

Faà di Bruno is Taylor Composition

A principled approach to the multi-variate Faà di Bruno formula in various incarnations

Heinrich Hartmann* 

doi:10.5281/zenodo.18203201

Abstract

We give an elementary, self-contained proof of the multi-variate Faa di Bruno formula as a composition theorem for Taylor Polynomials in the setting of real normed vector spaces: if $\phi : E \rightarrow F$ and $\psi : F \rightarrow G$ are k -fold Fréchet differentiable at the points $x \in E$, $y = \phi(x) \in F$ then the Taylor Polynomials compose in the natural way:

$$T_x^k(\psi \circ \phi) = \pi_{\leq k}(T_y^k(\psi) \circ T_x^k(\phi)).$$

From this composition principle we derive (i) the partition formula for higher Fréchet derivatives and (ii) the multi-index coefficient formula established by Constantine–Savits.

While many recent papers frame Faà di Bruno via Bell polynomials, trees, Hopf algebras, jets, or coordinate combinatorics, our approach isolates the analytic core—functoriality of pointwise Taylor approximation—under minimal assumptions. We avoid making finite-dimensionality or completeness assumptions and work with a weak notion of pointwise Fréchet differentiability. The familiar combinatorial coefficients arise mechanically from two standard operations: polarization and coefficient extraction from composed Taylor polynomials. The result is a short, self-contained account that covers all relevant versions of the Formula in a very general setting.

As an application we present a general higher-order Leibniz rule in both partition and multi-index form.

Outline

1	Introduction	1
2	Related Work	2
3	Polynomials	3
4	Fréchet Differentiability	5
5	Faa di Bruno - Composition Form	6
6	Faa di Bruno - Partition Form	7
7	Faa di Bruno - Multi-index Form	8
8	Leibniz Formula	10

1 Introduction

Faà di Bruno–type formulas describe how higher derivatives behave under composition. In the classical one-variable setting, they express $(f \circ g)^{(n)}$ as a polynomial in the derivatives of f and g with explicitly computable combinatorial coefficients (often packaged by Bell polynomials). In 1855 Faà di Bruno derived the identity (for $f, g : \mathbb{R} \rightarrow \mathbb{R}$)

$$\frac{d^n}{dx^n}(f \circ g)(x) = \sum_{k=1}^n f^{(k)}(g(x)) \sum_{(*)} \frac{n!}{m_1! \cdots m_n!} \prod_{j=1}^n \left(\frac{g^{(j)}(x)}{j!} \right)^{m_j}.$$

*Hartmann IT GmbH / HeinrichHartmann.com / Heinrich@HeinrichHartmann.com

Here the sum $(*)$ runs over all tuples $m_1, \dots, m_n \in \mathbb{N}_0$ constrained by the conditions $m_1 + \dots + m_n = k$ and $m_1 + 2m_2 + \dots + nm_n = n$.

This formula admits several equivalent repackagings in terms of Bell polynomials, partition sums, umbral expressions — each highlighting different facets of the combinatorial structure.

In several variables the underlying pattern remains the same, but the bookkeeping becomes substantially heavier: one must track multi-index exponents, manage products over vector-valued derivatives, and account for the richer symmetry of set partitions with block multiplicities. A complete explicit multivariate formula with fully worked-out combinatorial coefficients was only established by Constantine and Savits in 1996 - 141 years after Faà di Bruno’s original work.

This note takes a deliberately “analytic-first” viewpoint. We work with maps between real normed vector spaces and isolate the following question as the core of Faà di Bruno:

If ϕ is approximated near x_0 by a polynomial $T_{x_0}^k \phi$ of degree $\leq k$, and ψ is approximated near $y_0 = \phi(x_0)$ by a polynomial $T_{y_0}^k \psi$ of degree $\leq k$, what polynomial approximates $\psi \circ \phi$ near x_0 ?

There is one natural candidate: the polynomial composition $T_{y_0}^k \psi \circ T_{x_0}^k \phi$, truncated to degree $\leq k$. Our first main result confirms that this is correct under minimal hypotheses: pointwise k -fold Fréchet differentiability at a single point (no global smoothness assumptions, and no completeness or finite-dimensionality requirements). Concretely, we show that if ϕ and ψ admit pointwise Taylor decompositions of order k at x_0 and y_0 respectively, then $\psi \circ \phi$ does as well, with

$$T_{x_0}^k(\psi \circ \phi) = \pi_{\leq k}(T_{y_0}^k \psi \circ T_{x_0}^k \phi).$$

Once this “composition of Taylor polynomials” principle is established, the familiar combinatorial formulas fall out systematically:

- the **partition form** is obtained by polarization (passing from homogeneous polynomials to symmetric multilinear maps), yielding a formula for $D^k(\psi \circ \phi)(0)$ as a sum over set partitions of $[k]$;
- the **multi-index form** is obtained by coefficient extraction in coordinates: expand the Taylor polynomials in a monomial basis and compute the coefficient of x^ν in $T_{y_0}^k \psi(T_{x_0}^k \phi(x))$. This produces an explicit coefficient identity that recovers the multi-index Faà di Bruno formula (in particular the structure underlying the Constantine–Savits coefficients) as a corollary of a purely algebraic multinomial bookkeeping lemma.

As an application we include a higher-order Leibniz rule in two parallel forms: a partition identity for Fréchet derivatives and a multi-index identity for partial derivatives. The proof is deliberately routed through the Faà di Bruno formula in both incarnations.

2 Related Work

There is a large literature around Faà di Bruno formulas and their multivariate generalizations; we briefly summarize the most relevant references and how they compare to the present treatment.

More directly related to this paper:

- Faà di Bruno (1855, F. Faà di Bruno. “Sullo sviluppo delle funzioni”) [dB55] [dB57]. The original one-dimensional formulas introducing the identity that now bears his name, in Italian (1855) and French (1857). These short notes are the primary historical sources.
- Scott (1861, “Formulae of Successive Differentiation”) [Sco61] [Joh02]. In the one-variable case $g \circ f : \mathbb{R} \rightarrow \mathbb{R}$, Scott’s identity (as recounted by Johnson) gives a 1D Taylor-composition viewpoint via coefficient extraction: apply Taylor’s theorem to g in the increment $f(t+h) - f(t)$ and read off the coefficient of h^m .
- Constantine–Savits (1996, “A Multivariate Faà di Bruno Formula with Applications”) [CS96]. Foundational multivariate Faà di Bruno formula for arbitrary partial derivatives of a composition $h(x) =$

$f(g(x))$ with $f : \mathbb{R}^m \rightarrow \mathbb{R}$ and $g : \mathbb{R}^d \rightarrow \mathbb{R}^m$, expressed in full multiindex form using set partitions and multivariate Stirling numbers. Their proof is combinatorial and coordinate-based. The multiindex form stated here recovers their coefficient formula but is derived from a short argument based on jets and polarized Fréchet derivatives.

- Hernández Encinas–Muñoz Masqué (2003, “A short proof of the generalized Faà di Bruno’s formula”) [EM03]. Derive the multivariate Faà di Bruno formula by working at the level of r -jets (higher cotangent spaces) and using induced algebra morphisms on truncated local rings; in coordinates this amounts to composing truncated multivariate power series and applying the multinomial theorem. This is close in spirit to our “Taylor polynomials compose” viewpoint, but their formulation stays in the finite-dimensional multi-index setting and is phrased in the language of higher cotangent spaces rather than as an explicit functoriality statement for pointwise Taylor polynomials between general normed spaces.
- Levy (2006, “Why do partitions occur in Faà di Bruno’s chain rule for higher derivatives?”) [Lev06]. Gives what is, to our knowledge, the only explicit partition-form Faà di Bruno formula for n -th Fréchet derivatives between Banach spaces, and explains why partitions appear. The first part uses an algebraic description in terms of coalgebras on germs; the second part proves the Banach-space partition formula via iterated finite differences on parallelepipeds. The partition form proved in this paper agrees with Levy’s formula but is obtained from a very short jet-based computation.
- Duarte–Torres (2008, “A discrete Faà di Bruno’s formula”) [DT08]. Develop a discrete (finite-difference) Faà di Bruno formula for maps between linear spaces, valid for general functions of several variables, by building an algebra of symbolic finite-difference expressions and expanding $\Delta^\alpha(f \circ g)$ in terms of finite differences of f and g . Their analysis also yields a streamlined proof of the smooth Faà di Bruno formula when one passes from finite differences to derivatives; philosophically this is close to viewing higher derivatives as jets, but their primary emphasis is on the discrete setting.

3 Polynomials

We present a brief overview of the basic theory of Polynomials in an infinite dimensional setting of normed vector spaces following [Din99]. The reader who is only interested in the finite dimensional setting can safely skip this section.

(1) Definition (Polynomials [Din99]). Let E, F be real normed vector spaces.

- Write $\mathcal{L}(E, F)$ for the vector space of bounded linear maps $A : E \rightarrow F$, with operator norm:

$$\|A\| = \sup_{v \in E, \|v\| \leq 1} \|A(v)\|$$

- For $k \geq 0$ let $\mathcal{L}^{(k)}(E, F)$ be the space of bounded k -fold multi-linear maps $A_k : E^{\times k} \rightarrow F$, with operator norm:

$$\|A_k\| = \sup_{\substack{v_1, \dots, v_k \in E \\ \|v_i\| \leq 1}} \|A_k(v_1, \dots, v_k)\|.$$

- Let $\mathcal{L}^s({}^k E, F) \subset \mathcal{L}^{(k)}(E, F)$ be the subspace of symmetric multi-linear maps.
- A map $p : E \rightarrow F$ is called a k -homogeneous polynomial if there exists $A \in \mathcal{L}^s({}^k E, F)$ such that

$$p(x) = \frac{1}{k!} A[x, \dots, x] = \frac{1}{k!} A[x^{\otimes k}], \quad x \in E.$$

We write $\mathcal{P}^k(E, F) \subset \text{Map}(E, F)$ for the space of such maps.

- A polynomial map $p : E \rightarrow F$ is a finite sum of homogeneous polynomials:

$$p = \sum_{k=0}^n p_k \quad \text{with} \quad p_k \in \mathcal{P}^k(E, F).$$

We write $\mathcal{P}(E, F) := \sum_{k \geq 0} \mathcal{P}^k(E, F) \subset \text{Map}(E, F)$ for the space of all polynomial maps, and $\mathcal{P}_{\leq n}(E, F) := \sum_{k=0}^n \mathcal{P}^k(E, F)$ subspace of polynomial maps of degree $\leq n$.

- For $F = \mathbb{R}$ we write $\mathcal{P}(E) := \mathcal{P}(E, \mathbb{R})$.

(2) Lemma (Polarization). *Let $p : E \rightarrow F$ be a polynomial map, represented as a sum of multi-linear forms $p(x) = \sum_{k=0}^n p_k$, $p_k = \frac{1}{k!} A_k[x^{\otimes k}]$ with $A_k \in \mathcal{L}^s(kE, F)$.*

- *The top-degree multi-linear form $A_n \in \mathcal{L}^s(nE, F)$ can be recovered from $p : E \rightarrow F$ via:*

$$A_n[v_1, \dots, v_n] = \sum_{I \subset [n]} (-1)^{n-|I|} p\left(\sum_{i \in I} v_i\right)$$

for $v_i \in E$.

- *For a polynomial map $p : E \rightarrow F$, $p \neq 0$ the representation*

$$p(x) = \sum_{k=0}^n \frac{1}{k!} A_k[x^{\otimes k}]$$

in terms of $n \in \mathbb{N}_0$, $A_k \in \mathcal{L}^s(kE, F)$ is unique.

Proof. The polarization formula is Möbius inversion on the Boolean lattice $2^{[n]}$: By multilinearity,

$$p\left(\sum_{i \in I} v_i\right) = \sum_{k=0}^n \sum_{i_1, \dots, i_k \in I} \frac{1}{k!} A_k[v_{i_1}, \dots, v_{i_k}] = \sum_{J \subset I} A_{|J|}[v_J]$$

where $A_k[v_J]$ denotes evaluation on the $|J|$ -tuple indexed by J . The alternating sum $\sum_{J \subset [n]} (-1)^{n-|J|}$ inverts this: all terms with $J \subsetneq [n]$ cancel by the identity $\sum_{J \supsetneq I} (-1)^{n-|J|} = (1-1)^{n-|I|} = 0$, leaving only $J = [n]$. Uniqueness follows by induction: recover A_n via polarization, subtract, and repeat. \square

(3) Definition (Degree Truncation). Let $p : E \rightarrow F$ be a polynomial map, $p \neq 0$. By Lemma (2) the decomposition $p(x) = \sum_{k=0}^n \frac{1}{k!} A_k[x^{\otimes k}]$ is unique. We can hence define:

- $\deg(p) := n \in \mathbb{N}_0$ the degree of p . By convention $\deg(0) = -\infty$.
- $\pi_k(p) := p_k : E \rightarrow F$ the degree- k part of p .
- $\text{Pol}_k(p) := A_k \in \mathcal{L}^s(kE, F)$ the degree- k polarization of p .
- $\pi_{\leq k}(p) \in \mathcal{P}_{\leq k}(E, F)$ the (lower) degree truncation of p .
- $\pi_{> k}(p) \in \mathcal{P}_{> k}(E, F)$ the upper degree truncation of p .

We will make use of the following properties of polynomial maps:

(4) Lemma (Polynomial Properties). *Let $p : E \rightarrow F$, $q : F \rightarrow G$ be polynomial maps.*

1. *(Composition) The composition $q \circ p : E \rightarrow G$ is a polynomial with $\deg(q \circ p) \leq \deg(q) + \deg(p)$.*
2. *(Lipschitz) There exists $\varepsilon > 0, C > 0$ such that $\|\Delta_0 p(x)\| \leq C\|x\|$ for all $\|x\| < \varepsilon$.*
3. *(Upper Vanishing) We have $\|\pi_{> k}(p)(x)\|/\|x\|^k \rightarrow 0$ for $x \rightarrow 0$.*
4. *(Lower Vanishing) If $\|p(x)\|/\|x\|^k \rightarrow 0$ for $x \rightarrow 0$, and $k \geq \deg(p)$ then $p = 0$.*

Proof. Ad 1) Expanding $q(p(x))$ by multilinearity yields a sum of terms $B[A_1[x^{\otimes j_1}], \dots, A_\ell[x^{\otimes j_\ell}]]$ where $B \in \mathcal{L}(\ell F, G)$ and $A_i \in \mathcal{L}^s(j_i E, F)$ are bounded multilinear.

The composition $(u_1, \dots, u_k) \mapsto B[A_1[u_{I_1}], \dots, A_\ell[u_{I_\ell}]]$ is bounded multilinear, hence each term is a homogeneous polynomial of degree $j_1 + \dots + j_\ell \leq \ell \cdot \deg(p)$. Thus $q \circ p$ is a polynomial with $\deg(q \circ p) \leq \deg(q) + \deg(p)$.

Ad 2) Write $\Delta_0 p(x) = p(x) - p(0) = \sum_{j=1}^m \frac{1}{j!} A_j[x^{\otimes j}]$ with $A_j \in \mathcal{L}^s(jE, F)$ bounded. For $\|x\| < 1$ we have $\|x\|^j \leq \|x\|$, hence $\|\Delta_0 p(x)\| \leq \sum_{j=1}^m \frac{\|A_j\|}{j!} \|x\|^j \leq C\|x\|$ with $C = \sum_{j=1}^m \frac{\|A_j\|}{j!}$.

Ad 3) Each homogeneous component p_j of degree $j > k$ satisfies $\|p_j(x)\| \leq C_j \|x\|^j$, hence $\|p_j(x)\|/\|x\|^k = C_j \|x\|^{j-k} \rightarrow 0$ as $x \rightarrow 0$.

Ad 4) Assume $p \neq 0$, $n = \deg(p)$. Write $p = \sum_{j=m}^n p_j$ with $p_j = \pi_j(p)$ homogeneous, $p_m \neq 0$ the minimal non-vanishing degree.

Let $m = \min\{j : p_j \neq 0\} \leq \deg(p) \leq k$ be the bottom term, then $p = p_m + \pi_{>d}p$, and by (3) $\lim_{x \rightarrow 0} \|p(x)\|/\|x\|^k = \|p_d(x)\|/\|x\|^k$

If $p \neq 0$, let $j_0 := \min\{j : p_j \neq 0\} \leq \deg(p) \leq k$ and choose $v \in E$ with $p_{j_0}(v) \neq 0$. For $t \rightarrow 0^+$: $p(tv)/t^k = t^{j_0-k}p_{j_0}(v) + \pi_{>j_0}p(tv)$. Since $j_0 \leq k$, the leading term $t^{j_0-k}p_{j_0}(v)$ does not vanish as $t \rightarrow 0^+$, contradicting $\|p(x)\|/\|x\|^k \rightarrow 0$.

□

4 Fréchet Differentiability

We summarize the classical notion of Fréchet differentiability (following [Lan93]), and introduce a weaker point-wise notion of Fréchet differentiability.

(5) Definition (Fréchet differentiability [Lan93] XIII.6). Let E, F be normed vector spaces, $U_E \subset E$ an open subset and $\phi : U_E \rightarrow F$ a map.

- (Forward difference) For $x_0 \in U_E$ let

$$\Delta_{x_0}\phi(h) := \phi(x_0 + h) - \phi(x_0)$$

the forward difference, which is defined for $h \in E$ in a neighbourhood of 0.

- (Fréchet differentiable) We say that ϕ is Fréchet differentiable at a point $x_0 \in U_E$ if there exists a bounded linear map $L \in \mathcal{L}(E, F)$ such that

$$\lim_{h \rightarrow 0} \frac{\|\Delta_{x_0}\phi(h) - L(h)\|}{\|h\|} = 0.$$

The map L is uniquely determined and is called the derivative of ϕ at x_0 , written $D\phi(x_0)$.

- (Continuous Fréchet differentiable) We say that ϕ is continuously Fréchet differentiable on U_E if ϕ is Fréchet differentiable at every point $x \in U_E$ and the map: $x \mapsto D\phi(x) \in \mathcal{L}(E, F)$ is continuous.
- (Continuous k-times Fréchet differentiability) We recursively define ϕ to be k-times continuously Fréchet differentiable. For $k \geq 1$ we require $\phi : U_E \rightarrow F$ to be continuously Fréchet differentiable on U_E , and $D\phi : U_E \rightarrow \mathcal{L}(E, F)$ k-1-times continuously Fréchet differentiable as a map between normed vector spaces.
- We write $C^k(U_E, F)$ for the space of k-times continuously Fréchet differentiable functions.

(6) Proposition (Taylor Approximation [Lan93], XIII.6). Let $\phi : U_E \rightarrow F$ be C^k as above, and $x_0 \in U_E$.

1. The higher Fréchet differentials define symmetric multi-linear forms:

$$D^k\phi(x_0) = D(\dots(D\phi))(x_0) \in \mathcal{L}(E, \dots, \mathcal{L}(E, \mathcal{L}(E, F))) \cong \mathcal{L}^s({}^kE, F).$$

2. The associated degree-k polynomial:

$$T_{x_0}^k\phi(h) = \sum_{\ell=1}^k \frac{1}{\ell!} D^\ell\phi(x_0)[h^{\otimes \ell}] : E \rightarrow F$$

is called the degree-k Taylor polynomial of ϕ at x_0 . Note that we start this sum at $\ell = 1$ so that $T_{x_0}^k\phi(0) = 0$, by convention.

3. The Taylor Remainder $R_{x_0}^k\phi = \Delta_{x_0}\phi - T_{x_0}^k\phi$ satisfies:

$$\Delta_{x_0}\phi = T_{x_0}^k\phi + R_{x_0}^k\phi \quad \text{and} \quad \|R_{x_0}^k\phi(h)\|/\|h\|^k \rightarrow 0 \quad \text{as} \quad h \rightarrow 0.$$

We call this sum the *Taylor Decomposition* of ϕ at x_0 .

4. Let $\Delta_{x_0}\phi = T + R$ be any other decomposition with: $T \in \mathcal{P}_{\leq k}(E, F)$, $T(0) = 0$ a degree $\leq k$ polynomial and $R : E \rightarrow F$ with $\|R(h)\|/\|h\|^k \rightarrow 0$ for $h \rightarrow 0$. Then $T = T_{x_0}^k \phi$ and $R = R_{x_0}^k \phi$.

Proof. Properties 1-3 are standard and can be found, e.g. in [Lan93], XIII.6. For the uniqueness assume that $\Delta_{x_0}\phi = T + R = T' + R'$ are two Taylor decompositions of ϕ . Then $p = T - T' = R - R'$ is a polynomial of degree $\leq k$ with $\|p(h)\|/\|h\|^k \rightarrow 0$ for $h \rightarrow 0$. By Lemma (4) we must have $p = 0$. □

In the light of this Proposition, we make the following definition of point-wise Fréchet differentiability as approximation property for functions by polynomial maps:

(7) Definition (Pointwise Fréchet decomposition at 0).

- We say that a map $\phi : U_E \rightarrow F$ is k -fold *pointwise Fréchet differentiable* at x_0 if there exists a “Taylor” decomposition:

$$\Delta_{x_0}\phi = T + R$$

where a polynomial map $T \in \mathcal{P}_{\leq k}(E, F)$ with $T(0) = 0$ and a residual map $R : U_E \rightarrow F$ such that

$$\|R(x)\| / \|x\|^k \rightarrow 0 \quad \text{as } x \rightarrow 0.$$

By (6) this decomposition is unique if it exists, and we write: $T_{x_0}^k \phi = T$ and $R_{x_0}^k \phi = R$.

- We define the degree $1 \leq \ell \leq k$ Fréchet differentials as polarization of the Taylor polynomial via:

$$D^\ell \phi(x_0) = \text{Pol}_\ell T_{x_0}^k \phi \in \mathcal{L}^s(\ell E, F)$$

And set $D^0 \phi(x_0) = \phi(x_0) \in \mathcal{L}^s(0 E, F) = F$.

(8) Lemma (Lipschitz Continuity). *If $\phi : E \rightarrow F$ is pointwise Fréchet differentiable at x , then ϕ is Lipschitz continuous at x , i.e. there is $C > 0, \varepsilon > 0$ so that: $\|\Delta_{x_0}\phi(h)\| < C\|h\|$ for all $\|h\| < \varepsilon$.*

Proof. We may assume $x_0 = 0$. Let $\Delta_0\phi = P + R$ be a Taylor decomposition of order k at 0. By Lemma (4) there exist $\varepsilon > 0$ and $C_1 > 0$ such that $\|P(h)\| \leq C_1\|h\|$ for $\|h\| < \varepsilon$. Since $R(h) = o(\|h\|^k)$ we have $\|R(h)\| \leq C_2\|h\|^k \leq C_2\|h\|$ for $\|h\| < \min(\varepsilon, 1)$. Thus $\|\Delta_0\phi(h)\| \leq \|P(h)\| + \|R(h)\| \leq (C_1 + C_2)\|h\|$ for $\|h\|$ small. □

5 Faa di Bruno - Composition Form

(9) Theorem (Multivariate Faà di Bruno - Composition Form). *Let E, F, G be normed vector spaces. Let $U_E \subset E$ be open, and $x_0 \in U_E$ and $\phi : E \rightarrow F$ be k -fold pointwise Fréchet differentiable at x_0 . Let $U_F \subset F$ open with $y_0 = \phi(x_0) \in U_F$ and $\psi : U_F \rightarrow G$ be k -fold pointwise Fréchet differentiable at y_0 .*

Then:

- The composition $\psi \circ \phi : U_E \rightarrow G$ is pointwise k -fold Fréchet differentiable at x_0 .
- The Taylor Polynomials compose as polynomial maps in $\mathcal{P}(E, G)$:

$$T_{x_0}^k(\psi \circ \phi) = \pi_{\leq k}(T_{y_0}^k \psi \circ T_{x_0}^k \phi).$$

Proof. By translation, we may assume $x_0 = 0$ and $\phi(x_0) = y_0 = 0$ as well as $\psi(y_0) = 0$ throughout the proof. In particular $\Delta_{x_0}\phi = \phi$, $\Delta_{y_0}\psi = \psi$.

Choose Taylor decompositions $\phi = P + R$ and $\psi = Q + S$, with polynomials P, Q of degree $\leq k$ and remainders R, S vanishing to order k at 0. Now write $\psi \circ \phi = Q(\phi) + S(\phi) = \pi_{\leq k}(Q \circ P) + E$, with:

$$E := Q(\phi) + S(\phi) - Q(P) + \pi_{>k}(Q \circ P).$$

We have to show that $\|E(x)\|/\|x\|^k \rightarrow 0$ for $x \rightarrow 0$.

Clearly $\|\pi_{>k}(Q \circ P)\|/\|x\|^k \rightarrow 0$ for $x \rightarrow 0$ by Lemma (4).

To see that $\|S(\phi(x))\|/\|x\|^k \rightarrow 0$ we argue as follows: By the Lipschitz property of ϕ (Lemma (8)), we find $\delta > 0, C > 0$ so that $\|\phi(x)\| < C\|x\|$ for all $\|x\| < \delta$. Since S is a Taylor residual we have $\|S(y)\| < \eta(y)\|y\|^k$ with $\eta(y) \rightarrow 0$ as $y \rightarrow 0$. Hence $\|S(\phi(x))\|/\|x\|^k \leq \eta(\phi(x))C^k \rightarrow 0$ as $x \rightarrow 0$.

For the term $Q(P+R) - Q(P)$, we decompose $Q(y) = \sum_{\ell=0}^k Q_\ell(y) = \sum_{\ell} \frac{1}{\ell!} B_\ell[y^{\otimes \ell}]$ where $Q_\ell = \pi_\ell Q$ and $B_\ell = \text{Pol}_\ell Q$. Then

$$Q_\ell(P+R) - Q_\ell(P) = \sum_{\substack{i+j=\ell \\ j \geq 1}} \frac{1}{i!j!} B_\ell(P^{\otimes i}, R^{\otimes j}).$$

It suffices to show that each summand $B_\ell(P^{\otimes i}, R^{\otimes j})$ is in $o(\|x\|^k)$. To this end we note that:

- $\|B_\ell(P(x)^{\otimes i}, R(x)^{\otimes j})\| \leq \|B_\ell\| \cdot \|P(x)\|^i \|R(x)\|^j$ as $B_\ell \in \mathcal{L}^s(\ell F, G)$ bounded.
- $\|P(x)\| \leq C\|x\|$ for $\|x\| < \delta$ as P is Lipschitz (Lemma (4)) and $P(0) = 0$.
- $\|R(x)\| < \eta(x)\|x\|^k$ with $\eta(x) \rightarrow 0$ for $x \rightarrow 0$ as R is a Taylor remainder.

So that:

$$\|B_\ell(P^{\otimes i}(x), R^{\otimes j}(x))\|/\|x\|^k \leq \|B_\ell\| C^i \eta(x)^j \|x\|^{i+(j-1)k} \rightarrow 0 \quad \text{as } x \rightarrow 0$$

For the last step, note that $j \geq 1$ so the factor $\eta(x)^j \rightarrow 0$ as $x \rightarrow 0$, and $i + (j-1)k \geq 0$ hence $\|x\|^{i+(j-1)k}$ stays bounded (and even goes to zero for $j \geq 2$).

□

6 Faa di Bruno - Partition Form

(10) Definition. Let I be a finite set.

- An un-ordered (non-empty) partition π of I is a set $\pi = B_1, \dots, B_r$ of subsets $B_i \subset I$ ("blocks"), so that B_i are disjoint, $I = \cup_{i=1}^r B_i$ and $B_i \neq \emptyset$. We write $\pi \vdash I$ in this case and call $|\pi| = r$ the rank of π .
- An ordered partition π of I is a tuple $\pi = (B_1, \dots, B_r)$ of subsets $B_i \subset I$, so that B_i are disjoint and $I = \cup_{i=1}^r B_i$. We write $\pi \models I$ in this case.
Note that we allow empty blocks for ordered partitions, so that this datum is equivalent to a map $\sigma : I \rightarrow [r]$ with $|\pi| = r$ $B_i = \sigma^{-1}i$.
- For a set X , a tuple $v = (v_i)_{i \in I} \in X^I$ and a block $B \subset I$, we write $v_B := (v_b)_{b \in B} \in X^B$ for the restricted tuple.

(11) Lemma (Composition Polarization). *If $k = m_1 + \dots + m_r \in \mathbb{N}_0$ and $A_i \in \mathcal{L}^s(m_i E, F)$, and $B \in \mathcal{L}^s(r F, G)$. Let*

$$H(x) = B\left[\frac{1}{m_1!} A_1[x^{\otimes m_1}], \dots, \frac{1}{m_r!} A_r[x^{\otimes m_r}]\right]$$

then:

$$\text{Pol}_k(H)[u_1, \dots, u_k] = \sum_{\substack{(B_1, \dots, B_r) \models [k], \\ |B_i| = m_i}} B[A_1[u_{B_1}], \dots, A_r[u_{B_r}]].$$

Proof. Let $G[u_1, \dots, u_k]$ be the claimed expression for $\text{Pol}_k(H)[u_1, \dots, u_k]$. As both sides are symmetric, bounded k-multi-linear forms in u_1, \dots, u_k we only have to validate that $\frac{1}{k!} G[x^{\otimes k}] = H$. We find:

$$G[x^{\otimes k}] = \sum_{(B_1, \dots, B_r) \models [k], |B_i| = m_i} B[A_1[x^{\otimes m_1}], \dots, A_r[x^{\otimes m_r}]] = \frac{k!}{m_1! \dots m_r!} \cdot B[A_1[x^{\otimes m_1}], \dots, A_r[x^{\otimes m_r}]]$$

as there are $k!/m_1! \dots m_r!$ ordered partitions of a set of $[k]$ elements into subsets of order m_1, \dots, m_r . Re-arranging the factorials yields the claimed.

□

(12) Theorem (Multivariate Faà di Bruno - Partition Form). *Let $x_0 \in U_E \subset E$ and set $y_0 := \phi(x_0) \in F$. Assume that $\phi : U_E \rightarrow F$ is pointwise k -fold Fréchet differentiable at x_0 and that $\psi : F \rightarrow G$ is pointwise k -fold Fréchet differentiable at y_0 . For $u_1, \dots, u_k \in E$ we have:*

$$D^k(\psi \circ \phi)(x_0)[u_1, \dots, u_k] = \sum_{r=1}^k \sum_{\substack{\pi \vdash [k] \\ |\pi|=r}} D^r \psi(y_0)[D^{|\pi|} \phi(x_0)[u_{B_1}], \dots, D^{|\pi|} \phi(x_0)[u_{B_r}]].$$

Proof. Fix $u_1, \dots, u_k \in E$. By translation, we may again assume $x_0 = 0$ and $y_0 = 0$ throughout the proof. Fix Taylor decompositions $\phi = P + R$, $\psi = Q + S$ with P, Q polynomial of degree $\leq k$. By (9) we have $T_0^k \psi \circ \phi = \pi_{\leq k}(Q \circ P)$, and hence $D^k(\psi \circ \phi)(0) = \text{Pol}_k(Q \circ P)$.

Write the polynomial maps in polarized form as

$$P(x) = \sum_{m=1}^k \frac{1}{m!} P_m[x^{\otimes m}], \quad Q(y) = \sum_{r=1}^k \frac{1}{r!} Q_r[y^{\otimes r}].$$

Expanding Q_r by multilinearity and isolating the degree k term gives:

$$\pi_k Q(P(x)) = \sum_{r=1}^k \sum_{\substack{m_1, \dots, m_r \geq 1 \\ m_1 + \dots + m_r = k}} \frac{1}{r!} Q_r \left[\frac{1}{m_1!} P_{m_1}[x^{\otimes m_1}], \dots, \frac{1}{m_r!} P_{m_r}[x^{\otimes m_r}] \right].$$

Applying Pol_k on both sides making use of the composition formula from Lemma (11) now gives:

$$D^k(\psi \circ \phi)(0)[u_1, \dots, u_k] = \sum_{r=1}^k \sum_{\substack{m_1, \dots, m_r \geq 1 \\ m_1 + \dots + m_r = k}} \sum_{\substack{(B_1, \dots, B_r) \models [k] \\ |B_i| = m_i}} \frac{1}{r!} Q_r[P_{m_1}[u_{B_1}], \dots, P_{m_r}[u_{B_r}]].$$

Now observe that there are exactly $r!$ re-orderings of each ordered partition (B_1, \dots, B_r) , and these re-orderings yield the same summand as Q_r is symmetric. Hence the sum over $m_i, B \models [k]$ can be replaced by a single sum over un-ordered partitions $\pi \vdash [k]$ with given block sizes $|B_i| = m_i$. Furthermore we have $Q_r = D^r \psi(0)$, $P_m = D^m \phi(0)$ by definition. Substituting these into the above formulas yields the claim. □

7 Faa di Bruno - Multi-index Form

(13) Definition (Multi-Indices).

- Elements $\nu = (\nu_1, \dots, \nu_d) \in \mathbb{N}_0^d$ are called multi-indices of dimension d . We write $|\nu| = \sum_i \nu_i$ for the degree and $\nu! := \nu_1! \cdots \nu_d!$ for the factorial.
- If A is a commutative algebra and $v \in A^d$ is a d -tuple, then we write: $v^\nu = \prod_i v_i^{\nu_i} \in A$.
- If $E = \mathbb{R}^d$, write $x^1, \dots, x^d : E \rightarrow \mathbb{R}$ for the coordinate projections, then the set $x^\nu, \nu \in \mathbb{N}_0^d$ is a basis of Polynomial functions $\mathcal{P}(E)$.
- If $U \subset \mathbb{R}^d$ is open and $\phi : U \rightarrow F$ is pointwise k -fold Fréchet differentiable at $x_0 \in U$. Let $\nu \in \mathbb{N}_0^d$ with $k = |\nu|$. We define the partial derivative $\partial^\nu \phi(x_0) \in F$ via the basis expansion of $T_{x_0}^k \phi \in \mathcal{P}(\mathbb{R}^d, F) = \mathcal{P}(\mathbb{R}^d) \otimes F$ as coefficient of $x^\nu / \nu!$:

$$T_{x_0}^k \phi = \sum_{1 \leq |\nu| \leq k} \frac{x^\nu}{\nu!} \partial^\nu \phi(x_0).$$

(14) Theorem (Multivariate Faà di Bruno - Multi-index form). *Let $x_0 \in U \subset \mathbb{R}^d$ and set $y_0 := \phi(x_0) \in \mathbb{R}^n$. Assume that $\phi : U \rightarrow \mathbb{R}^n$ is pointwise k -fold Fréchet differentiable at x_0 and that $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is pointwise k -fold Fréchet differentiable at y_0 .*

For $\alpha \in \mathbb{N}_0^d$, $\alpha \neq 0$ the partial derivative of the composition are given by:

$$\frac{\partial^\alpha (f \circ \phi)(x_0)}{\alpha!} = \sum_{\substack{\beta \in \mathbb{N}_0^n \\ 1 \leq |\beta| \leq |\alpha|}} \partial^\beta f(y_0) \sum_{s \geq 1} \sum_{(*)} \prod_{r=1}^s \frac{1}{\beta_r!} \frac{(\partial^{\alpha_r} \phi(x_0))^{\beta_r}}{\alpha_r!^{|\beta_r|}},$$

Where $(*)$ runs over all $(\alpha_1, \dots, \alpha_s; \beta_1, \dots, \beta_s)$ with $\alpha_j \in \mathbb{N}_0^d$, $\beta_j \in \mathbb{N}_0^n$, $\alpha_j \neq 0$, $\beta_j \neq 0$ and $\alpha_1 < \dots < \alpha_s$ for a total order on \mathbb{N}_0^d satisfying the constraints: $\alpha_1 |\beta_1| + \dots + \alpha_s |\beta_s| = \alpha$ and $\beta_1 + \dots + \beta_s = \beta$.

Proof. We may assume that $x_0 = y_0 = f(y_0) = 0$, without loss of generality.

Step 1: Taylor Expansion) Write x^1, \dots, x^d for the coordinate functions on \mathbb{R}^d and y^1, \dots, y^n for the coordinates on \mathbb{R}^n . Expand the Taylor polynomials of ϕ and f as

$$P = T_{x_0}^k \phi = \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ 1 \leq |\alpha| \leq k}} \frac{x^\alpha}{\alpha!} p_\alpha, \quad F = T_{y_0}^k f = \sum_{\substack{\beta \in \mathbb{N}_0^n \\ 1 \leq |\beta| \leq k}} \frac{y^\beta}{\beta!} f_\beta,$$

with $p_\alpha = \partial^\alpha \phi(x_0) \in \mathbb{R}^n$ with components $p_\alpha^i = \partial^\alpha \phi^{(i)}(x_0)$ and $f_\beta = \partial^\beta f(y_0) \in \mathbb{R}$.

Step 2: Polynomial Composition) By Theorem (9) we have $\partial^\nu (f \circ \phi)(0)/\nu! = [x^\nu](F \circ P)$. We compute $[x^\nu](F \circ P)$ as follows: Let $m \geq 0$. For each coordinate y^i we have:

$$(y^i \circ P)^m = \left(\sum_{\alpha \in \mathbb{N}_0^d} \frac{p_\alpha^i}{\alpha!} x^\alpha \right)^m = \sum_{\alpha_1, \dots, \alpha_m \in \mathbb{N}_0^d} \prod_{j=1}^m \frac{p_{\alpha_j}^i}{\alpha_j!} x^{\alpha_j} = \sum_{\substack{\rho_i: \mathbb{N}_0^d \rightarrow \mathbb{N}_0 \\ \sum_\alpha \rho_i(\alpha) = m}} \left(\frac{m!}{\prod_{\alpha \in \mathbb{N}_0^d} \rho_i(\alpha)!} \right) \prod_{\alpha \in \mathbb{N}_0^d} \left(\frac{p_\alpha^i}{\alpha!} x^\alpha \right)^{\rho_i(\alpha)}.$$

Note that there are only finitely many maps $\rho : \mathbb{N}_0^d \rightarrow \mathbb{N}_0$ with $\sum_\alpha \rho(\alpha) = n$, and that each of those maps is finitely supported i.e. $\rho(\alpha) = 0$ except for finitely many $\alpha \in \mathbb{N}_0^d$. In the second step the map ρ_i counts the number of occurrences of α in a given tuple $(\alpha_1, \dots, \alpha_m)$, i.e. $\rho_i(\alpha) = \#\{j \mid \alpha_j = \alpha\}$. For each given map ρ_i we have $m! / \prod_{\alpha \in \mathbb{N}_0^d} \rho_i(\alpha)!$ tuples which give the same map.

Now let $\beta \in \mathbb{N}_0^n$ and apply this formula to each $m = \beta_i \in \mathbb{N}_0$. Write $\rho : \mathbb{N}_0^d \rightarrow \mathbb{N}_0^n$, $\rho(\alpha) = (\rho_1(\alpha), \dots, \rho_n(\alpha))$, so that:

$$\begin{aligned} (y \circ P)^\beta &= \prod_{i=1}^n (y^i \circ P)^{\beta_i} = \sum_{\substack{\rho: \mathbb{N}_0^d \rightarrow \mathbb{N}_0^n \\ \sum_\alpha \rho(\alpha) = \beta}} \left(\prod_{i=1}^n \frac{\beta_i!}{\prod_{\alpha \in \mathbb{N}_0^d} \rho_i(\alpha)!} \right) \prod_{i=1}^n \prod_{\alpha \in \mathbb{N}_0^d} \left(\frac{p_\alpha^i}{\alpha!} x^\alpha \right)^{\rho_i(\alpha)} \\ &= \beta! \sum_{\substack{\rho: \mathbb{N}_0^d \rightarrow \mathbb{N}_0^n \\ \sum_\alpha \rho(\alpha) = \beta}} \left(\prod_{\alpha \in \mathbb{N}_0^d} \frac{1}{\rho(\alpha)!} \frac{(p_\alpha)^{\rho(\alpha)}}{(\alpha!)^{|\rho(\alpha)|}} \right) x^{\sum_\alpha \alpha |\rho(\alpha)|}. \end{aligned}$$

Substituting into $F(P(x))$ and extracting the x^α coefficient gives:

$$\frac{\partial^\alpha (f \circ \phi)(0)}{\alpha!} = [x^\alpha] (F \circ P) = [x^\alpha] \sum_{\beta \in \mathbb{N}_0^n} \frac{f_\beta}{\beta!} (y \circ P)^\beta = \sum_{\beta \in \mathbb{N}_0^n} f_\beta \sum_{\rho(*)} \left(\prod_{\gamma \in \mathbb{N}_0^d} \frac{1}{\rho(\gamma)!} \frac{(p_\gamma)^{\rho(\gamma)}}{(\gamma!)^{|\rho(\gamma)|}} \right)$$

Here $(*)$ runs over all maps $\rho : \mathbb{N}_0^d \rightarrow \mathbb{N}_0^n$ satisfying $\sum_\gamma \rho(\gamma) = \beta$ and $\sum_\gamma \gamma |\rho(\gamma)| = \alpha$. We convert this sum into a tuple sum as follows: Fix a total order “ $<$ ” on \mathbb{N}_0^d and write $\text{Supp}(\rho) = \{\alpha_1 < \dots < \alpha_s\} \subset \mathbb{N}_0^d$. Setting $\beta_j = \rho(\alpha_j) \in \mathbb{N}_0^n$ for $j = 1, \dots, s$, the datum of a map $\rho : \mathbb{N}_0^d \rightarrow \mathbb{N}_0^n$ is equivalent to a tuple $(\alpha_1, \dots, \alpha_s; \beta_1, \dots, \beta_s)$ with $\alpha_r \in \mathbb{N}_0^d$, $\beta_r \in \mathbb{N}_0^n \setminus 0$ and $\alpha_1 < \dots < \alpha_s$. The constraints become: $\beta_1 + \dots + \beta_s = \beta$ and $|\beta_1|\alpha_1 + \dots + |\beta_s|\alpha_s = \alpha$.

Since $P(0) = 0$, we have $p_0 = 0$, so only tuples with $\alpha_j \neq 0$ for all j contribute. The constraint $|\beta_1|\alpha_1 + \dots + |\beta_s|\alpha_s = \alpha$ forces $1 \leq |\beta| \leq |\alpha|$: indeed $|\alpha| = \sum_j |\alpha_j| |\beta_j| \geq \sum_j |\beta_j| = |\beta|$ since $|\alpha_j| \geq 1$, and $\alpha \neq 0$ implies $s \geq 1$ hence $\beta \neq 0$. This concludes the proof. \square

As a by-product of the proof, we obtain the following map-form of the Faa di Bruno:

(15) Corollary (Map-Form Faa di Bruno). *In the setting of (14) we have the following formula:*

$$\frac{\partial^\alpha (f \circ \phi)(x_0)}{\alpha!} = \sum_{\beta \in \mathbb{N}_0^n} \partial^\beta f(y_0) \sum_{\rho(*)} \left(\prod_{\gamma \in \mathbb{N}_0^d} \frac{1}{\rho(\gamma)!} \frac{(\partial^\gamma \phi)^{\rho(\gamma)}}{(\gamma!)^{|\rho(\gamma)|}} \right)$$

Here $(*)$ runs over all maps $\rho : \mathbb{N}_0^d \setminus 0 \rightarrow \mathbb{N}_0^n$ satisfying $\sum_\gamma \rho(\gamma) = \beta$ and $\sum_\gamma \gamma |\rho(\gamma)| = \alpha$.

IRemark (Factorials and Symmetry). *Controlling the complexity of combinatorial expressions is a key technical difficulty when working with Faà di Bruno. An important insight in this presentation is that factorials are always associated with symmetrization: passing from multilinear maps to polynomials via diagonal evaluation, or quotienting out ordering information (e.g., from ordered to unordered partitions). Formulas are often simpler when symmetry is not yet imposed. We deliberately keep each factorial adjacent to the corresponding symmetrized quantity to make the combinatorics transparent.*

8 Leibniz Formula

As an example demonstrating the utility of the Faà di Bruno formula in composition form, we provide a general Leibniz rule.

(16) Proposition (Leibniz rule). *Let E be a real normed vector space and let $f_1, \dots, f_n : E \rightarrow \mathbb{R}$ be k -times Fréchet differentiable at a point $x \in E$. Denote by $\tilde{T}_x^k(f) = f(x) + T_x^k(f) = \sum_{m=0}^k \frac{1}{m!} D^m f[h^{\otimes m}]$ the augmented Taylor polynomial (including the degree=0) term. Then:*

1. (Taylor form).

$$\tilde{T}_x^k(f_1 \cdots f_n) = \pi_{\leq k}(\tilde{T}_x^k(f_1) \cdots \tilde{T}_x^k(f_n)).$$

2. (Partition form.) For $u_1, \dots, u_k \in E$,

$$D^k(f_1 \cdots f_n)(x)[u_1, \dots, u_k] = \sum_{(B_1, \dots, B_n) \models [k]} D^{|B_1|} f_1(x)[u_{B_1}] \cdots D^{|B_n|} f_n(x)[u_{B_n}]$$

Where we use the identity $D^0 f_i(x)[\] = f_i(x)$ for the empty block.

3. (Multi-index form.) If $E = \mathbb{R}^d$ and $\nu \in \mathbb{N}_0^d$ with $|\nu| = k$, then

$$\frac{\partial^\nu}{\nu!} (f_1 \cdots f_n)(x) = \sum_{\substack{\nu_1, \dots, \nu_n \in \mathbb{N}_0^d \\ \nu_1 + \dots + \nu_n = \nu}} \frac{\partial^{\nu_1} f_1(x)}{\nu_1!} \cdots \frac{\partial^{\nu_n} f_n(x)}{\nu_n!}$$

where the sum ranges over all tuples $(\nu_1, \dots, \nu_n) \in (\mathbb{N}_0^d)^n$ with $\nu_i \in \mathbb{N}_0^d$ with $\nu_1 + \dots + \nu_n = \nu$.

Proof. For $k = 0$ both expressions are trivially true. Hence we may assume $k \geq 1$.

Ad 1) Set $\phi = (f_1, \dots, f_n) : E \rightarrow \mathbb{R}^n$, and $\psi(y) = \prod_{i=1}^n y_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $c_i = f_i(x)$ so that $\psi \circ \phi = f_1 \cdots f_n$ and $c = \phi(x)$.

As $\psi(y) = \prod_{i=1}^n y_i$ is a polynomial itself, the Taylor polynomial at c can be calculated by translation and degree truncation: $T_c^k \psi(h) = \pi_{\leq k}(\psi(c+h) - \psi(c))$. Furthermore $T_x^k \phi = (T_x^k f_1, \dots, T_x^k f_n) \in \mathcal{P}(E, \mathbb{R}^n) = \mathcal{P}(E) \otimes \mathbb{R}^n$.

Now Faà di Bruno in Composition form (Theorem (9)) yields $T_x^k(f_1 \cdots f_n) = \pi_{\leq k}(\prod_{i=1}^n (c_i + T_x^k f_i) - \prod_{i=1}^n c_i)$. Adding $\prod_{i=1}^n c_i$ on both sides gives the claimed identity:

$$\tilde{T}_x^k(f_1 \cdots f_n) = \pi_{\leq k}(\tilde{T}_x^k(f_1) \cdots \tilde{T}_x^k(f_n)).$$

Ad 2) The partition formula is derived by applying Pol_k to both sides of (1). We hence find:

$$D^k(f_1 \cdots f_n)(x) = Pol_k \prod_{i=1}^n \sum_{m=0}^k \frac{1}{m!} D^m f_i(x) [h^{\otimes m}] = Pol_k \sum_{\substack{m_1, \dots, m_n \in \mathbb{N}_0 \\ m_1 + \dots + m_n = k}} \prod_{i=1}^n \frac{1}{m_i!} D^{m_i} f_i(x) [h^{\otimes m_i}].$$

The claim follows by the Polarization Lemma (11) applied to $B[a_1, \dots, a_n] = \prod a_i$.

Ad 3) Expanding the Taylor series in multi-index form $\tilde{T}_x^k f_i(h) = \sum_{0 \leq |\mu| \leq k} \frac{h^\mu}{\mu!} \partial^\mu f_i(x)$, we find:

$$\frac{\partial^\nu F(x)}{\nu!} = [h^\nu] \prod_{i=1}^n \sum_{0 \leq |\mu| \leq k} \frac{h^\mu}{\mu!} \partial^\mu f_i(x) = \sum_{\nu_1 + \dots + \nu_n = \nu} \frac{\partial^{\nu_1} f_1(x)}{\nu_1!} \cdots \frac{\partial^{\nu_n} f_n(x)}{\nu_n!}.$$

□

References

- [CS96] Gregory M. Constantine and Thomas H. Savits. A multivariate Faà di Bruno formula with applications. *Transactions of the American Mathematical Society*, 348(2):503–520, 1996. URL <https://www.ams.org/journals/tran/1996-348-02/S0002-9947-96-01501-2/>.
- [dB55] F. Faà di Bruno. Sullo sviluppo delle funzioni. *Annali di Scienze Matematiche e Fisiche*, 6:479–480, 1855.
- [dB57] F. Faà di Bruno. Note sur une nouvelle formule de calcul différentiel. *The Quarterly Journal of Pure and Applied Mathematics*, 1:359–360, 1857.
- [Din99] Seán Dineen. *Complex Analysis on Infinite Dimensional Spaces*. Springer Monographs in Mathematics. Springer, London, 1999. URL <https://doi.org/10.1007/978-1-4471-0869-6>.
- [DT08] J. Duarte and D. Torres. A discrete Faà di Bruno’s formula. *Journal of Difference Equations and Applications*, 14(10–11):1195–1213, 2008. URL <https://doi.org/10.1080/10236190701769598>.
- [EM03] Luis Hernández Encinas and Juan Muñoz Masqué. A short proof of the generalized Faà di Bruno’s formula. *Applied Mathematics Letters*, 16(7):975–979, 2003. URL [https://doi.org/10.1016/S0893-9659\(03\)90087-4](https://doi.org/10.1016/S0893-9659(03)90087-4).
- [Joh02] Warren P. Johnson. The curious history of Faà di Bruno’s formula. *The American Mathematical Monthly*, 109(3):217–234, March 2002. Also available via JSTOR (stable 2695352). URL <https://www.tandfonline.com/doi/abs/10.1080/00029890.2002.11919857>.

- [Lan93] Serge Lang. *Real and Functional Analysis*, volume 142 of *Graduate Texts in Mathematics*. Springer, New York, 3rd edition, 1993. URL <https://link.springer.com/book/10.1007/978-1-4612-0897-6>.
- [Lev06] Eliahu Levy. Why do partitions occur in Faa di Bruno's chain rule for higher derivatives?, 2006. arXiv preprint. URL <https://arxiv.org/abs/math/0602183>.
- [Sco61] George Scott. Formulae of successive differentiation. *The Quarterly Journal of Pure and Applied Mathematics*, 4:77–92, 1861. URL https://books.google.com/books/about/The_Quarterly_Journal_of_Pure_and_Applie.html?id=Czbx2jXVQMEC.