Elem. Math. 68 (2013) 33 – 38
0013-6018/13/010033-6
DOI 10.4171/EM/216

Elemente der Mathematik

The Faà di Bruno formula revisited

Raymond Mortini

Raymond Mortini, geboren 1958 in Luxemburg, studierte Mathematik mit Nebenfach Physik in Karlsruhe. Nach der Promotion und Habilitation wurde er zum ausserplanmässigen Professor in Karlsruhe ernannt. 1995 folgte er einem Ruf auf eine ordentliche Professur in Metz. 1993 wurde er mit dem Carl Freudenberg Preis ausgezeichnet. Seine Hauptarbeitsgebiete umfassen die Funktionentheorie, Funktionalanalysis und Operatorentheorie. Verschiedene Forschungsaufenthalte in den USA, Spanien und an seinem Lieblingsort, dem Mathematischen Forschungsinstitut Oberwolfach, prägen seinen Werdegang.

1 How one may (re)discover oneself this formula

Let f and g be two functions on $\mathbb{R}$ whose n-th derivatives exist. The first 4 derivatives of $f \circ g$ are easy to calculate:

$$
\begin{aligned}
(f \circ g)' &= (f' \circ g)g' \\
(f \circ g)^{(2)} &= (f^{(2)} \circ g)(g')^2 + (f' \circ g)g^{(2)} \\
(f \circ g)^{(3)} &= (f^{(3)} \circ g)(g')^3 + 3(f^{(2)} \circ g)g^{(2)}g' + (f' \circ g)g^{(3)} \\
(f \circ g)^{(4)} &= (f^{(4)} \circ g)(g')^4 + 6(f^{(3)} \circ g)g^{(2)}(g')^2 + 4(f^{(2)} \circ g)g^{(3)}g' + 3(f^{(2)} \circ g)(g^{(2)})^2 + (f' \circ g)g^{(4)}
\end{aligned}
$$

In der Leibnizschen Regel für die höheren Ableitungen eines Produktes $f \cdot g$ von zwei Funktionen treten die Binomialkoeffizienten auf. Bei Produkten von mehr als zwei Faktoren sind es entsprechend die Multinomialkoeffizienten. Die Kettenregel für die höheren Ableitungen der Zusammensetzung $f \circ g$ zweier Funktionen auf $\mathbb{R}$ weist hingegen eine kompliziertere Struktur auf. Die entsprechende Formel wurde erstmals vom vielseitigen italienischen Mathematiker Francesco Faà di Bruno 1855 publiziert und seither immer wieder untersucht. Die klassische Variante der Formel für die n-te Ableitung benützt eine etwas unbequeme Summation über die nichtnegativen Lösungen der diophantischen Gleichung $b_1 + 2b_2 + \cdots + nb_n = n$, die bekanntlich in der Kombinatorik eine wichtige Rolle spielt. In der vorliegenden Arbeit wird Faà di Brunos Formel auf eine intuitive Weise hergeleitet und eine einfachere Summationsreihenfolge vorgeschlagen.

The expressions get rapidly longer and more complicated. For example we have 42 summands for $n = 10$. Note that the number of terms equals the partition number $p(n)$, that is the number of ways (without order) of writing the integer n as a sum of strictly positive integers; by the Hardy-Ramanujan formula we have $p(n) \sim \frac{1}{4n\sqrt{3}} e^{\pi\sqrt{2n/3}}$.

Let

$$\boldsymbol{M}^j = \{\boldsymbol{k} = (k_1, \ldots, k_j) \in (\mathbb{N}^*)^j,\ k_1 \ge k_2 \ge \cdots \ge k_j \ge 1\}$$

be the set of ordered multi-indeces in $\mathbb{N}^* = \{1, 2, \ldots\}$. If g is a function defined on $\mathbb{R}$, and if $g^{(n)}$ is the n-th derivative of g, then we denote by $g^{(\boldsymbol{k})}$ the function $\prod_{i=1}^{j} g^{(k_i)}$, where $\boldsymbol{k} = (k_1, \ldots, k_j) \in \boldsymbol{M}^j$. Also, $g^{(0)}$ is, by convention, equal to the function g itself.

The classical Faà di Bruno formula from ca. 1850 gives an explicit formula for $(f \circ g)^{(n)}$:

$$(f \circ g)^{(n)}(x) = \sum \frac{n!}{b_1! b_2! \cdots b_n!} f^{(j)}(g(x)) \prod_{i=1}^{n} \left(\frac{g^{(i)}(x)}{i!} \right)^{b_i},$$

where the sum is taken over all different solutions in nonnegative integers $b_1, b_2, \ldots, b_n$ of

$$b_1 + 2b_2 + \cdots + nb_n = n \text{ and } j := b_1 + \cdots + b_n.$$

A nice historical survey on this appeared in [2]. See also [1].

Without being aware of that formula, I developed around 1976–1980 the following formula:

$$(f \circ g)^{(n)}(x) = \sum_{j=1}^{n} f^{(j)}(g(x)) \Bigg(\sum_{\substack{\boldsymbol{k} \in \boldsymbol{M}^j \\ |\boldsymbol{k}| = n}} C_{\boldsymbol{k}}^n \, g^{(\boldsymbol{k})}(x) \Bigg), \tag{1.1}$$

where

$$C_{\boldsymbol{k}}^n = \frac{\binom{n}{\boldsymbol{k}}}{\prod_i N(\boldsymbol{k}, i)!}.$$

Here $\binom{n}{\boldsymbol{k}}$ is the multinomial coefficient defined by $\binom{n}{\boldsymbol{k}} = \frac{n!}{k_1! k_2! \ldots k_j!}$, where $|\boldsymbol{k}| := k_1 + \cdots + k_j = n$, and $N(\boldsymbol{k}, i)$ is the number of times the integer i appears in the j-tuple $\boldsymbol{k}$ ($i \in \mathbb{N}^*$ and $\boldsymbol{k} \in (\mathbb{N}^*)^j$).

For example, the coefficient $C^6_{(4,1,1)}$ of the term $g^{(4)}(g')^2$ when looking at the 6-th order derivative of $f \circ g$ is

$$C^6_{(4,1,1)} = \frac{1}{2!} \; \frac{6!}{4! \cdot 1! \cdot 1!} = 15$$

and the coefficient $C^{10}_{(4,2,1,1,1,1)}$ of the term $g^{(4)} g'' (g')^4$ in the 10-th derivative is

$$C^{10}_{(4,2,1,1,1,1)} = \frac{1}{4!} \; \frac{10!}{4! \cdot 2! \cdot 1! \cdot 1! \cdot 1! \cdot 1!} = 3150.$$

The difference between our formula and the Faà di Buno formula is that we use a simpler summation order and do not consider exponents of the form $b_j = 0$. In particular, we do not need summation over those $(b_1, \dots, b_n)$ satisfying (the difficult to grasp) condition $\sum_{i=1}^n i b_i = n$. That these two formulas are equivalent though, immediately follows from a direct comparison of the coefficients. Indeed, for fixed j and $\boldsymbol{k} = (k_1, \dots, k_j) \in \boldsymbol{M}^j$, $|\boldsymbol{k}| = n$, we have:

$$\begin{aligned} C_{\boldsymbol{k}}^n &= \left(\frac{n!}{k_1! \cdots k_j!}\right) \frac{1}{N(\boldsymbol{k},1)!} \cdots \frac{1}{N(\boldsymbol{k},n)!} \\ &= \Bigg(n! \underbrace{\frac{1}{i_1! \cdots i_1!}}_{b_{i_1} \text{ times}} \cdots \underbrace{\frac{1}{i_\ell! \cdots i_\ell!}}_{b_{i_\ell} \text{ times}}\Bigg) \frac{1}{b_{i_1}! \cdots b_{i_\ell}!} = \frac{n!}{b_1! \cdots b_n!} \prod_{i=1}^n \left(\frac{1}{i!}\right)^{b_i} \end{aligned}$$

where the b_{i_m} are those exponents that are different from zero and where $\boldsymbol{k}$ has been represented in the canonical form $\boldsymbol{k} = (\underbrace{i_1, \dots, i_1}_{b_{i_1} \text{ times}}, \cdots, \underbrace{i_\ell, \dots, i_\ell}_{b_{i_\ell} \text{ times}})$ in decreasing order. Note that

$$\sum_{i=1}^n i b_i = \sum_{s=1}^{\ell} i_s b_{i_s} = \sum_{p=1}^{j} k_p = n$$

and that

$$\sum_{i=1}^n b_i = b_{i_1} + \cdots + b_{i_\ell} = j.$$

Next I would like to present the (intuitive) steps that led me to the discovery of the formula (1.1) above, at pre-PC times; the first (non-programmable) slide rule calculator SR50 had just appeared.

1) I calculated explicitely the derivatives $(f \circ g)^{(n)}$ up to the order 10 and wrote them down in a careful chosen order (see Figure 1);

2) An immediate guess is that

$$(f \circ g)^{(n)}(x) = \sum_{j=1}^n f^{(j)}(g(x)) \Bigg(\sum_{\substack{\boldsymbol{k} \in \boldsymbol{M}^j \\ |\boldsymbol{k}|=n}} c_{\boldsymbol{k}}^n \, g^{(\boldsymbol{k})}(x) \Bigg),$$

for some coefficients $c_{\boldsymbol{k}}^n$ to be determined.

3) Next I gave an inductive proof that this representation is correct; that needs the main step of the construction: where does the factor $g^{(k_1)} \cdots g^{(k_j)}$ with $k_1 + \cdots + k_j = n + 1$ comes from? So let us look at the ordered j-tuple $(k_1, \dots, k_j)$. This tuple is generated, through anti-differentiation, by the j j-tuples

$$(k_1 - 1, k_2, \dots, k_j),\ (k_1, k_2 - 1, k_3, \dots, k_j), \cdots\cdots, (k_1, \dots, k_j - 1);$$

that is,

$$\Big(g^{(k_1-1)}g^{(k_2)}\cdots g^{(k_j)}\Big)' = g^{(k_1)}g^{(k_2)}\cdots g^{(k_j)} + \cdots ,$$
$$\Big(g^{(k_1)}g^{(k_2-1)}\cdots g^{(k_j)}\Big)' = g^{(k_1)}g^{(k_2)}\cdots g^{(k_j)} + \cdots ,$$

etc.

The main difficulty being that the components k_j are not pairwise distinct. So their “multiplicities” had to be taken into account. This lead to the guess that one may have

$$c_1^1 = 1$$

$$c_{\boldsymbol{k}}^{n+1} = \sum_{i=1}^{j} \frac{N(\boldsymbol{k}-\mathbf{e}_i^j, k_i - 1)}{N(\boldsymbol{k}, k_i)}\, c^n_{\boldsymbol{k}-\mathbf{e}_i^j}$$

where for $i = 1, \dots, j$, $\mathbf{e}_i^j = (0, \dots, 0, \underbrace{1}_{i-th}, 0, \dots, 0)$, $\mathbf{e}_i^j \in \mathbb{N}^j$, $\boldsymbol{k} \in \boldsymbol{M}^j$, $|\boldsymbol{k}| = n+1$ and $1 \le j \le n+1$. Note that the j-tuple $\boldsymbol{k}-\mathbf{e}_i^j$ is not necessarily represented in the canonical form with decreasing coordinates. Also, if the i-th coordinate of $\boldsymbol{k}$ is one, then the i-th coordinate of $\boldsymbol{k}-\mathbf{e}_i^j$ is 0 and we identify $\boldsymbol{k}-\mathbf{e}_i^j$ with the associated $(j-1)$-tuple. For example in the case $\boldsymbol{k} = (3, 1, 1, 1)$ we have

$$20 = C^6_{(3,1,1,1)} = \frac{1}{1}C^5_{(2,1,1,1)} + \frac{1}{3}C^5_{(3,0,1,1)} + \frac{1}{3}C^5_{(3,1,0,1)} + \frac{1}{3}C^5_{(3,1,1,0)},$$

where (3, 0, 1, 1), (3, 1, 0, 1) and (3, 1, 1, 0) are identified with (3, 1, 1).

With these recursion formula the inductive proof went through.

4) Next one has to guess the explicite value of $c^n_{\boldsymbol{k}}$, $|\boldsymbol{k}| = n$. Now there are $\binom{n}{\boldsymbol{k}} = \frac{n!}{k_1!\dots k_j!}$ ways to choose k_1 objects out of n, then k_2 objects of the remaining ones, and so on. Due to the multiplicty, one has again to divide by $N(\boldsymbol{k}, i)!$.

This gives the guess that $c^n_{\boldsymbol{k}} = \dfrac{\binom{n}{\boldsymbol{k}}}{\prod_i N(\boldsymbol{k}, i)!}$.

5) These coefficients $c^n_{\boldsymbol{k}}$ actually satisfy the recursion relation above. Since $c_1^1 = 1$, and the fact that the recursion relation determines uniquely the next coefficients, we are done: $C^n_{\boldsymbol{k}} = c^n_{\boldsymbol{k}}$.

2 Further formulas and questions

Applying formula (1.1) for the function $f(x) = \log x$, $x > 0$ and $g(x) = e^x$ gives

$$\sum_{j=1}^{n} (-1)^{j-1}(j-1)! \sum_{\substack{\boldsymbol{k}\in \boldsymbol{M}^j \\ |\boldsymbol{k}|=n}} C^n_{\boldsymbol{k}} = 0, \quad n \ge 2\,;$$

1	$f' g'$															
1	$f'' g'^2$	**1**	$f' g''$													
1	$f^{(3)} g'^3$	**3**	$f'' g'' g'$	**1**	$f' g^{(3)}$											
1	$f^{(4)} g'^4$	**6**	$f^{(3)} g'' g'^2$	**4**	$f'' g^{(3)} g'$	**3**	$f'' g''^2$	**1**	$f' g^{(4)}$							
1	$f^{(5)} g'^5$	**10**	$f^{(4)} g'' g'^3$	**10**	$f^{(3)} g^{(3)} g'^2$	**15**	$f^{(3)} g''^{(2)} g'$	**5**	$f'' g^{(4)} g'$	**10**	$f'' g^{(3)} g''$			**1**	$f' g^{(5)}$	
1	$f^{(6)} g'^6$	**15**	$f^{(5)} g'' g'^4$	**20**	$f^{(4)} g^{(3)} g'^3$	**45**	$f^{(4)} g''^{(2)} g'^2$	**15**	$f^{(3)} g^{(4)} g'^2$	**60**	$f^{(3)} g^{(3)} g'' g'$	**15**	$f^{(3)} g''^3$	**6**	$f'' g^{(5)} g'$	$\dots$
1	$f^{(7)} g'^7$	**21**	$f^{(6)} g'' g'^5$	**35**	$f^{(5)} g^{(3)} g'^4$	**105**	$f^{(5)} g''^{(2)} g'^3$	**35**	$f^{(4)} g^{(4)} g'^3$	**210**	$f^{(4)} g^{(3)} g'' g'^2$	**105**	$f^{(4)} g''^3 g'$	**21**	$f^{(3)} g^{(5)} g'^2$	$\dots$
1	$f^{(8)} g'^8$	**28**	$f^{(7)} g'' g'^6$	**56**	$f^{(6)} g^{(3)} g'^5$	**210**	$f^{(6)} g''^{(2)} g'^4$	**70**	$f^{(5)} g^{(4)} g'^4$	**560**	$f^{(5)} g^{(3)} g'' g'^3$	**420**	$f^{(5)} g''^3 g'^2$	**56**	$f^{(4)} g^{(5)} g'^3$	$\dots$
1	$f^{(9)} g'^9$	**36**	$f^{(8)} g'' g'^7$	**84**	$f^{(7)} g^{(3)} g'^6$	**378**	$f^{(7)} g''^{(2)} g'^5$	**126**	$f^{(6)} g^{(4)} g'^5$	**1260**	$f^{(6)} g^{(3)} g'' g'^4$	**1260**	$f^{(6)} g''^3 g'^3$	**126**	$f^{(5)} g^{(5)} g'^4$	$\dots$
1	$f^{(10)} g'^{10}$	**45**	$f^{(9)} g'' g'^8$	**120**	$f^{(8)} g^{(3)} g'^7$	**630**	$f^{(8)} g''^{(2)} g'^6$	**210**	$f^{(7)} g^{(4)} g'^6$	**2520**	$f^{(7)} g^{(3)} g'' g'^5$	**3150**	$f^{(7)} g''^3 g'^4$	**252**	$f^{(6)} g^{(5)} g'^5$	$\dots$

Figure 1

whereas for $f(x) = x^n$ and $g(x) = e^x$ one obtains

$$\sum_{j=1}^{n} \binom{n}{j} j! \sum_{\substack{\boldsymbol{k} \in \boldsymbol{M}^j \\ |\boldsymbol{k}|=n}} C_{\boldsymbol{k}}^n = n^n.$$

In particular, $L := \sum_{\boldsymbol{k}:|\boldsymbol{k}|=n} C_{\boldsymbol{k}}^n \le n^n$. Is there an explicit expression for L? If one uses $f(x) = g(x) = e^x$, then

$$\left(e^{e^x}\right)^{(n)} |_{x=0} = eL.$$

One may also ask the following questions:

(1) What is $\sum_{\substack{\boldsymbol{k} \in \boldsymbol{M}^j \\ |\boldsymbol{k}|=n}} C_{\boldsymbol{k}}^n \quad (1 \le j \le n)$?

(2) What is $\max\{C_{\boldsymbol{k}}^n : |\boldsymbol{k}| = n\}$?

(3) Is there a formula for the number of partitions of n with fixed length j?

In our scheme (Figure 1), one can give easy formulas for the coefficients in each column. In fact, each element in a fixed column is a multiple of the first coefficient. More precisely, if $k_1 \ge k_2 \ge k_j > 1$, then we have:

$$C^{n+\ell}_{(k_1,\dots,k_j,\underbrace{1,\dots,1}_{\ell-\text{times}})} = C^n_{(k_1,\dots,k_j)} \binom{n+\ell}{\ell}, \quad \ell \in \mathbb{N}.$$

We observe that several columns coincide; for example $C^6_{(2,2,2)} = C^6_{(4,2)}$ and so the elements of the associated columns are the same.

There are actually infinitely many pairs of colums that coincide (just use that $C^{6+i}_{(2,2,2,i)} = C^{6+i}_{(4,2,i)}$ for every $i \geq 5$.)
Are there triples (or a higher number) of columns that coincide?

Acknowledgement. I thank Jérôme Noël for his help in finding the exact TEX-commands for creating the tabular form above. I also thank Claude Jung for valuable discussions around 1980.

References

[1] Craik, Alex. Prehistory of Faà di Bruno's formula, Amer. Math. Monthly 112 (2005), 119–130.

[2] Johnson, Warren. The curious history of Faà di Bruno's formula, Amer. Math. Monthly 109 (2002), 217–234.

Raymond Mortini
Département de Mathématiques
LMAM, UMR 7122, Université de Lorraine
Ile du Saulcy
F-57045 Metz, France
e-mail: mortini@univ-metz.fr