.

The fact that $<\pi^k, v^*> = v^*(k)$ can be deduced from a lemma about Hamilton-Jacobi equations by Davini, Fathi, Iturriaga and Zavidovique (2014).

Lemma 3.7. *Let w be any critical subsolution. For every $\lambda > 0$ and $x \in M$*

$$u_\lambda(x) \geq w(x) - \int_{TM} w(y)\, \mathrm{d}\tilde{\mu}_x^\lambda(y, v). \tag{3.6}$$

Translation

$$v_\lambda(k) \geq < \pi_\lambda^k, v^* >$$

The fact that $<\pi^k, v^*>= v^*(k)$ can be deduced from a lemma about Hamilton-Jacobi equations by Davini, Fathi, Iturriaga and Zavidovique (2014).

Lemma 3.7. *Let w be any critical subsolution. For every $\lambda > 0$ and $x \in M$*

$$u_\lambda(x) \geq w(x) - \int_{TM} w(y)\, \mathrm{d}\tilde{\mu}_x^\lambda(y, v). \tag{3.6}$$

Translation

$$v_\lambda(k) \geq < \pi_\lambda^k, v^* >$$

Further work

- Extend the result to n-stage games,

- Extend to the nonzero-sum case.

Thank you for your attention ! Questions ?