Zero-sum continuous-time Markov games with one-side stopping

Yurii Averboukh

Abstract

The paper is concerned with a variant of the continuous-time finite state Markov game of control and stopping where both players can affect transition rates, while only one player can choose a stopping time. We use the dynamic programming principle and reduce this problem to a system of ODEs with unilateral constraints. This system plays the role of the Bellman equation. We show that its solution provides the optimal strategies of the players. Additionally, we prove the existence and uniqueness theorem for the deduced system of ODEs with unilateral constraints.

MSC Classification (2020): 91A05, 91A25, 90C40, 60G40, 60J27.

Keywords: continuous-time Markov games, dynamic programming, verification theorem, stopping time.

1 Introduction

The continuous-time Markov processes can be controlled in two ways. First, the decision maker can affect transition rates. The second way involves the choice of the stopping time. The case where the transition rates are controlled refers to the Markov decision problem (see [7], [6] and reference therein for the detailed explanation of this field). The optimization of the Markov chain by the means of stopping time is studied within the optimal stopping theory [11, 12]. Certainly, one can mix both ways and study Markov decision processes with stopping [8]. The concept of Markov chain affected by several decision makers leads to such settings as the continuous time Markov games those examine the case when all players control transition rates [14], the Dynkin games [3, 4] where the players choose stopping times and the games of control and stopping where all players can affect the transition rates as well as the stopping time [5]. In the paper, we consider one-side stopping Markov game that implies that one player chooses a control and a stopping time, while the other one affects only the control. Our interest to this problem is motivated by the possible application of these games to the differential games with minimal cost studied in [2, 13]. In particular, it is shown that one can find a one-side stopping Markov game such that its value function approximates the value function of the original differential game with minimal cost [1].

We study the continuous-time finite state Markov chain governed by two decision makers with opposite purposes. Without loss of generality, we assume that the state
space is \( \mathcal{S} = \{1, \ldots, d\} \). The transition rates are described by the Kolmogorov matrix \( Q(t, u, v) = (Q_{i,j}(t, u, v))_{i,j=1}^{d} \), i.e., for each \( t, u \) and \( v \),
\[
Q_{i,j}(t, u, v) \geq 0 \quad \text{for} \quad i \neq j,
\]
while
\[
\sum_{j \in \mathcal{S}} Q_{i,j} = 0.
\]
Here \( u \) (respectively, \( v \)) is a control of the first (respectively, second) player. It is assumed that \( u \in U, v \in V \), while \( t \in [0, T] \). We assume that the stopping of the game is chosen by the first player. If \( t_0 \) is an initial time, \( X(t) \) is a state of the Markov chain at time \( t \), while \( \tau \) is a stopping time, then the quality of the examined one-side Markov game is evaluated by
\[
g(\tau, X(\tau)) + \int_{t_0}^{\tau} h(t', X(t'), u(t'), v(t')) dt'.
\]
Here \( E \) stands for the expectation. We assume that the first player tries to minimize functional (1), while the second player wishes to maximize it.

The universal tool for the study of Markov decision processes, Markov games as well as optimal stopping problems is the dynamic programming principle. It primary relies on the feedback strategies and reduces the problem of design of optimal strategies to a certain Bellman equation. In the case of the finite horizon Markov decision processes and the Markov games, the Bellman equation takes a form of system of ODEs. For the game theoretical case, this system was derived in paper by Zachrisson [14]. Thus, we call it a Zachrisson equation.

In the paper, we use the dynamic programming principle and show the players’ optimal strategies can be constructed based on the solution of the Zachrisson equation with unilateral constraints. Thus, the existence result for the one-side stopping Markov game is reduced to the existence theorem for the Zachrisson equation with unilateral constraints that is also obtained in the paper.

The rest of the paper is organized as follows. In Section 2 we formalize the concept of one-side stopping Markov game and introduce the classes of strategies. The main results are formulated in Section 3. Here, we, first, claim that, given a solution of the Zachrisson equation with unilateral constraints, one can design optimal players’ strategies. Additionally, in this section, we formulate the existence and uniqueness result for the Zachrisson equation with unilateral constraints. These statements imply the existence of the value function of the examined one-side stopping Markov game. Section 4 is concerned with the proof of the first result by the verification argument. The proof of the existence and uniqueness theorem for the Zachrisson equation with unilateral constraints is given in Section 5.

## 2 Preliminaries

In the paper, we identify every function \( \phi : \mathcal{S} \to \mathbb{R} \) with the corresponding vector in \( \mathbb{R}^d \) and assume that \( \phi(i) \) and \( \phi_i \) have the same meaning.
Given an initial time $t_0$, we fix the measurable space $(\Omega_{t_0}, \mathcal{F}_{t_0})$ letting $\Omega_{t_0}$ to be equal to the space of càdlàg functions from $[t_0, T]$ to $\mathcal{S}$ and choosing $\mathcal{F}_{t_0}$ to be a Borel $\sigma$-algebra on $\Omega_{t_0}$, i.e., $\Omega_{t_0} \triangleq D([t_0, T], \mathcal{S})$, and $\mathcal{F}_{t_0} \triangleq \mathcal{B}(D([t_0, T], \mathcal{S}))$.

If $s, r \in [t_0, T]$, $s < r$, then denote by $\mathcal{R}^{s,r}$ operator assigning to $\omega \in \Omega_{t_0} = D([t_0, T], \mathcal{S})$ its restriction on $[s, r]$.

Further, we put $\mathcal{F}_{t_0,t}$ to be the $\sigma$-algebra of sets $A \in \Omega_{t_0}$ such that $\mathcal{R}^{t_0,t}(A) \in \mathcal{B}(D([t_0, t], \mathcal{S}))$ and $\mathcal{R}^{t,t}((A) \in \{\varnothing, D([t, T], \mathcal{S})\}$. Obviously, $\mathcal{F}_{t_0} \triangleq \{\mathcal{F}_{t_0,t}\}_{t \in [t_0, T]}$ is a filtration on $\mathcal{F}_{t_0}$. We denote the set of stopping time w.r.t. the filtration $\mathcal{F}_{t_0}$ taking values in $[t_0, T]$ by $\mathcal{T}_{t_0}$.

If $\phi : [0, T] \times \mathcal{S} \to \mathbb{R}$ is differentiable at $s \in [0, T]$ for some $i \in \mathcal{S}$, and $\omega \in \Omega_{t_0}$ is such that $\omega(s) = i$, we use the notation

$$\frac{\partial}{\partial t} \phi(s, \omega(s))$$

for the derivative of the function $\phi$ w.r.t. the time variable at the point $s$, $i = \omega(s)$.

Below, we denote by $\mathcal{AC}$ the set of all functions $\phi : [0, T] \times \mathcal{S} \to \mathbb{R}$ such that mapping $t \mapsto \phi(t, i)$ is absolutely continuous for each $i \in \mathcal{S}$.

We consider two types of strategies. First, we will assume that the player has information only about current state and time. This leads to the concept of feedback strategies.

**Definition 2.1.** A measurable w.r.t. $\mathcal{F}_{t_0}$ the time mapping $u$ from $[0, T] \times \mathcal{S}$ to $U$ is called a first player’s feedback strategy.

We denote the set of all feedback strategies of the first player by $\mathcal{U}_{fb}$. As it was mentioned above the first player also chooses the stopping time. Within the feedback approach, it is controlled by the stopping rule defined as follows.

**Definition 2.2.** A càdlàg function from $[0, T] \times \mathcal{S}$ to $\{0, 1\}$ is called a stopping rule.

We will assume that the game stops at the first time when the stopping rule is equal to 1, i.e., if $\vartheta$ is a stopping rule, then the corresponding stopping time on $[t_0, T]$ is defined by the rule:

$$\bar{\tau}_{t_0}[\vartheta](\omega) \triangleq \min\{t \in [t_0, T] : \vartheta(t, \omega(t)) = 1\}.$$ 

The set all stopping rules is denoted by $\mathcal{T}_{fb}$.

The feedback strategies of the second player are defined in the same way.

**Definition 2.3.** We say that a mapping $v : [0, T] \times \mathcal{S} \to V$ is a second player’s strategy provided that $v$ is measurable w.r.t. the time variable.

The set of feedback strategies is denoted by $\mathcal{V}_{fb}$.

Since we consider the guaranteed results of the players, it is reasonable to follow the approach of [10] and enlarge of the set of strategies admitting strategies with memory.

**Definition 2.4.** If $t_0$ is an initial time, a $\mathcal{F}_{t_0}$-progressively measurable process with values in $U$ (respectively, $V$) is called a strategy with memory of the first (respectively, second) player on $[t_0, T]$. 

3
We denote the set of strategies with memory of the first (respectively, second) player on \([t_0, T]\) by \(U_{t_0}\) (respectively, \(V_{t_0}\)).

The link between feedback strategies and strategies with memory is straightforward. If \(u \in U_{fb}\), then \(\alpha[u]\) defined by the rule

\[
\alpha[u](t, \omega) \triangleq u(t, \omega(t))
\]

is a strategy with memory. Analogously, given \(v \in V_{fb}\), we define the corresponding strategy with memory \(\beta[v]\) setting \(\beta[v](t, \omega) \triangleq v(t, \omega(t))\).

The probability on \(\mathcal{F}_{t_0}\) generated by two strategies, initial time and initial state is defined in the following way.

**Definition 2.5.** Given two strategies with memory \(\alpha \in U_{t_0}\) and \(\beta \in V_{t_0}\), an initial time \(t_0 \in [0, T]\) and an initial state \(i_0 \in S\), we say that the probability \(\mathbb{P}^{\alpha, \beta}_{t_0, i_0}\) on \(\mathcal{F}_{t_0}\) is generated by \(\alpha, \beta, t_0\) and \(i_0\) provided that

- \(\mathbb{P}^{\alpha, \beta}_{t_0, i_0}(\{\omega : \omega(t_0) = i_0\}) = 1\).
- for every \(\phi \in \mathcal{AC}\), the process

\[
\phi(t, \omega(t)) - \int_{t_0}^{t} \left[ \frac{\partial}{\partial t} \phi(t', \omega(t')) + \sum_{j \in S} Q_{\omega(t'), j}(t', \alpha(t', \omega), \beta(t', \omega)) \phi(t', j) \right] dt'
\]

is \((\mathbb{P}^{\alpha, \beta}_{t_0, i_0}, \mathbb{F})\)-martingale.

Below \(\mathbb{E}^{\alpha, \beta}_{t_0, i_0}\) stands for the expectation corresponding to the probability \(\mathbb{P}^{\alpha, \beta}_{t_0, i_0}\).

If \(\alpha = \alpha[u]\) for some feedback strategy of the first player \(u\), we will write \(\mathbb{P}^{u, \beta}_{t_0, i_0}\) instead of \(\mathbb{P}^{\alpha[u], \beta}_{t_0, i_0}\). Analogously, \(\mathbb{P}^{\alpha, v}_{t_0, i_0} \triangleq \mathbb{P}^{\alpha, \beta[v]}_{t_0, i_0}\). Notice that the probability \(\mathbb{P}^{u, v}_{t_0, i_0}\) is always well-defined. Additionally, one can prove that \(\mathbb{P}^{\alpha, \beta}_{t_0, i_0}\) exists in the case when both \(\alpha\) and \(\beta\) depend on a finite number of time instants.

The quality of the players’ strategies is evaluated as follows. Given initial time \(t_0 \in [0, T]\), initial state \(i_0 \in S\), a first player’s strategy \(\alpha \in U_{t_0}\), a stopping time \(\tau \in \mathcal{T}_{t_0}\) and a strategy of the second player \(\beta \in V_{t_0}\), the corresponding payoff within the one-side stopping Markov game is equal to

\[
\mathbb{E}^{\alpha, \beta}_{t_0, i_0} G^{\alpha, \beta}_{t_0}(\tau),
\]

where we denote

\[
G^{\alpha, \beta}_{t_0}(t, \omega) \triangleq g(t, \omega(t)) + \int_{t_0}^{t} h(t', \omega(t')), \alpha(t', \omega), \beta(t', \omega)) dt'.
\]

As above, in the case when \(\alpha = \alpha[u]\) (respectively, \(\beta = \beta[v]\)), we will write \(G^{u, \beta}_{t_0}\) (respectively, \(G^{\alpha, v}_{t_0}\)).
3 Main result

We impose the following assumptions:

(H1) The control spaces $U$ and $V$ are compact;

(H2) the transition rates $Q_{i,j}$ and the payoff functions $g(\cdot, i), h(\cdot, i, \cdot, \cdot)$ are continuous for each $i$ and $j$;

(H3) the function $g$ is Lipschitz continuous w.r.t. the time variable;

(H4) for every $t \in [0, T], i \in S$ and each $w \in \mathbb{R}^d$,

$$\min_{u \in U} \max_{v \in V} \sum_{j \in S} \left[ Q_{i,j}(t, u, v) w_j + h(t, i, u, v) \right] = \max_{v \in V} \min_{u \in U} \sum_{j \in S} \left[ Q_{i,j}(t, u, v) w_j + h(t, i, u, v) \right].$$

Notice that assumption (H4) can be considered as an analog of the Isaacs’ condition in the theory of differential games [10].

We consider the upper and lower values of the game those corresponds to the upper and lower games. We assume that within the upper game the first player chooses his/her strategy (that includes both controls and stopping time) from the class of feedback strategies, while the response of the second player is from the class of strategies with memory. Formally, this means the advantage of the second player. In the lower game, we give the advantage to the first player. The second player’s choice is restricted to the class of feedback strategies, while the first player chooses his/her policy from the class of strategies with memory.

Therefore, the upper and lower values of the one-side stopping Markov game are defined by the rules:

$$\text{Val}^+(t_0, i_0) \triangleq \inf_{u \in U_{t_0}} \sup_{v \in V_{t_0}} \inf_{\tau \in T_{t_0}} \mathbb{E}^{t_0, i_0}_u G^{u, \beta}_{t_0} (\tau),$$

$$\text{Val}^-(t_0, i_0) \triangleq \sup_{v \in V_{t_0}} \inf_{u \in U_{t_0}} \inf_{\tau \in T_{t_0}} \mathbb{E}^{t_0, i_0}_v G^{\alpha, v}_{t_0} (\tau).$$

Notice that due to the fact that, for each $t_0 \in [0, T], i_0 \in S, u \in U_{t_0}, v \in V_{t_0}$, the probability $\mathbb{P}^{u,v}_{t_0, i_0}$ is well-defined, we have that all aforementioned values are finite. Moreover,

$$\text{Val}^+(t_0, i_0) \geq \text{Val}^-(t_0, i_0). \quad (3)$$

Below, we show that the one-side stopping Markov games has a value. Additionally, we will show that this value can be obtained as a solution of the Zachrisson equation with unilateral constraints. The latter object is defined as follows. For $w \in \mathbb{R}^d, i \in S$, set

$$H_i(t, w) \triangleq \min_{u \in U} \max_{v \in V} \sum_{j \in S} \left[ Q_{i,j}(t, u, v) w_j + h(t, i, u, v) \right]. \quad (4)$$
Notice that, due to assumption [H4],
\[ H_i(t, w) = \max_{v \in V} \min_{u \in U} \sum_{j \in S} \left[ Q_{i,j}(t, u, v) w_j + h(t, i, u, v) \right]. \]

Below, if \( \phi : [0, T] \times S \to \mathbb{R} \), the \( \phi(t) \) stands for the vector \( (\phi(t, 1), \ldots, \phi(t, d))^T \).

**Definition 3.1.** We say that a function \( \varphi \in \mathcal{AC} \) satisfies the Zachrisson equation with unilateral constraints if, for each \( i \in S \),

- (Z1) \( \varphi(T, i) = g(T, i) \);
- (Z2) \( \varphi(t, i) \leq g(t, i) \) for \( t \in [0, T] \);
- (Z3) for every point \( t \in [0, T] \) such that \( \frac{d}{dt} \varphi(t, i) \) is well-defined and \( \varphi(t, i) < g(t, i) \), one has \( \frac{d}{dt} \varphi(t, i) \leq -H_i(t, \varphi(t)) \);
- (Z4) if \( \varphi(\cdot, i) \) is differentiable at \( t \in [0, T] \), then \( \frac{d}{dt} \varphi(t, i) \geq -H_i(t, \varphi(t)) \).

**Theorem 3.2.** Let \( \varphi \) satisfy the Zachrisson equation with unilateral constraints. Then at every \( t_0 \in [0, T] \), \( i_0 \in S \),

\[ \text{Val}^+(t_0, i_0) = \text{Val}^-(t_0, i_0) = \varphi(t_0, i_0). \]

**Theorem 3.3.** There exists a unique function \( \varphi \) satisfying the Zachrisson equation with unilateral constraints.

These two theorems obviously yield the following.

**Corollary 3.4.** There exists the value of the one-side stopping Markov game.

Theorem 3.2 and 3.3 are proved in Sections 4 and 5 respectively.

### 4 Suboptimal strategies

To prove Theorem 3.2, we consider suboptimal strategies of the players. Let us start with the construction of suboptimal strategy of the first player in the one-side stopping Markov game.

Given a function \( \varphi \in \mathcal{AC} \), we define a strategy \( \hat{u}_\varphi \in U_{fs} \) and a stopping rule \( \hat{\nu}_{\varphi, t_0} \in T_{fs} \) by the rules:

\[ \hat{u}_\varphi(t, i) \in \text{Argmin}_{u \in U} \left[ \sum_{j \in S} Q_{i,j}(t, u, v) \varphi(t, j) + h(t, i, u, v) \right], \quad (5) \]

\[ \hat{\nu}_{\varphi}(t, i) = \begin{cases} 1, & \varphi(t, i) = g(t, i), \\ 0, & \varphi(t, i) \neq g(t, i). \end{cases} \]

The latter corresponds to the following stopping time with values on the time interval \([t_0, T] \)

\[ \hat{\tau}_{\varphi, t_0}(\omega) = \bar{\tau}_{t_0} [\hat{\nu}_\varphi](\omega) = \min \{ t \in [t_0, T] : \varphi(t, \omega(t)) = g(t, \omega(t)) \}. \quad (6) \]
Proposition 4.1. Assume that \( \varphi \) satisfies conditions \((Z1)\) \((Z3)\) of Definition 3.1. Then, for every \( t_0 \in [0, T] \), \( i_0 \in \mathcal{S} \) and each \( \beta \in \mathcal{U}_{t_0} \), the following inequality holds true:

\[
E_{t_0, i_0}^{\hat{u}_\varphi, \beta} \varphi(\tau_{\varphi, t_0}) \leq \varphi(t_0, i_0)
\]

(7)

Proof. First, due to assumption \((Z1)\), we have that the stopping time \( \tau_{\varphi, t_0} \) takes values in \([t_0, T]\).

By the optional sampling theorem \([9\text{ Theorem 7.29}]\) and the facts that process \((2)\) with \( \alpha = \alpha[\hat{u}_\varphi] \) is a \((\mathbb{P}_{t_0, i_0}, \mathbb{F}_t)\)-martingale, while \( E_{t_0, i_0}^{\hat{u}_\varphi, \beta}(\{\omega : \omega(t_0) = i_0\}) = 1 \), we have that

\[
\varphi(t_0, i_0) = E_{t_0, i_0}^{\hat{u}_\varphi, \beta} \left[ \phi(\tau_{\varphi, t_0}, \omega(\tau_{\varphi, t_0})) - \int_{t_0}^{\tau_{\varphi, t_0}} \left[ \frac{\partial}{\partial t} \phi(t', \omega(t')) + \sum_{j \in \mathcal{S}} Q_{\omega(t'), j}(t', \hat{u}_\varphi(t', \omega(t')), \beta(t', \omega)) \phi(t', j) \right] dt' \right].
\]

The assumption that \( \varphi \) satisfies condition \((Z3)\) implies that

\[
\frac{\partial}{\partial t} \varphi(t', \omega(t')) \leq -H(t', \varphi(t', \omega(t')))
\]

for a.e. \( t' \in [t_0, \tau_{\varphi, t_0}(\omega)] \). This and the definition of the stopping rule \( \hat{\varphi} \) (that, in its own turn, determines the stopping time \( \tau_{\varphi, t_0} \)) give that

\[
\varphi(t_0, i_0) \geq E_{t_0, i_0}^{\hat{u}_\varphi, \beta} g(\tau_{\varphi, t_0}, \omega(\tau_{\varphi, t_0})) + E_{t_0, i_0}^{\hat{u}_\varphi, \beta} \int_{t_0}^{\tau_{\varphi, t_0}} \left[ H_i(t', \varphi(t')) - \sum_{j \in \mathcal{S}} Q_{\omega(t'), j}(t', \hat{u}_\varphi(t', \omega(t')), \beta(t', \omega)) \phi(t', j) \right] dt'.
\]

Due to the definitions of \( H_i \) (see \((4)\)) and the strategy \( \hat{u}_\varphi \) (see \((3)\)), we have that

\[
H_i(t', \varphi(t')) - \sum_{j \in \mathcal{S}} Q_{\omega(t'), j}(t', \hat{u}_\varphi(t', \omega(t')), \beta(t', \omega)) \phi(t', j)
\]

\[
\geq h(t', i, \hat{u}_\varphi(t', \omega(t')), \beta(t', \omega)).
\]

Therefore,

\[
\varphi(t_0, i_0) \geq E_{t_0, i_0}^{\hat{u}_\varphi, \beta} g(\tau_{\varphi, t_0}, \omega(\tau_{\varphi, t_0})) + \int_{t_0}^{\tau_{\varphi, t_0}} h(t', i, \hat{u}_\varphi(t', \omega(t')), \beta(t', \omega)) dt'.
\]

Using the definition of the process \( G \), we obtain inequality \((7)\). \( \square \)

Now we consider the second player’s suboptimal strategy \( \tilde{v} \) that is defined on the base of a function \( \phi \in \mathcal{AC} \) as follows:

\[
\tilde{v}_\phi(t, i) \in \text{Argmax}_{\varphi(t, j) \in \mathbb{V}} \left[ \sum_{j \in \mathcal{S}} Q_{i, j}(t, u, v) \varphi(t, j) + h(t, i, u, v) \right].
\]

(8)
Proposition 4.2. Let \( \varphi \in AC \) satisfy condition (Z2), (Z4). Then, for every \( t_0 \in [0, T] \), \( i_0 \in S \), \( \alpha \in U_{t_0} \), \( \tau \in T_{t_0} \),
\[
E^{\alpha, \tilde{v}_\varphi}_{t_0, i_0} G^{\alpha, \tilde{v}_\varphi}_{t_0} (\tau) \geq \varphi(t_0, i_0). \tag{9}
\]

Proof. The proof is also by the verification arguments. Using the fact that process (2) with \( \phi = \varphi \) and \( \beta = \beta[\tilde{v}_\varphi] \) is a \((P^{\alpha, \tilde{v}_\varphi}, F_{t_0})\)-martingale, the equality \( P^{\alpha, \tilde{v}_\varphi}_{t_0, i_0} (\{ \omega : \omega(t_0) = i_0 \}) = 1 \) and optional sampling theorem [9, Theorem 7.29], we conclude that
\[
\varphi(t_0, i_0) = E^{\alpha, \tilde{v}_\varphi}_{t_0, i_0} \left[ g(\tau, \omega(\tau)) - \int_{t_0}^{\tau} h(t', \omega(t'), \alpha(t', \omega), \tilde{v}(t', \omega(t'))) dt' \right].
\]

Applying conditions (Z2), (Z4), the definitions of the Hamiltonian (see (4)) and the strategy \( \tilde{v}_\varphi \) (see (8)), we obtain the following estimate:
\[
\varphi(t_0, i_0) \leq E^{\alpha, \tilde{v}_\varphi}_{t_0, i_0} \left[ g(\tau, \omega(\tau)) - \int_{t_0}^{\tau} h(t', \omega(t'), \alpha(t', \omega), \tilde{v}(t', \omega(t'))) dt' \right].
\]

This gives the statement of the proposition. \( \square \)

Proof of Theorem 3.2. This theorem directly follows from Propositions 4.1, 4.2 and inequality (3). \( \square \)

5 Existence of solution for the Zachrisson equation with unilateral constraints

To prove the existence of the solution of the Zachrisson equation with unilateral constraints, we consider the sequence of functions \( \varphi_N : [0, T] \times S \to \mathbb{R} \) defined as follows.

First, we set \( t^+_N \triangleq T k / N \). Then, let
\[
\varphi_N(T, i) \triangleq g(T, i). \tag{10}
\]

Now, assume that the function \( \varphi_N \) is already defined for all \( t \in [t^+_N, T] \) and every \( i \in S \). Let \( \psi^k_N : [t^+_N, t^{k+1}_N] \times S \to \mathbb{R} \) solve the boundary value problem on \([t^+_N, t^{k+1}_N]\):
\[
\frac{d}{dt} \psi^k_N(t, i) = -H_i(t, \psi^k_N(t)), \quad \psi^k_N(t^{k+1}_N, i) = \varphi_N(t^{k+1}_N, i), \quad i \in S. \tag{11}
\]

Finally, set, for \( t \in [t^+_N, t^{k+1}_N] \),
\[
\varphi_N(t, i) \triangleq \psi^k_N(t, i) \wedge g(t, i). \tag{12}
\]

Now examine the properties of the sequence \( \{ \varphi_N \}_{N=1}^{\infty} \).

Lemma 5.1. There exists a constant \( C_1 \) such that, for every natural \( N, t \in [0, T], i \in S \),
\[
|\varphi_N(t, i)| \leq C_1.
\]
Proof. First, denote
\[ M_Q \triangleq \max\{|Q_{i,j}(t,u,v)| : i, j \in S, t \in [0,T], u \in U, v \in V\}, \tag{13} \]
\[ M_h \triangleq \max\{|h(t,i,u,v)| : t \in [0,T], i \in S, u \in U, v \in V\}, \tag{14} \]
\[ M_g \triangleq \max\{|g(t,i)| : t \in [0,T], i \in S\}. \tag{15} \]
Due to assumptions \((H1),(H2)\) the quantities \(M_Q, M_h\) and \(M_g\) are finite.

Put
\[ M(t) \triangleq (M_g + M_h(T-t))e^{dM_Q(T-t)}. \tag{16} \]

We will prove the estimate
\[ |\varphi_N(t,i)| \leq M(t) \tag{17} \]
by the backward induction.

By the definition of the function \(M\), \((17)\) holds true at \(t = T\).

Further, assume that \((17)\) is fulfilled on \([t^{k+1}_N,T]\). Let us show that the functions
\[ \xi_N^k(t) \triangleq \max_{i \in S} |\psi_N^k(t,i)| \tag{18} \]
satisfies on \([t^k_N,t^{k+1}_N]\) the following inequality:
\[ \xi_N^k(t) \leq M(t). \tag{19} \]

Indeed, by \((11)\) and \((4)\) we have that
\[ \xi_N^k(t) \leq M(t^{k+1}_N) + \int_{[t^k_N,t^{k+1}_N]} dM_Q \xi_N^k(t') dt' + M_h(t^{k+1}_N - t). \]

Applying the Gronwall’s inequality and definition of the function \(M\) (see \((16)\)), we conclude that \((19)\) holds true on \([t^k_N,t^{k+1}_N]\). The definition of \(\varphi_N\) on \([t^k_N,t^{k+1}_N]\) (see \((12)\)) and the fact that \(g\) is bounded by \(M_g\), while \(M(t) \geq M_g\) yield estimate \((17)\) on \([t^k_N,t^{k+1}_N]\).

To complete the proof, let
\[ C_1 \triangleq M(0). \]

\[ \square \]

Lemma 5.2. If \(k \in \{0, \ldots, N-1\}, \ s, r \in [t^k_N,t^{k+1}_N]\), then
\[ |\psi_N^k(s,i) - \psi_N^k(r,i)| \leq C_2 |s-r|, \]
where \(C_2\) is a constant depending only on \(g, h\) and \(Q\),

Proof. Since \(\psi_N^k\) satisfies ODE \((11)\), it suffices to estimate the right-hand side of this equation. First, due to the definition of the Hamiltonian and the constants \(M_Q, M_h\) (see \((11), (13), (14)\)), we have that
\[ |H_i(t,\psi_N^k)| \leq dM_Q \max_{i \in S} |\psi_N^k(t,i)| + M_h. \]

Using the estimate of \(\max_{i \in S} |\psi_N^k(t,i)|\) by the function \(M\) (see \((18), (19)\)) and the fact that \(M(t) \leq C_1\), we obtain the statement of the lemma with \(C_2 \triangleq dM_Q C_1 + M_h. \)

\[ \square \]
Lemma 5.3. There exists a constant $C_3$ such that, for every natural $N$, $s,r \in [0,T]$, $i \in S$,
\[
|\varphi_N(s,i) - \varphi_N(r,i)| \leq C_3|s-r|.
\] (20)

Proof. Recall (see assumption [H3]) the function $g$ is Lipschitz continuous w.r.t. the time variable. Let $L_g$ denote this Lipschitz constant. Put
\[
C_3 \triangleq L_g \lor C_2.
\] (21)

Further, let us consider time instants $t', t''$ such that they lie in the same time interval $[t^k_N, t^{k+1}_N]$ for some $k$. Without loss of generality, one can assume that $\varphi_N(t', i) > \varphi_N(t'', i)$. To estimate $|\varphi_N(t', i) - \varphi_N(t'', i)| = \varphi_N(t', i) - \varphi_N(t'', i)$, we consider two cases.

1. $\varphi_N(t'', i) < g(t'', i)$. Therefore, $\varphi_N(t'', i) = \psi^k_N(t'', i)$, while $\varphi_N(t', i) \leq \psi^k_N(t', i)$. The latter inequality directly follows from the definition of the function $\varphi_N$ (see (12)). By Lemma 5.2, we have that
\[
\varphi_N(t', i) - \varphi_N(t'', i) \leq \psi^k_N(t', i) - \psi^k_N(t'', i) \leq C_2|t' - t''| \leq C_3|t' - t''|.
\]

2. $\varphi_N(t'', i) = g(t'', i)$. As above, by definition of the function $\varphi_N$ (see (12)), $\varphi_N(t', i) \leq g(t', i)$. Hence,
\[
\varphi_N(t', i) - \varphi_N(t'', i) \leq g(t', i) - g(t'', i) \leq L_g|t' - t''| \leq C_3|t' - t''|.
\]

Therefore, we derive (20) for $s,r$ lying in some interval $[t^k_N, t^{k+1}_N]$. To complete the proof it suffices to sum up inequalities (20) on small time intervals. \hfill \square

Lemma 5.4. Let $t \in [0,T]$ and $i \in S$ be such that $\phi_N(t,i) < g(t,i)$ and $\varphi_N$ is differentiable at $t$. Then
\[
\frac{d}{dt}\phi_N(t,i) \leq -H_i(t,\varphi_N(t)).
\] (22)

Proof. First, let us consider the case where $t \in (t^k_N, t^{k+1}_N)$ for some $k$. Due to the assumption that $\phi_N(t,i) < g(t,i)$, we have that
\[
\phi_N(t,i) = \psi^k_N(t,i).
\]

Moreover, the inequality $\phi_N(t',i) < g(t',i)$ holds in some neighborhood of $t$. Thus,
\[
\frac{d}{dt}\phi_N(t,i) = \frac{d}{dt}\psi^k_N(t).
\]

By construction of the function $\psi^k_N$, we have that
\[
\frac{d}{dt}\phi_N(t,i) = \frac{d}{dt}\psi^k_N(t) = -H_i(t,\psi^k_N(t)).
\]

Further, notice that $\varphi_N(t,j) \leq \psi^k_N(t,j)$, $Q_{i,j}(t,u,v) \geq 0$ for $i \neq j$. This and the equality $\phi_N(t,i) = \psi^k_N(t,i)$ imply that, for every $u \in U$ and $v \in V$,
\[
\sum_{j \in S} Q_{i,j}(t,u,v)\varphi_N(t,j) + h_i(t,i,u,v) \leq \sum_{j \in S} Q_{i,j}(t,u,v)\psi^k_N(t,j) + h_i(t,i,u,v).
\] (23)
Therefore, using the definition of the Hamiltonian (see (4)), we obtain the inequality
\[
H_i(t, \varphi_N(t)) \leq H_i(t, \psi^k_N(t)).
\]
Combining this with (5), we conclude that (22) holds true for \( t \neq t^k_N \).

Now let us consider the case when \( t = t^{k+1}_N \) for some \( k \). By construction of the function \( \varphi_N \), we have that \( \varphi_N(t^{k+1}_N) = \psi^k_N(t^{k+1}_N) \). Additionally, due to the assumption \( \varphi_N(t^{k+1}_N, i) < g(t^{k+1}_N, i) \) and the continuity of these functions, we conclude that \( \varphi_N(t', i) < g(t', i) \) when \( t' \in (t^{k+1}_N - \delta, t^{k+1}_N] \). Hence, the construction of the function \( \varphi_N \) gives that \( \varphi_N(t', i) = \psi^k_N(t', i) \) on \( (t^{k+1}_N - \delta, t^{k+1}_N] \). Therefore,
\[
\frac{d}{dt} \varphi_N(t^{k+1}_N, i) = \frac{d}{dt} \psi^k_N(t^{k+1}_N, i) = -H_i(t^{k+1}_N, \psi^{k-1}_N(t^{k+1}_N)).
\]
Here \( \frac{d}{dt} \) stands for the left-hand derivative. This and (23) yield that (22) holds true in the case when \( t \in \{ t^k_N \}^{N-1}_{k=1} \).

**Lemma 5.5.** There exists a constant \( C_4 \) such that, for every \( s, r \in [0, T] \), \( s < r \), and \( i \in S \), the following inequality holds true:
\[
\varphi_N(r, i) - \varphi_N(s, i) \geq \int_s^r H_i(t, \varphi_N(t))dt - \frac{C_4}{N}
\]

**Proof.** Let \( m, n \in \{ 0, \ldots, N - 1 \} \) be such that \( s \in [t^m_N, t^{m+1}_N), r \in (t^n_N, t^{n+1}_N] \).
We have that
\[
\varphi_N(r, i) - \varphi_N(s, i) = \varphi_N(r, i) - \varphi_N(t^m_N) + \varphi_N(t^{m+1}_N, i) - \varphi_N(s, i)
\]
\[
+ \sum_{k=m+1}^{n-1} (\varphi_N(t^{k+1}_N, i) - \varphi_N(t^k_N, i))
\]
\[
\geq \varphi_N(r, i) - \varphi_N(t^m_N) + \varphi_N(t^{m+1}_N, i) - \varphi_N(s, i)
\]
\[
+ \sum_{k=m+1}^{n-1} (\psi^k_N(t^{k+1}_N, i) - \psi^k_N(t^k_N, i)).
\]

In the latter inequality, we used the definition of the functions \( \psi^k_N \) and \( \varphi_N \) (see (11) and (12)). Further, due to Lemma 5.3, \( |\varphi_N(r, i) - \varphi_N(t^m_N)|, |\varphi_N(t^{m+1}_N, i) - \varphi_N(s, i)| \leq C_3 T/N \). Using this and the fact that each function \( \psi^k_N(\cdot) \) satisfies (11), we conclude that, for each \( k \),
\[
\varphi_N(r, i) - \varphi_N(s, i) \geq -2C_3 \frac{T}{N} + \sum_{k=m+1}^{n-1} \int_{t^k_N}^{t^{k+1}_N} H_i(t, \psi^k_N(t))dt. \tag{24}
\]
By the equality \( \psi^k_N(t^{k+1}_N) = \varphi_N(t^{k+1}_N) \), we have that
\[
H_i(t, \psi^k_N(t)) = H_i(t, \varphi_N(t)) + (H_i(t, \psi^k_N(t)) - H_i(t, \psi^k_N(t^{k+1}_N)))
\]
\[
+ (H_i(t, \varphi_N(t^{k+1}_N)) - H_i(t, \varphi_N(t))).
\]

\[11\]
Lemmas 5.2, 5.3 and the fact that $H_i$ is Lipschitz continuous w.r.t. the second variable with the constant $dM_Q$ give that, for $t \in [t^k_N, t^{k+1}_N]$,

$$H_i(t, \psi^k_N(t)) \geq H_i(t, \phi^k_N(t)) - (C_3 + C_2)dM_Q \frac{T}{N}. \quad (25)$$

Finally, using the definition of Hamiltonian (see (4)) and Lemma 5.1, we conclude that

$$|H_i(t, \phi_N(t))| \leq C'.$$

Here $C'$ is a constant dependent only on the examined game. This, inequalities (24) and (25) and the fact that $r - t^m_N, t^{m+1}_N - s \leq T/N$ imply that

$$\phi_N(r, i) - \phi_N(s, i) \geq -(3C_3 + C_2 + C') \frac{T}{N} + \int_s^{t^{m+1}_N} H_i(t, \phi_N(t)) dt$$

$$+ \sum_{k=m+1}^{n-1} \int_{t^k_N}^{t^{k+1}_N} H_i(t, \phi_N(t)) dt + \int_{t^m_N}^r H_i(t, \phi_N(t)) dt$$

$$= -(3C_3 + C_2 + C') \frac{T}{N} + \int_s^r H_i(t, \phi_N(t)) dt.$$

This gives the statement of the lemma.

\[ \square \]

**Proof of Theorem 3.3.** As it was mentioned above, to prove the existence of the Zachriss equation with unilateral constraints, we use the sequence of functions $\{\phi_N\}_{N=1}^\infty$. Due to Lemma 5.1 these functions are uniformly bounded. Lemma 5.3 says that these functions are Lipschitz continuous with the same constant. By the Arzela–Ascoli theorem, there exist a sequence $\{\phi_{N_l}\}_{l=1}^\infty$ and a function $\phi$ such that $\phi_{N_l} \to \phi$ as $l \to \infty$. The limiting function $\phi$ is bounded by the constant $C_1$ and is Lipschitz continuous with the constant $C_3$. Hence,

$$\phi \in AC.$$

Further, by construction, each function $\phi_{N_l}$ satisfies the boundary condition

$$\phi_{N_l}(T) = g(T).$$

Passing to the limit, we obtain that $\phi$ satisfies condition (Z1). Analogously, (12) implies that

$$\phi_{N_l}(t, i) \leq g(t, i), \quad i \in S.$$ 

Therefore, condition (Z2) is fulfilled for the limiting function $\phi$.

Let $i \in S$ and let $t \in (0, T)$ be a point of differentiability of the functions $\phi(\cdot, i)$ such that $\phi(t, i) < g(t, i)$. Choose $\varepsilon$ satisfying

$$3\varepsilon \leq g(t, i) - \phi(t, i).$$

Due to the continuity of the functions $g$ and $\phi$, we have that there exists $\delta > 0$ such that, for $t' \in (t - \delta, t + \delta)$,

$$\phi(t', i) + 2\varepsilon < g(t', i).$$
Using the fact that \( \varphi_{N_l} \to \varphi \) as \( l \to \infty \), we have that, for sufficiently large \( l \) and every \( t' \in (t - \delta, t + \delta) \),
\[
\varphi_{N_l}(t', i) + \varepsilon < g(t', i).
\]
Lemma 5.4 implies that, if \( \hat{t} \in (t - \delta, t + \delta) \),
\[
\varphi_{N_l}(\hat{t}, i) - \varphi_{N_l}(t) \leq - \int_{t}^{\hat{t}} H_i(t', \varphi_{N_l}(t')) dt'.
\]
Passing to the limit, we conclude that
\[
\varphi(\hat{t}, i) - \varphi(t) \leq - \int_{t}^{\hat{t}} H_i(t', \varphi(t')) dt'
\]
when \( \hat{t} \in (t - \delta, t + \delta) \). Since \( t \in (0, T) \) is a point of differentiability of the function \( \varphi(\cdot, i) \), we arrive at condition \((Z3)\).

To prove that \( \varphi \) satisfies \((Z4)\), notice that Lemma 5.5 implies that, for every \( s, r \in [0, T], s < r \),
\[
\varphi_{N_l}(r) - \varphi_{N_l}(s) \geq - \int_{s}^{r} H_i(t, \varphi_{N_l}(t)) dt - C_4/N.
\]
Passing to the limit, we have
\[
\varphi(r) - \varphi(s) \geq - \int_{s}^{r} H_i(t, \varphi(t)) dt.
\]
This gives condition \((Z4)\) at every point of differentiability of the function \( \varphi \).

The uniqueness follows from Theorem 3.2. Indeed, each functions satisfying \((3.1)\) coincides with both upper and lower values.

\[\Box\]

**Acknowledgment.** The paper was prepared within the framework of the HSE University Basic Research Program in 2022.

**References**

[1] Y. V. Averboukh. Approximation of value function of differential game with minimal cost. *Vestn. Udmurt. Univ.: Mat. Mekhanika Komp’yuternye Nauki*, 31(4):536–561, 2021.

[2] E. N. Barron. Differential games with maximum cost. *Nonlinear Anal. Theory Methods Appl.*, 14(11):971–989, 1990.

[3] E. B. Dynkin. Game variant of a problem on optimal stopping. *Soviet Math. Dokl.*, 10:270–274, 1969.

[4] E. B. Frid. The optimal stopping rule for a two-person markov chain with opposing interests. *Theory Probab. Appl.*, 14(4):713–716, 1969.
[5] M. K. Ghosh and K. S. M. Rao. Zero-sum stochastic games with stopping and control. 
Oper. Res. Lett., 35(6):799–804, 2007.

[6] X. Guo, O. Hernández-Lerma, and T. Prieto-Rumeau. A survey of recent results on continuous-time Markov decision processes (with comments and rejoinder). 
Top, 14(2):177–261, 2006.

[7] X. P. Guo and O. Hernández-Lerma. Continuous-time Markov Decision Processes: 
Theory and Applications. Springer, New York, 2009.

[8] M. Horiguchi. Markov decision processes with a stopping time constraint. 
Math. Methods Oper. Res., 53(2):279–295, 2001.

[9] O. Kallenberg. Foundations of Modern Probability. Springer, New York, 2021.

[10] N. N. Krasovskii and A. I. Subbotin. Game-theoretical control problems. 
Springer, New York, 1988.

[11] E. Mordecki. Optimal stopping and perpetual options for Lévy processes. 
Finance Stoch., 6:473–493, 2002.

[12] A. N. Shiryaev. Optimal Stopping Rules. Springer, New York, 2008.

[13] A. I. Subbotin. Generalized solutions of first-order PDEs. The dynamical perspective. 
Birkhauser, Boston, 1995.

[14] L. Zachrisson. Markov games. Ann. Math. Stud., 52:211–253, 1964.

Yuri Averboukh: Higher School of Economics, 
11 Pokrovsky Bulvar, Moscow, Russia; 
Krasovskii Institute of Mathematics and Mechanics 
16 S. Kovalevskoi str., Yekaterinburg, Russia; 
e-mail: averboukh@gmail.com