The Formula of Faà di Bruno: Old and New

Ion Chiţescu
Faculty of Mathematics and Computer Science
Bucharest University
Faculty of Mathematics and Informatics
Spiru Haret University
Bucharest, Romania
ionchitescu@yahoo.com

Abstract
The famous formula of F. Faà di Bruno, concerning higher order derivatives of compounded functions, is introduced. Using this formula, a new formula for higher order derivatives of inverse functions is introduced. This last formula is used in order to obtain another new formula for higher order derivatives of functions depending on other functions. This type of formula can be used, e.g. for studying parametric representations.
Keywords: *generalized composition, generalized inverse, higher order derivatives.*
MSC 2010 Classification: 26A04, 26A24.

1. INTRODUCTION

The famous formula of Faà di Bruno, expressing the n^{th} derivative of compounded function $f \circ g$, has been studied and applied a lot. The present paper is dedicated to this formula and some of its applications.

After some preliminary facts (paragraph 2), the multi-indexes involved in Faà di Bruno′s formula are introduced in paragraph 3, the formula itself being introduced in paragraph 4. A readable proof can be found in [1].

Paragraph 5 is dedicated to the exposure of a new formula giving the expression of the n^{th} derivative of the inverse function. This is a result from [1] and uses Faà di Bruno′s formula.

The final paragraph 6 uses paragraph 5 giving a more general new formula, namely the formula for the n^{th} derivative of a function depending on other function, via the elimination of the parameter. This is a result of [2], where some other formulae are given. Here we present, supplementarily, how the formula is used for graphical representation of parametric curves, with application to the Descartes Folium.

The formula of Faà di Bruno appeared first in his papers [3] and [4], then in his treatises [5] and [6]. An interesting and polemic discussion about the history of this formula appears in [7], where the author (who does not love

Faà di Bruno at all) claims that the real author of the formula is someone else (not being sure who this someone else is).

We think it is interesting and important to give some information about the fascinating personality of Francesco Faà di Bruno (1825-1888). Descendant of an old italian noble family, son of a marquis, Francesco Faà di Bruno was first a military officer, with solid mathematical instruction, signing his mathematical works with the title *Cavaliere*.

Afterwards, he retired from the army, becoming a professional mathematician. Accepted by the great Cauchy, who was the supervisor of his doctoral thesis, F. Faà di Bruno became a prominent mathematician of his time and continued his career as an extraordinarius professor at University of Torino. His studies in the theory of elimination and the Theory of binary forms are now classical.

At the end of his life, F. Faà di Bruno discovered God (being strongly influenced in this direction by his professor Cauchy who was a fervent catholic) and become a priest, achieving a lot of humanitarian activities. Due to his achievements and his faith, the late pope John Paul 2^{nd} beatified him in 1988, a hundred years after his death.

2. PRELIMINARY FACTS

As usual, $\mathbb{N}$ is the set of natural numbers, $\mathbb{N}^* = \mathbb{N}\setminus\{0\}$ and $\mathbb{R}$ is the set of real numbers. A *real interval* is a set $I \subset \mathbb{R}$ which is a non degenerate interval.

If X and Y are non empty sets, $\varnothing \neq A \subset X$ and $f : A \to Y$ is injective, the *generalized inverse* of f is the function $h : f(A) \to A$ defined via $h(y) = x$, where $x \in A$ is uniquely defined by the condition $y = f(x)$. The bijection h will be denoted also via $h = f^{-1}$.

Let X, Y, Z be non empty sets and $\varnothing \neq A \subset X$, $\varnothing \neq B \subset Y$. Let $f : A \to Y$ and $g : B \to Z$ be two functions such that $f(A) \subset B$. The *generalized composition* of g and f is the function $g \circ f : A \to Z$, defined via $(g \circ f)(x) = g(f(x))$.

For a real interval I, $f : I \to \mathbb{R}$ and $x \in I$, we shall denote by $D^n f(x)$ the n^{th} derivative of f at x, where $n \in \mathbb{N}^*$. We shall also write $D^1 f(x) = f'(x)$ and $D^2 f(x) = f''(x)$. If $f : I \to H \subset \mathbb{R}$, we shall consider the function $F : I \to \mathbb{R}$ acting via $F(x) = f(x)$ and by definition $D^n f(x) = D^n F(x)$.

3. MULTIINDEXES OF TYPE (n, k)

Definition 1 *Let* $1 \leq k \leq n$ *be natural numbers. An element*

$$(k_1, k_2, \ldots, k_n) \in \mathbb{N}^*$$

will be called ***multiindex of type*** (n, k) *or* $(n, k)-$***multiindex*** *if it has the* *following two properties:*

$$\begin{cases} k_1 + k_2 + \ldots + k_n = k \\ k_1 + 2k_2 + \ldots + nk_n = n \end{cases}$$

We shall write $M(n,k)$ for the set of all $(n,k)-$multiindexes.
For any $1 \leq k \leq n$, the set $M(n,k)$ is not empty:

$$\begin{array}{l} (0,0,0,\ldots,0,0,1) \in M(n,1) \\ (1,0,0,\ldots,0,1,0) \in M(n,2) \\ (2,0,0,\ldots,1,0,0) \in M(n,3) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ (n-2,1,0,\ldots,0,0,0) \in M(n,n-1) \\ (n,0,0,\ldots,0,0,0) \in M(n,n) \end{array}$$

E.g. we have:

$$\begin{array}{ccc} M(1,1) = \{(1)\}\,; & M(2,1) = \{(0,1)\}\,; & M(2,2) = \{(2,0)\}\,; \\ M(3,1) = \{(0,0,1)\}\,; & M(3,2) = \{(1,1,0)\}\,; & M(3,3) = \{(3,0,0)\}\,. \end{array}$$

Notation: for $1 \leq k \leq n$ and $1 \leq i \leq n$

$$M(n,k,i) = \{(k_1,k_2,\ldots,k_n) \in M(n,k) | k_i \geq 1\}\,.$$

Definition 2 *For $1 \leq k \leq n$ the $(n,k)-$**derivation** is the function $D(n,k) : M(n,k) \to M(n+1,k+1)$, given via*

$$\begin{array}{l} D(1,1)((1)) = (2,0) \\ D(n,k)((k_1,k_2,\ldots,k_n) = (k_1+1,k_2,\ldots,k_n,0) \end{array}$$

Assertion 3 *Assume $2 \leq k \leq n$ and let $(k_1.k_2,\ldots,k_n) \in M(n,k,1)$. Then:*

(i) $k_n = 0$

(ii) $(k_1,k_2,\ldots,k_n) = D(n-1,k-1)((k_1-1,k_2,\ldots,k_{n-1}))$

Definition 4 *For $1 \leq k \leq n$ and $1 \leq i \leq n$, the $(n,k,i)-$**derivation** is the function*

$$D(n,k,i) : M(n,k,i) \to M(n+1,k)$$

defined as follows:

a) *For $i = 1$ (in this case $k_n = 0$, if $n \geq 2$):*

$$\begin{array}{l} D(1,1,1)((1)) = (0,1) \\ D(n,k,1)((k_1,k_2,\ldots,k_n)) = (k_1-1,k_2+1,k_3,\ldots,k_n,0) \end{array}$$

i.e. $D(n,k,1)((k_1,k_2,\ldots,k_{n-1},0)) = (k_1-1,k_2+1,k_3,\ldots,k_{n-1},0,0).$

b) *For* $1 < i < n$ *(in this case* $k_n = 0$*):*

$$D(n,k,i)((k_1,k_2,\ldots,k_n)) = (k_1,k_2,\ldots,k_i-1,k_{i+1}+1,k_{i+2},\ldots,k_n,0)$$

i.e.

$$\begin{aligned}&D(n,k,i)((k_1,k_2,\ldots,k_{n-1},0)) =\\ &= (k_1,k_2,\ldots,k_i-1,k_{i+1}+1,k_{i+2},\ldots,k_{n-1},0,0).\end{aligned}$$

c) *For* $i = n$ *(in this case* $k_n = 1$ *and* $k_p = 0$*, for all* $1 \le p \le n$*, hence* $k = 1$*):*

$$\begin{aligned}D(n,k,n)((k_1,k_2,\ldots,k_n)) &= D(n,1,n)((0,0,\ldots,0,1)) =\\ &= (0,0,\ldots,0,0,1).\end{aligned}$$

Assertion 5 *Let* $n \ge 2$ *and let* $u = (k_1.k_2,\ldots,k_n) \in M(n,k,i)$*, where* $1 < i \le n$*. Then:*

1) *For* $i = n$*. One must have* $k_n = 1$*, hence* $k = 1$ *and*

$$u = (0,0,\ldots,0,1) \in M(n,1,n).$$

It follows that $u = D(n-1,1,n-1)(v)$*, where*

$$v = (0,0,\ldots,0,1) \in M(n-1,1)$$

(and $v = (1)$*, if* $n = 2$*).*

2) *For* $i < n$*. One must have again* $k_n = 0$*.*
It follows that $u = D(n-1,k,i-1)(v)$*, where*

$$v = (k_1,k_2,\ldots,k_{i-1}+1,k_i-1,k_{i+1},\ldots,k_{n-1}) \in M(n-1,k).$$

It follows that all the elements of $M(n,k)$ can be obtained from

$$M(n-1,k-1)$$

and from

$$M(n-1,k)$$

using the derivation described above.

The following schema is constructed in this manner and gives:

$$\begin{aligned}M(4,1) &= \{(0,0,0,1)\}\\ M(4,2) &= \{(1,0,1,0),(0,2,0,0)\}\\ M(4,3) &= \{(2,1,0,0)\}\\ M(4,4) &= \{(4,0,0,0)\}\end{aligned}$$

$$
(1) - \left| \begin{array}{l}
(0,1) - \left| \begin{array}{l}
(0,0,1) - \left| \begin{array}{l} (0,0,0,1) \\ (1,0,1,0) \end{array} \right. \\
\\
(1,1,0) - \left| \begin{array}{l} (1,0,1,0) \\ (0,2,0,0) \\ (2,1,0,0) \end{array} \right.
\end{array} \right. \\
\\
(2,0) - \left| \begin{array}{l}
(1,1,0) - \left| \begin{array}{l} (1,0,1,0) \\ (0,2,0,0) \\ (2,1,0,0) \end{array} \right. \\
\\
(3,0,0) - \left| \begin{array}{l} (2,1,0,0) \\ (4,0,0,0) \end{array} \right.
\end{array} \right.
\end{array} \right.
$$

The elements of $M(n,k)$ have combinatorial interpretations. Namely, let us consider a set A having n elements and a partition π of A having exactly k sets (blocks). Then, the partition π is characterized by

$$(k_1, k_2, \ldots, k_n) \in M(n,k)$$

as follows: k_i is the number of blocks having exactly i elements. Indeed, it follows that

$$k_1 + k_2 + \ldots + k_n = k \text{ (total number of blocks)}$$

$$1k_1 + 2k_2 + \ldots nk_n = n \text{ (total number of elements).}$$

4. THE FORMULA OF FAÀ DI BRUNO

Let I be a real interval, $x \in I$ and $f : I \to \mathbb{R}$ be a function which is n times differentiable at x, where $n \in \mathbb{N}^*$.

Notation: For any natural $1 \leq k \leq p \leq n$, we define the *Faà di Bruno coefficient* $a(p,k)(x)$, via

$$
\begin{aligned}
&a(p,k)(x) = \\
&= \sum_{(k_1,k_2,\ldots,k_p)\in M(p,k)} \frac{p!}{k_1!k_2!\ldots k_p!} \left(\frac{D^1(f(x)}{1!}\right)^{k_1} \left(\frac{D^2(f(x)}{2!}\right)^{k_2} \ldots \left(\frac{D^p(f(x)}{p!}\right)^{k_p} = \\
&= \sum_{(k_1,k_2,\ldots,k_p)\in M(p,k)} \frac{p!}{k_1!(1!)^{k_1}\ldots k_p!(p!)^{k_p}} \left(D^1(f(x)\right)^{k_1} \ldots (D^p(f(x))^{k_p}
\end{aligned}
\tag{1}
$$

Remark. All the numbers

$$\frac{p!}{k_1!(1!)^{k_1}k_2!(2!)^{k_2}\ldots k_p!(p!)^{k_p}}$$

are natural numbers and have combinatorial interpretation.

Theorem 6. (Formula of Faà di Bruno) *Let I, J be real intervals and $f : I \to \mathbb{R}$, $g : J \to \mathbb{R}$ functions such that $f(I) \subset J$. Assume $x \in I$, f is n times differentiable at x and g is n times differentiable at $f(x)$, $n \in \mathbb{N}^*$.*

Then $g \circ f : I \to \mathbb{R}$ is n times differentiable at x and we have the formula (see (1)):

$$D^n(g \circ f)(x) = \sum_{k-1}^{n} D^k g(f(x)) a(n,k)(x).$$

Using directly the formula we get:

$$D^1(g \circ f)(x) = D^1 g(f(x)) D^1(f(x)$$

$$D^2(g \circ f)(x) = D^1 g(f(x)) D^2 f(x) + D^2 g(f(x)) \left(D^1 f(x)\right)^2$$

$$D^3(g \circ f)(x) = D^1 g(f(x)) D^3 f(x) +$$
$$+3D^2 g(f(x)) D^1 f(x) D^2 f(x) + D^3 g(f(x)) \left(D^1 f(x)\right)^3 .$$

5. HIGHER ORDER DERIVATIVES OF INVERSE FUNCTIONS

Theorem 7. *Let I be a real interval, $f : I \to \mathbb{R}$ an injective function such that $f(I) = J$ is a real interval and $x \in I$. Assume f is n times differentiable at x, $n \in \mathbb{N}^*$, $n \geq 2$ and $f'(x) \neq 0$. Write $f(x) = y$.*

Under these conditions, the generalized inverse $g = f^{-1} : J \to I$ is n times differentiable at y. For any natural $1 \leq m \leq n$ we have the formula

$$D^m g(y) = (-1)^{m+1} \frac{D_m(x)}{\left(D^1 f(x)\right)^{\frac{m(m+1)}{2}}}$$

where:

$$D_1(x) = 1$$
$$D_2(x) = a(2,1)(x) = D^2 f(x)$$
$$D_3(x) = \begin{vmatrix} a(2,1)(x) & a(2,2)(x) \\ a(3,1)(x) & a(3,2)(x) \end{vmatrix}$$

and, for $m \geq 4$

$$D_m(x) =$$
$$= \begin{vmatrix} a(2,1)(x) & a(2,2)(x) & 0 & \dots & 0 \\ a(3,1)(x) & a(3,2)(x) & a(3,3)(x) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a(m-1,1)(x) & a(m-1,2)(x) & a(m-1,3)(x) & \dots & a(m-1,m-1)(x) \\ a(m,1)(x) & a(m,2)(x) & a(m,3)(x) & \dots & a(m,m-1)(x) \end{vmatrix}$$

So, it is possible to compute $D^m g(y)$ using the expressions of $D^p f(x)$, $1 \leq p \leq m$.

A direct application of the above formula gives:

$$D^1g(y) = \frac{1}{D^1f(x)}$$
$$D^2g(y) = -\frac{D^2f(x)}{\left(D^1f(x)\right)^3}$$
$$D^3g(y) = \frac{3\left(D^2f(x)\right)^2 - D^3f(x)D^1f(x)}{\left(D^1f(x)\right)^5}$$
$$D^4g(y) = \frac{-D^4f(x)\left(D^1f(x)\right)^2 + 10D^1f(x)D^2f(x)D^3f(x) - 15\left(D^2f(x)\right)^3}{\left(D^1f(x)\right)^7}$$

6. **HIGHER ORDER DERIVATIVES OF FUNCTIONS DEPENDING ON OTHER FUNCTIONS**

In this last paragraph, we shall use the results of the preceding paragraph, obtaining new formulae in a more general framework.

Namely, we shall consider a real interval I, a function $f : I \to \mathbb{R}$ which is continuous and strictly monotone and another function $g : I \to \mathbb{R}$.

Considering the generalized inverse $h = f^{-1} : J = f(I) \to I$, we can define the function $F = g \circ h : J \to \mathbb{R}$ (hence $F(x) = g(f^{-1}(x))$ for any $x \in J$).

This construction can be presented in a more intuitive way as follows:

First variant Write, for $t \in I$: $f(t) = x$, $g(t) = y$. Then, if $x \in J$, we have $h(x) = t \in I$ anf $F(x) = g(h(x) = g(t) = y$, hence the equivalence

$$(x = f(t) \text{ and } y = g(t)) \Longleftrightarrow (y = F(x))$$

Second variant Again, we write

$$x = f(t),\, y = g(t)$$

and we eliminate t, obtaining $y = F(x)$.

In case $g(t) = t$ for any t, we get $F(x) = h(x)$ (the generalized inverse of f).

Our problem is to obtain formulae for $D^nF(x)$, using $D^pf(t)$, $D^pg(t)$, for $1 \leq p \leq n$.

To this end, we use Theorem 7 and obtain

Theorem 8. *Let I be real interval, $f : I \to \mathbb{R}$ a continuous and strictly monotone function and $g : I \to \mathbb{R}$. Assume that $t_0 \in I$ and f, g are n times differentiable at t_0, $n \in \mathbb{N}^*$. Assume also that $f'(t_0) \neq 0$.*

Define $F : J = f(I) \to \mathbb{R}$ via $F(x) = g(h(x)$, where $h = f^{-1} : J \to I$ is the generalized inverse of f.

Then F is n times differentiable at x_0 and one has the formula

$$D^nF(x_0) = n!(-1)^{n+1}\sum_{k=1}^{n} D^kg(t_0)C(n,k)(t_0)$$

where

$$C(n,k)(t_0) =$$
$$= (-1)^k \sum_{(k_1,\dots k_n)\in M(n,k)} \frac{1}{k_1!(1!)^{k_1}\dots k_n!(n!)^{k_n}} \left(\frac{D_1 f(t_0)}{f'(t_0)^{\frac{1\cdot 2}{2}}}\right)^{k_1} \dots \left(\frac{D_n f(t_0)}{f'(t_0)^{\frac{n(n+1)}{2}}}\right)^{k_n}$$

The reader can check that, in case $g(t) = t$ for any $t \in I$, one obtains

$$D^n F(x_0) = (-1)^{n+1} \frac{D_n(t_0)}{f'(t_0)^{\frac{n(n+1)}{2}}} = D^n h(x_0)$$

confirming Theorem 7.

Direct application of the formula gives

$$D^1 F(x_0) = F'(x_0) = \frac{g'(t_0)}{f'(t_0)}$$
$$D^2 F(x_0) = F''(x_0) = \frac{g''(t_0)f'(t_0) - g'(t_0)f''(t_0)}{f'(t_0)^3}$$

Remark. The results of the last paragraph can be used for the study of the parametric representration.

To be more precise, we consider a set $I \subset \mathbb{R}$ which can be a real interval or a finite union of real intervals and two continuous functions, $f, g : I \to \mathbb{R}$, giving the generalized path $\gamma : I \to \mathbb{R}^2$, $\gamma(t) = (f(t), g(t))$. So, we have the parametric representation

$$x = f(t),\ y = g(t).$$

In order to represent the image of γ, one divides I into a finite set of intervals such that, on each division interval, the function f is strictly monotone. We arrive at the framework of this last paragraph and we can construct the function $x \mapsto F(x) = g(f^{-1}(x))$. Representing the graph of F means to represent a portion of the image of γ, corresponding to the respective interval of strict monotonicity of f.

For instance, the Descartes Folium is the image of the path

$$\gamma : \mathbb{R}\backslash\{-1\} \to \mathbb{R}^2,$$

given via

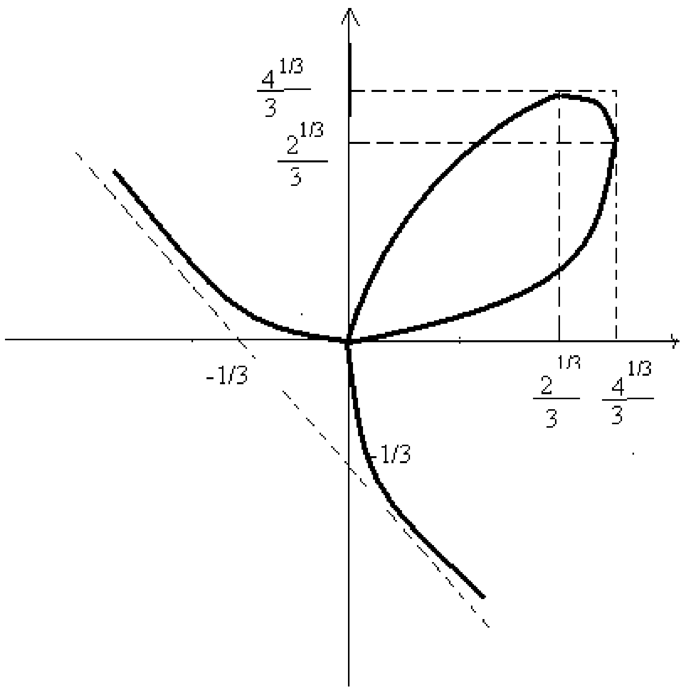
$$\gamma(t) = \left(\frac{t}{t^3+1}, \frac{t^2}{t^3+1}\right) = (f(t), g(t)).$$

Working for $t \in \mathbb{R}\backslash\left\{-1, 2^{-1/3}\right\}$, one has $f'(t) \neq 0$ on each of the three intervals thus obtained.

For $x = \frac{t}{t^3+1}$ one gets

$$F'(x) = \frac{2t - t^4}{1 - 2t^3}$$
$$F''(x) = \frac{2(1+t^3)^4}{(1-2t^3)^3}.$$

Studying the sign of $F'(x)$ and $F''(x)$ and using supplementary devices, one can represent the Descartes Folium as follows:



References

1. I. Chiţescu, *The Formula of Faà di Bruno and Higher Order Derivatives of Inverses* An. Univ. Buc., Ser. Mat. (2008), 133-148.
2. I. Chiţescu, *Higher Order Derivatives of Functions Depending On Other Functions* (in press).
3. Cavaliere Francesco Faà di Bruno, *Sullo sviluppo delle funzioni*, "Annali di Scienze Matematiche e Fisiche", 6 (1855), 479-480.
4. Cavaliere Francesco Faà di Bruno, *Note sur une nouvelle formule de calcul différentiel*, Quarterly J. Pure Appl. Math, 1 (1857), 359-360.
5. Cavaliere Francesco Faà di Bruno, *Théorie générale de l 'élimination Lieber et Faraguet*, Paris, 1859.
6. Cavaliere Francesco Faà di Bruno, *Théorie des formes binaires*, Librairie Brero, Torino, 1876.
7. W.P. Johnson, *The Curious History of Faà di Bruno's Formula*, Amer. Math. Month, 109 (2002), 217-234.