

# Two chain rules for divided differences and Faà di Bruno's formula

Michael S. Floater\* and Tom Lyche†

## Abstract

In this paper we derive two formulas for divided differences of a function of a function. Both formulas lead to other divided difference formulas, such as reciprocal and quotient rules. The two formulas can also be used to derive Faà di Bruno's formula and other formulas for higher derivatives of composite functions. We also derive a divided difference version of Faà di Bruno's determinant formula.

*Math Subject Classification:* 05A17, 05A18, 26A06, 26A24, 41A05, 65D05

*Keywords:* Chain rule, divided differences, Faà di Bruno's formula.

## 1 Introduction

Faà di Bruno's formula [3] is the higher chain rule for differentiation: it expresses an arbitrary derivative of a composite function  $g = f \circ \phi$  in terms of derivatives of  $f$  and  $\phi$ . Though not as well known as the Leibniz formula, it has appeared in various branches of mathematics, notably in mathematical statistics; see the survey paper of Johnson [8].

The formula has also played a useful role in geometric modelling. A parametric curve  $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^d$ , with  $d \geq 2$ , can be parameterized in many different ways. If  $\phi : [\alpha, \beta] \rightarrow \mathbb{R}$  is an increasing function such that  $\phi(\alpha) = a$  and  $\phi(\beta) = b$ , then the curve represented by  $\mathbf{f}$  can equally well be represented

---

\*Centre of Mathematics for Applications, Department of Informatics, University of Oslo, PO Box 1053, Blindern, 0316 Oslo, Norway, *email:* michael.f@ifi.uio.no

†Centre of Mathematics for Applications, Department of Informatics, University of Oslo, PO Box 1053, Blindern, 0316 Oslo, Norway, *email:* tom@ifi.uio.no

by  $\mathbf{g} : [\alpha, \beta] \rightarrow \mathbb{R}^d$  where  $\mathbf{g} = \mathbf{f} \circ \phi$ . Since it is frequently necessary to express the derivatives of the new representation,  $\mathbf{g}$ , in terms of the old,  $\mathbf{f}$ , it is hardly surprising that Faà di Bruno's formula will sooner or later play a role in any consideration of derivatives of  $\mathbf{g}$  of arbitrarily high order. For example, the formula was used by Goodman [6], Gregory [5] and Dyn and Micchelli [2] in dealing with the continuity of spline curves.

More recently, the formula has turned out to be an essential tool in the error analysis of various curve interpolation methods [11, 4]. In this application, the formula has been used to bound the size (the Euclidean norm) of a derivative  $\mathbf{g}^{(k)}(t)$  in terms of the sizes of  $\mathbf{f}'(\phi(t)), \mathbf{f}''(\phi(t)), \dots, \mathbf{f}^{(k)}(\phi(t))$  and the sizes (absolute values) of  $\phi'(t), \phi''(t), \dots, \phi^{(k)}(t)$ .

In fact, what was required in [11, 4] was a bound on certain *divided differences* of  $\mathbf{g}$ , and even though this task could be reduced there to bounding the size of a derivative of  $\mathbf{g}$ , there might be situations in future work of this kind where switching from divided differences to derivatives might not be valid (for example if  $\mathbf{g}$  is not smooth enough). Even if the switch is valid, more precise results might nevertheless follow from a more direct approach: bounding divided differences of  $\mathbf{g}$  in terms of divided differences of  $\mathbf{f}$  and  $\phi$ .

It was this that motivated this paper: to derive an explicit formula that expresses an arbitrary divided difference of a composite function  $g = f \circ \phi$  in terms of divided differences of  $f$  and  $\phi$ . No such formula seems to exist in the literature (see [13], [18], ([7], Chap. 6), and [1]) and we derive *two* natural formulas of this kind. In one the divided differences of  $f$  are at consecutive points while in the other the divided differences of  $\phi$  are at consecutive points. Both of these divided difference chain rules lead to other divided difference formulas, such as reciprocal and quotient rules. The two chain rules also have the attractive feature that they both provide a simple proof of Faà di Bruno's formula. We also derive a divided difference version of Faà di Bruno's determinant formula.

## 2 The first chain rule

As with derivatives there are several notations currently in use for divided differences. In this paper we will denote the divided difference of a real-valued function  $\phi$  at the real values  $t_0, t_1, \dots, t_n$  by  $[t_0, t_1, \dots, t_n]\phi$ ; see de Boor [1] for a discussion of divided differences and alternative notations. We have  $[t_i]\phi = \phi(t_i)$ , and when  $t_0$  and  $t_n$  are distinct we have  $[t_0, \dots, t_n]\phi =$

$([t_1, \dots, t_n]\phi - [t_0, \dots, t_{n-1}]\phi)/(t_n - t_0)$ . We can allow any of the  $t_i$  to be equal if  $\phi$  has sufficiently many derivatives to allow it. In particular, if all the  $t_i$  are equal to  $t$  say, then  $[t_0, t_1, \dots, t_n]\phi = \phi^{(n)}(t)/n!$ . We will make use of the product rule,

$$[t_0, \dots, t_n](\phi\psi) = \sum_{i=0}^n [t_0, \dots, t_i]\phi [t_i, \dots, t_n]\psi, \quad (1)$$

due to Popoviciu [12, 13] and Steffensen [18]. This formula is often called the Leibniz rule [1] since it generalizes the Leibniz formula for differentiation. Now define the composition  $g = f \circ \phi$  given by  $g(t) = f(\phi(t))$  and let  $\phi_i = \phi(t_i)$ . We will prove

**Theorem 1** *For  $n \geq 1$  and for  $\phi$  and  $f$  smooth enough,*

$$[t_0, \dots, t_n]g = \sum_{k=1}^n [\phi_0, \dots, \phi_k]f A_{k,n}\phi, \quad (2)$$

where

$$A_{k,n}\phi = [t_k, \dots, t_n]([t_0, \cdot]\phi \cdots [t_{k-1}, \cdot]\phi). \quad (3)$$

In order to get a more explicit expression for  $A_{k,n}\phi$ , we can use the product rule (1) extended to the product of  $r$  functions, namely,

$$[t_0, \dots, t_n](\psi_1\psi_2 \cdots \psi_r) = \sum_{0=\alpha_0 \leq \alpha_1 \leq \cdots \leq \alpha_r=n} \prod_{\beta=0}^{r-1} [t_{\alpha_\beta}, \dots, t_{\alpha_{\beta+1}}]\psi_{\beta+1}, \quad (4)$$

the sum being over integers  $\alpha_1, \dots, \alpha_{r-1}$  such that  $0 \leq \alpha_1 \leq \cdots \leq \alpha_{r-1} \leq n$ . Since  $A_{k,n}\phi$  is a divided difference of the product of the  $k$  functions  $[t_0, \cdot]\phi, \dots, [t_{k-1}, \cdot]\phi$ , the extended product rule (4) yields

$$A_{k,n}\phi = \sum_{k=\alpha_0 \leq \alpha_1 \leq \cdots \leq \alpha_k=n} \prod_{\beta=0}^{k-1} [t_\beta, t_{\alpha_\beta}, \dots, t_{\alpha_{\beta+1}}]\phi. \quad (5)$$

Using (5) the first three cases of the formula are

$$[t_0, t_1]g = [\phi_0, \phi_1]f [t_0, t_1]\phi \quad (6)$$

$$[t_0, t_1, t_2]g = [\phi_0, \phi_1, \phi_2]f [t_0, t_2]\phi [t_1, t_2]\phi + [\phi_0, \phi_1]f [t_0, t_1, t_2]\phi, \quad (7)$$

$$\begin{aligned} [t_0, t_1, t_2, t_3]g &= [\phi_0, \phi_1, \phi_2, \phi_3]f [t_0, t_3]\phi [t_1, t_3]\phi [t_2, t_3]\phi \\ &\quad + [\phi_0, \phi_1, \phi_2]f \left( [t_0, t_2]\phi [t_1, t_2, t_3]\phi + [t_0, t_2, t_3]\phi [t_1, t_3]\phi \right) \\ &\quad + [\phi_0, \phi_1]f [t_0, t_1, t_2, t_3]\phi. \end{aligned} \quad (8)$$

*Proof.* Assume initially that the  $t_i$  are distinct and consider the polynomial that interpolates  $f$  at  $\phi_0, \dots, \phi_n$  in Newton form,

$$p(x) = \sum_{k=0}^n (x - \phi_0) \cdots (x - \phi_{k-1}) [\phi_0, \dots, \phi_k] f, \quad (9)$$

valid whether or not the  $\phi_i$  are distinct as long as  $f$  is smooth enough. Then we find that  $g(t_i) = f(\phi_i) = p(\phi_i)$  for  $0 \leq i \leq n$ , and therefore

$$[t_0, \dots, t_n]g = [t_0, \dots, t_n](p \circ \phi).$$

Thus replacing  $x$  in (9) by  $\phi(t)$  and applying  $[t_0, \dots, t_n]$  gives (2) where

$$A_{k,n}\phi = [t_0, \dots, t_n]((\phi - \phi_0) \cdots (\phi - \phi_{k-1})). \quad (10)$$

But

$$(\phi(t) - \phi_0) \cdots (\phi(t) - \phi_{k-1}) = \omega_{k-1}(t)[t_0, t]\phi \cdots [t_{k-1}, t]\phi,$$

where  $\omega_{k-1}(t) = (t - t_0) \cdots (t - t_{k-1})$ , and so the product rule (1) gives

$$A_{k,n}\phi = \sum_{i=0}^n [t_0, \dots, t_i] \omega_{k-1} [t_i, \dots, t_n] ([t_0, \cdot]\phi \cdots [t_{k-1}, \cdot]\phi).$$

Since moreover  $\omega_{k-1}$  is a polynomial of degree  $k$  which is zero at  $t_0, \dots, t_{k-1}$ , and whose leading coefficient is 1, we have  $[t_0, \dots, t_i] \omega_{k-1} = \delta_{ik}$ , and this gives equation (3). Finally, we can let the  $t_i$  coalesce in the formula provided  $\phi$  (and therefore also  $g$ ) is smooth enough.  $\square$

We note that Popoviciu already made some progress in [13] towards finding the expression (5). He found  $A_{2,n}\phi$  using equation (10) but for general  $k$  only deduced the number of terms in  $A_{k,n}\phi$ . Due to equation (3),  $A_{k,n}$  is invariant to permutations of the values  $t_0, \dots, t_{k-1}$  and to permutations of  $t_k, \dots, t_n$ . For  $n \geq 3$  this leads to alternatives to equation (5). For example, by permuting  $t_0$  and  $t_1$  in (3) before applying the product rule gives the expression

$$A_{2,3}\phi = [t_1, t_2]\phi [t_0, t_2, t_3]\phi + [t_1, t_2, t_3]\phi [t_0, t_3]\phi,$$

as an alternative to the coefficient of  $[\phi_0, \phi_1, \phi_2]f$  in (8).

### 3 The second chain rule

While the divided differences of  $f$  in (2) are over consecutive points  $\phi_i$ , the divided differences of  $\phi$  in (5) are not at consecutive points  $t_i$ . We now give an alternative formula in which the opposite is true.

**Theorem 2** For  $n \geq 1$  and for  $\phi$  and  $f$  smooth enough,

$$[t_0, \dots, t_n]g = \sum_{k=1}^n \sum_{0=i_0 < \dots < i_k=n} [\phi_{i_0}, \dots, \phi_{i_k}]f \prod_{j=0}^{k-1} [t_{i_j}, \dots, t_{i_{j+1}}]\phi. \quad (11)$$

This formula is fairly easy to remember. We sum over the  $\binom{n-1}{k-1}$  choices of  $k-1$  strictly increasing integers  $\{i_1, \dots, i_{k-1}\}$  from the set  $\{1, 2, \dots, n-1\}$ . Summing over  $k$  we see that the divided difference of  $g$  contains  $2^{n-1}$  terms, the same number as in the first chain rule (2) using (5). The  $f$  differences are of order  $k$  and the first and last point  $\phi_0$  and  $\phi_n$  are always included. For the product terms making up the  $\phi$  differences we simply fill the gaps between each  $t_{i_j}$  and  $t_{i_{j+1}}$ . The first three cases are

$$[t_0, t_1]g = [\phi_0, \phi_1]f [t_0, t_1]\phi, \quad (12)$$

$$\begin{aligned} [t_0, t_1, t_2]g &= [\phi_0, \phi_2]f [t_0, t_1, t_2]\phi \\ &\quad + [\phi_0, \phi_1, \phi_2]f [t_0, t_1]\phi [t_1, t_2]\phi, \end{aligned} \quad (13)$$

$$\begin{aligned} [t_0, t_1, t_2, t_3]g &= [\phi_0, \phi_3]f [t_0, t_1, t_2, t_3]\phi \\ &\quad + [\phi_0, \phi_1, \phi_3]f [t_0, t_1]\phi [t_1, t_2, t_3]\phi \\ &\quad + [\phi_0, \phi_2, \phi_3]f [t_0, t_1, t_2]\phi [t_2, t_3]\phi \\ &\quad + [\phi_0, \phi_1, \phi_2, \phi_3]f [t_0, t_1]\phi [t_1, t_2]\phi [t_2, t_3]\phi. \end{aligned} \quad (14)$$

*Proof.* It is enough to prove (11) for distinct  $t_0, \dots, t_n$  and  $\phi_0, \dots, \phi_n$  since a divided difference is a continuous function of its arguments and also of the function we apply it to. We use induction on  $n$ . Using (12) it is easily seen that (11) holds for  $n=1$ . Suppose that (11) holds for divided differences of order  $n$  and consider  $[t_0, \dots, t_{n+1}]g$ . Using the last two points to reduce the order of the divided difference we have

$$[t_0, \dots, t_{n+1}]g = \frac{[t_0, \dots, t_{n-1}, t_{n+1}]g - [t_0, \dots, t_n]g}{t_{n+1} - t_n}. \quad (15)$$

By the induction hypothesis

$$\begin{aligned}
[t_0, \dots, t_{n-1}, t_{n+1}]g &= \sum_{k=1}^n \sum_{0=i_0 < \dots < i_{k-1} < n} [\phi_{i_0}, \dots, \phi_{i_{k-1}}, \phi_{n+1}]f \times \\
&\quad [t_{i_{k-1}}, \dots, t_{n-1}, t_{n+1}]\phi \prod_{j=0}^{k-2} [t_{i_j}, \dots, t_{i_{j+1}}]\phi, \\
[t_0, \dots, t_{n-1}, t_n]g &= \sum_{k=1}^n \sum_{0=i_0 < \dots < i_{k-1} < n} [\phi_{i_0}, \dots, \phi_{i_{k-1}}, \phi_n]f \times \\
&\quad [t_{i_{k-1}}, \dots, t_{n-1}, t_n]\phi \prod_{j=0}^{k-2} [t_{i_j}, \dots, t_{i_{j+1}}]\phi.
\end{aligned} \tag{16}$$

Consider the difference between the two sums in (16). For each choice of  $(i_0, \dots, i_{k-1})$  we have a common product term multiplying a difference of the form  $A_1 B_1 - A_0 B_0$  where

$$A_j := [\phi_{i_0}, \dots, \phi_{i_{k-1}}, \phi_{n+j}]f, \quad B_j = [t_{i_{k-1}}, \dots, t_{n-1}, t_{n+j}]\phi, \quad j = 0, 1.$$

Now

$$\begin{aligned}
\frac{A_1 B_1 - A_0 B_0}{t_{n+1} - t_n} &= \frac{A_1(B_1 - B_0) + (A_1 - A_0)B_0}{t_{n+1} - t_n} \\
&= [\phi_{i_0}, \dots, \phi_{i_{k-1}}, \phi_{n+1}]f [t_{i_{k-1}}, \dots, t_{n+1}]\phi \\
&\quad + [\phi_{i_0}, \dots, \phi_{i_{k-1}}, \phi_n, \phi_{n+1}]f [t_{i_{k-1}}, \dots, t_n]\phi [t_n, t_{n+1}]\phi.
\end{aligned}$$

We insert these expressions in (15) and (16) and obtain, with  $i_k = n + 1$ ,

$$\begin{aligned}
&[t_0, \dots, t_{n+1}]g \\
&= \sum_{k=1}^n \sum_{0=i_0 < \dots < i_{k-1} < n} [\phi_{i_0}, \dots, \phi_{i_{k-1}}, \phi_{n+1}]f \prod_{j=0}^{k-1} [t_{i_j}, \dots, t_{i_{j+1}}]\phi \\
&\quad + \sum_{k=2}^{n+1} \sum_{0=i_0 < \dots < i_{k-1} = n} [\phi_{i_0}, \dots, \phi_{i_{k-1}}, \phi_{n+1}]f \prod_{j=0}^{k-1} [t_{i_j}, \dots, t_{i_{j+1}}]\phi \\
&= \sum_{k=1}^{n+1} \sum_{0=i_0 < \dots < i_k = n+1} [\phi_{i_0}, \dots, \phi_{i_k}]f \prod_{j=0}^{k-1} [t_{i_j}, \dots, t_{i_{j+1}}]\phi
\end{aligned}$$

which is (11) with  $n$  replaced by  $n + 1$ . □

## 4 Reciprocal and quotient rules

We can use the two chain rules to obtain further formulas that might be of interest. It is easily shown by induction that the divided difference of the function  $f(x) = 1/x$  is  $[t_0, \dots, t_n]f = (-1)^n/(t_0 \cdots t_n)$ ; see ([17], p. 20) and [1]. Thus, applying the first chain rule (2) with this  $f$  gives a reciprocal rule for divided differences of order  $n \geq 1$ ,

$$[t_0, \dots, t_n] \frac{1}{\phi} = \sum_{k=1}^n \frac{(-1)^k A_{k,n} \phi}{\phi_0 \cdots \phi_k}.$$

Combining this with the product rule (1) gives a quotient rule for divided differences. For smooth enough functions  $\phi$  and  $\psi$  we get

$$[t_0, \dots, t_n] \frac{\psi}{\phi} = \frac{1}{\phi_0} [t_0, \dots, t_n] \psi + \sum_{r=1}^n \sum_{k=1}^r \frac{(-1)^k A_{k,r} \phi}{\phi_0 \cdots \phi_k} [t_r, \dots, t_n] \psi.$$

Alternatively we could apply the second chain rule (11) giving, for  $n \geq 1$ , the reciprocal rule

$$[t_0, \dots, t_n] \frac{1}{\phi} = \sum_{k=1}^n (-1)^k \sum_{0=i_0 < \dots < i_k=n} \frac{\prod_{j=0}^{k-1} [t_{i_j}, \dots, t_{i_{j+1}}] \phi}{\phi_{i_0} \phi_{i_1} \cdots \phi_{i_k}},$$

and the quotient rule

$$\begin{aligned} [t_0, \dots, t_n] \frac{\psi}{\phi} &= \frac{1}{\phi_0} [t_0, \dots, t_n] \psi \\ &+ \sum_{r=1}^n \sum_{k=1}^r (-1)^k \sum_{0=i_0 < \dots < i_k=r} \frac{\prod_{j=0}^{k-1} [t_{i_j}, \dots, t_{i_{j+1}}] \phi [t_r, \dots, t_n] \psi}{\phi_{i_0} \phi_{i_1} \cdots \phi_{i_k}}. \end{aligned}$$

## 5 Faà di Bruno's formula

Equation (6) is familiar because for  $t_0 \neq t_1$  and  $\phi_0 \neq \phi_1$  it is simply

$$\frac{g(t_1) - g(t_0)}{t_1 - t_0} = \frac{f(\phi_1) - f(\phi_0)}{\phi_1 - \phi_0} \frac{\phi_1 - \phi_0}{t_1 - t_0}, \quad (17)$$

from which the chain rule

$$g'(t) = f'(\phi(t)) \phi'(t) \quad (18)$$

follows as the limiting case  $t_0 = t_1 = t$ . The limiting cases  $t_i = t$  of equations (7) and (8) (or (13) and (14)) are, after cancellation of factorials,

$$\begin{aligned} g''(t) &= f'(\phi(t))\phi''(t) + f''(\phi(t))(\phi'(t))^2. \\ g'''(t) &= f'(\phi(t))\phi'''(t) + 3f''(\phi(t))\phi'(t)\phi''(t) + f'''(\phi(t))(\phi'(t))^3. \end{aligned}$$

These would normally be found by differentiating (18). It was Faà di Bruno who first found the general formula (without proof) in 1857 in [3]:

$$g^{(n)}(t) = \sum \frac{n!}{b_1! \cdots b_n!} f^{(k)}(\phi(t)) \left( \frac{\phi^{(1)}(t)}{1!} \right)^{b_1} \cdots \left( \frac{\phi^{(n)}(t)}{n!} \right)^{b_n}, \quad (19)$$

where the sum is over all  $k = 1, \dots, n$  and solutions  $b_1, \dots, b_n \geq 0$  to

$$b_1 + 2b_2 + \cdots + nb_n = n, \quad b_1 + \cdots + b_n = k. \quad (20)$$

Various approaches to proving this formula have been proposed by Riordan in [14] and ([15], pp. 34–36), Jordan ([9], Sec. 12), Knuth ([10], p. 481), Roman [16], and Johnson [8]. A nice survey of these and earlier approaches can be found in [8]. In addition to these, the formula was also rediscovered, and a rigorous proof given, by Goodman [6]. The ‘modern’ proof which has emerged through the work of Riordan, Goodman, and Johnson is to use set partitions. A *partition*  $\pi$  of a set  $S$  is a collection of disjoint subsets of  $S$  whose union is  $S$ . The subsets are known as the *blocks* of the partition. Let  $P_n$  be the collection of all partitions of the set  $\{1, 2, \dots, n\}$ . For example,  $P_2 = \{\pi_1, \pi_2\}$  where

$$\pi_1 = \{\{1\}, \{2\}\}, \quad \pi_2 = \{\{1, 2\}\},$$

and  $P_3 = \{\pi_1, \dots, \pi_5\}$  where

$$\begin{aligned} \pi_1 &= \{\{1\}, \{2\}, \{3\}\}, & \pi_2 &= \{\{1, 2\}, \{3\}\}, & \pi_3 &= \{\{1, 3\}, \{2\}\}, \\ \pi_4 &= \{\{2, 3\}, \{1\}\}, & \pi_5 &= \{\{1, 2, 3\}\}. \end{aligned}$$

One can then show that

$$g^{(n)}(t) = \sum_{\pi \in P_n} f^{(\#\pi)}(\phi(t)) \prod_{B \in \pi} \phi^{(\#B)}(t), \quad (21)$$

where  $\pi$  is a partition of the set  $\{1, 2, \dots, n\}$ ,  $B$  is a block of  $\pi$ , and  $\#$  denotes set cardinality. Unlike (19), this ‘set partition’ formula can be proved



relatively easily by induction on  $n$ . By differentiating (21), and using the fact that every partition in  $P_{n+1}$  can be constructed from a partition in  $P_n$  by adding the element  $n + 1$  either to one of its blocks or as a new singleton block, one establishes the formula with  $n$  replaced by  $n + 1$ . Once the set partition formula has been established, a combinatorial argument can be used to collect together repeated terms to arrive at (19).

We can now instead deduce Faà di Bruno's formula (19) from either of the two chain rules for  $[t_0, \dots, t_n]g$ . The special case where all the  $t_i$  are equal to  $t$  in the first chain rule (2),(5) gives

$$\frac{g^{(n)}(t)}{n!} = \sum_{k=1}^n \frac{f^{(k)}(\phi(t))}{k!} c_{k,n}(t), \quad (22)$$

where

$$c_{k,n}(t) = \sum_{k=\alpha_0 \leq \dots \leq \alpha_k = n} \prod_{\beta=0}^{k-1} \frac{\phi^{(\alpha_{\beta+1} - \alpha_{\beta} + 1)}(t)}{(\alpha_{\beta+1} - \alpha_{\beta} + 1)!}. \quad (23)$$

Alternatively, letting all the  $t_i$  equal  $t$  in the second chain rule (11) also gives equation (22), but now with

$$c_{k,n}(t) = \sum_{0=i_0 < \dots < i_k = n} \prod_{j=0}^{k-1} \frac{\phi^{(i_{j+1} - i_j)}(t)}{(i_{j+1} - i_j)!}. \quad (24)$$

Clearly both formulas (23) and (24) can be rewritten as

$$c_{k,n}(t) = \sum_{\substack{j_1 + \dots + j_k = n \\ j_1, \dots, j_k \geq 1}} \left( \frac{\phi^{(j_1)}(t)}{j_1!} \right) \dots \left( \frac{\phi^{(j_k)}(t)}{j_k!} \right). \quad (25)$$

Now two terms in the sum in (25) are the same whenever their corresponding sequences  $(j_1, \dots, j_k)$  contain the same number of ones, the same number of twos, and so on up to the  $n$ 's. Thus, since the number of positive integer solutions to the equation  $j_1 + \dots + j_k = n$  containing  $b_1$  ones,  $b_2$  twos, and so on up to  $b_n$   $n$ 's is the multinomial coefficient  $k!/(b_1! \dots b_n!)$ , it follows that

$$c_{k,n}(t) = \sum \frac{k!}{b_1! \dots b_n!} \left( \frac{\phi^{(1)}(t)}{1!} \right)^{b_1} \dots \left( \frac{\phi^{(n)}(t)}{n!} \right)^{b_n}, \quad (26)$$

where the sum is over all solutions  $b_1, \dots, b_n \geq 0$  to (20). Substituting this into (22) gives Faà di Bruno's formula (19).

## 6 A Determinant Formula

The second chain rule (11) can also be expressed in the form of a determinant, analogous to Faà di Bruno's determinant formula [3].

**Theorem 3** *If  $f$  and  $\phi$  are sufficiently smooth functions,  $g = f \circ \phi$ , and  $n \geq 1$  then*

$$[t_0, \dots, t_n]g = \begin{vmatrix} (0, 1) & (0, 2) & (0, 3) & \cdots & (0, n-1) & (0, n) \\ -1 & (1, 2) & (1, 3) & \cdots & (1, n-1) & (1, n) \\ 0 & -1 & (2, 3) & \cdots & (2, n-1) & (2, n) \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & (n-2, n-1) & (n-2, n) \\ 0 & 0 & 0 & \cdots & -1 & (n-1, n) \end{vmatrix}, \quad (27)$$

where a product  $(i, \dots, j)(j, \dots, k)$  is to be interpreted as  $(i, \dots, j, \dots, k)$  and for  $0 \leq i_0 < i_1 < \dots < i_k \leq n$  we define

$$(i_0, \dots, i_k) := [\phi_{i_0}, \dots, \phi_{i_k}]f \prod_{j=0}^{k-1} [t_{i_j}, \dots, t_{i_{j+1}}]\phi.$$

We illustrate with the first few cases of  $n$ . For  $n = 2$  we have

$$\begin{aligned} \begin{vmatrix} (0, 1) & (0, 2) \\ -1 & (1, 2) \end{vmatrix} &= (0, 1)(1, 2) + (0, 2) = (0, 1, 2) + (0, 2) \\ &= [\phi_0, \phi_1, \phi_2]f [t_0, t_1]\phi [t_1, t_2]\phi + [\phi_0, \phi_2]f [t_0, t_1, t_2]\phi, \end{aligned}$$

while for  $n = 3$  we expand the determinant with respect to the first row and obtain

$$\begin{aligned} \begin{vmatrix} (0, 1) & (0, 2) & (0, 3) \\ -1 & (1, 2) & (1, 3) \\ 0 & -1 & (2, 3) \end{vmatrix} &= (0, 1)(1, 2)(2, 3) + (0, 1)(1, 3) + (0, 2)(2, 3) + (0, 3) \\ &= (0, 1, 2, 3) + (0, 1, 3) + (0, 2, 3) + (0, 3) \end{aligned}$$

which gives (14).

*Proof.* We define  $D_{0,n} = 1$  and for  $k = 0, 1, \dots, n-1$

$$D_{n-k,k} := \begin{vmatrix} (k, k+1) & (k, k+2) & \cdots & (k, n) \\ -1 & (k+1, k+2) & \cdots & (k+1, n) \\ 0 & -1 & \cdots & (k+2, n) \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & (n-1, n) \end{vmatrix}.$$

Then  $D_{n,0}$  is the determinant in (27) and by expanding this determinant with respect to the first row it is easy to see that  $D_{n,0} = \sum_{k=1}^n (0, k) D_{n-k,k}$ . The proof follows by induction on the size of the determinants. Using the current notation equation (11) implies that

$$[t_r, \dots, t_n]g = \sum_{k=0}^{n-1-r} \sum_{r < i_1 < \dots < i_k < n} (r, i_1, \dots, i_k, n), \quad r = 0, \dots, n-1. \quad (28)$$

Suppose that  $D_{n-r,r} = [t_r, \dots, t_n]g$  for  $r = 1, \dots, n-1$ . Note that the orders of these determinants are  $< n$ . Thus using (28) with a slight change of notation we assume that

$$D_{n-i_1, i_1} = \sum_{k=0}^{n-i_1} \sum_{i_1 < i_2 < \dots < i_k < n} (i_1, i_2, \dots, i_k, n), \quad i_1 = 1, \dots, n-1.$$

Then

$$\begin{aligned} D_{n,0} &= \sum_{i_1=1}^n (0, i_1) D(n-i_1, i_1) = \sum_{i_1=1}^n (0, i_1) \sum_{k=0}^{n-i_1} \sum_{i_1 < i_2 < \dots < i_k < n} (i_1, i_2, \dots, i_k, n) \\ &= \sum_{k=0}^{n-1} \sum_{i_1=1}^{n-k} (0, i_1) \sum_{i_1 < i_2 < \dots < i_k < n} (i_1, i_2, \dots, i_k, n) \\ &= \sum_{k=0}^{n-1} \sum_{1 \leq i_1 < i_2 < \dots < i_k < n} (0, i_1, i_2, \dots, i_k, n) = [t_0, \dots, t_n]g. \end{aligned}$$

The last equality follows from (28) with  $r = 0$ .  $\square$

## 7 Other formulas

Johnson, in ([8], Sec. 3) lists four other formulas for arbitrary derivatives of composite functions from the literature, namely those of T.A. (believed to

be an artillery captain named J. F. C. Tiburce Abadie), Scott, Meyer, and Hoppe. The first two formulas follow easily from our two divided difference formulas.

The special case where all the  $t_i$  are equal to  $t$  in  $A_{k,n}\phi$  in (3) gives

$$c_{k,n}(t) = A_{k,n}\phi = \underbrace{[t, \dots, t]}_{n-k+1} ([t, \cdot]\phi)^k = \frac{1}{(n-k)!} \frac{d^{n-k}}{dx^{n-k}} \left( \frac{\phi(x) - \phi(t)}{x-t} \right)^k \Big|_{x=t},$$

which, when substituted into (22), and after writing  $x = t + h$ , gives T.A.'s formula:

$$g^{(n)}(t) = \sum_{k=1}^n \binom{n}{k} f^{(k)}(\phi(t)) \frac{d^{n-k}}{dh^{n-k}} \left( \frac{\phi(t+h) - \phi(t)}{h} \right)^k \Big|_{h=0}.$$

On the other hand, recall equation (25). Since the Leibniz rule applied to  $\phi^k$  gives

$$\frac{d^n}{dt^n} (\phi(t)^k) = n! \sum_{\substack{j_1 + \dots + j_k = n \\ j_1, \dots, j_k \geq 0}} \left( \frac{\phi^{(j_1)}(t)}{j_1!} \right) \dots \left( \frac{\phi^{(j_k)}(t)}{j_k!} \right),$$

it follows that

$$c_{k,n}(t) = \frac{1}{n!} \frac{d^n}{dt^n} (\phi(t)^k) \Big|_{\phi(t)=0},$$

which gives Scott's formula

$$g^{(n)}(t) = \sum_{k=1}^n \frac{f^{(k)}(\phi(t))}{k!} \left\{ \frac{d^n}{dt^n} (\phi(t)^k) \Big|_{\phi(t)=0} \right\}. \quad (29)$$

From (29) one can easily derive Meyer's formula

$$g^{(n)}(t) = \sum_{k=1}^n \frac{f^{(k)}(\phi(t))}{k!} \left\{ \frac{d^n}{dh^n} (\phi(t+h) - \phi(t))^k \right\} \Big|_{h=0}$$

and subsequently Hoppe's formula

$$g^{(n)}(t) = \sum_{k=1}^n \frac{f^{(k)}(\phi(t))}{k!} \sum_{j=0}^k \binom{k}{j} (-\phi(t))^{k-j} \frac{d^n}{dt^n} (\phi(t))^j.$$

## 8 Final remarks

We found two chain rules for divided differences; (2,5) and (11). The formulas are the same for  $n = 1$  and the same for  $n = 2$  after permuting  $t_1$  and  $t_2$ . For  $n \geq 3$  they are distinct. After we had derived the two formulas we found alternative proofs. The first formula (2) can quite easily be proved by induction on  $n$  using the same recursion (15) used in the proof of the second formula (11). On the other hand, the second chain rule (11) can be derived using polynomials in the spirit of the proof of the first rule (2). Since this latter observation could be useful for deriving multivariate versions of (11), we will now give the derivation.

Observe that it is enough to prove (11) when  $f$  is a polynomial of degree at most  $n$ , for if not we interpolate  $f$  with a polynomial  $p$  at  $\phi_0, \dots, \phi_n$  and prove the formula for  $p$ . Further, due to the linearity of divided differences, it is sufficient to prove the formula for the monomial  $f(x) = x^r$  for all  $r$ ,  $1 \leq r \leq n$ . In this case, we can use the extended product rule (4) to obtain

$$[t_0, \dots, t_n]g = [t_0, \dots, t_n](\phi^r) = \sum_{0=\alpha_0 \leq \alpha_1 \leq \dots \leq \alpha_r=n} \prod_{\beta=0}^{r-1} [t_{\alpha_\beta}, \dots, t_{\alpha_{\beta+1}}] \phi. \quad (30)$$

Now by counting repeated terms, each sequence  $(\alpha_0, \alpha_1, \dots, \alpha_r)$  can be written as

$$(\alpha_0, \alpha_1, \dots, \alpha_r) = \underbrace{(i_0, \dots, i_0)}_{1+\mu_0} \underbrace{(i_1, \dots, i_1)}_{1+\mu_1} \dots \underbrace{(i_k, \dots, i_k)}_{1+\mu_k}$$

for some  $k$  with  $1 \leq k \leq r$ , and where  $0 = i_0 < i_1 < \dots < i_k = n$  and  $\mu_0 + \dots + \mu_k = r - k$ . The corresponding product in (30) then becomes

$$\prod_{\beta=0}^{r-1} [t_{\alpha_\beta}, \dots, t_{\alpha_{\beta+1}}] \phi = \phi_{i_0}^{\mu_0} \phi_{i_1}^{\mu_1} \dots \phi_{i_k}^{\mu_k} \prod_{j=0}^{k-1} [t_{i_j}, \dots, t_{i_{j+1}}] \phi.$$

It follows that

$$\begin{aligned} [t_0, \dots, t_n]g &= \sum_{k=1}^r \sum_{0=i_0 < \dots < i_k=n} \sum_{\mu_0 + \dots + \mu_k = r-k} \phi_{i_0}^{\mu_0} \dots \phi_{i_k}^{\mu_k} \prod_{j=0}^{k-1} [t_{i_j}, \dots, t_{i_{j+1}}] \phi \\ &= \sum_{k=1}^n \sum_{0=i_0 < \dots < i_k=n} [\phi_{i_0}, \dots, \phi_{i_k}] f \prod_{j=0}^{k-1} [t_{i_j}, \dots, t_{i_{j+1}}] \phi \end{aligned}$$

for  $f(x) = x^r$ , which proves equation (11) for  $f(x) = x^r$  and therefore for arbitrary  $f$ .

We are planning a second paper dealing with multivariate generalizations of the two divided difference chain rules.

## References

- [1] C. de Boor, Divided differences, *Surveys in approximation theory* **1** (2005), 46–69.
- [2] N. Dyn and C. Micchelli, Piecewise polynomial spaces and geometric continuity of curves, *Numer. Math.* **54** (1988), 319–337.
- [3] C. F. Faà di Bruno, Note sur une nouvelle formule de calcul différentiel, *Quarterly J. Pure Appl. Math* **1** (1857), 359–360.
- [4] M. S. Floater, Arc length estimation and the convergence of parametric polynomial interpolation, *BIT* **45** (2005), 679–694.
- [5] J. Gregory, Geometric continuity, in *Mathematical Methods in Computer Aided Geometric Design*, T. Lyche and L. L. Schumaker (eds.), Academic Press, Boston, 1989, pp. 353–371.
- [6] T. N. T. Goodman, Properties of  $\beta$ -splines, *J. Approx. Theory* **44** (1985), 132–153.
- [7] E. Isaacson and H. B. Keller, *Analysis of numerical methods*, Wiley, 1966.
- [8] W. P. Johnson, The curious history of Faà di Bruno’s formula, *Amer. Math. Monthly* **109** (2002), 217–234.
- [9] C. Jordan, *Calculus of finite differences*, Chelsea, New York, 1947.
- [10] D. Knuth, *The art of computer programming, Vol I*, Addison Wesley, 1975.
- [11] K. Mørken and K. Scherer, A general framework for high-accuracy parametric interpolation, *Math. Comp.* **66** (1997), 237–260.

- [12] T. Popoviciu, Sur quelques propriétés des fonctions d'une ou de deux variables réelles, dissertation, presented at the Faculté des Sciences de Paris, published by Institutul de Arte Grafice "Ardealul" (Cluj, Romania), 1933.
- [13] T. Popoviciu, Introduction à la théorie des différences divisées, Bull. Math. Soc. Roumaine Sciences **42** (1940), 65–78.
- [14] J. Riordan, Derivatives of composite functions, Bull. Amer. Math. Soc. **52** (1946), 664–667.
- [15] J. Riordan, *An introduction to combinatorial analysis*, John Wiley, New York, 1958.
- [16] S. Roman, The formula of Faà di Bruno, Amer. Math. Monthly **87** (1980), 805–809.
- [17] J. F. Steffensen, *Interpolation*, Baltimore, 1927.
- [18] J. F. Steffensen, Note on divided differences, Danske Vid. Selsk. Math.-Fys. Medd **17** (1939), 1–12.