



Proceedings of the Estonian Academy of Sciences,  
2017, **66**, 3, 309–322

<https://doi.org/10.3176/proc.2017.3.04>

Available online at [www.eap.ee/proceedings](http://www.eap.ee/proceedings)

---

MATHEMATICS

---

## On the differentiation of a composite function with a generalized vector argument on homogeneous time scales

Vadim Kaparin\* and Ülle Kotta

Department of Software Science, School of Information Technologies, Tallinn University of Technology, Akadeemia tee 21, 12618 Tallinn, Estonia; [kotta@cc.ioc.ee](mailto:kotta@cc.ioc.ee)

Received 21 November 2016, accepted 23 January 2017, available online 30 June 2017

© 2017 Authors. This is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>).

**Abstract.** The paper proves a theorem on the differentiation of a composite function with a generalized vector argument. The theorem is formulated in terms of the delta derivative, which in the case of homogeneous time scales incorporates both the ordinary derivative and the difference operator. The term “generalized vector argument” implies that a composite function is allowed to depend not only on some variables but also on their delta derivatives. A formula in the theorem shows how the higher-order delta and partial derivatives of a composite function commute. Moreover, it enables reducing the order of the delta derivative, making computations simpler and more efficient. The computational efficiency of the formula was analysed on the basis of experiments in the symbolic computation software *Mathematica*.

**Key words:** delta derivative, partial derivative, composite function, time scale calculus.

### 1. INTRODUCTION

Various areas of mathematics, physics, control theory, etc. require the differentiation of a composite function with a vector argument, whose components are functions of a common variable (for instance, time variable  $t$ ). In order to solve some theoretical problems, it is useful to have a rule for the commutation of the higher-order total and partial derivatives of a function. The theorem proved in this paper offers such a rule (see Section 3). In order to make the theorem applicable to homogeneous time scales, we formulated it in terms of the delta derivative, merging both the ordinary derivative (in the case of a continuous time scale) and the difference operator (in the case of a uniformly sampled discrete time scale). The special case of this theorem, when a composite function depends only on some variables but not on their derivatives, was proved in [7] for the ordinary derivative, and then applied in [6] and [8] to prove the main results. Afterwards, the extension of the theorem to the case of homogeneous time scales was presented in [4] as a supplementary result. Unlike [7] and [4], in this paper we address a more general case when a vector argument contains besides some variables also their derivatives of different order. The main motivation to derive the theorem for such a case was the necessity to prove the main result of [5]. A secondary benefit of the theorem can be found in its computational efficiency. In other words, the

---

\* Corresponding author, [vkaparin@cc.ioc.ee](mailto:vkaparin@cc.ioc.ee)

formula from the theorem allows one to reduce the time required for the computation of certain combinations of derivatives (see Sections 4 and 5).

Note that though the theorem is stated in a single language of time scale formalism, the proof is given separately for the continuous- and discrete-time cases, since, to the best of our knowledge, a single *explicit* formula for the *higher-order* delta derivative of a composite function is absent. To prove the continuous-time part, we used Mishkov’s formula [10], which explicitly expresses the higher-order total derivative of a composite function with a vector argument. In the discrete-time case basic tools from the time scale calculus [3] were employed.

## 2. PRELIMINARIES

In order to represent the main result of this paper for both the ordinary derivative and the difference operator simultaneously, we employ the time scale formalism, which is briefly recalled in this section. A more thorough introduction to the time scale calculus can be found in [3].

A *time scale*  $\mathbb{T}$  is an arbitrary nonempty (topologically) closed subset of the real numbers  $\mathbb{R}$ . Though the variety of time scales is wide (see, for example, [1,3]), this paper is focused on only two of them, i.e., the continuous time scale,  $\mathbb{T} = \mathbb{R}$ , and the discrete time scale,  $\mathbb{T} = \tau\mathbb{Z} := \{\tau k : k \in \mathbb{Z}\}$  for  $\tau > 0$ . The most important notions of the time scale calculus are the *forward jump operator*  $\sigma$ , the *delta derivative*  $\Delta$ , and the *graininess function*  $\mu$ . Applications of  $\sigma$  and  $\Delta$  to the function  $\xi : \mathbb{T} \rightarrow \mathbb{R}$  as well as the values of the graininess function  $\mu$  are presented in Table 1 for two cases  $\mathbb{T} = \mathbb{R}$  and  $\mathbb{T} = \tau\mathbb{Z}$ . A time scale  $\mathbb{T}$  is called *homogeneous* if  $\mu \equiv \text{const}$  and, as can be seen from Table 1, both time scales  $\mathbb{T} = \mathbb{R}$  and  $\mathbb{T} = \tau\mathbb{Z}$  possess this property. Note that for the simplicity of exposition we omit the variable  $t$ , i.e., use instead of  $\xi(t)$  the shorter notation  $\xi$ . Moreover, we denote by  $\xi^{(n)}$  the delta derivative of an arbitrary order  $n$ , whereas  $\xi^{\sigma^n}$  stands for the  $n$ -fold application of the forward jump operator, so  $\xi^{(n)} := (\xi^{(n-1)})^\Delta$  and  $\xi^{\sigma^n} := (\xi^{\sigma^{n-1}})^\sigma$ .

**Proposition 2.1.** For  $F, G : \mathbb{T} \rightarrow \mathbb{R}$  the delta derivative and forward jump operator satisfy the following properties:

- (i)  $F^\sigma = F + \mu F^\Delta$ ,
- (ii)  $(\alpha F + \beta G)^\Delta = \alpha F^\Delta + \beta G^\Delta$ , for  $\alpha, \beta \in \mathbb{R}$ ,
- (iii)  $(FG)^\Delta = F^\sigma G^\Delta + F^\Delta G$ ,
- (iv) on a homogeneous time scale operators  $\Delta$  and  $\sigma$  commute, i.e.,

$$(F^\sigma)^\Delta = (F^\Delta)^\sigma.$$

Observe that the generalization of item (i) above yields

$$F^{(n)} = \frac{1}{\mu^n} \sum_{k=0}^n (-1)^k C_n^k F^{\sigma^{n-k}} \tag{1}$$

for  $\mu \neq 0$  and

$$F^{\sigma^n} = \sum_{i=0}^n C_n^i \mu^i F^{(i)}, \tag{2}$$

where  $C_n^k$  and  $C_n^i$  are the binomial coefficients.

**Table 1.** Basic types of operators/functions

$\mathbb{T}$	$\xi^\sigma(t)$	$\xi^\Delta(t)$	$\mu$
$\mathbb{R}$	$\xi(t)$	$\frac{d\xi(t)}{dt}$	0
$\tau\mathbb{Z}$	$\xi(t + \tau)$	$\frac{\xi(t+\tau) - \xi(t)}{\tau}$	$\tau$

### 3. MAIN RESULT

The main result of this paper is the theorem below, which shows how the partial derivative of the higher-order delta derivative of a composite function can be expressed through the higher-order delta derivatives of the partial derivatives of this function. A composite function with a vector argument, consisting of an arbitrary but finite number of variables and their delta derivatives, is considered.

**Theorem 3.1.** *Let  $\Phi : \mathbb{R}^{\varsigma_1 + \dots + \varsigma_r + r} \rightarrow \mathbb{R}$  be a composite function of a vector argument  $\Xi := [\xi_1^{(0)}, \dots, \xi_1^{(\varsigma_1)}, \dots, \xi_r^{(0)}, \dots, \xi_r^{(\varsigma_r)}]$  with  $\xi_j^{(s_j)} : \mathbb{T} \rightarrow \mathbb{R}$  for all  $j = 1, \dots, r$  and  $s_j = 0, \dots, \varsigma_j$ , such that delta derivatives of  $\Phi(\Xi)$  are defined up to and including an order  $b$ ; then*

$$\frac{\partial(\Phi(\Xi))^{(b)}}{\partial \xi_l^{(a)}} = \sum_{\gamma=\max(0, a-b)}^{\min(\varsigma_l, a)} C_b^{b-a+\gamma} \left( \left( \frac{\partial \Phi(\Xi)}{\partial \xi_l^{(\gamma)}} \right)^{(b-a+\gamma)} \right)^{\sigma^{a-\gamma}}, \tag{3}$$

where  $l \in \{1, \dots, r\}$ ;  $C_b^{b-a+\gamma}$  denotes the binomial coefficient, and  $a, b$  are non-negative integers, such that  $b + \varsigma_l \geq a$ .

*Proof.* Since there is no single explicit formula for the higher-order delta derivative of a composite function on time scales, two cases  $\mathbb{T} = \mathbb{R}$  and  $\mathbb{T} = \tau\mathbb{Z}$  will be considered separately.

*Case 1 ( $\mathbb{T} = \mathbb{R}$ ).* Note that throughout this part of the proof the symbol  $\langle n \rangle$  stands for the  $n$ th-order ordinary derivative with respect to  $t$ . According to Mishkov’s formula [10], the  $b$ th derivative of the composite function  $\Phi(\Xi)$  can be computed by the formula  $(\Phi(\Xi))^{(b)} = \sum_b \mathcal{A}_b \mathcal{B}_b^\Phi \mathcal{C}_b$ , where

$$\begin{aligned} \mathcal{A}_b &:= \frac{b!}{\prod_{i=1}^b (i!)^{k_i} \prod_{j=1}^r \prod_{i=1}^b \prod_{s=0}^{\varsigma_j} q_{i,j,s}}, \\ \mathcal{B}_b^\Phi &:= \frac{\partial^k \Phi(\Xi)}{\partial \xi_1^{(0)p_{1,0}} \dots \partial \xi_1^{(\varsigma_1)p_{1,\varsigma_1}} \dots \partial \xi_r^{(0)p_{r,0}} \dots \partial \xi_r^{(\varsigma_r)p_{r,\varsigma_r}}}, \\ \mathcal{C}_b &:= \prod_{j=1}^r \prod_{i=1}^b \prod_{s=0}^{\varsigma_j} (\xi_j^{(i+s)})^{q_{i,j,s}}, \end{aligned} \tag{4}$$

and the sum  $\sum_b$  is taken over all non-negative integer solutions of the Diophantine equations

$$\begin{cases} \sum_{i=1}^b ik_i = b, \\ \sum_{j=1}^r \sum_{s=0}^{\varsigma_j} q_{i,j,s} = k_i, \quad i = 1, \dots, b. \end{cases} \tag{5}$$

Moreover, integers  $p_{j,s}$  and  $k$  in  $\mathcal{B}_b^\Phi$  stand for the order of the partial derivative with respect to  $\xi_j^{(s)}$  and the order of the mixed partial derivative of  $\Phi(\Xi)$ , respectively, and satisfy the relations

$$\begin{aligned} p_{j,s} &= \sum_{i=1}^b q_{i,j,s}, \quad j = 1, \dots, r; \quad s = 0, \dots, \varsigma_j, \\ k &= \sum_{j=1}^r \sum_{s=0}^{\varsigma_j} p_{j,s} = \sum_{i=1}^b k_i. \end{aligned} \tag{6}$$

Using the product rule for differentiation, one obtains

$$\frac{\partial (\Phi(\Xi))^{(b)}}{\partial \xi_l^{(a)}} = M + N, \tag{7}$$

where

$$M := \sum_b \mathcal{A}_b \mathcal{B}_b^\Phi \frac{\partial \mathcal{C}_b}{\partial \xi_l^{(a)}}, \quad N := \sum_b \mathcal{A}_b \frac{\partial \mathcal{B}_b^\Phi}{\partial \xi_l^{(a)}} \mathcal{C}_b.$$

In order to express explicitly  $\partial \mathcal{C}_b / \partial \xi_l^{(a)}$  in  $M$ , let us change in  $\mathcal{C}_b$  the multiplication index  $i$  into  $t = i + s$ . It is easy to observe that  $t$  varies from 1 to  $b + \zeta_j$ . In this case  $s = t - i$ , whose minimal and maximal values are  $t - b$  and  $t - 1$ , respectively. On the other hand,  $s$  changes from 0 to  $\zeta_j$ . Therefore, we define the range of  $s$  from  $\max(0, t - b)$  to  $\min(\zeta_j, t - 1)$ . As a result, one may use the following identity:

$$\prod_{i=1}^b \prod_{s=0}^{\zeta_j} \alpha_{i+s,i,s} = \prod_{t=1}^{b+\zeta_j} \prod_{s=\max(0,t-b)}^{\min(\zeta_j,t-1)} \alpha_{t,t-s,s},$$

which leads to

$$\mathcal{C}_b = \prod_{j=1}^r \prod_{t=1}^{b+\zeta_j} \prod_{s=\max(0,t-b)}^{\min(\zeta_j,t-1)} \left( \xi_j^{(t)} \right)^{q_{t-s,j,s}}.$$

Next, separating the multiplier with  $j = l$  and  $t = a$ , one may rewrite the product above as

$$\mathcal{C}_b = \prod_{\gamma=\max(0,a-b)}^{\min(\zeta_l,a-1)} \left( \xi_l^{(a)} \right)^{q_{a-\gamma,l,\gamma}} \mathcal{C}_b \mathcal{C}_b, \substack{j=l \ j \neq l \\ t \neq a}$$

where, in order to avoid confusion,  $\gamma$  replaces the index  $s$ , independent from  $j$  and  $t$ . Using the product of powers property, we obtain

$$\mathcal{C}_b = \left( \xi_l^{(a)} \right)^\Lambda \mathcal{C}_b \mathcal{C}_b, \substack{j=l \ j \neq l \\ t \neq a}$$

where  $\Lambda := \sum_{\gamma=\max(0,a-b)}^{\min(\zeta_l,a-1)} q_{a-\gamma,l,\gamma}$ . Thus, the partial derivative of  $\mathcal{C}_b$  with respect to  $\xi_l^{(a)}$  reads

$$\frac{\partial \mathcal{C}_b}{\partial \xi_l^{(a)}} = \frac{\Lambda}{\xi_l^{(a)}} \left( \xi_l^{(a)} \right)^\Lambda \mathcal{C}_b \mathcal{C}_b = \frac{\Lambda}{\xi_l^{(a)}} \mathcal{C}_b,$$

allowing one to rewrite  $M$  as

$$M = \sum_{\gamma=\max(0,a-b)}^{\min(\zeta_l,a-1)} \sum_b \frac{q_{a-\gamma,l,\gamma}}{\xi_l^{(a)}} \mathcal{A}_b \mathcal{B}_b^\Phi \mathcal{C}_b. \tag{8}$$

Now it is easy to observe that in  $M$  only addends of the sum  $\sum_b$  with  $q_{a-\gamma,l,\gamma} \neq 0$  will matter. Note that, according to (5),  $q_{a-\gamma,l,\gamma} \neq 0$  implies  $k_{a-\gamma} \neq 0$ . Taking into account that in this case  $k_{a-\gamma} - 1$  and  $q_{a-\gamma,l,\gamma} - 1$  are non-negative, it is eligible to introduce the following notations:

$$\bar{k}_i := \begin{cases} k_i - 1 & \text{if } i = a - \gamma, \\ k_i & \text{otherwise,} \end{cases} \tag{9}$$

$$\bar{q}_{i,j,s} := \begin{cases} q_{i,j,s} - 1 & \text{if } i = a - \gamma, \quad j = l; \quad s = \gamma, \\ q_{i,j,s} & \text{otherwise,} \end{cases}$$

for  $i = 1, \dots, b; j = 1, \dots, r; s = 0, \dots, \zeta_j$  and rewrite (5) as

$$\begin{cases} \sum_{i=1}^b i\bar{k}_i = b - a + \gamma, \\ \sum_{j=1}^r \sum_{s=0}^{\zeta_j} \bar{q}_{i,j,s} = \bar{k}_i, \quad i = 1, \dots, b. \end{cases} \tag{10}$$

Note that (10) is satisfied only under the following conditions:

$$\begin{aligned} \bar{k}_i &= 0, & b - a + \gamma < i \leq b, \\ \bar{q}_{i,j,s} &= 0, & b - a + \gamma < i \leq b; \quad j = 1, \dots, r; \quad s = 0, \dots, \zeta_j. \end{aligned} \tag{11}$$

As a consequence, (10) yields the following Diophantine equations:

$$\begin{cases} \sum_{i=1}^{b-a+\gamma} i\bar{k}_i = b - a + \gamma, \\ \sum_{j=1}^r \sum_{s=0}^{\zeta_j} \bar{q}_{i,j,s} = \bar{k}_i, \quad i = 1, \dots, b - a + \gamma, \end{cases} \tag{12}$$

leading to the conclusion

$$\sum_{q_{a-\gamma,l,\gamma} \neq 0} \alpha = \sum_{b-a+\gamma} \alpha, \tag{13}$$

where the sum  $\sum_{b-a+\gamma}$  is taken over all non-negative integer solutions of (12).

Next, for  $j = 1, \dots, r$  and  $s = 0, \dots, \zeta_j$  we denote

$$\begin{aligned} \bar{p}_{j,s} &:= \begin{cases} p_{j,s} - 1 & \text{if } j = l; \quad s = \gamma, \\ p_{j,s} & \text{otherwise,} \end{cases} \\ \bar{k} &:= k - 1, \end{aligned} \tag{14}$$

which together with (9) and (11) allows one to rewrite (6) as

$$\begin{aligned} \bar{p}_{j,s} &= \sum_{i=1}^{b-a+\gamma} \bar{q}_{i,j,s}, \quad j = 1, \dots, r; \quad s = 0, \dots, \zeta_j, \\ \bar{k} &= \sum_{j=1}^r \sum_{s=0}^{\zeta_j} \bar{p}_{j,s} = \sum_{i=1}^{b-a+\gamma} \bar{k}_i. \end{aligned} \tag{15}$$

Furthermore, applying (9), (11), and (14) to (4) and taking into account that by direct computations  $b!/(a-\gamma)! = C_b^{b-a+\gamma}(b-a+\gamma)!$ , we obtain

$$\begin{aligned} \mathcal{A}_{q_{a-\gamma,l,\gamma} \neq 0} &= \frac{C_b^{b-a+\gamma}}{(\bar{q}_{a-\gamma,l,\gamma} + 1)} \mathcal{A}_{b-a+\gamma}, \\ \mathcal{B}_{q_{a-\gamma,l,\gamma} \neq 0}^\Phi &= \mathcal{B}_{b-a+\gamma}^\Phi, \\ \mathcal{C}_{q_{a-\gamma,l,\gamma} \neq 0} &= \xi_l^{(a)} \mathcal{C}_{b-a+\gamma}, \end{aligned} \tag{16}$$

where

$$\begin{aligned} \mathcal{A}_{b-a+\gamma} &:= \frac{(b-a+\gamma)!}{\prod_{i=1}^{b-a+\gamma} (i!)^{\bar{k}_i} \prod_{j=1}^r \prod_{i=1}^{b-a+\gamma} \prod_{s=0}^{\zeta_j} \bar{q}_{i,j,s}!}, \\ \mathcal{B}_{b-a+\gamma}^{\bar{\Phi}} &:= \frac{\partial^{\bar{k}} \bar{\Phi}(\Xi)}{\partial \xi_1^{(0)\bar{p}_{1,0}} \dots \partial \xi_1^{(\zeta_1)\bar{p}_{1,\zeta_1}} \dots \partial \xi_r^{(0)\bar{p}_{r,0}} \dots \partial \xi_r^{(\zeta_r)\bar{p}_{r,\zeta_r}}}, \\ \mathcal{C}_{b-a+\gamma} &:= \prod_{j=1}^r \prod_{i=1}^{b-a+\gamma} \prod_{s=0}^{\zeta_j} \left( \xi_j^{(i+s)} \right)^{\bar{q}_{i,j,s}}, \end{aligned} \tag{17}$$

and

$$\bar{\Phi}(\Xi) := \frac{\partial \Phi(\Xi)}{\partial \xi_l^{(\gamma)}}. \tag{18}$$

As a result, recalling that  $q_{a-\gamma,l,\gamma} = \bar{q}_{a-\gamma,l,\gamma} + 1$  and using (13), (16), we express (8) as

$$M = \sum_{\gamma=\max(0,a-b)}^{\min(\zeta_l,a-1)} C_b^{b-a+\gamma} \sum_{b-a+\gamma} \mathcal{A}_{b-a+\gamma} \mathcal{B}_{b-a+\gamma}^{\bar{\Phi}} \mathcal{C}_{b-a+\gamma}.$$

Now, taking into account (12), (15), (17), and (18), one may again apply Mishkov’s formula to obtain

$$M = \sum_{\gamma=\max(0,a-b)}^{\min(\zeta_l,a-1)} C_b^{b-a+\gamma} \left( \frac{\partial \Phi(\Xi)}{\partial \xi_l^{(\gamma)}} \right)^{(b-a+\gamma)}.$$

Next, the application of Mishkov’s formula to  $N$  yields

$$N = \left( \frac{\partial \Phi(\Xi)}{\partial \xi_l^{(a)}} \right)^{(b)} = C_b^{b-a+\gamma} \left( \frac{\partial \Phi(\Xi)}{\partial \xi_l^{(\gamma)}} \right)^{(b-a+\gamma)}$$

for  $\gamma = a$ . Thus, observing that  $N$  is different from zero only for  $\zeta_l \geq a$ , we may rewrite (7) as

$$\frac{\partial (\Phi(\Xi))^{(b)}}{\partial \xi_l^{(a)}} = \sum_{\gamma=\max(0,a-b)}^{\min(\zeta_l,a)} C_b^{b-a+\gamma} \left( \frac{\partial \Phi(\Xi)}{\partial \xi_l^{(\gamma)}} \right)^{(b-a+\gamma)},$$

confirming (3) for  $\mathbb{T} = \mathbb{R}$  ( $\sigma^{a-\gamma} = \text{id}$ ).

Case 2 ( $\mathbb{T} = \tau\mathbb{Z}$ ). Note that in this case the symbol  $\langle n \rangle$  means the  $n$ -fold application of the difference operator. Using (1) for  $n = b$ , the definition of the operator  $\sigma$ , and the chain rule for the partial derivative with respect to  $\xi_l^{(a)}$ , one obtains

$$\frac{\partial (\Phi(\Xi))^{(b)}}{\partial \xi_l^{(a)}} = \frac{1}{\mu^b} \sum_{k=0}^b (-1)^k C_b^k \sum_{\gamma=0}^{\zeta_l} \frac{\partial (\Phi(\Xi))^{\sigma^{b-k}}}{\partial (\xi_l^{(\gamma)})^{\sigma^{b-k}}} \cdot \frac{\partial (\xi_l^{(\gamma)})^{\sigma^{b-k}}}{\partial \xi_l^{(a)}},$$

which, according to (2) for  $n = b - k$ , yields

$$\frac{\partial (\Phi(\Xi))^{(b)}}{\partial \xi_l^{(a)}} = \frac{1}{\mu^b} \sum_{k=0}^b (-1)^k C_b^k \sum_{\gamma=0}^{\zeta_l} \frac{\partial (\Phi(\Xi))^{\sigma^{b-k}}}{\partial (\xi_l^{(\gamma)})^{\sigma^{b-k}}} \sum_{i=0}^{b-k} C_{b-k}^i \mu^i \frac{\partial \xi_l^{(\gamma+i)}}{\partial \xi_l^{(a)}}.$$

In order to find  $\partial \xi_l^{\langle \gamma+i \rangle} / \partial \xi_l^{\langle a \rangle}$ , it is necessary to express the sum  $\gamma + i$  in terms of a single index. For that purpose we, first, use the identity

$$\sum_{k=0}^b \sum_{i=0}^{b-k} \alpha_{k,i} = \sum_{i=0}^b \sum_{k=0}^{b-i} \alpha_{k,i},$$

leading to

$$\frac{\partial (\Phi(\Xi))^{(b)}}{\partial \xi_l^{\langle a \rangle}} = \frac{1}{\mu^b} \sum_{i=0}^b \sum_{k=0}^{b-i} (-1)^k C_b^k \sum_{\gamma=0}^{\zeta_l} \frac{\partial (\Phi(\Xi))^{\sigma^{b-k}}}{\partial (\xi_l^{\langle \gamma \rangle})^{\sigma^{b-k}}} C_{b-k}^i \mu^i \frac{\partial \xi_l^{\langle \gamma+i \rangle}}{\partial \xi_l^{\langle a \rangle}},$$

and then we change the summation index  $i$  into  $\iota = \gamma + i$ . It is easy to observe that  $\iota$  varies from 0 to  $b + \zeta_l$ . In this case  $\gamma = \iota - i$ , whose minimal and maximal values are  $\iota - b$  and  $\iota$ , respectively. On the other hand,  $\gamma$  changes from 0 to  $\zeta_l$ . Therefore, we define the range of  $\gamma$  from  $\max(0, \iota - b)$  to  $\min(\zeta_l, \iota)$ . As a result, one may use the identity

$$\sum_{i=0}^b \sum_{\gamma=0}^{\zeta_l} \alpha_{\gamma+i,i,\gamma} = \sum_{\iota=0}^{b+\zeta_l} \sum_{\gamma=\max(0,\iota-b)}^{\min(\zeta_l,\iota)} \alpha_{\iota,\iota-\gamma,\gamma}$$

to obtain

$$\frac{\partial (\Phi(\Xi))^{(b)}}{\partial \xi_l^{\langle a \rangle}} = \frac{1}{\mu^b} \sum_{\iota=0}^{b+\zeta_l} \sum_{\gamma=\max(0,\iota-b)}^{\min(\zeta_l,\iota)} \sum_{k=0}^{b-\iota+\gamma} (-1)^k C_b^k \frac{\partial (\Phi(\Xi))^{\sigma^{b-k}}}{\partial (\xi_l^{\langle \gamma \rangle})^{\sigma^{b-k}}} C_{b-k}^{\iota-\gamma} \mu^{\iota-\gamma} \frac{\partial \xi_l^{\langle \iota \rangle}}{\partial \xi_l^{\langle a \rangle}}.$$

Taking into account that  $\partial \xi_l^{\langle \iota \rangle} / \partial \xi_l^{\langle a \rangle}$  equals 0 for every  $\iota$ , except for  $\iota = a$  when it equals 1, and using the relation  $\mu^{a-\gamma} / \mu^b = 1 / \mu^{b-a+\gamma}$ , one may rewrite the equality above as

$$\frac{\partial (\Phi(\Xi))^{(b)}}{\partial \xi_l^{\langle a \rangle}} = \sum_{\gamma=\max(0,a-b)}^{\min(\zeta_l,a)} \frac{1}{\mu^{b-a+\gamma}} \sum_{k=0}^{b-a+\gamma} (-1)^k C_b^k C_{b-k}^{a-\gamma} \frac{\partial (\Phi(\Xi))^{\sigma^{b-k}}}{\partial (\xi_l^{\langle \gamma \rangle})^{\sigma^{b-k}}}.$$

Finally, applying the identity  $C_b^k C_{b-k}^{a-\gamma} = C_b^{b-a+\gamma} C_{b-a+\gamma}^k$  and the properties

$$\frac{\partial (F(\xi))^{\sigma^i}}{\partial \xi^{\sigma^i}} = \left( \frac{\partial F(\xi)}{\partial \xi} \right)^{\sigma^i}, \quad F^{\sigma^{i+j}} = (F^{\sigma^i})^{\sigma^j},$$

we obtain

$$\frac{\partial (\Phi(\Xi))^{(b)}}{\partial \xi_l^{\langle a \rangle}} = \sum_{\gamma=\max(0,a-b)}^{\min(\zeta_l,a)} C_b^{b-a+\gamma} \left( \frac{1}{\mu^{b-a+\gamma}} \sum_{k=0}^{b-a+\gamma} (-1)^k C_{b-a+\gamma}^k \left( \frac{\partial \Phi(\Xi)}{\partial \xi_l^{\langle \gamma \rangle}} \right)^{\sigma^{b-a+\gamma-k}} \right)^{\sigma^{a-\gamma}},$$

which, according to (1) for  $n = b - a + \gamma$ , confirms (3). This completes the proof. □

The following corollary provides, in a sense, an inverse relation to that presented in Theorem 3.1, i.e., it shows how the delta derivative of the partial derivative of a composite function can be expressed through the partial derivatives of the delta derivatives of this function.

**Corollary 3.2.** *Under the assumption of Theorem 3.1 for all  $b \leq w$  the following holds:*

$$\left( \frac{\partial \Phi(\Xi)}{\partial \xi_l^{\langle v \rangle}} \right)^{\langle w \rangle} = \sum_{\zeta=0}^{\min(v,w)} (-1)^\zeta C_w^\zeta \left( \frac{\partial \Phi(\Xi)^{\langle w-\zeta \rangle}}{\partial \xi_l^{\langle v-\zeta \rangle}} \right)^{\sigma^\zeta}, \tag{19}$$

where  $w$  is a non-negative integer and  $v = 0, \dots, \zeta_l$ .

*Proof.* Taking into account that  $w - \zeta$  and  $v - \zeta$  are always non-negative and  $w + \zeta_l \geq v$ , one may apply formula (3) to  $a = v - \zeta$  and  $b = w - \zeta$  to rewrite (19) as

$$\left(\frac{\partial \Phi(\Xi)}{\partial \xi_l^{(v)}}\right)^{\langle w \rangle} = \sum_{\zeta=0}^{\min(v,w)} (-1)^\zeta C_w^\zeta \sum_{\gamma=\max(0,v-w)}^{v-\zeta} C_{w-\zeta}^{w-v+\gamma} \left(\left(\frac{\partial \Phi(\Xi)}{\partial \xi_l^{(\gamma)}}\right)^{\langle w-v+\gamma \rangle}\right)^{\sigma^{v-\gamma}}.$$

Using the relation  $C_w^\zeta C_{w-\zeta}^{w-v+\gamma} = C_{v-\gamma}^\zeta C_w^{v+\gamma}$  and changing the summation order according to the identity

$$\sum_{\zeta=0}^{\min(v,w)} \sum_{\gamma=\max(0,v-w)}^{v-\zeta} \alpha_{\zeta,\gamma} = \sum_{\gamma=\max(0,v-w)}^v \sum_{\zeta=0}^{v-\gamma} \alpha_{\zeta,\gamma},$$

we obtain

$$\left(\frac{\partial \Phi(\Xi)}{\partial \xi_l^{(v)}}\right)^{\langle w \rangle} = \sum_{\gamma=\max(0,v-w)}^v C_w^{v+\gamma} \left(\left(\frac{\partial \Phi(\Xi)}{\partial \xi_l^{(\gamma)}}\right)^{\langle w-v+\gamma \rangle}\right)^{\sigma^{v-\gamma}} \sum_{\zeta=0}^{v-\gamma} (-1)^\zeta C_{v-\gamma}^\zeta.$$

Separation of the last addend of the sum  $\sum_{\gamma=\max(0,v-w)}^v$  yields

$$\left(\frac{\partial \Phi(\Xi)}{\partial \xi_l^{(v)}}\right)^{\langle w \rangle} = \left(\frac{\partial \Phi(\Xi)}{\partial \xi_l^{(v)}}\right)^{\langle w \rangle} + \mathcal{O}_v, \tag{20}$$

where

$$\mathcal{O}_v := \begin{cases} 0 & \text{if } v = 0, \\ \sum_{\gamma=\max(0,v-w)}^{v-1} C_w^{v+\gamma} \left(\left(\frac{\partial \Phi(\Xi)}{\partial \xi_l^{(\gamma)}}\right)^{\langle w-v+\gamma \rangle}\right)^{\sigma^{v-\gamma}} \sum_{\zeta=0}^{v-\gamma} (-1)^\zeta C_{v-\gamma}^\zeta & \text{if } v > 0. \end{cases}$$

Using the binomial theorem  $(\alpha + \beta)^{v-\gamma} = \sum_{\zeta=0}^{v-\gamma} C_{v-\gamma}^\zeta \alpha^\zeta \beta^{v-\gamma-\zeta}$  for  $\alpha = -1$ ,  $\beta = 1$ , and  $v - \gamma > 0$ , we get

$$\sum_{\zeta=0}^{v-\gamma} (-1)^\zeta C_{v-\gamma}^\zeta = 0,$$

implying  $\mathcal{O}_v = 0$ . Thus, the identity (20) confirms (19). □

#### 4. EXAMPLES

The examples in this section illustrate the application of formula (3).

**Example 4.1.** Consider the composite function  $\Phi_1(u, y, y^{(1)}, y^{(2)}) = uy^{(1)} + yy^{(2)}$  defined on the continuous time scale  $\mathbb{T} = \mathbb{R}$  (the symbol  $\langle i \rangle$  stands for the  $i$ th-order ordinary derivative with respect to  $t$  and  $\sigma^i = \text{id}$ ). Assume that we need to compute the partial derivative with respect to  $y^{(3)}$  of the 5th-order total derivative of the function  $\Phi_1(\cdot)$ . For this purpose the ordinary steps are the computation of the total derivative

$$\begin{aligned} (\Phi_1(\cdot))^{(5)} &= 10u^{(3)}y^{(3)} + 5y^{(2)}u^{(4)} + 10u^{(2)}y^{(4)} + 15y^{(3)}y^{(4)} \\ &\quad + y^{(1)}u^{(5)} + 5u^{(1)}y^{(5)} + 11y^{(2)}y^{(5)} + uy^{(6)} + 5y^{(1)}y^{(6)} + yy^{(7)} \end{aligned}$$



and then differentiation with respect to  $y^{(3)}$ , yielding

$$\frac{\partial (\Phi_1(\cdot))^{(5)}}{\partial y^{(3)}} = 10u^{(3)} + 15y^{(4)}.$$

On the other hand, one may use formula (3) as follows:

$$\begin{aligned} \frac{\partial (\Phi_1(\cdot))^{(5)}}{\partial y^{(3)}} &= C_5^2 \left( \frac{\partial \Phi_1(\cdot)}{\partial y} \right)^{(2)} + C_5^3 \left( \frac{\partial \Phi_1(\cdot)}{\partial y^{(1)}} \right)^{(3)} + C_5^4 \left( \frac{\partial \Phi_1(\cdot)}{\partial y^{(2)}} \right)^{(4)} \\ &= 10y^{(4)} + 10u^{(3)} + 5y^{(4)} = 10u^{(3)} + 15y^{(4)}. \end{aligned}$$

Observe that formula (3) does not require the computation of  $(\Phi_1(\cdot))^{(5)}$ . Moreover, it allows one to compute partial derivatives before total derivatives and reduces the maximal order of the total derivative. As a result, the computations become simpler and more efficient.

**Example 4.2.** Consider the composite function  $\Phi_2(u, y, y^{(1)}, y^{(2)}) = uy^{(1)} + yy^{(2)}$  defined on a homogeneous time scale<sup>1</sup>  $\mathbb{T}$ . Assume that we need to compute the partial derivative with respect to  $y^{(4)}$  of the 2nd-order total delta derivative of the function  $\Phi_2(\cdot)$ . Usually, one needs to compute first the total delta derivative

$$\begin{aligned} (\Phi_2(\cdot))^{(2)} &= y^{(1)}u^{(2)} + 2u^{(1)}y^{(2)} + 2\mu u^{(2)}y^{(2)} + (y^{(2)})^2 + uy^{(3)} + 2\mu u^{(1)}y^{(3)} \\ &\quad + 2y^{(1)}y^{(3)} + \mu^2 u^{(2)}y^{(3)} + 2\mu y^{(2)}y^{(3)} + yy^{(4)} + 2\mu y^{(1)}y^{(4)} + \mu^2 y^{(2)}y^{(4)} \end{aligned}$$

and then the partial derivative with respect to  $y^{(4)}$ , which yields

$$\frac{\partial (\Phi_2(\cdot))^{(2)}}{\partial y^{(4)}} = y + 2\mu y^{(1)} + \mu^2 y^{(2)}.$$

Alternatively, formula (3) leads to

$$\frac{\partial (\Phi_2(\cdot))^{(2)}}{\partial y^{(4)}} = C_2^0 \left( \frac{\partial \Phi_2(\cdot)}{\partial y^{(2)}} \right)^{\sigma^2} = y^{\sigma^2} = y + 2\mu y^{(1)} + \mu^2 y^{(2)}.$$

Like in the case of Example 4.1, one may observe that the application of formula (3) significantly simplifies the computations.

## 5. COMPUTATIONAL EFFICIENCY

In this section the computational efficiency of formula (3) is examined on the basis of experiments performed via the symbolic computation program *Mathematica*. The objective of the experiments is the comparison of the computation times for both sides of formula (3) with different variations in parameters. For that purpose we employed the *Mathematica* function `Timing`, which returns the time (in seconds) used by CPU for the evaluation of an expression. Furthermore, in order to avoid the influence of previous computations, internal system caches of stored results were cleared between the experiments by means of the command `ClearSystemCache`. The technical

<sup>1</sup> Note that due to the computational complexity of the single formula for the delta derivative of a composite function on time scales [3] it is easier to perform computations separately for continuous and discrete time scales. However, in this example the function  $\Phi_2(\cdot)$  contains only products of variables. Therefore, the rules from Proposition 2.1 are enough to perform computations in a unified manner.

**Table 2.** Average computation times [s] of the left-hand side (L) and the right-hand side (R) of formula (3) for  $\mathbb{T} = \mathbb{R}$ ,  $\zeta_1 = \dots = \zeta_r = 2$  and various values of  $a, b, r$ 

$b$	$a$	$r = 2$		$r = 3$		$r = 4$	
		L	R	L	R	L	R
0	0	0.000053	0.000076	0.000061	0.000082	0.000070	0.000095
	1	0.000066	0.000087	0.000070	0.000088	0.000076	0.000095
	2	0.000074	0.000099	0.000087	0.000102	0.000095	0.000131
1	0	0.000342	0.000203	0.000520	0.000284	0.000763	0.000404
	1	0.000325	0.000226	0.000528	0.000326	0.000773	0.000428
	2	0.000332	0.000237	0.000543	0.000322	0.000735	0.000376
	3	0.000222	0.000105	0.000309	0.000127	0.000427	0.000130
2	0	0.002030	0.000894	0.004706	0.001841	0.008979	0.003272
	1	0.002102	0.000978	0.004808	0.002004	0.009048	0.003462
	2	0.002108	0.001070	0.005141	0.002111	0.009324	0.003439
	3	0.001292	0.000253	0.002658	0.000339	0.005239	0.000388
	4	0.001038	0.000102	0.002402	0.000117	0.004755	0.000120
3	0	0.013657	0.005440	0.040496	0.016051	0.100210	0.040032
	1	0.013987	0.006068	0.041588	0.017764	0.103942	0.042604
	2	0.014221	0.006497	0.041898	0.017717	0.107050	0.042283
	3	0.009361	0.000875	0.027026	0.001984	0.063336	0.003640
	4	0.008009	0.000244	0.023754	0.000300	0.054955	0.000415
	5	0.006769	0.000094	0.021688	0.000105	0.050361	0.000101

specifications of the computer, used in the experiments, are the following: processor Intel Core i7-3740QM, 2.7 GHz, 4 cores, 256 KB L2 cache (per core), 6 MB L3 cache; 16 GB of 1600 MHz DDR3 Random-access memory; operating system OS X 10.11.3; computation software *Mathematica* 9.0.1.

The experiments were performed separately for a continuous time scale ( $\mathbb{T} = \mathbb{R}$ ) and its discrete counterpart ( $\mathbb{T} = \tau\mathbb{Z}$ ) due to the computational complexity of the single formula for the delta derivative of a composite function on time scales [3]. In the case  $\mathbb{T} = \mathbb{R}$  we used the built-in *Mathematica* function D to compute derivatives, whereas in the case  $\mathbb{T} = \tau\mathbb{Z}$  the functions DeltaD and ForwardShift from the package NLControl (see [2]) were employed as the difference and forward jump operators, respectively. The results of the experiments are presented in Tables 2–5. Note that for the sake of reliability, each experiment was repeated ten times and the average value of the obtained results was entered into the table.

Tables 2 and 3 consist of the results for  $\mathbb{T} = \mathbb{R}$ , whereas the results for  $\mathbb{T} = \tau\mathbb{Z}$  are displayed in Tables 4 and 5. Both in Tables 2 and 4 the number of variables  $r$  in a vector argument  $\Xi$  varies, while the maximal orders of their derivatives  $\zeta_1, \dots, \zeta_r$  are fixed. In contrast, Tables 3 and 5 contain the results for vector arguments with a fixed number of variables  $r = 1$  and various maximal orders of derivatives  $\zeta_1$ . For ease of comparison, cells with worse results have grey background. Moreover, light grey means that a result is less than two times worse, grey denotes the results that are twice or more but less than ten times worse, and dark grey indicates that a result is ten or more times worse.

Both Tables 2 and 3 show that for  $\mathbb{T} = \mathbb{R}$  the right-hand side of formula (3) is more efficient than its left-hand counterpart in almost all cases, except for the trivial one, when  $b = 0$ . Nevertheless, even in this case the differences between results are not very significant. Next, one can observe that the bigger the values of the parameters  $b, r, \zeta_1, \dots, \zeta_r$ , the more efficient is the right-hand side of (3). Moreover, its efficiency extremely increases for  $a > \zeta_l, l \in \{1, \dots, r\}$ . Finally, note that in the worst of the presented cases (Table 3:  $b = 0, a = 0, \zeta_1 = 2$ ) the right-hand side of (3) is 1.56 times slower than its traditional analogue, while in the best case (Table 2:  $b = 3, a = 5, r = 4$ ) it is almost 500 times faster than the rival.

**Table 3.** Average computation times [s] of the left-hand side (L) and the right-hand side (R) of formula (3) for  $\mathbb{T} = \mathbb{R}$ ,  $r = 1$  and various values of  $a, b, \zeta_1$

$b$	$a$	$\zeta_1 = 2$		$\zeta_1 = 3$		$\zeta_1 = 4$	
		L	R	L	R	L	R
0	0	0.000044	0.000069	0.000045	0.000065	0.000048	0.000066
	1	0.000052	0.000071	0.000056	0.000074	0.000059	0.000085
	2	0.000065	0.000088	0.000073	0.000083	0.000073	0.000094
	3			0.000082	0.000097	0.000083	0.000096
	4					0.000087	0.000103
1	0	0.000160	0.000128	0.000206	0.000156	0.000257	0.000163
	1	0.000179	0.000161	0.000218	0.000178	0.000260	0.000195
	2	0.000188	0.000185	0.000242	0.000196	0.000285	0.000212
	3	0.000131	0.000084	0.000233	0.000204	0.000301	0.000225
	4			0.000158	0.000094	0.000306	0.000253
	5					0.000185	0.000098
2	0	0.000518	0.000285	0.000907	0.000437	0.001242	0.000566
	1	0.000550	0.000355	0.000962	0.000518	0.001344	0.000692
	2	0.000628	0.000395	0.001006	0.000528	0.001412	0.000747
	3	0.000435	0.000159	0.001036	0.000630	0.001414	0.000784
	4	0.000336	0.000084	0.000643	0.000185	0.001442	0.000853
	5			0.000516	0.000086	0.000906	0.000214
	6					0.000785	0.000108
3	0	0.001968	0.001025	0.003959	0.001924	0.007193	0.003336
	1	0.002155	0.001201	0.004668	0.002420	0.007605	0.003901
	2	0.002181	0.001369	0.004676	0.002430	0.007883	0.003971
	3	0.001420	0.000355	0.004639	0.002347	0.008109	0.003982
	4	0.001287	0.000152	0.003140	0.000566	0.007729	0.002973
	5	0.001184	0.000081	0.002521	0.000191	0.005106	0.000732
	6			0.002361	0.000091	0.004585	0.000231
	7					0.004090	0.000109

Analysis of Tables 4 and 5 reveals that, like in the case  $\mathbb{T} = \mathbb{R}$ , in the discrete-time case  $\mathbb{T} = \tau\mathbb{Z}$  the right-hand side of formula (3) is not efficient for  $b = 0$  and is extremely efficient for  $a > \zeta_l, l \in \{1, \dots, r\}$ . However, the situation is different for  $a \leq \zeta_l, l \in \{1, \dots, r\}$ , when the right-hand side of (3) yields almost the same but (in most cases) slightly worse results than its left-hand counterpart. The reason for such discrepancy between the cases  $\mathbb{T} = \mathbb{R}$  and  $\mathbb{T} = \tau\mathbb{Z}$  is the forward jump operator  $\sigma^{a-\gamma}$ , which requires additional computational time for  $\mathbb{T} = \tau\mathbb{Z}$  and vanishes for  $\mathbb{T} = \mathbb{R}$  as  $\sigma = \text{id}$  in this case. For this reason the right-hand side of (3) is frequently faster than the left-hand one for  $a = 0$ , when  $\sigma^{a-\gamma} = \text{id}$ . Finally, one may observe that in the worst of the presented cases (Table 5:  $b = 0, a = 0, \zeta_1 = 2$ ) the right-hand side of (3) is 1.69 times slower than its traditional analogue, while in the best case (Table 4:  $b = 3, a = 5, r = 4$ ) it is 140 times faster than the rival.

## 6. CONCLUSIONS

The theorem in the paper provides the commutation rule for the higher-order delta and partial derivatives of a composite function with a generalized vector argument. The applicability of the theorem to theoretical research in the field of nonlinear control systems has already been confirmed in [4–6,8], whose proofs of the main results rely on the suggested commutation rule. Due to the unified formulation of the theorem on homogeneous time

**Table 4.** Average computation times [s] of the left-hand side (L) and the right-hand side (R) of formula (3) for  $\mathbb{T} = \tau\mathbb{Z}$ ,  $\zeta_1 = \dots = \zeta_r = 2$  and various values of  $a, b, r$ 

$b$	$a$	$r = 2$		$r = 3$		$r = 4$	
		L	R	L	R	L	R
0	0	0.000080	0.000134	0.000090	0.000147	0.000095	0.000152
	1	0.000079	0.000133	0.000095	0.000140	0.000103	0.000158
	2	0.000082	0.000131	0.000092	0.000148	0.000102	0.000162
1	0	0.007141	0.007390	0.010098	0.009886	0.013027	0.012799
	1	0.007421	0.007648	0.010079	0.010208	0.012972	0.012917
	2	0.007305	0.007330	0.009827	0.009943	0.012853	0.012522
	3	0.007121	0.000201	0.009314	0.000262	0.013014	0.000297
2	0	0.036564	0.035992	0.051326	0.049728	0.065659	0.064578
	1	0.035993	0.037032	0.050713	0.052529	0.065522	0.067529
	2	0.035646	0.036563	0.050799	0.052248	0.065570	0.067461
	3	0.035917	0.007455	0.050597	0.010483	0.065084	0.013046
	4	0.035634	0.000350	0.050537	0.000430	0.065531	0.000542
3	0	0.081602	0.080837	0.118095	0.117639	0.153833	0.153824
	1	0.081111	0.086094	0.117980	0.125055	0.152707	0.160879
	2	0.082174	0.089330	0.117103	0.126280	0.152490	0.163512
	3	0.083010	0.038566	0.117600	0.053632	0.152013	0.069758
	4	0.081971	0.008012	0.117231	0.011007	0.152981	0.014747
	5	0.081106	0.000620	0.116588	0.000828	0.153230	0.001088

scales, it can be used for both continuous-time control systems and delta-domain models, introduced in [9] as the sampled-data models of continuous-time systems, expressed in terms of the difference (delta) operator (i.e., delta derivative in the time scale formalism). Moreover, we tend to believe that the presented theorem has potential benefit for various areas of physics and mathematical analysis, where the differentiation of a composite function is employed. As a supplementary advantage, the formula in the theorem provides a more efficient way of computing certain combinations of the higher-order delta and partial derivatives than the straightforward one.

The presented result is limited only to the case of homogeneous time scales and its formulation and proof for an arbitrary time scale remain a topic for future research. Note, however, that such kind of generalization is not an easy task due to the absence of necessary mathematical tools, unified for all time scales. Therefore, one of the approaches in this situation is gradual extension of results. For example, one of the possible future steps in this direction can be the study of the problem on a larger class of regular time scales, which comprises non-homogeneous time scales with certain properties [1].

## ACKNOWLEDGEMENTS

This work was supported by the Estonian Research Council, personal research funding grant PUT481. The publication costs of this article were covered by the Estonian Academy of Sciences.

**Table 5.** Average computation times [s] of the left-hand side (L) and the right-hand side (R) of formula (3) for  $\mathbb{T} = \tau\mathbb{Z}$ ,  $r = 1$  and various values of  $a, b, \zeta_1$

$b$	$a$	$\zeta_1 = 2$		$\zeta_1 = 3$		$\zeta_1 = 4$	
		L	R	L	R	L	R
0	0	0.000069	0.000117	0.000071	0.000120	0.000075	0.000123
	1	0.000070	0.000118	0.000074	0.000124	0.000076	0.000128
	2	0.000075	0.000119	0.000074	0.000124	0.000079	0.000132
	3			0.000077	0.000124	0.000080	0.000130
	4					0.000080	0.000130
1	0	0.004483	0.004533	0.005043	0.005207	0.006452	0.006407
	1	0.004545	0.004549	0.004983	0.005308	0.006072	0.006458
	2	0.004521	0.004575	0.004923	0.005139	0.006502	0.006954
	3	0.004505	0.000170	0.005521	0.005863	0.006238	0.006472
	4			0.005499	0.000183	0.006391	0.006722
	5					0.006447	0.000207
2	0	0.021232	0.020691	0.026917	0.026434	0.031223	0.030869
	1	0.021868	0.022223	0.026369	0.027353	0.031741	0.033307
	2	0.021638	0.022654	0.026853	0.027787	0.030976	0.033024
	3	0.021208	0.004848	0.027023	0.027985	0.031722	0.032879
	4	0.020886	0.000239	0.025254	0.005919	0.031201	0.033274
	5			0.026037	0.000276	0.030747	0.006490
	6					0.031818	0.000323
3	0	0.046192	0.046174	0.059380	0.058439	0.069883	0.069811
	1	0.046156	0.050937	0.058061	0.063677	0.071080	0.075143
	2	0.046334	0.052445	0.059123	0.065035	0.070029	0.078227
	3	0.046622	0.022958	0.058873	0.064086	0.070392	0.078743
	4	0.046141	0.005024	0.058500	0.028165	0.070003	0.077608
	5	0.046084	0.000359	0.058415	0.006560	0.070802	0.033692
	6			0.058311	0.000478	0.070428	0.007383
	7					0.070839	0.000582

**REFERENCES**

1. Bartosiewicz, Z., Kotta, Ü., Pawłuszewicz, E., and Wyrwas, M. Control systems on regular time scales and their differential rings. *Math. Control Signals Systems*, 2011, **22**, 185–201.
2. Belikov, J., Kotta, Ü., and Tönso, M. NLControl: symbolic package for study of nonlinear control systems. In *The IEEE Multi-Conference on Systems and Control*. IEEE, Hyderabad, India, 2013, 322–327.
3. Bohner, M. and Peterson, A. *Dynamic Equations on Time Scales: An Introduction with Applications*. Birkhäuser, Boston, 2001.
4. Ciulkin, M., Kaparin, V., Kotta, Ü., and Pawłuszewicz, E. Linearization by input–output injections on homogeneous time scales. *Proc. Estonian Acad. Sci.*, 2014, **63**, 387–397.
5. Ciulkin, M., Pawłuszewicz, E., Kaparin, V., and Kotta, Ü. Linearization by generalized input–output injections on homogeneous time scales. In *The 21st International Conference on Methods and Models in Automation and Robotics*. IEEE, Miedzydroje, Poland, 2016, 48–53.
6. Kaparin, V. and Kotta, Ü. Necessary and sufficient conditions in terms of differential-forms for linearization of the state equations up to input-output injections. In *The 8th UKACC International Conference on Control* (Burnham, K. J. and Ersanilli, V. E., eds), IET, Coventry, UK, 2010, 507–511.
7. Kaparin, V. and Kotta, Ü. Theorem on the differentiation of a composite function with a vector argument. *Proc. Estonian Acad. Sci.*, 2010, **59**, 195–200.
8. Kaparin, V. and Kotta, Ü. Transformation of nonlinear state equations into the observer form: necessary and sufficient conditions in terms of one-forms. *Kybernetika*, 2015, **51**, 36–58.

9. Middleton, R. H. and Goodwin, G. C. *Digital Control and Estimation: A Unified Approach*. Prentice Hall, Englewood Cliffs, 1990.
10. Mishkov, R. L. Generalization of the formula of Faa di Bruno for a composite function with a vector argument. *Int. J. Math. Math. Sci.*, 2000, **24**, 481–491.

## Üldistatud vektorargumendiga liitfunktsiooni diferentseerimisest homogeensetel ajaskaaladel

Vadim Kaparin ja Ülle Kotta

On tõestatud teoreem üldistatud vektorargumendiga liitfunktsiooni diferentseerimise kohta. Teoreem on formuleeritud deltatuletise kaudu, mis homogeensetel ajaskaaladel ühildab tavalise ajalise tuletise ja diferentsoperaatori. Termin “üldistatud vektorargument” all mõistetakse liitfunktsiooni, mis sõltub mitte ainult muutujatest, vaid ka nende deltatuletistest. Teoreemis esitatud valem näitab, kuidas kommuteeruvad liitfunktsiooni kõrgemat järku delta- ja osatuletised. Peale selle võimaldab valem alandada deltatuletise järku, mis lihtsustab arvutusi. Valemi arvutusliku efektiivsuse analüüs on esitatud *Mathematica* tarkvaras tehtud katsete alusel.