Proceedings of the Estonian Academy of Sciences,
2010, **59**, 3, 195–200
doi: 10.3176/proc.2010.3.01
Available online at www.eap.ee/proceedings

MATHEMATICS

Theorem on the differentiation of a composite function with a vector argument

Vadim Kaparin* and Ülle Kotta

Institute of Cybernetics at Tallinn University of Technology, Akadeemia tee 21, 12618 Tallinn, Estonia; kotta@cc.ioc.ee

Received 9 November 2009, revised 5 January 2010, accepted 8 January 2010

Abstract. The paper provides a theorem on the differentiation of a composite function with a vector argument. The theorem shows how the partial derivative of the total derivative of the composite function can be expressed through the total derivative of the partial derivative of the composite function. The proof of the theorem is based on Mishkov's formula, which is the generalization of the well-known Faà di Bruno's formula for a composite function with a vector argument.

Key words: differential calculus, partial derivative, total derivative, composite function.

1. INTRODUCTION

The theorem proved in this paper was required as an intermediate result in solving the problem of the transformation of the nonlinear control system, described by state equations, into the observer form and finding the necessary conditions for the possibility of such transformation [2]. The deduction of the necessary conditions involves frequent application of the differentiation of the composite functions with respect to time argument and taking the partial derivatives of the differentiated composite function with respect to one of the variables or its derivatives. The goal of this paper is to present and prove a formula (commutation rule) which allows changing the order of taking the total higher-order derivatives of the composite function and their partial derivatives with respect to one of the variables or its derivative. Since this result may be useful in the solution of other nonlinear control problems, we propose it as a separate contribution. For example, probably the main result, provided in the paper, can be applied for observer design in [4].

The main tool for proving the theorem (commutation rule) is Mishkov's theorem [3] which provides the explicit formula for the nth derivative of a composite function with a vector argument. Mishkov's formula is a straightforward generalization of the well-known Faà di Bruno's formula [1] which gives an explicit equation for the nth-order derivative of the composite function with a scalar argument.

* Corresponding author, vkaparin@cc.ioc.ee

2. MAIN RESULT

The following theorem shows how the partial derivative of the total derivative of the composite function can be expressed through the total derivative of the partial derivative of this function. The composite function with the vector argument with an arbitrary number of components is considered.

Theorem 1. *Assume that $f(\xi_1(t),\xi_2(t),\ldots,\xi_r(t))$ is a composite function for which derivatives up to order $a+b$ are defined; then*

$$\frac{\partial\,(f(\xi_1(t),\xi_2(t),\ldots,\xi_r(t)))^{(a+b)}}{\partial\xi_l^{(a)}(t)}=C_{a+b}^{b}\left(\frac{\partial f(\xi_1(t),\xi_2(t),\ldots,\xi_r(t))}{\partial\xi_l(t)}\right)^{(b)},$$

where $l=1,2,\ldots,r$, C_{a+b}^{b} is the binomial coefficient and a,b are nonnegative integers.

Proof. In the proof we omit the variable t of $\xi_i(t)$, i.e. use instead of $\xi_i(t)$ a shorter notation ξ_i, which allows the bulky formulas to be written in a more compact form. According to Mishkov's formula [3], the $(a+b)$th derivative of the composite function with a vector argument can be computed by the formula

$$(f(\xi_1,\xi_2,\ldots,\xi_r))^{(a+b)}=\sum_0\sum_1\sum_2\cdots\sum_{a+b}\frac{(a+b)!}{\prod\limits_{i=1}^{a+b}(i!)^{k_i}\prod\limits_{i=1}^{a+b}\prod\limits_{j=1}^{r}q_{i,j}!}\frac{\partial^k f}{\partial\xi_1^{p_1}\partial\xi_2^{p_2}\cdots\partial\xi_r^{p_r}}\prod_{i=1}^{a+b}\prod_{j=1}^{r}\left(\xi_j^{(i)}\right)^{q_{i,j}}, \quad (1)$$

where the respective sums are taken over all nonnegative integer solutions of the Diophantine equations as follows:

$$\sum_0\to k_1+2k_2+\cdots+(a+b)k_{a+b}=a+b, \quad (2)$$

$$\sum_i\to q_{i,1}+q_{i,2}+\cdots+q_{i,r}=k_i, \quad (3)$$

for $i=1,\ldots,a+b$, and p_j and k on the right-hand side of (1) satisfy the relations

$$\begin{aligned} p_j&=q_{1,j}+q_{2,j}+\cdots+q_{a+b,j}, \quad j=1,2,\ldots,r,\\ k&=p_1+p_2+\cdots+p_r=k_1+k_2+\cdots+k_{a+b}. \end{aligned} \quad (4)$$

In taking the partial derivative of sum (1) with respect to $\xi_l^{(a)}$, only addends of sum (1) with $q_{a,l}\neq0$ will matter. Denote by $h(\cdot)$ and $g(\cdot)$ the parts of sum (1) corresponding to $q_{a,l}\neq0$ and $q_{a,l}=0$, respectively; then

$$(f(\xi_1,\xi_2,\ldots,\xi_r))^{(a+b)}=h(\cdot)+g(\cdot). \quad (5)$$

Note that it is possible to state that $h(\cdot)$ equals the expression in the right-hand side of (1) where, in addition to the restrictions expressed by (2), (3), and (4), the condition $q_{a,l}\neq0$ has to be satisfied. Note also that if $q_{a,l}\neq0$, then $k_a\neq0$. We prove the formula separately for the cases $a>b$ and $a\leq b$.

First, consider the case when $a>b$. Since $k_a\neq0$ and $q_{a,l}\neq0$, in order to satisfy (2) and (3), the following must hold

$$\begin{aligned} &k_a=1, \quad k_i=0, \quad b<i\leq a+b,\ i\neq a,\\ &q_{a,l}=1, \quad q_{a,j}=0, \quad j=1,2,\ldots,r,\ j\neq l,\\ &q_{i,j}=0, \quad b<i\leq a+b,\ i\neq a,\ j=1,2,\ldots,r. \end{aligned} \quad (6)$$

As a result, under the condition $q_{a,l}\neq0$, one can rewrite (2) as follows:

$$\sum_0\ \to\ k_1+2k_2+\cdots+bk_b=b, \quad (7)$$

and in (3), now $i=1,\ldots,b$.

Using (6) and changing the notations, taking $\bar{p}_j = p_j$ for $j = 1,2,\ldots,r$, $j \neq l$, $\bar{p}_l = p_l - 1$ and $\bar{k} = k - 1$, equations (4) may be rewritten as

$$\begin{aligned}\bar{p}_j &= q_{1,j} + q_{2,j} + \cdots + q_{b,j}, \quad j = 1,2,\ldots,r,\\ \bar{k} &= \bar{p}_1 + \bar{p}_2 + \cdots + \bar{p}_r = k_1 + k_2 + \cdots + k_b.\end{aligned} \tag{8}$$

Note also that under conditions (6)

$$\prod_{i=1}^{a+b} (i!)^{k_i} = a! \prod_{i=1}^{b} (i!)^{k_i}, \qquad \prod_{i=1}^{a+b}\prod_{j=1}^{r} q_{i,j}! = \prod_{i=1}^{b}\prod_{j=1}^{r} q_{i,j}!,$$

and

$$\prod_{i=1}^{a+b}\prod_{j=1}^{r} \left(\xi_j^{(i)}\right)^{q_{i,j}} = \xi_l^{(a)} \prod_{i=1}^{b}\prod_{j=1}^{r} \left(\xi_j^{(i)}\right)^{q_{i,j}}.$$

Taking into account the above equations and the fact that the partial derivative of $g(\cdot)$ in (5) with respect to $\xi_l^{(a)}$ equals 0, we obtain, in new variables $\bar{p}_j$ and $\bar{k}$:

$$\begin{aligned}\frac{\partial \left(f(\xi_1,\xi_2,\ldots,\xi_r)\right)^{(a+b)}}{\partial \xi_l^{(a)}} = \sum_0 \sum_1 \sum_2 \cdots \sum_b \frac{(a+b)!}{a! \prod_{i=1}^{b} (i!)^{k_i} \prod_{i=1}^{b}\prod_{j=1}^{r} q_{i,j}!}\\ \times \frac{\partial^{\bar{k}+1} f}{\partial \xi_1^{\bar{p}_1} \cdots \partial \xi_{l-1}^{\bar{p}_{l-1}} \partial \xi_l^{\bar{p}_l+1} \partial \xi_{l+1}^{\bar{p}_{l+1}} \cdots \partial \xi_r^{\bar{p}_r}} \prod_{i=1}^{b}\prod_{j=1}^{r} \left(\xi_j^{(i)}\right)^{q_{i,j}}.\end{aligned} \tag{9}$$

Note that in (9) all the partial derivatives with respect to ξ_j are of order $\bar{p}_j$ except with respect to ξ_l when the order of the partial derivative is $\bar{p}_l + 1$. In order to unify the orders, denote $\bar{f} := \frac{\partial f}{\partial \xi_l}$. We also multiply the right-hand side of equation (9) by $\frac{b!}{b!}$ to obtain

$$\begin{aligned}\frac{\partial \left(f(\xi_1,\xi_2,\ldots,\xi_r)\right)^{(a+b)}}{\partial \xi_l^{(a)}} = C_{a+b}^{b} \sum_0 \sum_1 \sum_2 \cdots \sum_b \frac{b!}{\prod_{i=1}^{b} (i!)^{k_i} \prod_{i=1}^{b}\prod_{j=1}^{r} q_{i,j}!}\\ \times \frac{\partial^{\bar{k}} \bar{f}}{\partial \xi_1^{\bar{p}_1} \partial \xi_2^{\bar{p}_2} \cdots \partial \xi_r^{\bar{p}_r}} \prod_{i=1}^{b}\prod_{j=1}^{r} \left(\xi_j^{(i)}\right)^{q_{i,j}}.\end{aligned} \tag{10}$$

It is easy to observe now that, according to Mishkov's formula, the sum on the right-hand side of (10) together with the conditions (3) for $i = 1,\ldots,b$, (7) and (8), is the bth-order total derivative of the function $\bar{f}$. Consequently,

$$\frac{\partial \left(f(\xi_1,\xi_2,\ldots,\xi_r)\right)^{(a+b)}}{\partial \xi_l^{(a)}} = C_{a+b}^{b} \bar{f}^{(b)} = C_{a+b}^{b} \left(\frac{\partial f(\xi_1,\xi_2,\ldots,\xi_r)}{\partial \xi_l}\right)^{(b)}. \tag{11}$$

Second, consider the case $a \leq b$. Since $k_a \neq 0$, in order to satisfy (2) and (3), the following must hold:

$$\begin{aligned}k_i &= 0, \quad b < i \leq a+b,\\ q_{i,j} &= 0, \quad b < i \leq a+b,\ j = 1,2,\ldots,r.\end{aligned} \tag{12}$$

Therefore, it is possible to rewrite condition (2) as

$$\sum_0 \ \rightarrow \ k_1 + \cdots + (a-1)k_{a-1} + a(k_a - 1) + (a+1)k_{a+1} + \cdots + bk_b = b, \tag{13}$$

and in (3), now $i = 1, \ldots, b$.

Again, in order to unify the notation in (13), one can take $\bar{k}_i = k_i$ for $i = 1, 2, \ldots, b$, $i \neq a$ and $\bar{k}_a = k_a - 1$. This allows (13) to be rewritten as follows:

$$\sum_0 \ \rightarrow \ \bar{k}_1 + 2\bar{k}_2 + \cdots + b\bar{k}_b = b, \tag{14}$$

and (3) as

$$\begin{aligned} \sum_i \ &\rightarrow \ q_{i,1} + q_{i,2} + \cdots + q_{i,r} = \bar{k}_i, \quad i = 1, \ldots, b, \ i \neq a, \\ \sum_a \ &\rightarrow \ q_{a,1} + q_{a,2} + \cdots + q_{a,r} = \bar{k}_a + 1. \end{aligned} \tag{15}$$

Since $q_{a,l} \geq 1$, we can denote $\bar{q}_{a,l} := q_{a,l} - 1$ and the remaining q's as $\bar{q}_{i,j} := q_{i,j}$. Thereby (15) can be rewritten in unified notation as

$$\sum_i \ \rightarrow \ \bar{q}_{i,1} + \bar{q}_{i,2} + \cdots + \bar{q}_{i,r} = \bar{k}_i, \tag{16}$$

for $i = 1, \ldots, b$. Changing notations, taking $\bar{p}_j = p_j$ for $j = 1, 2, \ldots, r$, $j \neq l$, $\bar{p}_l = p_l - 1$ and $\bar{k} = k - 1$, equations (4) may be rewritten as

$$\begin{aligned} &\bar{p}_j = \bar{q}_{1,j} + \bar{q}_{2,j} + \cdots + \bar{q}_{b,j}, \quad j = 1, 2, \ldots, r, \\ &\bar{k} = \bar{p}_1 + \bar{p}_2 + \cdots + \bar{p}_r = \bar{k}_1 + \bar{k}_2 + \cdots + \bar{k}_b. \end{aligned} \tag{17}$$

Taking (12) into account and using variables $\bar{k}_i$ and $\bar{q}_{i,j}$, we have

$$\begin{aligned} &\prod_{i=1}^{a+b} (i!)^{k_i} = a! \prod_{i=1}^{b} (i!)^{\bar{k}_i}, \qquad \prod_{i=1}^{a+b} \prod_{j=1}^{r} q_{i,j}! = (\bar{q}_{a,l} + 1) \prod_{i=1}^{b} \prod_{j=1}^{r} \bar{q}_{i,j}!, \\ &\prod_{i=1}^{a+b} \prod_{j=1}^{r} \left(\xi_j^{(i)}\right)^{q_{i,j}} \\ &= \left(\xi_l^{(1)}\right)^{\bar{q}_{1,l}} \cdots \left(\xi_l^{(a-1)}\right)^{\bar{q}_{a-1,l}} \left(\xi_l^{(a)}\right)^{\bar{q}_{a,l}+1} \left(\xi_l^{(a+1)}\right)^{\bar{q}_{a+1,l}} \cdots \left(\xi_l^{(b)}\right)^{\bar{q}_{b,l}} \prod_{i=1}^{b} \prod_{\substack{j=1 \\ j \neq l}}^{r} \left(\xi_j^{(i)}\right)^{\bar{q}_{i,j}}. \end{aligned} \tag{18}$$

Furthermore, on the basis of (18) and the fact that the partial derivative of $g(\cdot)$ in (5) with respect to $\xi_l^{(a)}$ equals 0, we obtain, in new variables $\bar{p}_j$ and $\bar{k}$

$$\begin{aligned} \frac{\partial \left(f(\xi_1, \xi_2, \ldots, \xi_r)\right)^{(a+b)}}{\partial \xi_l^{(a)}} = \sum_0 \sum_1 \sum_2 \cdots \sum_b & \frac{(a+b)!}{a! \prod_{i=1}^{b} (i!)^{\bar{k}_i} \prod_{i=1}^{b} \prod_{j=1}^{r} \bar{q}_{i,j}!} \\ & \times \frac{\partial^{\bar{k}+1} f}{\partial \xi_1^{\bar{p}_1} \cdots \partial \xi_{l-1}^{\bar{p}_{l-1}} \partial \xi_l^{\bar{p}_l+1} \partial \xi_{l+1}^{\bar{p}_{l+1}} \cdots \partial \xi_r^{\bar{p}_r}} \prod_{i=1}^{b} \prod_{j=1}^{r} \left(\xi_j^{(i)}\right)^{\bar{q}_{i,j}}. \end{aligned}$$

Like in case $a > b$ we denote $\bar{f} = \frac{\partial f}{\partial \xi_l}$ and multiply the right-hand side of the equality given above by $\frac{b!}{b!}$ to obtain

$$\frac{\partial \left(f(\xi_1, \xi_2, \ldots, \xi_r)\right)^{(a+b)}}{\partial \xi_l^{(a)}} = C_{a+b}^{b} \sum_0 \sum_1 \sum_2 \cdots \sum_b \frac{b!}{\prod_{i=1}^{b} (i!)^{\bar{k}_i} \prod_{i=1}^{b} \prod_{j=1}^{r} \bar{q}_{i,j}!} \times \frac{\partial^{\bar{k}} \bar{f}}{\partial \xi_1^{\bar{p}_1}, \partial \xi_2^{\bar{p}_2} \cdots \partial \xi_r^{\bar{p}_r}} \prod_{i=1}^{b} \prod_{j=1}^{r} \left(\xi_j^{(i)}\right)^{\bar{q}_{i,j}}. \quad (19)$$

Again it is not difficult to observe that according to Mishkov's formula, the sum on the right-hand side of equation (19), together with the conditions (14), (16), and (17), is the bth-order total derivative of the function $\bar{f}$. Consequently, (11) holds again, and this completes the proof. □

Some useful corollaries of the theorem are given below.

Corollary 1. *Under the assumptions of Theorem* 1

$$\frac{\partial \left(f(\xi_1(t), \xi_2(t), \ldots, \xi_r(t))\right)^{(m+n)}}{\partial \xi_l(t)} = \left(\frac{\partial \left(f(\xi_1(t), \xi_2(t), \ldots, \xi_r(t))\right)^{(m)}}{\partial \xi_l(t)} \right)^{(n)},$$

where m and n are nonnegative integers.

Corollary 2. *Under the assumptions of Theorem* 1

$$\frac{\partial \left(f(\xi_1(t), \xi_2(t), \ldots, \xi_r(t))\right)^{(n)}}{\partial \xi_l(t)} = \left(\frac{\partial f(\xi_1(t), \xi_2(t), \ldots, \xi_r(t))}{\partial \xi_l(t)} \right)^{(n)},$$

where n is a nonnegative integer.

3. EXAMPLE

The example in this section illustrates the statement of Theorem 1. Consider the composite function $f(x(t), y(t))$ and assume that we need to take the partial derivative with respect to $\ddot{y}(t)$ of the 3rd-order total derivative of the function. Direct computations yield

$$\frac{\partial \left(f(x(t), y(t))\right)^{(3)}}{\partial \ddot{y}(t)} = 3 \frac{\partial^2 f(x(t), y(t))}{\partial y(t)^2} \dot{y}(t) + 3 \frac{\partial^2 f(x(t), y(t))}{\partial x(t) \partial y(t)} \dot{x}(t).$$

On the other hand, taking the partial derivative of $f(x(t), y(t))$ with respect to $y(t)$ and the total derivative of the obtained result, one gets

$$\left(\frac{\partial f(x(t), y(t))}{\partial y(t)} \right)^{(1)} = \frac{\partial^2 f(x(t), y(t))}{\partial y(t)^2} \dot{y}(t) + \frac{\partial^2 f(x(t), y(t))}{\partial x(t) \partial y(t)} \dot{x}(t).$$

Multiplying both sides of the above equality by C_3^1, we have

$$C_3^1 \left(\frac{\partial f(x(t), y(t))}{\partial y(t)} \right)^{(1)} = 3 \frac{\partial^2 f(x(t), y(t))}{\partial y(t)^2} \dot{y}(t) + 3 \frac{\partial^2 f(x(t), y(t))}{\partial x(t) \partial y(t)} \dot{x}(t).$$

It is not difficult to check that

$$\frac{\partial \left(f(x(t), y(t))\right)^{(3)}}{\partial \ddot{y}(t)} = C_3^1 \left(\frac{\partial f(x(t), y(t))}{\partial y(t)} \right)^{(1)}.$$

4. CONCLUSIONS

The paper shows how to commute the operations of taking higher-order total and partial derivatives of composite functions with vector arguments. The formula, provided in the paper, may be applicable not only in differential calculus. As already mentioned in the introduction, the theorem was a useful tool in deriving solvability conditions of a certain problem in nonlinear control theory. With high probability it may be useful in dealing with other nonlinear control problems where the derivatives of the composite functions with a vector argument often show up.

ACKNOWLEDGEMENT

This work was partially supported by the Estonian Science Foundation through grant No. 6922.

REFERENCES

1. Johnson, W. P. The curious history of Faa di Bruno's Formula. *Amer. Math. Monthly*, 2002, **109**, 217–234.
2. 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. *UKACC International Conference on CONTROL 2010 7–10 September, Coventry, UK* (submitted for publication).
3. Mishkov, R. L. Generalization of the formula of Faa di Bruno for a composite function with a vector argument. *Internat. J. Math. Math. Sci.*, 2000, **24**(7), 481–491.
4. Mishkov, R. L. Nonlinear observer design by reduced generalized observer canonical form. *Internat. J. Control*, 2005, **78**, 172–185.

Teoreem vektorargumendiga liitfunktsiooni diferentseerimisest

Vadim Kaparin ja Ülle Kotta

On tõestatud teoreem vektorargumendiga liitfunktsiooni diferentseerimise kohta. Teoreemis esitatud valem näitab, kuidas liitfunktsiooni täistuletise osatuletist saab väljendada tema osatuletise täistuletise kaudu. Teoreemi tõestus põhineb Mishkovi valemil, mis omakorda kujutab endast tuntud Faà di Bruno valemi üldistust vektorargumendiga liitfunktsiooni jaoks. Näide illustreerib teoreetilist tulemust.