Nonlinear Dynamics and Systems Theory, **17** (3) (2017) 303–314

Approximation of the Optimal Control Problem on an Interval with a Family of Optimization Problems on Time Scales

O. Lavrova *[1], V. Mogylova [2], O. Stanzhytskyi,[1] and O. Misiats [3]

[1] *Taras Shevchenko National University of Kyiv, Kyiv, Ukraine*
[2] *National Technical University of Ukraine "Kyiv Polytechnic Institute", Kyiv, Ukraine*
[3] *Courant Institute of Mathematical Sciences, New York*

Received: December 15, 2016; Revised: July 12, 2017

Abstract: In this paper we consider a family of optimization problems defined on variable time scales $\mathbb{T}_\lambda$, which depend on the parameter λ We prove that the family of value functions $V_\lambda(t_0, x)$ of the optimal control problem on $[t_0, t_1]_{\mathbb{T}_\lambda}$ converges locally uniformly in $\mathbb{R}^d$ to the value function $V(t_0, x)$ of the optimal control problem on $[t_0, t_1]$, provided $\sup_{t\in[t_0,t_1]_{\mathbb{T}_\lambda}} \mu_\lambda(t) \to 0$ as $\lambda \to 0$, where $\mu_\lambda(t)$ is the graininess function of $\mathbb{T}_\lambda$.

Keywords: *time scale; value function; right-scattered point; right-dense point; graininess function.*

Mathematics Subject Classification (2010): 34K35, 34N99, 39A13, 49K15, 93C15, 93C56.

1 Introduction

This work is devoted to the study of the limiting behavior of the optimal control problem for dynamic equations, defined on a family of time scales $\mathbb{T}_\lambda$, in the regime when the graininess function μ_λ converges to zero as $\lambda \to 0$. At the same time the segment of the time scale $[t_0, t_1]_{\mathbb{T}_\lambda} = [t_0, t_1] \cap \mathbb{T}_\lambda$ approaches $[t_0, t_1]$ e.g. in the Hausdorff metric. The natural question that arises is how the optimal control problem on the time scale is related to the corresponding control problem on the interval $[t_0, t_1]$.

The answer to the above question is well understood for Eulerian time scales (according to classification [6]) that is, if $\mathbb{T}_\lambda = \lambda\mathbb{Z}_+$, $\lambda > 0$, and the equation on time scales becomes a difference equation.

* Corresponding author: mailto:lavrova_olia@ukr.net

The results listed above are based on Euler polygonal method, which guarantees that the corresponding solutions of differential and difference equations on finite time intervals are close to each other, provided the steps are small. This method works really well if the right-hand sides of differential equations are continuous. In this case, both solutions are smooth, which makes it relatively easy to estimate the difference between them. However, the right-hand sides of the optimal control problem, considered in this paper, depend on the control parameter function $u(t)$. Generally speaking, $u(t)$ is only measurable. This makes the solution of the differential equation only absolutely continuous. In turn, this significantly complicates the estimates for the difference between corresponding solutions. Estimates of this type were obtained with convex analysis techniques in the works [9] – [11]. Using these estimates, the authors showed that the value function for the difference equation approximation converges to the corresponding value function for continuous differential equation as the approximation step goes to zero.

Our work extends the result [9] – [11] on the limiting behavior of the value function to the case of general time scales. However, we use different methods since the topological structure of the time scale we are considering may be complex. The main difficulty in our work is to establish the *uniform* convergence of solutions of the Cauchy problem on $[t_0, t_1] \cap_{\mathbb{T}_\lambda}$ to the solution of the corresponding Cauchy problem on $[t_0, t_1]$. This makes our analysis significantly different from [12], where only special pointwise convergence was obtained. More sophisticated approach is necessary because, in contrast with [12], the right-sides of our equations are not piecewise continuous, as well as we are dealing with much more general time scales (as opposed to the Eulerian time scale in [9] – [11]).

This paper is organized as follows. In Section 2 we provide some definitions and preliminary results on time scales calculus, and state the main result. The main result on the convergence of the family of the value functions to the value function of the limit problem is proved in Section 3.

2 Preliminaries and Main Result

2.1 Basic notions of time scales theory

The time scales theory was introduced by S. Hilger in his PhD thesis [13] (1988) as a unified theory for both discrete and continuous analysis. This theory was further developed by a number of authors, see [4] and references therein. For reader's convenience, we present several notions from this theory, which are used in this paper.

Time scale $\mathbb{T}$ is a non-empty closed subset of $\mathbb{R}$; $A_{\mathbb{T}} := A \cap \mathbb{T}$ for $A \subset \mathbb{R}$; $\sigma : \mathbb{T} \to \mathbb{T}$, $\sigma(t) := \inf\{s \in \mathbb{T} : s > t\}$ is the *forward jump operator*; $\rho : \mathbb{T} \to \mathbb{T}$, $\rho(t) = \sup\{s \in \mathbb{T} : s < t\}$ is the *backward jump operator* (here $\inf \emptyset := \sup \mathbb{T}$ and $\sup \emptyset := \inf \mathbb{T}$); $\mu : \mathbb{T} \to [0, \infty)$, $\mu(t) := \sigma(t) - t$ is called *the graininess function.* A point $t \in \mathbb{T}$ is called *left-dense* (LD) (*left-scattered* (LS), *right-dense* (RD) or *right-scattered* (RS)) if $\rho(t) = t$ ($\rho(t) < t$, $\sigma(t) = t$ or $\sigma(t) > t$); $\mathbb{T}^k := \mathbb{T} \setminus \{M\}$ if $\mathbb{T}$ has a left-scattered maximum M, $\mathbb{T}^k := \mathbb{T}$ otherwise.

A function $f : \mathbb{T} \to \mathbb{R}^d$ is called Δ*–differentiable* at $t \in \mathbb{T}^k$ if the limit

$$f^\Delta(t) = \lim_{s \to t} \frac{f(\sigma(t)) - f(s)}{\sigma(t) - s}$$

exists in $\mathbb{R}^d$. The properties of the Lebesgue Δ – measure and the Lebesgue Δ – integrability are described, e.g. in [3].

2.2 Control theory on time scales

Let $\mathbb{T}$ be a time scale, such that $\sup \mathbb{T} = +\infty$, $t_0, t_1 \in \mathbb{T}$, and $U \subset \mathbb{R}^m$ is a compact set.

▷ An *optimal control problem on the time scale* $\mathbb{T}$ is the problem of the type

$$\begin{cases} x^\Delta = f(t,x,u), \\ x(t_0) = x_0, \\ J(u) = \int_{[t_0,t_1)_\mathbb{T}} L(s,x(s),u(s))\Delta s + \Psi(x(t_1)) \to \inf, u \in \mathcal{U}(t_0), \end{cases} \tag{2.1}$$

where $f : [t_0,t_1]_\mathbb{T} \times \mathbb{R}^d \times U \to \mathbb{R}^d$, $L : [t_0,t_1]_\mathbb{T} \times \mathbb{R}^d \times U \to \mathbb{R}^1$ and $\Psi : \mathbb{R}^d \to \mathbb{R}^1$.

▷ $\mathcal{U}(t) := L^\infty([t,t_1]_\mathbb{T}, U)$, i.e. the set of bounded, Δ–measurable functions [5, Chapter 5.7] defined on $[t,t_1]_\mathbb{T}$ and taking values in U for each $t \in [t_0,t_1)_\mathbb{T}$, is called *the set of admissible controls.*

▷ The *Bellman function* (or *the value function*) is

$$V(t_0,x_0) := \inf_{u(\cdot)\in\mathcal{U}(t_0)} J(t_0,x_0,u). \tag{2.2}$$

2.3 Main result

Let $\Lambda \subset \mathbb{R}$, such that 0 is a limit point of Λ, be the set of indices. Consider the family of time scales $\mathbb{T}_\lambda, \lambda \in \Lambda$, such that $\sup \mathbb{T}_\lambda = \infty$. For any $t_0, t_1 \in \mathbb{T}_\lambda$, denote $[t_0,t_1]_{\mathbb{T}_\lambda} = [t_0,t_1] \cap \mathbb{T}_\lambda$ and $\mu_\lambda = \sup_{t\in[t_0,t_1]_{\mathbb{T}_\lambda}} \mu(t)$. Assume

$$\mu_\lambda(t) \to 0 \text{ as } \lambda \to 0. \tag{2.3}$$

In this case $\mathbb{T}_\lambda$ converges (e.g. in the Hausdorff metric) to a continuous time scale $\mathbb{T}_0$ (here we use the classification from [6]), and hence $[t_0,t_1]_{\mathbb{T}_\lambda}$ becomes $[t_0,t_1]$ in the limit $\lambda \to 0$. For every $\mathbb{T}_\lambda$ consider the optimal control problem on the time scale $[t_0,t_1]_{\mathbb{T}_\lambda}$:

$$\begin{cases} x^\Delta = f(t,x,u), \\ x(t_0) = x, \\ J_\lambda(u) = \int_{[t_0,t_1)_{\mathbb{T}_\lambda}} L(t,x(t),u(t))\Delta t \to \inf, u \in \mathcal{U}(t_0). \end{cases} \tag{2.4}$$

Along with (2.4), consider the corresponding continuous optimal control problem on the interval $[t_0,t_1]$:

$$\begin{cases} \frac{dx(t)}{dt} = f(t,x(t),u(t)), \\ x(t_0) = x, \\ J(u) = \int_{t_0}^{t_1} L(t,x(t),u(t))dt \to \inf, u \in \mathcal{U}(t_0). \end{cases} \tag{2.5}$$

Denote $V_\lambda(t_0,x)$ and $V(t_0,x)$ to be the corresponding Bellman functions for these problems, given by (2.2). Our main result is the following theorem.

Theorem 2.1 *Let $\mathbb{T}_\lambda$ be such that (2.3) holds. In addition, assume that*

1) The functions f, f_x and L are continuous on $[t_0,t_1] \times \mathbb{R}^d \times U$;

2) f and L are globally Lipschitz in x, with Lipschitz constant $K > 0$.

Then

$$V_\lambda(t_0,\cdot) \to V(t_0,\cdot) \text{ in } C_{loc}(\mathbb{R}^d),\ \lambda \to 0. \tag{2.6}$$

3 Proof of Theorem 2.1.

Without loss of generality, we assume that $t_0 = 0$ and $t_1 = 1$. Consider an arbitrary time scale $\mathbb{T}_\lambda$ and an arbitrary admissible control $u_\lambda(t)$ on it. Let $x_\lambda(t)$ be a corresponding admissible trajectory. Denote $\tilde{u}_\lambda(t)$ to be the extension of $u_\lambda(t)$ to the entire interval $[0, 1]$:

$$\tilde{u}_\lambda(t) := \begin{cases} u_\lambda(t), t \in [0,1]_{\mathbb{T}_\lambda}, \\ u_\lambda(r), t \in [r, \sigma(r)),\ \ r \in \mathrm{RS}. \end{cases} \tag{3.1}$$

This control is admissible for the problem (2.5). The proof of the main result will heavily rely on the following two lemmas.

Lemma 3.1 *Let $x(t)$ be a solution of*

$$\begin{cases} \frac{dx}{dt} = f(t, x, \tilde{u}_\lambda(t)), \\ x(0) = x_0. \end{cases}$$

Then

$$\left| \int_{[0,1)_{\mathbb{T}_\lambda}} L(t, x_\lambda(t), u_\lambda(t))\Delta t - \int_0^1 L(t, x(t), \tilde{u}_\lambda(t))dt \right| \to 0,\ \lambda \to 0. \tag{3.2}$$

Proof. Fix $\varepsilon > 0$. Our goal is to show that the expression in (3.2) can be made less than ε for all sufficiently small λ. Using Gronwall inequality and its analogue for time scales [4], one can show that for any $r > 0$ there is $C(r) > 0$ such that

$$|x_\lambda(t)| \le C(r),\ t \in [0,1]_{\mathbb{T}_\lambda}, |x(t)| \le C(r),\ t \in [0,1],\ |x_0| \le r. \tag{3.3}$$

The estimates (3.3) are uniform for all admissible controls, since U is compact. Therefore, there is a constant $C_1(r) > 0$ such that

$$\begin{gathered} |L(t, x_\lambda(t), u_\lambda(t))| \le C_1(r),\ |f(t, x_\lambda(t), u_\lambda(t))| \le C_1(r), \\ |f_x(t, x_\lambda(t), u_\lambda(t))| \le C_1(r),\ \forall t \in [0,1]_{\mathbb{T}_\lambda} \text{ and } \lambda \in \Lambda. \end{gathered} \tag{3.4}$$

Then we have

$$\begin{aligned} \int_{[0,1)_{\mathbb{T}_\lambda}} L(t, x_\lambda(t), u_\lambda(t))dt = & \int_{[0,1)_{\mathbb{T}} \setminus \mathrm{RS}} L(t, x_\lambda(t), u_\lambda(t))dt \\ & + \sum_{r \in \mathrm{RS}} L(r, x_\lambda(r), u_\lambda(r))\mu(r). \end{aligned} \tag{3.5}$$

In view of (3.4),

$$\sum_{r \in \mathrm{RS}} L(r, x_\lambda(r), u_\lambda(r))\mu(r) \le C_1 \sum_{r \in \mathrm{RS}} \mu_\lambda(r),$$

which holds true regardless the sums are finite or infinite. Then

$$\begin{aligned} \sum_{r \in \mathrm{RS}} L(r, x_\lambda(r), u_\lambda(r))\mu(r) = & \sum_{k=1}^{N} L(r_k, x_\lambda(r_k), u_\lambda(r_k))\mu(r_k) \\ & + \sum_{k \ge N+1} L(r_k, x_\lambda(r_k), u_\lambda(r_k))\mu(r_k), \end{aligned} \tag{3.6}$$

where $N = N(\lambda) \geq 1$ is chosen so that

$$\sum_{k=N(\lambda)+1} \mu(r_k) \leq \frac{\mu_\lambda}{2}. \tag{3.7}$$

We now remove the right-scattered points, which appear in the sum (3.7), from the time scale. By construction, their total Δ–measure does not exceed $\frac{\mu_\lambda}{2}$. Denote $A = \bigcup_{k=N(\lambda)+1}[r_k, \sigma(r_k))$. Clearly, $|A| \leq \frac{\mu_\lambda}{2}$, where $|A|$ stands for Lebesgue measure of A. Denote $B := [0,1]\backslash A$.

Next, in the same way as it was done in (3.1), we may define a piecewise-constant extension of $x_\lambda(t)$ to the entire interval $[0,1]$. This extension is denoted with $\tilde{x}_\lambda(t)$. Similarly, the function $L(t,x,u)$, which is defined only for $t \in \mathbb{T}_\lambda$, may be extended to $\widetilde{L}(t,x,u)$, defined for $t \in [0,1]$. Clearly, this extension satisfies the same bound $|\widetilde{L}(t,x,u)| \leq C$. Therefore, using the results from [7, Theorem 2.9.],

$$\int_{[0,1]_{\mathbb{T}_\lambda}} L(t, x_\lambda(t), u_\lambda(t))\Delta t = \int_0^1 \widetilde{L}(t, \tilde{x}_\lambda(t), \tilde{u}_\lambda(t))dt.$$

Consequently,

$$\left| \int_{[0,1)_{\mathbb{T}_\lambda}} L(t, x_\lambda(t), u_\lambda(t))\Delta t - \int_0^1 L(t, x(t), \tilde{u}_\lambda(t))dt \right| \\ \leq C\mu_\lambda + \int_B \left| (\widetilde{L}(t, \tilde{x}_\lambda(t), \tilde{u}_\lambda(t)) - L(t, x(t), \tilde{u}_\lambda(t))) \right| dt. \tag{3.8}$$

Let us estimate the last integral in (3.8). The set B consists of a finite number of right-scattered points $(r_1, \dots, r_N)$ and possibly intervals between them, consisting of limit points. In view of (3.3) and the compactness of U, the functions $f(t,x,u)$ and $L(t,x,u)$, without loss of generality, are defined on a compact set, hence they are uniformly continuous. Therefore, there exists $\varepsilon_1 = \varepsilon_1(\varepsilon) > 0$ such that

$$|L(t,x,u) - L(s,x,u)| < \varepsilon, \ |f(t,x,u) - f(s,x,u)| < \varepsilon, \text{ if } |t-s| < \varepsilon_1. \tag{3.9}$$

In view of (2.3), we can choose λ small enough so that $\mu_\lambda < \varepsilon_1$. Denoting $B_1 = B \setminus \bigcup_{i=1}^N [r_i, \sigma(r_i))$, we have

$$\int_B \widetilde{L}(t, \tilde{x}_\lambda(t), \tilde{u}_\lambda(t))dt = \int_{B_1} L(t, x_\lambda(t), u_\lambda(t))dt + \sum_{i=1}^N \int_{r_i}^{\sigma(r_i)} L(r_i, \tilde{x}_\lambda(t), \tilde{u}_\lambda(t))dt.$$

Hence,

$$\int_B |\widetilde{L}(t, \tilde{x}_\lambda(t), \tilde{u}_\lambda(t)) - L(t, x(t), \tilde{u}_\lambda(t))|dt \leq K \int_B |x_\lambda(t) - x(t)|dt + \varepsilon, \tag{3.10}$$

where we used (3.9) and the Lipschitz property of L. Now, we estimate the difference $|\tilde{x}_\lambda(t) - x(t)|$. Without loss of generality, assume that the time scale $\mathbb{T}_\lambda$ has the following structure (Figure 1).

Here

1) the solid line indicates the line segments, which consist of limit points;

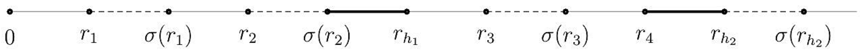

Figure 1: The structure of time scale.

2) the dashed line indicates the line segments $[r_i, \sigma(r_i))$, i.e. r_i are the remaining right-scattered points;

3) the boldface solid line indicates the set of points which were removed, i.e. the set A.

The argument is similar for other structures of time scales.

1) For $t \in [0, r_1]$ we have $u_\lambda(t) = \tilde{u}_\lambda(t)$, therefore $\tilde{x}_\lambda(t) = x(t)$.

2) For $t \in [r_1, \sigma(r_1))$, clearly $\tilde{x}_\lambda(t) = x_\lambda(r_1) = x(r_1)$ and $\tilde{u}_\lambda(t) = u_\lambda(r_1)$. It follows from the integral representation of the solution

$$x(t) = x(r_1) + \int_{r_1}^{t} f(s, x(s), u_\lambda(r_1))ds$$

that $x \in C^2[r_1, \sigma(r_1))$. Hence, using Taylor's expansion with the remainder in the Lagrange form, we obtain

$$x(t) = x(r_1) + f(r_1, x(r_1), u_\lambda(r_1))(t - r_1)$$
$$+ f'_x(s_1, x(s_1), u_\lambda(r_1)) \cdot f(s_1, x(s_1), u_\lambda(r_1))\frac{(t - r_1)^2}{2}, \tag{3.11}$$

for some $s_1 \in [r_1, \sigma(r_1)]$. Here f'_x is the Jacobian matrix. It follows from (3.4) that

$$\max_{t \in [t_0, t_1]} |f'_x(t, x(t), u_\lambda(t))f(t, x(t), u_\lambda(t))| \leq C_1^2. \tag{3.12}$$

Thus, when $t \in [r_1, \sigma(r_1))$, we obtain

$$|x(t) - \tilde{x}_\lambda(t)| \leq \int_{r_1}^{\sigma(r_1)} |f(t, x(t), u_\lambda(r_1))|\, dt \leq C_1\mu(r_1). \tag{3.13}$$

But when $t = \sigma(r_1)$ we have

$$\tilde{x}_\lambda(\sigma(r_1)) = x_\lambda(r_1) + f(r_1, x_\lambda(r_1), u_\lambda(r_1))\mu(r_1) = x(r_1) + f(r_1, x(r_1), u_\lambda(r_1))\mu(r_1).$$

Thus, from (3.11) and (3.12) we get

$$|x(\sigma(r_1)) - \tilde{x}_\lambda(\sigma(r_1))| \leq C_1^2 \frac{\mu_\lambda^2(r_1)}{2}. \tag{3.14}$$

3) For $t \in [\sigma(r_1), r_2]$, it follows from (3.14) and Gronwall inequality

$$|\tilde{x}_\lambda(t) - x(t)| \leq \frac{\mu_\lambda^2(r_1)}{2} C_1^2 e^{K(r_2 - \sigma(r_1))}. \tag{3.15}$$

4) For $t \in [r_2, \sigma(r_2))$ we may argue the same way as for $t \in [r_1, \sigma(r_1))$ to get

$$|x(t) - \tilde{x}_\lambda(t)| \leq \frac{\mu_\lambda^2(r_1)}{2} C_1^2 e^{K(r_2 - \sigma(r_1))} + \mu_\lambda(r_2) C_1. \tag{3.16}$$

$$|x(\sigma(r_2)) - \tilde{x}_\lambda(\sigma(r_2))| \leq \frac{\mu_\lambda^2(r_1)}{2} C_1^2 [(1 + K\mu_\lambda(r_2)) e^{K(r_2 - \sigma(r_1))} + 1].$$

5) On the line segment $t \in [\sigma(r_2), r_{h_1}]$ we have

$$|\tilde{x}(r_{h_1}) - \tilde{x}(\sigma(r_2))| \leq C_1(r_{h_1} - r_2) = C_1\mu_1,\ |x(r_{h_1}) - x(\sigma(r_2))| \leq C_1\mu_1,$$

$$|\tilde{x}_\lambda(r_{h_1}) - x(r_{h_1})| \leq 2C_1\mu_1 + (1 + K\mu_\lambda(r_2)) \frac{\mu_\lambda^2(r_1)}{2} C_1^2 e^{K(r_2 - \sigma(r_1))} + \frac{\mu_\lambda^2(r_2)}{2} C_1^2. \tag{3.17}$$

Continuing this procedure to the remaining intervals of Figure 1 for $t \in [r_{h_1}, r_5]$, we have the following estimate

$$\begin{aligned}
&|\tilde{x}_\lambda(t) - x(t)| \leq ((1 + K\mu_\lambda(r_{h_2}))(1 + K\mu_\lambda(r_3))2C_1\mu_1 \\
&* e^{K((r_4 - \sigma(r_3)) + (r_3 - r_{h_1}))} + \frac{1}{2}(1 + K\mu_\lambda(r_{h_2})) e^{K((r_4 - \sigma(r_3))} \mu_\lambda^2(r_1) C_1^2 \\
&* e^{K((r_2 - \sigma(r_1)) + (r_3 - r_{h_1}))} (1 + K\mu_\lambda(r_2))(1 + K\mu_\lambda(r_3)) + \frac{1}{2}(1 + K\mu_\lambda(r_{h_2})) \\
&* e^{K((r_4 - \sigma(r_3))} \mu_\lambda^2(r_2) C_1^2 e^{K(r_3 - \sigma(r_{h_1}))} (1 + K\mu_\lambda(r_3)) + \frac{1}{2}(1 + K\mu_\lambda(r_{h_2})) \\
&* e^{K((r_4 - \sigma(r_3))} \mu_\lambda^2(r_3) C_1^2 + 2C_1\mu_2(1 + K\mu_\lambda(r_{h_2})) + \frac{1}{2}\mu_\lambda^2(r_{h_2}) C_1^2) e^{K(r_5 - \sigma(r_{h_2}))}.
\end{aligned} \tag{3.18}$$

Denote

$$\Pi := (1 + K\mu_\lambda(r_1))(1 + K\mu_\lambda(r_2))(1 + K\mu_\lambda(r_3))(1 + K\mu_\lambda(r_{h_2})) \ldots (1 + K\mu_\lambda(r_N)),$$

where the product is taken over all right-scattered points of Figure 1. Then

$$\ln \Pi \leq K(\mu_\lambda(r_1) + \mu_\lambda(r_2) + \mu_\lambda(r_3) \ldots) \leq K.$$

Note also that the sum of all powers of e that appear in (3.18) does not exceed K, since those arguments involve the lengths of disjoint subintervals of $[0, 1]$. Consequently, for $t \notin [r_k, \sigma(r_k))$ we have

$$|\tilde{x}_\lambda(t) - x(t)| \leq \mu_\lambda \left(\Pi e^K C_1 + \frac{1}{4} \Pi C_1^2 e^K \right) \to 0,\ \lambda \to 0. \tag{3.19}$$

For $t \in [r_k, \sigma(r_k))$, arguing as in (3.19), we get

$$|\tilde{x}_\lambda(t) - x(t)| \leq \mu_\lambda \left(\Pi e^K \left(C_1 + \frac{C_1^2}{4} \right) + 3C_1 \right). \tag{3.20}$$

Therefore, $|\tilde{x}_\lambda(t) - x(t)| \to 0, \lambda \to 0$, uniformly for $t \in [0, 1]$. Combining (3.8) and (3.10) for any time scale $\mathbb{T}_\lambda$ and admissible control $u_\lambda(t)$ for (2.4), there is an admissible control $\tilde{u}_\lambda(t)$ for (2.5), such that

$$|J_\lambda(u_\lambda) - J(\tilde{u}_\lambda)| \to 0,\ \lambda \to 0. \tag{3.21}$$

This completes the proof of Lemma 3.1. □

Lemma 3.2 *For any admissible control $u(\cdot)$ for the problem (2.5) and for every time scale $\mathbb{T}_\lambda$ there is an admissible control $u_{ts}^\lambda(\cdot)$ for the problem (2.4) such that*

$$|J(u) - J_\lambda(u_{ts}^\lambda)| \to 0,\ \lambda \to 0. \tag{3.22}$$

Proof. Let $u_{ts}^\lambda(\cdot)$ be an arbitrary admissible control for the problem (2.4) and $x_{ts}^\lambda(\cdot)$ be the corresponding trajectory. Similarly, let $x(\cdot)$ be an admissible trajectory of the problem (2.5) which corresponds to the admissible control $u(\cdot)$. Then

$$\begin{aligned}\int_{[0,1)_{\mathbb{T}_\lambda}} L(t, x_{ts}^\lambda(t), u_{ts}^\lambda(t))\Delta t = &\int_{[0,1)_{\mathbb{T}}\setminus \mathrm{RS}} L(t, x_{ts}^\lambda(t), u_{ts}^\lambda(t))\Delta t \\ &+ \sum_{r\in \mathrm{RS}} L(r, x_{ts}^\lambda(r), u_{ts}^\lambda(r))\mu(r).\end{aligned} \tag{3.23}$$

For $|x_0| \le R$ and $t \in [0,1]_{\mathbb{T}_\lambda}$ the estimates (3.4) hold for any fixed $R > 0$. Hence

$$\sum_{r\in \mathrm{RS}} L(r, x_{ts}^\lambda(r), u_{ts}^\lambda(r))\mu(r) \le C_1 \sum_{r\in \mathrm{RS}} \mu_\lambda(r), \tag{3.24}$$

uniformly for all $u_{ts}^\lambda(\cdot)$. In particular, the sum in (3.23) is convergent, similarly to (3.6). Once again, for every $\lambda > 0$ we choose $N(\lambda) \ge 1$ such that $\sum_{k=N+1}\mu(r_k) \le \frac{\mu_\lambda}{2}$. As before, denote $A = \bigcup_{k=N+1}[r_k, \sigma(r_k))$. Its Lebesgue measure is small: $|A| \le \frac{\mu_\lambda}{2}$. Introduce $B := [0,1]_{\mathbb{T}_\lambda}\setminus A$. In other word, B contains only finitely many right-scattered points $r_1, \dots, r_N$. Now, for any admissible $u(\cdot)$ and $u_{ts}^\lambda(\cdot)$ we write

$$\begin{aligned}&\left|\int_0^1 L(t, x(t), u(t))dt - \int_{[0,1)_{\mathbb{T}_\lambda}} L(t, x_{ts}^\lambda(t), u_{ts}^\lambda(t))\Delta t\right| \\ &\le \mu_\lambda + \left|\int_{[0,1]\setminus A} L(t, x(t), u(t))dt - \int_B L(t, x_{ts}^\lambda(t), u_{ts}^\lambda(t))dt\right|.\end{aligned} \tag{3.25}$$

Fix $\varepsilon > 0$. By Luzin's theorem, there is function $u_\varepsilon(t)$, which is continuous on $[0,1]$ and such that $|A_\varepsilon| < \varepsilon$, where $A_\varepsilon := \{t \in [0,1] : u(t) \neq u_\varepsilon(t)\}$, $\lambda(A_\varepsilon) < \varepsilon$. Denote $B_\varepsilon = [0,1]\setminus A_\varepsilon$. Since f and L are uniformly continuous on the compact set $[0,1]\times \overline{B(0,C_1)}\times U$, for any $0 < \varepsilon_1 < \varepsilon$ there is $0 < \varepsilon_2 = \varepsilon_2(\varepsilon_1)$ such that if $|u - u_1| < \varepsilon_2$, then

$$|f(t,x,u) - f(t,x,u_1)| + |L(t,x,u) - L(t,x,u_1)| < \varepsilon_1 \tag{3.26}$$

for any $t \in [0,1]$ and $|x| \le C_1$. Without loss of generality, assume that $\varepsilon_2 < \varepsilon_1$. Note that $u_\varepsilon(t)$ is uniformly continuous on $[0,1]$. Therefore, one can find $0 < \varepsilon_3 < \varepsilon_2$ such that if $|t-s| < \varepsilon_3$, then $|u_\varepsilon(t) - u_\varepsilon(s)| < \varepsilon_2$. Note that, for sufficiently small λ, $\mu_\lambda < \varepsilon_3$.

We are now in position to construct a new admissible control u_c^λ, which would take into account the structure of $\mathbb{T}_\lambda$. The construction is done separately on each of the intervals as follows:

- ▷ for $t \in A$, i.e. for $t \in [r_i, \sigma(r_i)), i \ge N(\lambda)+1$, set $u_c^\lambda(t) := u(r_i)$;
- ▷ for $t \in [\sigma(r_i), r_{i+1}), i = 1, ..., N(\lambda)-1$, set $u_c^\lambda(t) = u(t)$;
- ▷ if $t \in [r_i, \sigma(r_i))$ and $B_\varepsilon \cap [r_i, \sigma(r_i)) = \emptyset$, define $u_c^\lambda(t) := u(r_i)$, $1 \le i \le N$;

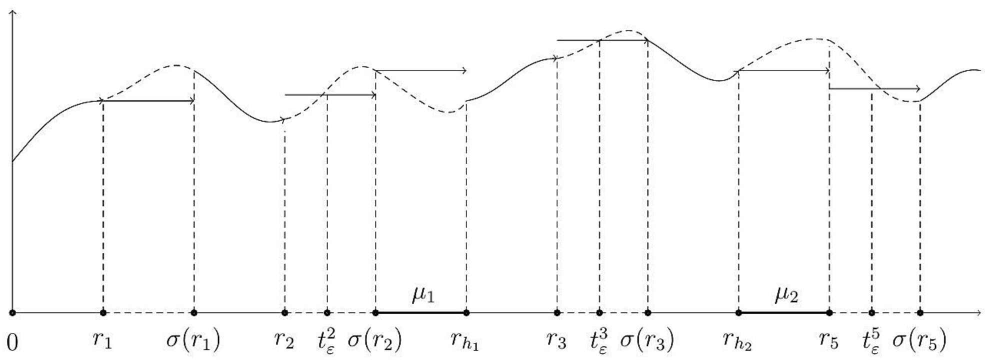

Figure 2: The construction of control.

▷ finally, if $t \in [r_i, \sigma(r_i))$ and $B_\varepsilon \cap [r_i, \sigma(r_i)) \neq \emptyset$, for such $1 \leq i \leq N$ set $u_c^\lambda(t) := u_\varepsilon(t_\varepsilon^i)$, where t_ε^i is an arbitrary point of the set $B_\varepsilon \cap [r_i, \sigma(r_i))$. Since $t_\varepsilon^i \in B_\varepsilon$, then $u_\varepsilon(t_\varepsilon^i) = u(t_\varepsilon^i) \in U$, therefore the control $u_c^\lambda(t)$ on $[r_i, \sigma(r_i))$ is admissible.

Figure 2 visualizes this construction. Here, on the x-axis, the intervals $[r_i, \sigma(r_i))$, $0 \leq i \leq N$, are denoted with dashed lines, the set A is comprised of solid boldface intervals, and the intervals, on which $u(t)$ is continuous, are denoted with solid thin lines. In addition, the graph of $u(t)$ is dashed, and the graph of the new control $u_c^\lambda(t)$ is solid.

In what follows, we are going to analyze the time scale, depicted in Figure 2. The analysis is similar in other cases. Let $x_c^\lambda(t)$ be an admissible trajectory for (2.5). By construction, $u_c^\lambda(t)$ is an extension (in the sense of (3.1)) of some admissible control $u_{ts}^\lambda(t)$ on the time scale $\mathbb{T}_\lambda$. Then, it follows from (3.21) that

$$|J_\lambda(u_{ts}^\lambda) - J(u_c^\lambda)| \to 0,\ \lambda \to 0. \tag{3.27}$$

Let us show that

$$|J(u) - J(u_c^\lambda)| \to 0,\ \lambda \to 0. \tag{3.28}$$

We have

$$\left|\int_0^1 (L(t,x(t),u(t)) - L(t,x_c^\lambda(t),u_c^\lambda(t)))dt\right| \leq \left|\int_{A_\varepsilon} (L(t,x(t),u(t)) - L(t,x_c^\lambda(t),u_c^\lambda(t)))dt\right| + \left|\int_{B_\varepsilon} (L(t,x(t),u(t)) - L(t,x_c^\lambda(t),u_c^\lambda(t)))dt\right|. \tag{3.29}$$

The first term in the right-hand side of (3.29) can be bounded by $C_1(R)|A_\varepsilon| \leq C_1(R)\varepsilon$. We now estimate the second term in the right-hand side of (3.29):

$$\left|\int_{B_\varepsilon} (L(t,x(t),u(t)) - L(t,x_c^\lambda(t),u_c^\lambda(t)))dt\right| \leq K\int_{B_\varepsilon} |x(t) - x_c^\lambda(t)|dt + \int_{B_\varepsilon} \left|L(t,x_c^\lambda(t),u(t)) - L(t,x_c^\lambda(t),u_c^\lambda(t))dt\right|. \tag{3.30}$$

Next,

$$\int_{B_\varepsilon} \left|L(t,x_c^\lambda(t),u(t)) - L(t,x_c^\lambda(t),u_c^\lambda(t))\right| dt \le C_1(R)\mu_\lambda + \int_{B_\varepsilon\cap\bar{A}} \left|L(t,x_c^\lambda(t),u(t)) - L(t,x_c^\lambda(t),u_c^\lambda(t))\right| dt \tag{3.31}$$

and

$$\int_{B_\varepsilon\cap\bar{A}} |L(t,x_c^\lambda(t),u(t)) - L(t,x_c^\lambda(t),u_c^\lambda(t))|dt = \sum_{i=1}^{N-1}\int_{[\sigma(r_i),r_{i+1})\cap B_\varepsilon} |L(t,x_c^\lambda(t),u(t)) - L(t,x_c^\lambda(t),u_c^\lambda(t))|dt + \sum_{i=1}^{N}\int_{[r_i,\sigma(r_i))\cap B_\varepsilon} |L(t,x_c^\lambda(t),u(t)) - L(t,x_c^\lambda(t),u_c^\lambda(t))|dt. \tag{3.32}$$

By construction of u_c^λ, the first term in the right-hand side of (3.32) is zero, and some of the terms in the second sum may vanish if there are no points from the set B_ε in the interval $[r_i,\sigma(r_i))$. Since $\mu_\lambda < \varepsilon_3$, by uniform continuity of $u_\varepsilon(t)$ and (3.26), we have

$$\sum_{i=1}^{N}\int_{[r_i,\sigma(r_i))\cap B_\varepsilon} |L(t,x_c^\lambda(t),u(t)) - L(t,x_c^\lambda(t),u_c^\lambda(t))|dt \le \varepsilon_1\sum_{i=1}^{N}\mu(r_i) \le \varepsilon_1. \tag{3.33}$$

Then from (3.30), (3.31) and (3.33) we have

$$\int_{B_\varepsilon} |L(t,x(t),u(t)) - L(t,x_c^\lambda(t),u_c^\lambda(t))|dt \le K\int_{B_\varepsilon} |x(t)-x_c^\lambda(t)|dt + C_1(R)\mu_\lambda + \varepsilon_1. \tag{3.34}$$

It remains to estimate the difference $|x(t) - x_c^\lambda(t)|$ in (3.34). We are going to do this in the setting of Figure 2, the analysis in the general case is analogous.

1) For $t\in[0,r_1]$, $u(t) = u_c^\lambda(t)$, therefore $x(t) = x_c^\lambda(t)$.

2) For $t\in(r_1,\sigma(r_1)]$, we have

$$\begin{aligned}|x(t) - x_c^\lambda(t)| &\le \int_{[r_1,\sigma(r_1))\cap A_\varepsilon} \left|f(t,x_c^\lambda(t),u(t)) - f(t,x_c^\lambda(t),u(r_1))\right| dt \\ &+ \int_{r_1}^t K|x(s) - x_c^\lambda(s)|ds + \int_{[r_1,\sigma(r_1))\cap B_\varepsilon} \left|f(t,x_c^\lambda,u_\varepsilon(t)) - f(t,x_c^\lambda(t),u_\varepsilon(t_\varepsilon^1))\right| dt \\ &\le \int_{r_1}^t K|x(s) - x_c^\lambda(s)|ds + 2C_1(R)\,|[r_1,\sigma(r_1))\cap A_\varepsilon| + \varepsilon_1\mu(r_1),\end{aligned}$$

where we used the uniform continuity of f on $[0,1]\times\overline{B(0,C_1)}\times U$. Then by Gronwall inequality, we obtain

$$|x(t) - x_c^\lambda(t)| \le (\varepsilon_1\mu(r_1) + 2C_1(R)|[r_1,\sigma(r_1))\cap A_\varepsilon|)e^{K\mu(r_1)} = \delta_1 e^{K\mu(r_1)}, \tag{3.35}$$

where $\delta_1 = \varepsilon_1\mu(r_1) + 2C_1(R)|[r_1,\sigma(r_1))\cap A_\varepsilon|$.

3) For $t \in [\sigma(r_1), r_2)$ we have $|x(t) - x_c^\lambda(t)| \leq \delta_1 e^{K\mu(r_1)} e^{K(r_2-\sigma(r_1))}$.

Continuing this procedure to the remaining intervals of Figure 2 for $t \in [r_2, \sigma(r_5))$ we have the following estimate

$$\begin{aligned}
&|x(t) - x_c^\lambda(t)| \leq (\varepsilon_1\mu(r_1) + 2C_1(R)\,|[r_1, \sigma(r_1)) \cap A_\varepsilon| \\
&\times e^{\{K(\mu(r_1)+(r_2-\sigma(r_1))+\mu(r_2)+\mu(r_3)+(r_{h_2}-\sigma(r_3))+\mu(r_5)+(r_3-r_{h_1}))\}} \\
&+ (2C_1(R)\,|[r_2, \sigma(r_2)) \cap A_\varepsilon| + \varepsilon_1\mu(r_2))e^{\{K(\mu(r_2)+\mu(r_3)+(r_{h_2}-\sigma(r_3))+\mu(r_5)+(r_3-r_{h_1}))\}} \\
&+ 2C_1(R)\mu_1 e^{\{K((r_3-r_{h_1})+\mu(r_3)+(r_{h_2}-\sigma(r_3))+\mu(r_5))\}} + (2C_1(R)\,|[r_3, \sigma(r_3)) \cap A_\varepsilon| \\
&+ \varepsilon_1\mu(r_3))e^{\{K(\mu(r_3)+(r_{h_2}-\sigma(r_3))+\mu(r_5))\}} + 2C_1(R)\,|[r_5, \sigma(r_5)) \cap A_\varepsilon|\, e^{K\mu(r_5)} \\
&+ 2C_1(R)\mu_2 e^{K\mu(r_5)} + \varepsilon_1\mu(r_5)e^{K\mu(r_5)}. \qquad (3.36)
\end{aligned}$$

Once again, the sum of all powers of e in (3.36) does not exceed K, since it is the sum of lengths of disjoint subintervals of $[0,1]$. Altogether, for $t \in [0,1]$ we have

$$|x(t) - x_c^\lambda(t)| \leq (\varepsilon_1 + 2C_1(R)\varepsilon + C_1(R)\mu_\lambda)e^K. \qquad (3.37)$$

From (3.29)–(3.34) we get

$$\begin{aligned}
&\left|\int_0^1 (L(t, x(t), u(t)) - L(t, x_c^\lambda(t), u_c^\lambda(t)))\right| dt \\
&\qquad \leq K(\varepsilon_1 + 2C_1(R)\varepsilon + C_1(R)\mu_\lambda)e^K + C_1(R)\mu_\lambda + \varepsilon_1 + C_1(R)\varepsilon. \qquad (3.38)
\end{aligned}$$

Since ε and ε_1 can be chosen arbitrarily small, we have $|J(u) - J(u_c^\lambda)| \to 0, \lambda \to 0$, hence the proof of Lemma 3.2 follows from (3.27). □

We now return to the proof of Theorem 2.1. In Lemma 3.1 we have shown that for an arbitrary time scale $\mathbb{T}_\lambda$ and an arbitrary admissible control for the problem (2.4) $u_\lambda(t)$, there is an admissible control $\tilde{u}_\lambda$ for the problem (2.5), such that $|J_\lambda(u_\lambda) - J(\tilde{u}_\lambda)| = \varphi(\lambda) \to 0$, $\lambda \to 0$. Consequently, $J(\tilde{u}_\lambda) \leq J_\lambda(u_\lambda) + \varphi(\lambda)$. Using the definition of the value function, we have $V(0,x) \leq J_\lambda(u_\lambda) + \varphi(\lambda)$. We may take the infimum over all admissible controls to get $V(0,x) \leq V_\lambda(0,x) + \varphi(\lambda)$. There exists a uniformly converging subsequence $V_{\lambda_n}(0,x)$: $V_{\lambda_n}(0,x) \rightrightarrows V_0(0,x)$, $|x| \leq r$, with $\lambda_n \to 0$ as $n \to \infty$. Passing to the limit as $\lambda_n \to 0$, we have $V(0,x) \leq V_0(0,x)$.

Let us show that inequality $V(0,x) < V_0(0,x)$ is impossible. By contradiction, assume $V(0,x) < V_0(0,x)$. Then there are $\delta > 0$ and $n_0 \geq 1$ such that for $\lambda_n \leq \lambda_{n_0}$ we have $V_{\lambda_n}(0,x) > V(0,x) + \delta$. However, for such $\delta > 0$ we may construct an admissible control $u(t)$ for the system (2.5), such that $J(u) + \frac{\delta}{2} < V_{\lambda_n}(0,x)$. For such $u(t)$ we now apply Lemma 3.2 to construct an admissible control $u_{ts}^{\lambda_n}$, such that (3.22) holds. Then for sufficiently small λ_n we have $J_{\lambda_n}(u_{ts}^{\lambda_n}) < V_{\lambda_n}(0,x)$, which leads to a contradiction. Therefore, $V(0,x) = V_0(0,x)$, i.e. any convergent sequence $V_{\lambda_n}(0,x)$ has $V(0,x)$ as its limit. Since the family $V_\lambda(0,x)$ is compact, we have $V_\lambda(0,x) \to V_0(0,x)$, $\lambda \to 0$, which proves Theorem 2.1. □

References

[1] Zhan, Z., Wei, W. and Xu, H. Hamilton-Jacobi-Bellman equations on time scales. *Math. Comput. Modelling* **49** (2009) 2019–2028.

[2] Lastivka, L. and Lavrova, O. The method of dynamic programming for systems of differential equations on time scales. *Bulletin of Kyiv Shevchenko National University* (2) (2014) 71–76. (in Ukrainian)

[3] Bourdin, L. and Trelat, E. General Cauchy-Lipschitz theory for Δ-Cauchy problems with Caratheodory dynamics on time scales. *Journal of Difference Equations and Applications* **20** (4) (2014) 526–547.

[4] Bohner, M. and Peterson, A. *Dynamic Equations on Time Scales. An Introduction with Applications.* Birkhäuser Boston Inc., Boston, MA, 2001.

[5] Bohner, M. and Peterson, A. *Advances in Dynamic Equations on Time Scales.* Birkhäuser Boston Inc., Boston, MA, 2003.

[6] Hall, K. and Oberste-Vorth, R. Totally discrete and Eulerian time scales. In: *Difference Equations, Special Functions and Orthogonal Polynomials* (S. Elaydi, J. Cushing, R. Lasser, A. Ruffing, V. Papageorgion, W. Van Assche, eds.). World Scientific, 2007, 462–470.

[7] Bohner, M., Stanzhytskyi, O. and Bratochkina, A. Stohastic dynamic equastions on general time scales. *Electronic Journal of Differential Equations* **2013** (57) (2013) 1–15.

[8] Fleming, W. H. and Soner, H. M. *Controlled Markov Processes and Viscosity Solution. Springer Science and Business Media*, Inc., 2006.

[9] I. Capuzzo Dolcetta. On a discrete approximation of the Hamilton-Jacobi equation of dynamic programming. *Appl. Math. Optim.* **10** (1983) 1831–1847.

[10] I. Capuzzo Dolcetta and Ishii, H. Approximate solutions of the Bellman equation of deterministic control theory. *Appl. Math. Optim.* **11** (1984) 161–181.

[11] Gonzalez, R. and Tidball, M. *On a Discrete Time Approximation of the Hamilton-Jacobi Equation of Dynamic Programming.* Report de recherche, INRIA, 1991.

[12] Lawrence, B. and Oberste-Vorth, R. Solutions of dynamic equations with varying time scales. In: *Difference Equations, Special Functions and Orthogonal Polynomials* (S. Elaydi, J. Cushing, R. Lasser, A. Ruffing, eds). World Scientific, 2007, 452–461.

[13] Hilger, S. *Ein Maßkettenkalkül mit Anwendungen auf Zentrumsmannigfaltigkeiten.* PhD thesis, Universität Würzburg, 1988.