



A DISCRETE FAÀ DI BRUNO'S FORMULA

PEDRO DUARTE AND MARIA JOANA TORRES

ABSTRACT. We derive a discrete Faà di Bruno's formula that rules the behaviour of finite differences under composition of functions with vector values and arguments.

1. INTRODUCTION

The Faà di Bruno's formula gives an expression for the n -th derivative ($n \geq 1$) of the composition $f \circ g$ of two functions f and g in terms of derivatives of f and g . This formula has many versions, depending on the type of derivatives considered. To simplify assumptions we shall always assume that f and g are both smooth functions, where smooth function means a function of class C^∞ . Assume first that f and g are real valued functions of one real argument. Faà di Bruno gave the formula, in this case, in two short papers [5], [6] of 1855 and 1857, without proof or stated sources. The formula is

$$(1.1) \quad (f \circ g)^{(n)}(x) = \sum \frac{n!}{b_1! \cdots b_n!} f^{(b_1 + \cdots + b_n)}(g(x)) \left(\frac{g'(x)}{1!} \right)^{b_1} \cdots \left(\frac{g^{(n)}(x)}{n!} \right)^{b_n},$$

where the sum is taken over all solutions $(b_1, \dots, b_n) \in \mathbb{N}^n$ of the equation

$$b_1 + 2b_2 + \cdots + nb_n = n.$$

We note that $\mathbb{N} = \{0, 1, \dots\}$ starts with zero, so that $b_i \geq 0$ for every $i = 1, \dots, n$. Faà di Bruno also gave an alternative version of (1.1) in the form of a determinant.

According to Lukacs [21], the need for such a formula was already recognized as early as volume 1 of the second edition of Lacroix's *Traité du calcul différentiel et du calcul intégral* of 1810 [20]. In [18], Johnson points out that Lacroix actually solves the problem in volume 3, which came out in 1819.

Received by the editors November 21 2010, and in revised form September 26, 2012.
2000 *Mathematics Subject Classification.* 26B12, 65D25.

Key words and phrases. Finite difference, chain rule, Faà di Bruno's formula.

The first author was partially supported by FCT, Financiamento Base 2008 - ISFL/1/209.

The second author was partially supported by the Research Centre of Mathematics of the University of Minho through the FCT Pluriannual Funding Program.

Johnson also signals two other appearances of the formula previous to Faà di Bruno. More recently, Craik in his paper [4], disagrees with Johnson's evaluation of Arbogast's *Du Calcul des Dérivations* of 1800 [1], and argues this is the first reference to the formula, fifty-five years before Faà di Bruno's formulation. He also examines some early work by British authors on Faà di Bruno's formula not mentioned by Johnson. As Craik puts it, the formula "might as appropriately be named after one of the ten or more authors who obtained versions of it before Faà di Bruno". Craik suggests that only the determinantal formulation should be attributed to Faà di Bruno.

Assume next that f and g are functions between Euclidean spaces of possibly different dimensions, the domain of f containing the range of g . Let e_1, \dots, e_k be vectors in the Euclidean domain of g . The list $e = (e_1, \dots, e_k)$ will be referred as a k -multi-vector. Given $\alpha = (\alpha_1, \dots, \alpha_k) \in \mathbb{N}^k$ and a point x in the domain of g , we denote by $D^\alpha g_x(e)$ the derivative $D_{e_1}^{\alpha_1} \circ \dots \circ D_{e_k}^{\alpha_k} g(x)$ of order $|\alpha| = \alpha_1 + \dots + \alpha_k$. Given $\alpha \in \mathbb{N}^k$, let

$$[\alpha] = \{ \beta = (\beta_1, \dots, \beta_k) \in \mathbb{N}^k \mid \beta \leq \alpha \} = \{0, \beta^1, \beta^2, \dots, \beta^m\},$$

where $\beta \leq \alpha$ means that $\beta_i \leq \alpha_i$ for all $i = 1, \dots, k$. Then, Faà di Bruno's formula takes the form (see [3], [8], [13], [16], [22]):

(1.2)

$$D^\alpha (f \circ g)_x(e) = \sum \frac{\alpha!}{b_1! \dots b_m!} D^{(b_1, \dots, b_m)} f_{g(x)} \left(\frac{D^{\beta^1} g_x(e)}{(\beta^1)!}, \dots, \frac{D^{\beta^m} g_x(e)}{(\beta^m)!} \right),$$

where the sum is taken over all solutions $(b_1, \dots, b_m) \in \mathbb{N}^m$ of the equation

$$b_1 \beta^1 + b_2 \beta^2 + \dots + b_m \beta^m = \alpha,$$

and $\alpha! = \alpha_1! \alpha_2! \dots \alpha_k!$ when $\alpha = (\alpha_1, \dots, \alpha_k)$.

Assume now the multi-index α belongs to the discrete cube $I^k = \{0, 1\}^k$. In this case we have $\alpha! = 1$, $\beta^i! = 1$, and $b_i \in \{0, 1\}$, which implies $b_i! = 1$, for all $i = 1, \dots, m$. Then formula (1.2) reduces to

$$D^\alpha (f \circ g)_x(e) = \sum D^{(b_1, \dots, b_m)} f_{g(x)} \left(D^{\beta^1} g_x(e), \dots, D^{\beta^m} g_x(e) \right),$$

where the sum is taken over all solutions $(b_1, \dots, b_m) \in \mathbb{N}^m$ of the equation

$$b_1 \beta^1 + b_2 \beta^2 + \dots + b_m \beta^m = \alpha.$$

We can simplify this formula a little more.

Definition 1.1. Given a multi-index $\alpha \in I^k$, a partition of α with size r is any subset with r elements $\{\alpha^1, \dots, \alpha^r\} \subseteq I^k - \{0\}$ such that $\alpha^1 + \dots + \alpha^r = \alpha$. More generally, given $\alpha \in I^k$, \mathcal{P}_α denotes the set of all partitions $\xi = \{\alpha^1, \dots, \alpha^r\}$ of α .

Of course the size r must range between 1 and $|\alpha|$, since the smallest partition $\{\alpha\}$ has size 1, and $|\alpha^i| \geq 1$ for each α^i in some partition. Then

$$(1.3) \quad D^\alpha (f \circ g)_x(e) = \sum_{\alpha^1 + \dots + \alpha^r = \alpha} D^r f_{g(x)} \left(D^{\alpha^1} g_x(e), \dots, D^{\alpha^r} g_x(e) \right),$$

where the sum is taken over all partitions of α with size ranging from 1 to $|\alpha|$. Formula (1.3) is at the same time a special case and an extension of (1.2). Given $\alpha \in \mathbb{N}^k$, take \tilde{e} to be the r -multi-vector with α_i components equal to e_i , for each $i = 1, \dots, k$, where $r = |\alpha|$. Set $\tilde{\alpha} \in I^r = \{0, 1\}^r$ to be the multi-index with all components equal to 1. Then $D^\alpha(f \circ g)_x(e) = D^{\tilde{\alpha}}(f \circ g)_x(\tilde{e})$. Grouping and counting equal terms we can derive (1.2) from (1.3).

The Faà di Bruno's formula is useful in different domains of analysis, statistics, and computational calculus; for example, see [7], [9], [11], [14], [15], [17], [21] and for a discussion of the computational aspects, see also [10].

In [12] a discrete version of Faà di Bruno's formula is derived in terms of divided differences. In this paper we deal with *finite differences* instead of divided differences which allows us to state a discrete version of Faà di Bruno's formula with functorial behaviour. The history of finite difference calculus goes back a long way, parallel to that of infinitesimal calculus. We refer to [2] and [19] for classical treatments of this calculus. We briefly recall some basic definitions in order to state our discrete version of Faà di Bruno's formula that behaves functorially.

Given two Euclidean spaces X and Y , let Y^X denote the space of all functions $f : X \rightarrow Y$. For each vector $u \in X$, let τ_u be the *translation operator* defined by $(\tau_u f)(x) = f(x + u)$. We define the *difference operator along vector u* , $\Delta_u : Y^X \rightarrow Y^X$, by $\Delta_u = \tau_u - \text{id}$, which corresponds to setting

$$\Delta_u f(x) = f(x + u) - f(x).$$

Notice these operators always commute, i.e., $\Delta_u \circ \Delta_v = \Delta_v \circ \Delta_u$ for all vectors $u, v \in X$, since $\tau_v \circ \tau_u = \tau_{u+v} = \tau_u \circ \tau_v$. More generally, we will consider compositions of difference operators along possibly repeated vectors.

Definition 1.2. *Given a multi-vector $u = (u_1, \dots, u_k) \in X^k$, a finite difference operator of order k , along the multi-vector u , is the composition operator $\Delta_u^k : Y^X \rightarrow Y^X$ defined by $\Delta_u^k = \Delta_{u_1} \circ \Delta_{u_2} \circ \dots \circ \Delta_{u_k}$. More generally, given $\alpha \in \mathbb{N}^k$, $\Delta_u^\alpha = (\Delta_{u_1})^{\alpha_1} \circ \dots \circ (\Delta_{u_k})^{\alpha_k}$ denotes a difference operator of order $|\alpha|$.*

Next we introduce an algebra of symbolic finite difference expressions. Let x, y, \dots be symbols representing points, u_1, u_2, \dots be symbols representing vectors, and f, g, \dots be symbols representing functions. We denote by \mathcal{D} the set of all symbolic finite difference expressions, which we define as the smallest set of expressions such that:

- (1) if x is a symbol representing a point then $x \in \mathcal{D}$,
- (2) if u is a symbol representing a vector then $u \in \mathcal{D}$,
- (3) if $t \in \mathcal{D}$ and f is a symbol representing a function then $f(t) \in \mathcal{D}$,
- (4) if $t, s \in \mathcal{D}$ then $t + s \in \mathcal{D}$,
- (5) if $\alpha \in \mathbb{N}^k$, $s, t_1, \dots, t_k \in \mathcal{D}$ and f is a symbol representing a function then $\Delta_{(t_1, \dots, t_k)}^\alpha f(s) \in \mathcal{D}$.

We consider as equal all terms which formally can be proved to be equal using property transformation rules of finite differences. Of course, depending on the interpretation given to the point, vector and function symbols, many terms in \mathcal{D} will be formal but meaningless expressions. We define recursively the *order* of a term $\text{ord} : \mathcal{D} \rightarrow \mathbb{N}$:

- (1) if x is a symbol representing a point then $\text{ord}(x) = 0$,
- (2) if u is a symbol representing a vector then $\text{ord}(u) = 1$,
- (3) if $t \in \mathcal{D}$ and f is a symbol representing a function then $\text{ord}(f(t)) = 0$,
- (4) if $t, s \in \mathcal{D}$ then $\text{ord}(t + s) = \min\{\text{ord}(t), \text{ord}(s)\}$,
- (5) if $\alpha \in \mathbb{N}^k$, $s, t_1, \dots, t_k \in \mathcal{D}$ and f is a symbol representing a function then $\text{ord}(\Delta_{(t_1, \dots, t_k)}^\alpha f(s)) = \alpha_1 \text{ord}(t_1) + \dots + \alpha_k \text{ord}(t_k)$.

Given a term $t = \Delta_{(t_1, \dots, t_k)}^\alpha f(s) \in \mathcal{D}$ which is meaningful for some interpretation of its symbols (all functions being smooth), if all vectors u_i in t are small of order ε , then t is small of order $\varepsilon^{\text{ord}(t)}$. The following theorem is our main result. Let X, Y and Z stand for Euclidean spaces.

Theorem 1.3. *Given maps $f \in Y^X$ and $g \in X^Z$, a multi-vector $u \in Z^k$ and $\alpha \in I^k$,*

$$(1.4) \quad \Delta_u^\alpha (f \circ g)(x) = \sum_{\alpha^1 + \dots + \alpha^n = \alpha} \Delta_{\Delta_u^{\alpha^1} g(x), \dots, \Delta_u^{\alpha^n} g(x)}^n f(g(x)) + \dots,$$

where the ellipsis stand for higher order terms and the sum is taken over all partitions of α with size ranging from 1 to $|\alpha|$.

The next theorem gives an explicit formula for (1.4).

Theorem 1.4. *Given $\alpha \in I^k$, there are recursively defined sets $\mathcal{A}_0^\xi, \mathcal{A}_{\alpha^1}^\xi, \dots, \mathcal{A}_{\alpha^r}^\xi$ associated with each partition $\xi = \{\alpha^1, \dots, \alpha^r\}$ of α , such that*

- (1) the sets $\mathcal{A}_0^\xi, \mathcal{A}_{\alpha^1}^\xi, \dots, \mathcal{A}_{\alpha^r}^\xi$ are pairwise disjoint,
- (2) $\beta \in \mathcal{A}_\beta^\xi \subset [\alpha]$, for $\beta = 0, \alpha^1, \dots, \alpha^r$,
- (3) $|\gamma| > |\beta|$, for every $\gamma \in \mathcal{A}_\beta^\xi - \{\beta\}$, and

for any given maps $f \in Y^X$ and $g \in X^Z$, and any multi-vector $u \in Z^k$,

$$(1.5) \quad \Delta_u^\alpha (f \circ g)(x) = \sum_{\xi = \{\alpha^1, \dots, \alpha^r\} \in \mathcal{P}_\alpha} \Delta_{u_{\alpha^1}^\xi, \dots, u_{\alpha^r}^\xi}^r f(u_0^\xi),$$

where $u_\beta^\xi = \sum_{\gamma \in \mathcal{A}_\beta^\xi} \Delta_u^\gamma g(x)$, for each $\beta = 0, \alpha^1, \dots, \alpha^r$.

See Theorem 3.8 in Section 3. Note that $u_0^\xi = g(x) + \dots$, with remainder $\sum_{\gamma \in \mathcal{A}_0^\xi - \{0\}} \Delta_u^\gamma g(x)$, and for each i , $u_{\alpha^i}^\xi = \Delta_u^{\alpha^i} g(x) + \dots$, with remainder $\sum_{\gamma \in \mathcal{A}_{\alpha^i}^\xi - \{\alpha^i\}} \Delta_u^\gamma g(x)$. By Theorem 1.4(3) both these remainders are terms of higher order.

2. THE INFINITESIMAL FORMULA

Let $D \subseteq X$ be an open domain and $f \in Y^D$. The tangent space $T(D)$, and the tangent map $Tf : T(D) \rightarrow T(Y)$, are defined to be $T(D) = D \times X$, respectively $Tf(x, u) = (f(x), D_u f(x))$. The chain rule shows that the construct T is a functor, which essentially means that $T(f \circ g) = Tf \circ Tg$. Inductively, we can define higher order tangent spaces and tangent maps, by setting $T^k(D) = T(T^{k-1}(D))$ and $T^k(f) = T(T^{k-1}(f))$. Then the iterated correspondence T^k becomes also a functor. The tangent map of order k , $T^k(f)$ can be explicitly expressed in terms of higher order directional derivatives of f . As we shall see the pattern of these expressions rules the behaviour of higher order derivatives under composition. To get an explicit expression for $T^k(f)$, we need some special notation to denote elements in $T^k(D)$. We shall call any family $\bar{u} = (u_\alpha)_{\alpha \in I^k}$ of vectors in X indexed over the discrete cube $I^k = \{0, 1\}^k$ a k -cuboid of X . Notice that any k -cuboid \bar{u} can be thought of as a pair of $(k-1)$ -cuboids, obtained by restricting its indices to the two opposite faces $\{k=0\}$ and $\{k=1\}$ of the discrete cube I^k . More precisely, \bar{u} is identified with the pair (\bar{u}_0, \bar{u}_1) , where $\bar{u}_0 = (u_{(\alpha, 0)})_{\alpha \in I^{k-1}}$ and $\bar{u}_1 = (u_{(\alpha, 1)})_{\alpha \in I^{k-1}}$. Therefore, we can and shall identify the tangent space $T^k(X)$ with the set of all k -cuboids of X :

$$T^k(X) = \{ \bar{u} = (u_\alpha)_{\alpha \in I^k} \mid u_\alpha \in X \text{ for all } \alpha \in I^k \}.$$

The k -tangent space $T^k(D)$, to an open domain D , is the set of all k -cuboids $\bar{u} \in T^k(X)$ whose base point u_0 belongs to D . The k -tangent space over a point $x \in D$ is the set $T_x^k(D)$ of all $\bar{u} \in T^k(D)$ such that $u_0 = x$. We shall use the multi-index derivative notation

$$D_u^\alpha f(x) = (D_{u_1})^{\alpha_1} \circ \dots \circ (D_{u_k})^{\alpha_k} f(x).$$

Because multi-indices are cumbersome to write, we shall adopt the following writing convention. Given a cuboid $\bar{u} \in T^k(X)$, u_{i_1, \dots, i_n} stands for the component u_α , where α is the multi-index $(\alpha_1, \dots, \alpha_k)$ defined by $\alpha_j = 1$, if $j \in \{i_1, \dots, i_n\}$, $\alpha_j = 0$ otherwise. Given a multi-index $u = (u_1, \dots, u_k) \in X^k$ we shall denote by $\langle\langle x; u \rangle\rangle$, the k -cuboid \bar{w} such that $w_0 = x, w_1 = u_1, \dots, w_k = u_k$, and $w_{i_1, \dots, i_n} = 0$ for all $n \geq 2$. With this notation it is very easy to check that

Proposition 2.1. *If $f \in Y^X$ is a map of class C^k , then for all $x \in X$, $u \in X^k$,*

$$T^k(f)(\langle\langle x; u \rangle\rangle) = (D_u^\alpha f(x))_{\alpha \in I^k}.$$

The notation

$$\sum_{\alpha^1 + \dots + \alpha^n = \alpha} A_{\alpha^1, \dots, \alpha^n}$$

will always denote a sum taken over all partitions of α (see Definition 1.1) with size ranging from 1 to $|\alpha|$.

Proposition 2.2. *Given a map $f \in Y^X$ of class C^k , and a k -cuboid $\bar{u} \in T_x^k(X)$, writing $T^k f \bar{u} = (T^\alpha f \bar{u})_{\alpha \in I^k}$, we have for each $\alpha \in I^k$*

$$(2.1) \quad T^\alpha f \bar{u} = \sum_{\alpha^1 + \dots + \alpha^n = \alpha} D_{u_{\alpha^1}, \dots, u_{\alpha^n}}^n f(x).$$

Proof. This proposition is proved by induction. \square

Let us call the sum $|\alpha^1| + \dots + |\alpha^n|$ the *order* of the term $D_{u_{\alpha^1}, \dots, u_{\alpha^n}}^n f(x)$. Then, it is clear that the expression (2.1) is homogeneous, as all summands have order $|\alpha|$.

We give a couple of examples:

$$T^{(1,1)}(f)(\bar{u}) = D_{u_{1,2}} f(x) + D_{u_1, u_2}^2 f(x),$$

and

$$\begin{aligned} T^{(1,1,1)}(f)(\bar{u}) &= D_{u_{1,2,3}} f(x) \\ &+ D_{u_1, u_2, 3}^2 f(x) + D_{u_2, u_1, 3}^2 f(x) + D_{u_3, u_1, 2}^2 f(x) \\ &+ D_{u_1, u_2, u_3}^3 f(x). \end{aligned}$$

From Proposition 2.1, we see that the tangent map pattern (2.1) rules the behaviour of higher order derivatives under composition.

Proposition 2.3. *Given maps $f \in Y^X$ and $g \in X^Z$ of class C^k , and a multi-vector $u \in Z^k$, we have for each $\alpha \in I^k$*

$$D_u^\alpha (f \circ g)(x) = \sum_{\alpha^1 + \dots + \alpha^n = \alpha} D_{D_u^{\alpha^1} g(x), \dots, D_u^{\alpha^n} g(x)}^n f(g(x)).$$

Proof. We have that

$$\begin{aligned} D_u^\alpha (f \circ g)(x) &= T^\alpha (f \circ g) \langle\langle x; u \rangle\rangle \\ &= T^\alpha (f) T^k (g) \langle\langle x; u \rangle\rangle \\ &= T^\alpha (f) \left(D_u^\beta g(x) \right)_{\beta \in I^k} \\ &= \sum_{\alpha^1 + \dots + \alpha^n = \alpha} D_{D_u^{\alpha^1} g(x), \dots, D_u^{\alpha^n} g(x)}^n f(x). \end{aligned}$$

\square

3. THE DISCRETE FORMULA

In order to characterize finite difference operators, consider as before the discrete cube $I^k = \{0, 1\}^k$ as a set of multi-indices. Given $\alpha \in I^k$, we write $\alpha = (\alpha_1, \dots, \alpha_k)$, where each α_i represents a binary digit, $\alpha_i = 0$ or $\alpha_i = 1$. The set I^k is partially ordered by the relation

$$\alpha \leq \beta \Leftrightarrow \alpha_i \leq \beta_i, \text{ for all } i = 1, \dots, k.$$

We also write $|\alpha| = \alpha_1 + \dots + \alpha_k$ and $\alpha \cdot u = \alpha_1 u_1 + \dots + \alpha_k u_k$. A simple computation shows that

Proposition 3.1. *Given $f \in Y^X$, for all $x \in X$, and $u \in X^k$,*

$$\Delta_u^\alpha f(x) = \sum_{\beta \leq \alpha} (-1)^{|\alpha| - |\beta|} f(x + \beta \cdot u).$$

A key property of the difference operator Δ is the following kind of additivity.

Proposition 3.2. *Given $f \in Y^X$, for all $x \in X$, and $u, v \in X$,*

$$(3.1) \quad \Delta_{u+v} f(x) = \Delta_u f(x) + \Delta_v f(x + u).$$

We define now a discrete equivalent of the k -tangent map to $f : X \rightarrow Y$. This will be a mapping $T_k(f) : T^k(X) \rightarrow T^k(Y)$. The construct T_k will again be a functor. For that purpose we introduce the *difference operator* $\Delta : T^k(X) \rightarrow T^k(X)$,

$$\Delta \bar{u} = (\Delta^{\alpha \bar{u}})_{\alpha \in I^k}, \text{ where } \Delta^{\alpha \bar{u}} = \sum_{\beta \leq \alpha} (-1)^{|\alpha| - |\beta|} u_\beta.$$

Notice that $\Delta^0 \bar{u} = u_0$. The operator Δ is invertible. Its inverse is the *sum operator* $\Delta^{-1} : T^k(X) \rightarrow T^k(X)$,

$$\Delta^{-1} \bar{u} = \left(\sum_{\beta \leq \alpha} u_\beta \right)_{\alpha \in I^k}.$$

The correspondence $f \in Y^X \rightsquigarrow f_* : T^k(X) \rightarrow T^k(Y)$,

$$f_*(x_\alpha)_{\alpha \in I^k} = (f(x_\alpha))_{\alpha \in I^k},$$

is obviously a functor. Thus, defining $T_k(f) = \Delta \circ f_* \circ \Delta^{-1}$, the correspondence $f \rightsquigarrow T_k(f)$ is conjugated to $f \rightsquigarrow f_*$. Therefore, T_k behaves functorially too.

Proposition 3.3. *Given $f \in Y^X$, for all $x \in X$, and $u \in X^k$,*

$$T_k(f) (\langle\langle x; u \rangle\rangle) = (\Delta_u^\alpha f(x))_{\alpha \in I^k}.$$

Proof. It is enough to notice that $\Delta^{-1} \langle\langle x; u \rangle\rangle = (x + \alpha \cdot u)_{\alpha \in I^k}$. \square

We shall say that any component u_α , of a k -cuboid \bar{u} , has order $|\alpha|$. Then, we define recursively the *order of a finite difference term* $\Delta_{u_1, \dots, u_n}^n f(x)$ to be the sum of the orders of terms u_1, \dots, u_n . Each term u_i can either be some cuboid component, as in the proposition below, or else another finite difference term, as in Theorem 1.3. In both cases, the formulas for $T_\alpha f \bar{u}$ and $\Delta_u^\alpha (f \circ g)(x)$, respectively, have a main part which is a sum of order $|\alpha|$ terms, plus a remainder consisting of terms with order greater than $|\alpha|$.

Proposition 3.4. *Given $f \in Y^X$ and $\bar{u} \in T_x^k(X)$, and writing $T_k f \bar{u} = (T_\alpha f \bar{u})_{\alpha \in I^k}$, we have for each $\alpha \in I^k$*

$$(3.2) \quad T_\alpha f \bar{u} = \sum_{\alpha^1 + \dots + \alpha^n = \alpha} \Delta_{u_{\alpha^1}, \dots, u_{\alpha^n}}^n f(x) + \dots,$$

where the ellipsis stand for a sum of higher order terms.

Proof. This proposition follows from Theorem 3.8 in this section, using the property (3.1) to expand differences. \square

We recall that given $\alpha \in I^k$, \mathcal{P}_α denotes the set of all partitions $\xi = \{\alpha^1, \dots, \alpha^n\}$ of α . The multi-index obtained from α adding digit 1 at the end will be denoted by $\alpha \diamond 1$. Therefore $\alpha \diamond 1$ belongs to I^{k+1} and has order $|\alpha \diamond 1| = |\alpha| + 1$. Next lemma relates \mathcal{P}_α and $\mathcal{P}_{\alpha \diamond 1}$.

Lemma 3.5. *Given $\xi = \{\alpha^1, \dots, \alpha^n\} \in \mathcal{P}_\alpha$, consider the partitions of $\alpha \diamond 1$*

$$\tilde{\xi}_0 = \{0 \diamond 1, \alpha^1 \diamond 0, \dots, \alpha^n \diamond 0\}, \text{ of size } n + 1,$$

$$\tilde{\xi}_i = \{\alpha^1 \diamond 0, \dots, \alpha^i \diamond 1, \dots, \alpha^n \diamond 0\}, \text{ of size } n, \text{ for } 1 \leq i \leq n.$$

These partitions exhaust $\mathcal{P}_{\alpha \diamond 1}$ without repetitions,

$$\mathcal{P}_{\alpha \diamond 1} = \{\tilde{\xi}_i \mid 0 \leq i \leq n \text{ and } \xi \in \mathcal{P}_\alpha\}.$$

Given two k -cuboids $\bar{u}, \bar{v} \in T^k(X)$, we shall denote by $\bar{w} = \llbracket \bar{u}, \bar{v} \rrbracket$ the unique $(k+1)$ -cuboid such that $\bar{w}_{\alpha \diamond 0} = u_\alpha$ and $\bar{w}_{\alpha \diamond 1} = v_\alpha$.

Lemma 3.6. *Given $f \in Y^X$, for all $\bar{u}, \bar{v} \in T^k(X)$,*

$$(1) \Delta^{-1} \llbracket \bar{u}, \bar{v} \rrbracket = \llbracket \Delta^{-1} \bar{u}, \Delta^{-1}(\bar{u} + \bar{v}) \rrbracket,$$

$$(2) \Delta \llbracket \bar{u}, \bar{v} \rrbracket = \llbracket \Delta \bar{u}, \Delta \bar{v} - \Delta \bar{u} \rrbracket,$$

$$(3) T_{k+1}(f) \llbracket \bar{u}, \bar{v} \rrbracket = \llbracket T_k(f) \bar{u}, T_k(f)(\bar{u} + \bar{v}) - T_k(f) \bar{u} \rrbracket.$$

Lemma 3.7. *Given $f \in Y^X$, for all $x \in X$, $w \in X$, and $u, v \in X^k$,*

$$\begin{aligned} & \Delta_{(u_1+v_1), \dots, (u_n+v_n)}^n f(x+w) - \Delta_{u_1, \dots, u_n}^n f(x) = \\ & = \Delta_{w, u_1, \dots, u_n}^{n+1} f(x) + \sum_{i=1}^n \Delta_{u_1, \dots, u_{i-1}, v_i, (u_{i+1}+v_{i+1}), \dots, (u_n+v_n)}^n f(x+w+u_i). \end{aligned}$$

Proof. Follows by repeated application of (3.1). \square

We call *maximum order* of a partition $\xi = \{\alpha^1, \dots, \alpha^n\} \in \mathcal{P}_\alpha$ to the number

$$\text{maxord}(\xi) := \max\{|\alpha^1|, \dots, |\alpha^n|\}.$$

Theorem 3.8. *Given $\alpha \in I^k$, there are recursively defined sets, $\mathcal{A}_0^\xi, \mathcal{A}_{\alpha^1}^\xi, \dots, \mathcal{A}_{\alpha^r}^\xi$, associated with each partition $\xi = \{\alpha^1, \dots, \alpha^r\}$ of α , such that*

(1) *the sets $\mathcal{A}_0^\xi, \mathcal{A}_{\alpha^1}^\xi, \dots, \mathcal{A}_{\alpha^r}^\xi$ are pairwise disjoint,*

(2) *$\beta \in \mathcal{A}_\beta^\xi \subset [\alpha]$, for $\beta = 0, \alpha^1, \dots, \alpha^r$,*

(3) *$0 < \gamma < \alpha$ and $|\gamma| < \text{maxord}(\xi)$, for every $\gamma \in \mathcal{A}_0^\xi - \{0\}$,*

(4) *$\alpha^i < \gamma < \alpha$ and $|\gamma| \leq \text{maxord}(\xi)$, for every $\gamma \in \mathcal{A}_{\alpha^i}^\xi - \{\alpha^i\}$, and*

for any given map $f \in Y^X$ and any multi-vector $u \in Z^k$,

$$(3.3) \quad T_\alpha f \bar{u} = \sum_{\xi=\{\alpha^1, \dots, \alpha^r\} \in \mathcal{P}_\alpha} \Delta_{u_{\alpha^1}^\xi, \dots, u_{\alpha^r}^\xi}^r f(u_0^\xi),$$

where $u_\beta^\xi = \sum_{\gamma \in \mathcal{A}_\beta^\xi} u_\gamma$, for each $\beta = 0, \alpha^1, \dots, \alpha^r$.

Proof. This proposition is proved by induction on k . It is obvious when $\alpha \in I^k$ with $k \leq 2$. Assume it holds when $k \leq n$. Any given multi-index $\alpha \in I^{k+1}$ is of the form $\alpha \diamond 0$ or $\alpha \diamond 1$. Since the first case follows by the induction hypothesis, we now restrict our attention to the second case. Given $\bar{w} \in T^{k+1}(X)$, write it as a pair $\bar{w} = \llbracket \bar{u}, \bar{v} \rrbracket$ of k -cuboids $\bar{u}, \bar{v} \in T^k(X)$. Using Lemma 3.7 we deduce

$$\begin{aligned} T_{\alpha \diamond 1}(f)(\bar{w}) &= T_{\alpha \diamond 1}(f)\llbracket \bar{u}, \bar{v} \rrbracket = T_\alpha(f)(\bar{u} + \bar{v}) - T_\alpha(f)\bar{u} \\ &= \sum_{\xi=\{\alpha^1, \dots, \alpha^n\} \in \mathcal{P}_\alpha} \Delta_{u_{\alpha^1}^\xi + v_{\alpha^1}^\xi, \dots, u_{\alpha^n}^\xi + v_{\alpha^n}^\xi}^n f(u_0^\xi + v_0^\xi) - \Delta_{u_{\alpha^1}^\xi, \dots, u_{\alpha^n}^\xi}^n f(u_0^\xi) \\ &= \sum_{\xi=\{\alpha^1, \dots, \alpha^n\} \in \mathcal{P}_\alpha} \Delta_{v_0^\xi, u_{\alpha^1}^\xi, \dots, u_{\alpha^n}^\xi}^{n+1} f(u_0^\xi) \\ &\quad + \sum_{i=1}^n \Delta_{u_{\alpha^1}^\xi, \dots, u_{\alpha^{i-1}}^\xi, v_{\alpha^i}^\xi, (u_{\alpha^{i+1}}^\xi + v_{\alpha^{i+1}}^\xi), \dots, (u_{\alpha^n}^\xi + v_{\alpha^n}^\xi)}^n f(u_0^\xi + v_0^\xi + u_{\alpha^i}^\xi). \end{aligned}$$

To finish the proof we just need to establish a one-to-one correspondence between summands above and partitions in $\mathcal{P}_{\alpha \diamond 1}$. Using the notation introduced in Lemma 3.5, the partition $\tilde{\xi}_0$ is associated with the first summand, while the partitions $\tilde{\xi}_i$ ($1 \leq i \leq n$) are associated each with one of the subsequent n summands. Making these identifications we arrive at the equation between the previous sum and the following one:

$$\begin{aligned} T_{\alpha \diamond 1}(f)(\bar{w}) &= \sum_{\xi=\{\alpha^1, \dots, \alpha^n\} \in \mathcal{P}_\alpha} \left(\Delta_{w_{0 \diamond 1}^{\tilde{\xi}_0}, w_{\alpha^1 \diamond 0}^{\tilde{\xi}_0}, \dots, w_{\alpha^n \diamond 0}^{\tilde{\xi}_0}}^{n+1} f(w_{0 \diamond 0}^{\tilde{\xi}_0}) + \right. \\ &\quad \left. + \sum_{i=1}^n \Delta_{w_{\alpha^1 \diamond 0}^{\tilde{\xi}_i}, \dots, w_{\alpha^{i-1} \diamond 0}^{\tilde{\xi}_i}, w_{\alpha^i \diamond 1}^{\tilde{\xi}_i}, \dots, w_{\alpha^n \diamond 1}^{\tilde{\xi}_i}}^n f(w_{0 \diamond 0}^{\tilde{\xi}_i}) \right), \end{aligned}$$

which, by Lemma 3.5, yields

$$T_{\alpha \diamond 1}(f)(\bar{w}) = \sum_{\tilde{\xi}=\{\beta^1, \dots, \beta^r\} \in \mathcal{P}_{\alpha \diamond 1}} \Delta_{w_{\beta^1}^{\tilde{\xi}}, \dots, w_{\beta^r}^{\tilde{\xi}}}^r f(w_0^{\tilde{\xi}}).$$

We assume that $u_\beta^\xi = \sum_{\gamma \in \mathcal{A}_\beta^\xi} u_\gamma$, and similar relations hold for $v_\beta^\xi, w_{\beta \diamond 0}^\xi$, etc. Then, matching terms, we arrive at the following recursive equations on the sets $\mathcal{A}_\beta^{\tilde{\xi}_i}$:

$$\begin{cases} \mathcal{A}_{0\Diamond 0}^{\tilde{\xi}_0} &= \mathcal{A}_0^\xi \Diamond 0, \\ \mathcal{A}_{0\Diamond 1}^{\tilde{\xi}_0} &= \mathcal{A}_0^\xi \Diamond 1, \\ \mathcal{A}_{\alpha^i\Diamond 0}^{\tilde{\xi}_0} &= \mathcal{A}_{\alpha^i}^\xi \Diamond 0, \end{cases}$$

and

$$\begin{cases} \mathcal{A}_{\alpha^i\Diamond 1}^{\tilde{\xi}_i} = \mathcal{A}_{\alpha^i}^\xi \Diamond 1, & \text{if } j = i, \\ \mathcal{A}_{\alpha^j\Diamond 0}^{\tilde{\xi}_i} = \mathcal{A}_{\alpha^j}^\xi \Diamond 0, & \text{if } j < i, \\ \mathcal{A}_{\alpha^j\Diamond 0}^{\tilde{\xi}_i} = \mathcal{A}_{\alpha^j}^\xi \Diamond 0 \cup \mathcal{A}_{\alpha^j}^\xi \Diamond 1 & \text{if } j > i, \\ \mathcal{A}_{0\Diamond 0}^{\tilde{\xi}_i} = \mathcal{A}_0^\xi \Diamond 0 \cup \mathcal{A}_0^\xi \Diamond 1 \cup \mathcal{A}_{\alpha^i}^\xi \Diamond 0, \end{cases}$$

where $\mathcal{A} \Diamond i = \{ \beta \Diamond i \mid \beta \in \mathcal{A} \}$ for $i = 0, 1$. These recursive relations ensure that $T_{\alpha\Diamond 1}(f)(\bar{w})$ has the correct development (3.3). It is now easy to check inductively that these sets satisfy the conditions (1), (2), (3), and (4). \square

We prove now Theorem 1.3 and Theorem 1.4 stated in the Introduction.

Proof of Theorem 1.3. The statement of Theorem 1.3 can be obtained as an elementary corollary of Theorem 3.8 via Proposition 3.4. Indeed we have that

$$\begin{aligned} \Delta_u^\alpha(f \circ g)(x) &= T_\alpha(f \circ g)\langle\langle x; u \rangle\rangle \\ &= T_\alpha(f) T_k(g)\langle\langle x; u \rangle\rangle \\ &= T_\alpha(f) \left(\Delta_u^\beta g(x) \right)_{\beta \in I^k} \\ &= \sum_{\alpha^1 + \dots + \alpha^n = \alpha} \Delta_{\Delta_u^{\alpha^1} g(x), \dots, \Delta_u^{\alpha^n} g(x)}^n f(g(x)) + \dots \end{aligned}$$

where the ellipsis stand for a sum of higher order terms. \square

Proof of Theorem 1.4. The statement of Theorem 1.4 can be obtained as an elementary corollary of Theorem 3.8: we just need to set $u_\gamma = \Delta_u^\gamma g(x)$ in Theorem 3.8 and note that $T_\alpha f \bar{u} = \Delta_u^\alpha(f \circ g)(x)$. \square

An algorithm that produces explicit expressions for $T_\alpha f \bar{u}$ has been devised, which was also used to confirm the correction of the above recursive definitions.¹ The following formulas were computer generated by this package.

$$\begin{aligned} T_{(1,1)} f(\bar{u}) &= \Delta_{u_1,2} f(u_0 + u_2 + u_1) \\ &\quad + \Delta_{u_1,2}^2 f(u_0), \end{aligned}$$

¹The source code can be retrieved from <http://ptmat.fc.ul.pt/~pduarte/Research/FiniteDifferences/index.html>.

$$\begin{aligned}
T_{(1,1,1)}f(\bar{u}) &= \Delta_{u_1,2,3}f(u_0 + u_3 + u_2 + u_{2,3} + u_1 + u_{1,3} + u_{1,2}) \\
&\quad + \Delta_{u_1,u_2,3}^2f(u_0 + u_3 + u_2) \\
&\quad + \Delta_{u_1,3,u_2+u_{2,3}}^2f(u_0 + u_3 + u_1) \\
&\quad + \Delta_{u_1,2,u_3+u_{2,3}+u_{1,3}}^2f(u_0 + u_2 + u_1) \\
&\quad + \Delta_{u_1,u_2,u_3}^3f(u_0),
\end{aligned}$$

$$\begin{aligned}
\Delta_v^2(f \circ g)(x) &= \Delta_{\Delta_{v_1,v_2}^2g(x)}f(g(x) + \Delta_{v_1}g(x) + \Delta_{v_2}g(x)) \\
&\quad + \Delta_{\Delta_{v_1}g(x),\Delta_{v_2}g(x)}^2f(g(x)),
\end{aligned}$$

and

$$\begin{aligned}
\Delta_v^3(f \circ g)(x) &= \Delta_{\Delta_{v_1,v_2,v_3}^3g(x)}f(g(x) + \Delta_{v_1}g(x) + \Delta_{v_2}g(x) + \Delta_{v_3}g(x)) \\
&\quad + \Delta_{v_1,v_2}^2g(x) + \Delta_{v_1,v_3}^2g(x) + \Delta_{v_2,v_3}^2g(x) \\
&\quad + \Delta_{\Delta_{v_1}g(x),\Delta_{v_2,v_3}^2g(x)}^2f(g(x) + \Delta_{v_2}g(x) + \Delta_{v_3}g(x)) \\
&\quad + \Delta_{\Delta_{v_1,v_3}^2g(x),\Delta_{v_2}g(x)+\Delta_{v_2,v_3}^2g(x)}^2f(g(x) + \Delta_{v_1}g(x) + \Delta_{v_3}g(x)) \\
&\quad + \Delta_{\Delta_{v_1,v_2}^2g(x),\Delta_{v_3}g(x)+\Delta_{v_1,v_3}^2g(x)+\Delta_{v_2,v_3}^2g(x)}^2f(g(x) + \Delta_{v_1}g(x) \\
&\quad \quad + \Delta_{v_2}g(x)) \\
&\quad + \Delta_{\Delta_{v_1}g(x),\Delta_{v_2}g(x),\Delta_{v_3}g(x)}^3f(g(x)).
\end{aligned}$$

REFERENCES

1. L. F. A. Arbogast, *Du calcul des dérivations*, Levrault, Strasbourg, 1800.
2. G. Boole, *A treatise on the calculus of finite differences*, 2 ed., Dover Publications, New York.
3. G. M. Constantine and T. H. Savits, *A multivariate Faà di Bruno formula with applications*, Trans. Amer. Math. Soc. **348** (1996), no. 2, 503–520.
4. A. Craik, *Prehistory of Faà di Bruno's formula*, Amer. Math. Monthly **112** (2005), 119–130.
5. C. F. Faà di Bruno, *Sullo sviluppo delle funzioni*, Annali di Scienze Matematiche e Fisiche **6** (1855), 479–480.
6. ———, *Note sur une nouvelle formule de calcul différentiel*, Quarterly J. Pure Appl. Math. **1** (1857), 359–360.
7. N. Dyn and C. Micchelli, *Piecewise polynomial spaces and geometric continuity of curves*, Numer. Math. **54**, no. 3, 319–337.
8. L. H. Encinas and J. M. Masque, *A short proof of the generalized Faà di Bruno's formula*, Appl. Math. Lett. **16** (2003), no. 6, 975–979.
9. G. Falqui, C. Reina, and A. Zampa, *Krichever maps, Faà di Bruno polynomials, and cohomology in KP theory*, Lett. Math. Phys **42** (1997), no. 4, 349–361.
10. H. Flanders, *From Ford to Faà*, Amer. Math. Monthly **108** (2001), no. 6, 559–561.
11. M. Floater, *Arc length estimation and the convergence of polynomial curve interpolation*, BIT **45** (2005), no. 4, 679–694.

12. M. Floater and T. Lyche, *Two chain rules for divided differences and Faà di Bruno's formula*, Math. Comp. **76** (2007), no. 258, 867–877.
13. L. E. Fraenkel, *Formulae for high derivatives of composite functions*, Math. Proc. Cambridge Philos. Soc. **83** (1978), no. 2, 159–165.
14. T. N. T. Goodman, *Properties of β -splines*, J. Approx. Theory **44** (1985), no. 2, 132–153.
15. J. Gregory, *Geometric continuity*, Mathematical Methods in Computer Aided Geometric Design, Academic Press, Boston, 1989, pp. 353–371.
16. H. Gzyl, *Multidimensional extension of Faà di Bruno's formula*, J. Math. Anal. Appl. **116** (1986), no. 2, 450–455.
17. C. Hespel, *Iterated derivatives of the output of a nonlinear dynamical systems and Faà di Bruno formula*, Math. Comp. Simulation **42** (1996), 641–657.
18. W. P. Johnson, *The curious history of Faà di Bruno's formula*, Amer. Math. Monthly **109** (2002), no. 3, 217–234.
19. C. Jordan, *Calculus of finite differences*, 3 ed., Chelsea Publishing Co., Inc., New York, 1965.
20. S. F. Lacroix, *Traité du calcul différentiel et du calcul intégral*, Duprat, Paris, 1810-19, 3 vols.
21. E. Lukacs, *Applications of Faà di Bruno's formula in mathematical statistics*, Amer. Math. Monthly **62** (1955), 340–348.
22. T-W. Ma, *Higher chain formula proved by combinatorics*, Electron. J. Combin. **16** (2009), #N21.
23. K. Mørken and K. Scherer, *A general framework for high-accuracy parametric interpolation*, Math. Comp. **66** (1997), no. 217, 237–260.

CMAF, DEPARTAMENTO DE MATEMÁTICA, FACULDADE DE CIÊNCIAS DA
UNIVERSIDADE DE LISBOA, PORTUGAL
E-mail address: pmduarte@fc.ul.pt

CMAT, DEPARTAMENTO DE MATEMÁTICA E APLICAÇÕES, UNIVERSIDADE DO MINHO,
CAMPUS DE GUALTAR, 4700-057 BRAGA, PORTUGAL
E-mail address: jtorres@math.uminho.pt