A discrete Faà di Bruno's formula

P. Duarte[a], M. J. Torres[b]

[a] *CMAF, Departamento de Matemática, Faculdade de Ciências da Universidade de Lisboa, Portugal (e-mail: pedromiguel.duarte@gmail.com).*

Supported by Fundação para a Ciência e a Tecnologia, Financiamento Base 2008 - ISFL/1/209.

[b] *CMAT, Departamento de Matemática, Universidade do Minho, Campus de Gualtar, 4700-057 Braga, Portugal (e-mail: jtorres@math.uminho.pt).*

Abstract

We derive some formulas that rule the behaviour of finite differences under composition of functions with vector values and arguments.

Key words: Finite difference, chain rule, Faà di Bruno formula.

2000 MSC: 26B12, 65D25

1. Introduction

The Faà di Bruno formula [2] 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. Functions of class C^∞ will be called smooth. To simplify assumptions we shall always assume that f and g are both smooth functions. This formula has many versions, depending on the type of derivatives considered. Assume first that f and g are real valued functions of one real argument. Then the formula takes the following form

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

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

$$b_1 + 2\, b_2 + \ldots + n\, b_n = n \ .$$

We note that $\mathbb{N} = \{0, 1, \ldots\}$ starts with zero, so that $b_i \geq 0$ for every $i = 1, \ldots, n$.

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 $e_1, \ldots, e_k$ be vectors in the euclidean domain of g. The list $e = (e_1, \ldots, e_k)$ will be referred as a k-multi-vector. Given $\alpha = (\alpha_1, \ldots, \alpha_k) \in \mathbb{N}^k$ and a point x in the domain of g, we denote by $D^\alpha g_x(e)$ the derivative $D_{e_1}^{\alpha_1} \circ \cdots \circ D_{e_k}^{\alpha_k} g(x)$ of order $|\alpha| = \alpha_1 + \ldots + \alpha_k$. Given $\alpha \in \mathbb{N}^k$, let

$$[\alpha] = \{\, \beta \in \mathbb{N}^k \,:\, \beta \leq \alpha \,\} = \{0, \beta^1, \beta^2, \ldots, \beta^m\} \,.$$

Then, Faà di Bruno's formula takes the form

$$D^\alpha (f \circ g)_x(e) = \sum \frac{\alpha!}{b_1! \cdots b_m!} D^{(b_1, \ldots, b_m)} f_{g(x)} \left(\frac{D^{\beta^1} g_x(e)}{(\beta^1)!}, \ldots, \frac{D^{\beta^m} g_x(e)}{(\beta^m)!} \right) \tag{2}$$

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

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

and $\alpha! = \alpha_1! \, \alpha_2! \cdots \alpha_k!$ when $\alpha = (\alpha_1, \ldots, \alpha_k)$. 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 $i = 1, \ldots, m$. Then formula (2) reduces to

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

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

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

We can simplify this formula a little more.

Definition 1. *Given a multi-index $\alpha \in I^k$, we call* partition of α with size r *to any subset with r elements $\{\alpha^1 \ldots, \alpha^r\} \subseteq I^k - \{0\}$ such that $\alpha^1 + \cdots + \alpha^r = \alpha$.*

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

$$D^\alpha(f \circ g)_x(e) = \sum_{\alpha^1+\cdots+\alpha^r=\alpha} D^m f_{g(x)} \left(D^{\alpha^1} g_x(e), \ldots, D^{\alpha^r} g_x(e)\right) \qquad (3)$$

where the sum is taken over all partitions of α with size ranging from 1 to $|\alpha|$. Formula (3) is at the same time a special case and an extension of (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 $i = 1, \ldots, k$, 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 (2) from (3).

After having written this article we found the reference [3] where a discrete version of this formula is derived in terms of divided differences. Here we deal with *finite differences* instead of *divided differences*, which means our formula holds for general functions of several variables. The history of Finite Difference Calculus goes back a long way, parallel to that of Infinitesimal Calculus. We refer to [4] for a modern treatment of this calculus. We briefely recall some basic definitions in order to state our discrete version of Faà di Bruno's formula. Given a vector $u \in X$, in some euclidean space X, let τ_u be the *translation operator* defined by $(\tau_u f)(x) = f(x+u)$. This operator acts on every space of functions $f \in Y^X$. The *difference operator along vector* u, $\Delta_u : Y^X \to Y^X$, is defined by $\Delta_u = \mathrm{id} - \tau_u$, which corresponds to set $(\Delta_u f)(x) = f(x+u) - f(x)$. 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 $\tau_v \circ \tau_u = \tau_{u+v} = \tau_u \circ \tau_v$.

Definition 2. *We call finite difference operator of order k to any composition of k difference operators along possibly repeated vectors. Given a multi-vector $u = (u_1, \cdots, u_k) \in X^k$, we denote by $\Delta_u^k : Y^X \to Y^X$ the composite operator $\Delta_u^k = \Delta_{u_1} \circ \Delta_{u_2} \circ \cdots \circ \Delta_{u_k}$. More generally, given $\alpha \in \mathbb{N}^k$, $\Delta_u^\alpha = (\Delta_{u_1})^{\alpha_1} \circ \ldots \circ (\Delta_{u_k})^{\alpha_k}$ denotes a difference operator of order $|\alpha|$.*

Next we introduce an algebra of symbolic finite difference expressions. Let $x, y, \ldots$ be symbols representing points, $u_1, u_2, \ldots$ be symbols representing vectors, and $f, g, \ldots$ 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 $\alpha \in \mathbb{N}^k$, $s, t_1, \ldots, t_k \in \mathcal{D}$ and f is a symbol representing a function then $\Delta^{\alpha}_{(t_1,\ldots,t_k)} f(s) \in \mathcal{D}$.

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 $\mathrm{ord} : \mathcal{D} \to \mathbb{N}$:

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

Given a term $t = \Delta^{\alpha}_{(t_1,\ldots,t_k)} f(s) \in \mathcal{D}$ 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 $\epsilon^{\mathrm{ord}(t)}$. The following theorem is our main result. Let X, Y and Z stand for linear spaces.

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

$$\Delta^{\alpha}_{u}(f \circ g)(x) = \sum_{\alpha^1 + \cdots + \alpha^n = \alpha} \Delta^{n}_{\Delta^{\alpha^1}_{u} g(x), \ldots, \Delta^{\alpha^n}_{u} g(x)} f\,(\,g(x)\,) \;+\; \cdots \quad , \tag{4}$$

where the ellipsis stand for higher order terms.

Next theorem gives an explicit formula for (4).

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

1. *the sets $\mathcal{A}^{\xi}_0$, $\mathcal{A}^{\xi}_{\alpha^1}$, $\ldots$, $\mathcal{A}^{\xi}_{\alpha^r}$ are pairwise disjoint;*
2. *$\beta \in \mathcal{A}^{\xi}_{\beta} \subset [\alpha]$, for $\beta = 0, \alpha^1, \ldots, \alpha^r$;*

3. $|\gamma| > |\beta|$, *for every* $\gamma \in \mathcal{A}^{\xi}_{\beta} - \{\beta\}$; *and*

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

$$\Delta^{\alpha}_{u}(f \circ g)(x) = \sum_{\xi=\{\alpha^1,\ldots,\alpha^r\}\in\mathcal{P}_\alpha} \Delta^{r}_{u^{\xi}_{\alpha^1},\ldots,u^{\xi}_{\alpha^r}} f(u^{\xi}_0) \ , \tag{5}$$

where $u^{\xi}_{\beta} = \sum_{\gamma\in\mathcal{A}^{\xi}_{\beta}} \Delta^{\gamma}_{u} g(x)$, *for each* $\beta = 0, \alpha^1, \ldots, \alpha^r$.

See theorem 1. Note that $u^{\xi}_0 = g(x) + \cdots$, with remainder $\sum_{\gamma\in\mathcal{A}^{\xi}_0-\{0\}} \Delta^{\gamma}_{u} g(x)$, and for each i, $u^{\xi}_{\alpha^i} = \Delta^{\alpha^i}_{u} g(x) + \cdots$, with remainder $\sum_{\gamma\in\mathcal{A}^{\xi}_{\alpha^i}-\{\alpha^i\}} \Delta^{\gamma}_{u} g(x)$. By 3. both these remainders are terms of higher order.

2. The Infinitesimal Formula

Let $D \subseteq X$ be an open domain. The tangent space $T(D)$, and the tangent map $Tf : T(D) \to T(Y)$, are defined to be $T(D) = D \times X$, respectively $Tf(x,u) = (f(x), D_u f(x))$. The chain rule shows that 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 *k-cuboid* of X to any family $\overline{u} = (u_\alpha)_{\alpha\in I^k}$ of vectors in X indexed over the discrete cube $I^k = \{0,1\}^k$. Notice that any k-cuboid $\overline{u}$ can be thought of as a pair of $(k-1)$-cuboids, corresponding to restrict its indices to the two opposite faces $\{\alpha_k = 0\}$ and $\{\alpha_k = 1\}$ of the discrete cube I^k. Therefore, we can and shall identify the tangent space $T^k(X)$ with the set of all k-cuboids of X:

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

The k-tangent space $T^k(D)$, to an open domain D, is the set of all k-cuboids $\overline{u} \in T^k(X)$ whose base point u_0 belongs to D. The k-tangent space over a

point $x \in D$ is the set $T^k_x(D)$ of all $\overline{u} \in T^k(D)$ such that $u_0 = x$. We shall use the multi-index derivative notation

$$D^\alpha_u f(x) = (D_{u_1})^{\alpha_1} \circ \ldots \circ (D_{u_k})^{\alpha_k} f(x) \ .$$

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

Proposition 1. *If $f \in Y^X$ is a class C^k map, then for all $x \in X$, $u \in X^k$,*

$$T^k(f)\left(\langle\!\langle x; u \rangle\!\rangle\right) = \left(D^\alpha_u f(x)\right)_{\alpha \in I^k} \ .$$

The following kind of notation

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

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

Proposition 2. *Given a class C^k map $f \in Y^X$ and a k-cuboid $\overline{u} \in T^k_x(X)$, writing $T^k f\, \overline{u} = (T^\alpha f\, \overline{u})_{\alpha \in I^k}$, we have for each $\alpha \in I^k$*

$$T^\alpha f\, \overline{u} = \sum_{\alpha^1 + \cdots + \alpha^n = \alpha} D^n_{u_{\alpha^1},\ldots,u_{\alpha^n}} f(x) \ . \tag{6}$$

Proof. This proposition is proved by induction. □

Let us call *order* of a term $D^n_{u_{\alpha^1},\ldots,u_{\alpha^n}} f(x)$ to the sum $|\alpha^1| + \ldots + |\alpha^n|$. Then, it is clear that expression (6) is homogeneous: all summands have order $|\alpha|$.

A couple of examples

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

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

From proposition 1, we see that the tangent map pattern (6) rules the behaviour of higher order derivatives under composition.

Proposition 3. *Given class C^k maps $f \in Y^X$ and $g \in X^Z$, and a multi-vector $u \in Z^k$,*

$$
D^\alpha_u (f \circ g)(x) = \sum_{\alpha^1 + \cdots + \alpha^n = \alpha} D^n_{D^{\alpha^1}_u g(x), \ldots, D^{\alpha^n}_u g(x)} f\left(g(x)\right) \ .
$$

Proof.

$$
\begin{aligned}
D^\alpha_u (f \circ g)(x) &= T^\alpha (f \circ g) \langle\!\langle x; u \rangle\!\rangle \\
&= T^\alpha(f)\, T^k(g) \langle\!\langle x; u \rangle\!\rangle \\
&= T^\alpha(f) \left(D^\beta_u g(x) \right)_{\beta \in I^k} \\
&= \sum_{\alpha^1 + \cdots + \alpha^n = \alpha} D^n_{D^{\alpha^1}_u g(x), \ldots, D^{\alpha^n}_u g(x)} f(x)
\end{aligned}
$$

□

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 $\alpha = (\alpha_1, \cdots, \alpha_k)$, 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

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

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

Proposition 4. *Given $f \in Y^X$, for all $x \in X$, and $u \in X^k$,*

$$
\Delta^\alpha_u f(x) = \sum_{\beta \leq \alpha} (-1)^{|\alpha| - |\beta|} f\left(x + \beta \cdot u\right)
$$

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

Proposition 5. *Given $f \in Y^X$, for all $x \in X$, and $u, v \in X$,*

$$\Delta_{u+v} f(x) = \Delta_u f(x) + \Delta_v f(x+u) \ . \tag{7}$$

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

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

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

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

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

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

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

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

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

We shall say that any component u_α, of a k-cuboid $\overline{u}$, has order $|\alpha|$. Then, we define recursively the *order of a finite difference term* $\Delta^n_{u_1,\ldots,u_n} f(x)$ to be the sum of the orders of terms $u_1, \ldots, u_n$. Each term u_i can either be some cuboid component, as in proposition below, or else another finite difference term, as in theorem A. In both cases the formulas for $T_\alpha f \overline{u}$ 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 $> |\alpha|$.

Proposition 7. *Given $f \in Y^X$ and $\overline{u} \in T_x^k(X)$, writing $T_k f\, \overline{u} = (T_\alpha f\, \overline{u})_{\alpha \in I^k}$, we have for each $\alpha \in I^k$*

$$T_\alpha f\, \overline{u} = \sum_{\alpha^1 + \cdots + \alpha^n = \alpha} \Delta^n_{u_{\alpha^1}, \ldots, u_{\alpha^n}} f(\, x\,) \; + \; \cdots \tag{8}$$

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

Proof. This proposition follows from theorem 1, using the property (7) to expand differences. □

Given $\alpha \in I^k$, we shall denote by $\mathcal{P}_\alpha$ the set of all partitions $\xi = \{\alpha^1, \ldots, \alpha^n\}$ 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 $\mathcal{P}_\alpha$ with $\mathcal{P}_{\alpha\Diamond 1}$.

Lema 1. *Given $\xi = \{\alpha^1, \ldots, \alpha^n\} \in \mathcal{P}_\alpha$, consider the partitions of $\alpha\Diamond 1$*

$$\tilde{\xi}_0 = \{0\Diamond 1,\, \alpha^1\Diamond 0,\, \ldots,\, \alpha^n\Diamond 0\} \quad \textit{with size } n+1,$$
$$\tilde{\xi}_i = \{\alpha^1\Diamond 0,\, \ldots,\, \alpha^i\Diamond 1,\, \ldots, \alpha^n\Diamond 0\} \quad \textit{with size } n, \quad \textit{for } \; 1 \le i \le n.$$

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

$$\mathcal{P}_{\alpha\Diamond 1} = \{\tilde{\xi}_i \; : \; 0 \le i \le n, \quad \textit{and} \;\; \xi \in \mathcal{P}_\alpha \; \} \; .$$

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

Lema 2. *For all $\overline{u}$, $\overline{v} \in T^k(X)$,*

1. $\Delta^{-1}[\![\overline{u}, \overline{v}]\!] = [\![\Delta^{-1}\overline{u},\, \Delta^{-1}(\overline{u} + \overline{v})]\!]$.
2. $\Delta[\![\overline{u}, \overline{v}]\!] = [\![\Delta\overline{u},\, \Delta\overline{v} - \Delta\overline{u}]\!]$.
3. $T_{k+1}(f)[\![\overline{u}, \overline{v}]\!] = [\![T_k(f)\overline{u},\, T_k(f)(\overline{u} + \overline{v}) - T_k(f)\overline{u}]\!]$.

Lema 3.

$$\Delta^n_{(u_1+v_1),\,\ldots,\,(u_n+v_n)} f(x+w) - \Delta^n_{u_1,\,\ldots,\,u_n} f(x) =$$
$$\Delta^{n+1}_{w,u_1,\,\ldots,\,u_n} f(x) + \sum_{i=1}^{n} \Delta^n_{u_1,\,\ldots,\,u_{i-1},\,v_i,\,(u_{i+1}+v_{i+1}),\,\ldots,\,(u_n+v_n)} f(x + w + u_i)$$

Proof. Follows by an iterated application of (7). $\square$

We call *maximum order* of a partition $\xi = \{\alpha^1, \ldots, \alpha^n\} \in \mathcal{P}_\alpha$ to the number $\operatorname{maxord}(\xi) = \max\{|\alpha^1|, \ldots, |\alpha^n|\}$.

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

1. *the sets $\mathcal{A}_0^\xi, \mathcal{A}_{\alpha^1}^\xi, \ldots, \mathcal{A}_{\alpha^r}^\xi$ are pairwise disjoint;*
2. *$\beta \in \mathcal{A}_\beta^\xi \subset [\alpha]$, for $\beta = 0, \alpha^1, \ldots, \alpha^r$;*
3. *$0 < \gamma < \alpha$ and $|\gamma| < maxord(\xi)$, for every $\gamma \in \mathcal{A}_0^\xi - \{0\}$;*
4. *$\alpha^i < \gamma < \alpha$ and $|\gamma| \leq maxord(\xi)$, for every $\gamma \in \mathcal{A}_{\alpha^i}^\xi - \{\alpha^i\}$; and*

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

$$T_\alpha f\, \overline{u} = \sum_{\xi=\{\alpha^1,\ldots,\alpha^r\}\in\mathcal{P}_\alpha} \Delta^r_{u^\xi_{\alpha^1},\ldots,u^\xi_{\alpha^r}} f(u_0^\xi)\ , \tag{9}$$

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

Proof. This proposition is proved by induction in 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 induction hypothesis, we now restrict our attention to the second case. Given $\overline{w} \in T^{k+1}(X)$, write it as a pair $\overline{w} = [\![\overline{u}, \overline{v}]\!]$, of k-cuboids $\overline{u}$, $\overline{v} \in T^k(X)$. Using lemma 3 we deduce

$$\begin{aligned}
T_{\alpha\Diamond 1}(f)(\overline{w}) &= T_{\alpha\Diamond 1}(f)[\![\overline{u}, \overline{v}]\!] = T_\alpha(f)(\overline{u} + \overline{v}) - T_\alpha(f)\overline{u} \\
&= \sum_{\xi=\{\alpha^1,\ldots,\alpha^n\}\in\mathcal{P}_\alpha} \Delta^n_{u^\xi_{\alpha^1}+v^\xi_{\alpha^1},\ldots,u^\xi_{\alpha^n}+v^\xi_{\alpha^n}} f(u_0^\xi + v_0^\xi) \ - \ \Delta^n_{u^\xi_{\alpha^1},\ldots,u^\xi_{\alpha^n}} f(u_0^\xi) \\
&= \sum_{\xi=\{\alpha^1,\ldots,\alpha^n\}\in\mathcal{P}_\alpha} \Delta^{n+1}_{v_0^\xi, u^\xi_{\alpha^1},\ldots,u^\xi_{\alpha^n}} f(u_0^\xi) + \\
&\quad + \sum_{i=1}^n \Delta^n_{u^\xi_{\alpha^1},\ldots,u^\xi_{\alpha^{i-1}}, v^\xi_{\alpha^i}, (u^\xi_{\alpha^{i+1}}+v^\xi_{\alpha^{i+1}}),\ldots,(u^\xi_{\alpha^n}+v^\xi_{\alpha^n})} f(u_0^\xi + v_0^\xi + u^\xi_{\alpha^i})
\end{aligned}$$

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 1, 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:

$$\sum_{\xi=\{\alpha^1,\ldots,\alpha^n\}\in\mathcal{P}_\alpha} \left(\Delta^{n+1}_{w^{\tilde{\xi}_0}_{0\Diamond 1}, w^{\tilde{\xi}_0}_{\alpha^1\Diamond 0}, \ldots, w^{\tilde{\xi}_0}_{\alpha^n\Diamond 0}} f(w^{\tilde{\xi}_0}_{0\Diamond 0}) \; + \; \sum_{i=1}^{n} \Delta^{n}_{w^{\tilde{\xi}_{\alpha^i}}_{\alpha^1\Diamond 0}, \ldots, w^{\tilde{\xi}_{\alpha^i}}_{\alpha^{i-1}\Diamond 0}, w^{\tilde{\xi}_{\alpha^i}}_{\alpha^i\Diamond 1} \ldots, w^{\tilde{\xi}_{\alpha^i}}_{\alpha^n\Diamond 1}} f(w^{\tilde{\xi}_{\alpha^i}}_{0\Diamond 0}) \right)$$

which, by lemma 1, is equal to

$$\sum_{\tilde{\xi}=\{\beta^1,\ldots,\beta^r\}\in\mathcal{P}_{\alpha\Diamond 1}} \Delta^{r}_{w^{\tilde{\xi}}_{\beta^1}, \ldots, w^{\tilde{\xi}}_{\beta^r}} f(w^{\tilde{\xi}}_0) \; .$$

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

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

and

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

where $\mathcal{A} \diamond i = \{ \beta \diamond i \; : \; \beta \in \mathcal{A} \}$ for $i = 0, 1$. These recursive relations ensure that $T_{\alpha\Diamond 1}(f)(\overline{w})$ has the correct development (9). It is now easy to check inductively that these sets satisfy the conditions 1., 2., 3. and 4. $\square$

Theorem A in the introduction is a corollary of theorem 1. Its proof is completly analogous to that of proposition 3 and, therefore, will be omitted.

An algorithm that produces explicit expressions for $T_\alpha f\overline{u}$ has been devised, which was also used to confirm the correction of the above recursive definitions. Its source code can be retrieved from http://ptmat.ptmat.fc.ul.pt/

~pduarte/Research/FiniteDifferences/index.html. The following formulas were computer generated by this package.

$$\begin{aligned} T_{(1,1)}f(\overline{u}) = \Delta_{u_{1,2}} f\left(u_0 + u_2 + u_1\right) + \\ \Delta^2_{u_1,u_2} f\left(u_0\right) \end{aligned}$$

$$\begin{aligned} T_{(1,1,1)}f(\overline{u}) = \Delta_{u_{1,2,3}} f\left(u_0 + u_3 + u_2 + u_{2,3} + u_1 + u_{1,3} + u_{1,2}\right) + \\ \Delta^2_{u_1,u_{2,3}} f\left(u_0 + u_3 + u_2\right) + \\ \Delta^2_{u_{1,3},u_2+u_{2,3}} f\left(u_0 + u_3 + u_1\right) + \\ \Delta^2_{u_{1,2},u_3+u_{2,3}+u_{1,3}} f\left(u_0 + u_2 + u_1\right) + \\ \Delta^3_{u_1,u_2,u_3} f\left(u_0\right) \end{aligned}$$

$$\begin{aligned} \Delta^2_v(f \circ g)(x) = \Delta_{\Delta^2_{v_1,v_2} g(x)} f\left(g(x) + \Delta_{v_1} g\left(x\right) + \Delta_{v_2} g\left(x\right)\right) + \\ \Delta^2_{\Delta_{v_1} g(x),\Delta_{v_2} g(x)} f\left(g(x)\right) \end{aligned}$$

$$\begin{aligned} \Delta^3_v(f \circ g)(x) = \Delta_{\Delta^3_{v_1,v_2,v_3} g(x)} f\left(g(x) + \Delta_{v_1} g\left(x\right) + \Delta_{v_2} g\left(x\right) + \Delta_{v_3} g\left(x\right) \ + \right. \\ \left. \Delta^2_{v_1,v_2} g\left(x\right) + \Delta^2_{v_1,v_3} g\left(x\right) + \Delta^2_{v_2,v_3} g\left(x\right)\right) \ + \\ \Delta^2_{\Delta_{v_1} g(x),\Delta^2_{v_2,v_3} g(x)} f\left(g(x) + \Delta_{v_2} g\left(x\right) + \Delta_{v_3} g\left(x\right)\right) \ + \\ \Delta^2_{\Delta^2_{v_1,v_3} g(x),\Delta_{v_2} g(x)+\Delta^2_{v_2,v_3} g(x)} f\left(g(x) + \Delta_{v_1} g\left(x\right) + \Delta_{v_3} g\left(x\right)\right) \ + \\ \Delta^2_{\Delta^2_{v_1,v_2} g(x),\Delta_{v_3} g(x)+\Delta^2_{v_1,v_3} g(x)+\Delta^2_{v_2,v_3} g(x)} f\left(g(x) + \Delta_{v_1} g\left(x\right) + \Delta_{v_2} g\left(x\right)\right) \ + \\ \Delta^3_{\Delta_{v_1} g(x),\Delta_{v_2} g(x),\Delta_{v_3} g(x)} f\left(g(x)\right) \end{aligned}$$

References

[1] G. Boole, A treatise on the calculus of finite differences, Cambridge, 1860.

[2] C. F. Faà di Bruno, Note sur une nouvelle formule du calcul differentiel, Quart. J. Math., 1 (1857) 359-360.

[3] M. Floater, T. Lyche, Two chain rules for divided differences and Faà di Bruno's formula, Math. Comp. 76 , no. 258 (2007) 867-877.

[4] C. Jordan, Calculus of Finite Differences, 3rd ed., Chelsea Publishing Co., Inc., New York, 1965.