

# One simple remark concerning the uniform value\*

Dmitry Khlopin  
*khlopin@imm.uran.ru*

August 27, 2018

## Abstract

The paper is devoted to dynamic games. We consider a general enough framework, which is not limited to e.g. differential games and could accommodate both discrete and continuous time. Assuming common dynamics, we study two game families with total payoffs that are defined either as the Cesàro average (long run average game family) or Abel average (discounting game family) of the running costs. We study a robust strategy that would provide a near-optimal total payoff for all sufficiently small discounts and for all sufficiently large planning horizons. Assuming merely the Dynamic Programming Principle, we prove the following Tauberian theorem: if a strategy is uniformly optimal for one of the families (when discount goes to zero for discounting games, when planning horizon goes to infinity in long run average games) and its value functions converge uniformly, then, for the other family, this strategy is also uniformly optimal and its value functions converge uniformly to the same limit.

**Keywords:** Dynamic programming principle, dynamic games, uniform value, Abel mean, Cesàro mean

**MSC2010** 91A25, 49L20, 49N70, 91A23, 40E05

In dynamic optimization, most often, the potential infinity of the planning horizon is emulated by considering the running cost averaged with respect to the uniform or exponential distributions (Cesàro average or Abel average, respectively). In this paper, we consider the correspondence between the value function of dynamic games with these total payoff families. The theorems that describe the connection between the Cesàro and Abel averages are long known as Tauberian.

---

\*Krasovskii Institute of Mathematics and Mechanics, Russian Academy of Sciences, 16, S.Kovalevskaja St., 620990, Yekaterinburg, Russia; Institute of Mathematics and Computer Science, Ural Federal University, 4, Turgeneva St., 620083, Yekaterinburg, Russia

The existence of a limit of the value function when the running cost is averaged with respect to the uniform or exponential distributions means that the value functions are robust with respect to the choice of discount (the planning horizon) as long as it is small (large) enough. In particular, it is exactly this value (asymptotic approach) that is viewed as the game's value for the infinite planning horizon in stochastic statements. Furthermore, in these statements, often enough (see [1, 15]), one could also find a robust strategy that would provide a near-optimal total payoff for all sufficiently small discounts and for all sufficiently large planning horizons [21, 23, 7, 11] (uniform approach).

Within the framework of the asymptotic approach, under mild assumptions, there holds the following Tauberian theorem: the uniform convergence of the value functions for the running costs averaged with respect to the uniform and/or exponential distributions guarantees, for the other distribution, the uniform convergence of its value functions to the same limit. Such a result was proved for stochastic games with finite numbers of states and actions [15] and for discrete-time optimal control problems [12]. It has been relatively recently transferred to the general control problems [16], differential games [8], and a broad class of stochastic games [22]. Then, a Tauberian theorem for all two-person zero-sum games satisfying the Dynamic Programming Principle was proved [10].

Surprisingly, the Tauberian theorem proved in [10] lends itself well to the proof of the corresponding Tauberian theorem for the uniform approach. This result is the main contribution of this article.

### Dynamic system

Assume the following items are given:

- a nonempty set  $\Omega$ , the data space;
- a nonempty set  $\mathbb{K}$  of maps from  $[0, \infty)$  to  $\Omega$ ;
- a running cost  $g: \Omega \mapsto [0, 1]$ .

For each process  $z \in \mathbb{K}$ , let the map  $t \mapsto g(z(t))$  be Borel measurable. Now, for all positive  $\lambda, T > 0$ , consider the payoffs

$$v_T(z) \triangleq \frac{1}{T} \int_0^T g(z(t)) dt \quad \forall z \in \mathbb{K};$$

$$w_\lambda(z) \triangleq \lambda \int_0^\infty e^{-\lambda t} g(z(t)) dt \quad \forall z \in \mathbb{K}.$$

### On game value maps

Denote by  $\mathfrak{U}$  the set of all bounded maps from  $\Omega$  to  $\mathbb{R}$ ; denote by  $\mathfrak{C}$  a non-empty set of maps from  $\mathbb{K}$  to  $\mathbb{R}$ . Hereinafter assume that the set  $\mathfrak{C}$  contains all conceivable payoffs, and the set  $\mathfrak{U}$  contains all value functions for all games with payoffs  $c \in \mathfrak{C}$ .

Let  $\mathfrak{C}$  satisfy the following condition:

$$Ac + B \in \mathfrak{C} \text{ for all } A \geq 0, B \in \mathbb{R} \text{ if } c \in \mathfrak{C}. \quad (1a)$$

Hereinafter we assume that  $v_T, w_\lambda \in \mathfrak{C}$  for all positive  $\lambda, T$ .

A map  $V$  from  $\mathfrak{C}$  to  $\mathfrak{U}$  is called a *game value map* if the following conditions hold:

$$V[Ac + B] = AV[c] + B \text{ for all } c \in \mathfrak{C}, A \geq 0, B \in \mathbb{R}, \quad (1b)$$

$$V[c_1](\omega) \leq V[c_2](\omega) \text{ for all } \omega \in \Omega \text{ if } c_1(z) \leq c_2(z) \text{ for all } z \in \mathbb{K}. \quad (1c)$$

### On Dynamic Programming Principle

Fix a game value map  $V$ . For all positive  $\lambda, T, h > 0$  and the game value map  $V$ , define payoffs  $\zeta_{h,T}: \mathbb{K} \rightarrow \mathbb{R}$ ,  $\xi_{h,\lambda}: \mathbb{K} \rightarrow \mathbb{R}$  as follows:

$$\zeta_{h,T}(z) \triangleq \frac{1}{T+h} \int_0^h g(z(t)) dt + \frac{T}{T+h} V[v_T](z(h)) \quad \forall z \in \mathbb{K};$$

$$\xi_{h,\lambda}(z) \triangleq \lambda \int_0^h e^{-\lambda t} g(z(t)) dt + e^{-\lambda h} V[w_\lambda](z(h)) \quad \forall z \in \mathbb{K}.$$

**Definition 1** *Let us say that the payoffs  $v_T (T > 0)$  (resp.,  $w_\lambda (\lambda > 0)$ ) enjoy the weak Dynamic Programming Principle with respect to the game value map  $V$  if the payoffs  $\zeta_{h,T} (T, h \in \mathbb{N})$  (resp., payoffs  $\xi_{h,\lambda} (\lambda > 0, h \in \mathbb{N})$ ) lie in  $\mathfrak{C}$  and, for every  $\varepsilon > 0$ , there exists natural  $N$  such that, for all natural  $h, T > N$  and positive  $\lambda < 1/N$ ,*

$$|V[v_{T+h}](\omega) - V[\zeta_{h,T}](\omega)| < \varepsilon \quad \left( |V[w_\lambda](\omega) - V[\xi_{h,\lambda}](\omega)| < \varepsilon \right) \quad \forall \omega \in \Omega. \quad (2)$$

In particular, the family of payoffs  $v_T (T > 0)$  (resp.,  $w_\lambda (\lambda > 0)$ ) enjoys the weak Dynamic Programming Principle if

$$V[v_{T+h}] = V[\zeta_{h,T}], \quad \left( V[w_\lambda] = V[\xi_{h,\lambda}] \right) \quad \forall h, T \in \mathbb{N}, \lambda > 0.$$

Note that the technical requirements  $\zeta_{h,T} \in \mathfrak{C}$  and  $\xi_{h,\lambda} \in \mathfrak{C}$  can always be provided for by extending  $V$  [10, Lemma 1]. The key requirement in the Dynamic Programming Principle is the uniform approximativity of the value functions  $V[v_{T+h}]$  and  $V[w_\lambda]$  with  $V[\zeta_{h,T}]$  and  $V[\xi_{h,\lambda}]$ , respectively.

The cornerstone result for this article's main theorem was proved in [10]:

**Theorem 1** *Let there be given a game value map  $V: \mathfrak{C} \rightarrow \mathfrak{U}$ . Let  $v_T, w_\lambda \in \mathfrak{C}$  for all  $\lambda, T > 0$ . Assume that payoffs  $v_T (T > 0)$  and payoffs  $w_\lambda (\lambda > 0)$  enjoy the weak Dynamic Programming Principle.*

*Then, the following two statements are equivalent:*

- (i) *The family of functions  $V[v_T] (T > 0)$  converges uniformly on  $\Omega$  as  $T \uparrow \infty$ .*
- (ii) *The family of functions  $V[w_\lambda] (\lambda > 0)$  converges uniformly on  $\Omega$  as  $\lambda \downarrow 0$ .*

*Moreover, when at least one of these statements hold, then, for both families, the corresponding limits of the value functions exist, are uniform in  $\omega \in \Omega$ , and coincide.*

Note that in this theorem the requirement of the dynamic programming principle could not be omitted, see [9]. The limits must remain uniform unless additional assumptions are made, even in control problems and the stochastic statement; see the counterexample in e.g. [16] and [22], respectively.

### On strategies

Assume that a strategy set  $\mathfrak{S}$  is given, and, for each strategy  $s \in \mathfrak{S}$ , we construct a game value map  $V_s: \mathfrak{C} \rightarrow \mathfrak{U}$ . Consider the game value map

$$V_{\text{best}}[c](\omega) = \sup_{s \in \mathfrak{S}} V_s[c](\omega) \quad \forall c \in \mathfrak{C}, \omega \in \Omega. \quad (3)$$

**Definition 2** *Let us say that a strategy  $s \in \mathfrak{S}$  is called uniformly optimal for the payoff family  $v_T (T > 0)$  (resp.,  $w_\lambda (\lambda > 0)$ ) iff*

$$\limsup_{T \uparrow \infty} \sup_{\omega \in \Omega} |V_{\text{best}}[v_T] - V_s[v_T]| = 0 \quad \left( \limsup_{\lambda \downarrow 0} \sup_{\omega \in \Omega} |V_{\text{best}}[w_\lambda] - V_s[w_\lambda]| = 0 \right).$$

An unexpected pleasure is the fact that the assumptions of the following theorem allow the use of any strategy, a strategy of whatever kind: if only it satisfies the Dynamic Programming Principle, we get a uniformly optimal strategy for both payoff families.

**Theorem 2** *Let there be given game value maps  $V_s: \mathfrak{C} \rightarrow \mathfrak{U}$  for each strategy  $s \in \mathfrak{S}$ , and let the game value map  $V_{\text{best}}$  be defined by the rule (3). Let there be given a strategy  $s_* \in \mathfrak{S}$ .*

*Assume that the payoffs  $v_T (T > 0)$  and payoffs  $w_\lambda (\lambda > 0)$  enjoy the weak Dynamic Programming Principle with respect to  $V_{s_*}$  and  $V_{\text{best}}$ . Then, the following conditions are equivalent:*

- (v) *the strategy  $s_*$  is uniformly optimal for the payoff family  $v_T (T > 0)$ , in addition, the functions  $V_{\text{best}}[v_T]$  converge uniformly in  $\Omega$  as  $T \uparrow \infty$ ;*

(w) the strategy  $s_*$  is uniformly optimal for the payoff family  $w_\lambda(\lambda > 0)$ , in addition, the functions  $V_{\text{best}}[w_\lambda]$  converge uniformly in  $\Omega$  as  $\lambda \downarrow 0$ ;

(eq) all limits in

$$\lim_{T \uparrow \infty} V_{\text{best}}[v_T](\omega), \lim_{\lambda \downarrow 0} V_{\text{best}}[w_\lambda](\omega), \lim_{T \uparrow \infty} V_{s_*}[v_T](\omega), \lim_{\lambda \downarrow 0} V_{s_*}[w_\lambda](\omega) \quad (4)$$

exist, are uniform in  $\omega \in \Omega$ , and coincide.

Proof.

Note that (eq)  $\Rightarrow$  (v) and (eq)  $\Rightarrow$  (w) hold by the definition. So, it would suffice to prove (v)  $\Rightarrow$  (eq), (w)  $\Rightarrow$  (eq). We will prove (v)  $\Rightarrow$  (eq); the proof of the last implication (w)  $\Rightarrow$  (eq) is similar.

Let the strategy  $s_*$  be uniformly optimal for the payoff family  $v_T(T > 0)$ , and let the functions  $V_{\text{best}}[v_T]$  converge uniformly in  $\Omega$  as  $T \uparrow \infty$  to a function  $U_* \in \mathfrak{U}$ . Then, by the definition of the uniformly optimal strategy, the functions  $V_{s_*}[v_T]$  also converge to  $U_*$  uniformly in  $\Omega$  as  $T \uparrow \infty$ . Applying Theorem 1 for the game value map  $V_{s_*}$ , we see that  $V_{s_*}[w_\lambda]$  converge to  $U_*$  uniformly in  $\Omega$  as  $\lambda \downarrow 0$ . Applying this theorem for the game value map  $V_{\text{best}}$ , we have that  $V_{\text{best}}[w_\lambda]$  converge to  $U_*$  uniformly in  $\Omega$  as  $\lambda \downarrow 0$ . Then, in view of the definition of the uniformly optimal strategy, we see that  $s_*$  is uniformly optimal for the payoff family  $w_\lambda(\lambda > 0)$ .  $\square$

Question 1. Is the uniform optimality of a strategy for the payoffs  $v_T(T > 0)$  and for the payoffs  $w_\lambda(\lambda > 0)$  still equivalent without the additional requirement on the uniform convergence of arbitrary value functions?

Question 2. Could the assumption on Dynamic Programming Principle be relaxed?

Question 3. Can there be a topology for  $\mathfrak{U}$  that is different from the uniform topology?

A partial answer to Question 2 will be considered for the following important case.

### The case of one player

For all  $\tau \in [0, \infty)$ ,  $z', z'' \in \mathbb{K}$  with the property  $z'(\tau) = z''(0)$  and them only, let us define their concatenation—likewise, a mapping from  $[0, \infty)$  to  $\Omega$ —as follows:

$$(z' \diamond_\tau z'')(t) \triangleq \begin{cases} z'(t), & 0 \leq t \leq \tau; \\ z''(t - \tau), & t > \tau. \end{cases} \quad \forall t \geq 0.$$

Following [16], assume that

$$\Gamma(\omega) \triangleq \{z \in \mathbb{K} \mid z(0) = \omega\} \neq \emptyset \quad \forall \omega \in \Omega;$$

also, assume that the set  $\mathbb{K}$  is closed with respect to concatenation.

Let  $\mathfrak{S}$  be the set of all selectors  $\Omega \in \omega \mapsto s[\omega] \in \mathbb{K}$  of the set-valued map  $\omega \mapsto \Gamma(\omega)$ , i.e., the set of all mappings  $\Omega \in \omega \mapsto s[\omega] \in \mathbb{K}$  such that  $s[\omega](0) = \omega$  for all  $\omega \in \Omega$ .

**Definition 3** Let us say that a strategy  $s_* \in \mathfrak{S}$  is stationary-like (does not change when shifted in time) if

$$s_*[\omega](t+1) = s_*[s_*[\omega](1)](t) \quad \forall \omega \in \Omega, t \geq 0.$$

Let  $\mathfrak{C}$  and  $\mathfrak{U}$  be the sets of all scalar bounded maps from  $\mathbb{K}$  and  $\Omega$ , respectively. For all  $s \in \mathfrak{S}$ , define  $V_s: \mathfrak{C} \rightarrow \mathfrak{U}$  by the following rule:

$$V_s[c](\omega) \triangleq c(s[\omega]) \quad \forall \omega \in \Omega, c \in \mathfrak{C}.$$

Let us also define  $V_{\text{best}}: \mathfrak{C} \rightarrow \mathfrak{U}$  by (3). Then,

$$V_{\text{best}}[c](\omega) = \sup_{z \in \mathbb{K}, z(0) = \omega} c(z) \quad \forall \omega \in \Omega, c \in \mathfrak{C}.$$

It is easy to see that  $V_{\text{best}}$  and  $V_s$ , for all  $s \in \mathfrak{S}$ , are game value maps.

**Theorem 3** Assume that the set  $\mathbb{K}$  is closed with respect to concatenation and that  $\Gamma(\omega)$  is non-empty for all  $\omega \in \Omega$ . Let a strategy  $s_* \in \mathfrak{S}$  be stationary-like.

Then, conditions (v), (w), (eq) are equivalent.

Proof.

It is easy to see that, since the strategy  $s_* \in \mathfrak{S}$  is stationary-like, the payoffs  $v_T (T > 0)$  and payoffs  $w_\lambda (\lambda > 0)$  enjoy the weak Dynamic Programming Principle with respect to  $V_{s_*}$ .

Since the set  $\mathbb{K}$  is closed with respect to concatenation, we have

$$\begin{aligned} V_{\text{best}}[c](\omega) &= \sup_{s \in \mathfrak{S}} c(s[\omega]) \geq \sup_{s_0, s_1 \in \mathfrak{S}} c(s_0[\omega] \diamond_1 s_1[s_0[\omega](1)]), \\ V_{\text{best}}[c](\omega) &= \sup_{s \in \mathfrak{S}} c(s[\omega]) \geq \sup_{s_0, s_1 \in \mathfrak{S}} c(s_0[\omega] \diamond_n s_1[s_0[\omega](n)]), \quad \forall n \in \mathbb{N}. \end{aligned}$$

Then, for all positive  $\lambda$  and natural  $n, T$ , for all  $\omega \in \Omega$ , we have

$$\begin{aligned} V_{\text{best}}[w_\lambda](\omega) &\geq \sup_{s \in \mathfrak{S}} \left[ \lambda \int_0^n e^{-\lambda t} g(s[\omega](t)) dt + e^{-\lambda n} V_{\text{best}}[w_\lambda](s[\omega](n)) \right], \\ V_{\text{best}}[v_{T+n}](\omega) &\geq \sup_{s \in \mathfrak{S}} \left[ \frac{1}{T+n} \int_0^n g(s[\omega](t)) dt + \frac{T}{T+n} V_{\text{best}}[v_T](s[\omega](n)) \right]. \end{aligned}$$

Thus, with respect to the game value map  $V_{\text{best}}$ ,  $V_{\text{best}}[w_\lambda]$  is a subsolution (see [10, Definition 1]) for the payoffs  $w_\lambda (\lambda > 0)$ , and  $V_{\text{best}}[v_T]$  is a subsolution for the payoffs  $v_T (T > 0)$ .

So, it would again suffice to prove (v)  $\Rightarrow$  (eq), (w)  $\Rightarrow$  (eq).

Let the strategy  $s_*$  be uniformly optimal for the payoff family  $v_T (T > 0)$  (resp.,  $w_\lambda (\lambda > 0)$ ), and let its value functions  $V_{\text{best}}[v_T]$  converge uniformly in  $\Omega$  to a function  $U_* \in \mathfrak{U}$ . Then,

by the definition of the uniformly optimal strategy, the value functions  $V_{s_*}[v_T]$  (resp.,  $V_{s_*}[w_\lambda]$ ) also converge uniformly in  $\Omega$  to  $U_* \in \mathfrak{U}$ . Since the payoffs  $v_T$  ( $T > 0$ ) and payoffs  $w_\lambda$  ( $\lambda > 0$ ) enjoy the Dynamic Programming Principle with respect to  $V_{s_*}$ , applying Theorem 1 for the game value map  $V_{s_*}$ , we see that  $V_{s_*}[w_\lambda]$  and  $V_{s_*}[v_T]$  also converge to  $U_*$  uniformly in  $\Omega$ . So, the lower limit of  $V_{\text{best}}[v_T] - U_*$  (resp., of  $V_{\text{best}}[w_\lambda] - U_*$ ) is nonnegative.

On the other hand, applying [10, Proposition 3] to the subsolution  $V_{\text{best}}[w_\lambda]$  (applying [10, Proposition 4] to the subsolution  $V_{\text{best}}[v_T]$ ) and for the game value map  $V_{\text{best}}$  we find that, for every positive  $\varepsilon$ , there exists a natural  $N$  such that  $V_{\text{best}}[w_\lambda](\omega) \leq V_{\text{best}}[v_T](\omega) + \varepsilon$  ( $V_{\text{best}}[v_T](\omega) \leq V_{\text{best}}[w_\lambda](\omega) + \varepsilon$ ) for all positive  $T > N$ ,  $\lambda = 1/N$ , and for all  $\omega \in \Omega$ . So, the upper limit of  $U_* - V_{\text{best}}[v_T]$  (resp., of  $U_* - V_{\text{best}}[w_\lambda]$ ) is also nonnegative.

Thus,  $V_{\text{best}}[w_\lambda]$  and  $V_{\text{best}}[v_T]$  converge to  $U_*$  uniformly in  $\Omega$ . Since, see above,  $V_{s_*}[w_\lambda]$  and  $V_{s_*}[v_T]$  also converge to  $U_*$  uniformly in  $\Omega$ , by the definition,  $s_*$  is a uniformly optimal strategy for both the payoffs  $v_T$  ( $T > 0$ ) and  $w_\lambda$  ( $\lambda > 0$ ).  $\square$

## References

- [1] Bewley, T., Kohlberg, E.: The asymptotic theory of stochastic games. *Math. Oper. Res.* 1, 197-208 (1976)
- [2] Blackwell D (1962) Discrete dynamic programming, *Ann Math Statist*
- [3] Cannarsa, P., Quincampoix, M.: Vanishing Discount Limit and Nonexpansive Optimal Control and Differential Games. *SIAM J. Control Optim.* 53(4), 1789-1814 (2015)
- [4] Escobedo-Trujillo, B.A., Jasso-Fuentes, H., Lopez-Barrientos J.D.: Blackwell-Nash equilibria in zero-sum stochastic differential games, preprint (2017)
- [5] Gaitsgory V, Quincampoix M (2013) On sets of occupational measures generated by a deterministic control system on an infinite time horizon. *Nonlinear Anal-Theor* 88:27-41
- [6] Grüne, L.: On The Relation Between Discounted And Average Optimal Value Functions. *J. Diff. Eq.* 148, 65-99 (1998)
- [7] Jaśkiewicz, A., Nowak, A.S.: Non-Zero-Sum Stochastic Games. Eds: T.Başar, G.Zaccour. *Handbook of Dynamic Game Theory*, Springer (in print)
- [8] Khlopin DV (2015) Uniform Tauberian theorem for differential games. *Automat Rem Contr+*, 2016 77(4):734–750
- [9] Khlopin DV On uniform Tauberian theorems for dynamic games. *Sb. Math.* 209(1), 122-144 (2018)

- [10] Khlopin DV (2018) Tauberian Theorem for Value Functions. *Dyn. Games Appl*, 8(2):401–422
- [11] Laraki R, Renault J Acyclic Gambling Games. arXiv preprint arXiv:1702.0686 (2017)
- [12] Lehrer, E., Sorin, S.: A uniform Tauberian theorem in dynamic programming. *Math. Oper. Res.* 17(2), 303-307 (1992)
- [13] Li, X., Quincampoix, M., Renault, J.: Limit value for optimal control with general means, *Discrete Contin Dyn Syst. Series A*, 36, 2113-2132 (2016)
- [14] Li X, Venel X (2016) Recursive games: uniform value, Tauberian theorem and the Mertens conjecture “ $\max \min = \lim v_n = \lim v_\lambda$ ”. *Int J of Game Theory*, 45(1):155-189
- [15] Mertens JF, Neyman A. (1981) Stochastic Games. *Int. J. Game Theory* 10(2), 53-66
- [16] Oliu-Barton, M., Vigerel, G.: A uniform Tauberian theorem in optimal control. In: *Advances in Dynamic Games*. pp.199-215 Birkhäuser, Boston (2013)
- [17] Oliu-Barton M. The Splitting Game: Uniform Value and Optimal Strategies. *Dyn. Games Appl*, 8(1):157–179 (2018)
- [18] Quincampoix M, Renault J (2011) On the existence of a limit value in some non expansive optimal control problems. *SIAM J Control Optim* 49(5):2118-2132
- [19] Renault, J.: General limit value in dynamic programming. *J. Dyn. and Games* 1(3), 471-484 (2013)
- [20] Rosenberg D, Solan E, Vieille N (2002) Blackwell optimality in markov decision processes with partial observation. *Ann Stat* 30:1178–1193
- [21] Solan, E., Ziliotto, B. (2016). Stochastic games with signals. In: *Advances in Dynamic and Evolutionary Games* (pp. 77-94). Birkhäuser, Cham.
- [22] Ziliotto, B.: A Tauberian theorem for nonexpansive operators and applications to zero-sum stochastic games. *Math. Oper. Res.* 41(4), 1522-1534 (2016)
- [23] Ziliotto, B.: General limit value in zero-sum stochastic games. *Int. J. Game Theory* **45**, 353-374 (2016) doi:10.1007/s00182-015-0509-3