Reprinted from the AMERICAN MATHEMATICAL MONTHLY
Vol. 87, No. 10, December 1980

THE FORMULA OF FAA DI BRUNO

STEVEN ROMAN

Department of Mathematics, University of South Florida, Tampa, FL 33620

1. Introduction. Almost every calculus student is familiar with the formula of Leibniz for the nth derivative of the product of two functions

$$D^n f(t)g(t)=\sum_{k=0}^{n}\binom{n}{k}D^k f(t)D^{n-k}g(t).$$

A much less well known formula is that of Faà di Bruno for the nth derivative of the composition $f(g(t))$ (see Theorem 2). It is the purpose of this paper to give a new proof of this formula.

Several proofs of this formula have appeared in the literature. For example, in [1] there is a

The author received his Ph.D in mathematics in 1975 at the University of Washington under Branko Grünbaum. Since then he has been Instructor of Applied Mathematics at the Massachusetts Institute of Technology, Lecturer of Mathematics at the University of California at Santa Barbara, and Assistant Professor of Mathematics at the University of South Florida. He is currently Associate Professor of Mathematics at California State University at Fullerton. His research interests include combinatorics, graph theory, and applied mathematics. —*Editors* brief sketch of a proof using Taylor series. However, the omitted details are quite cumbersome. In [2] there is a proof involving the Bell polynomials. In [3] there is a proof which relies on the old-style umbral calculus developed in the mid 19th century [3], [4]. However, this technique is sometimes not mathematically rigorous and must resort to justification by other means. The umbral calculus has taken great strides in the past decade [5]–[15] and is now a completely rigorous theory. We shall use this theory to prove the formula of Faà di Bruno.

2. The Umbral Calculus. For our purposes the basic ideas of the umbral calculus may be summarized as follows. Let P be the algebra of polynomials in a single variable x over a field C, usually the real or complex numbers. Let P^* be the dual vector space of linear functionals on P. We use the notation $\langle L|p(x)\rangle$, borrowed from Physics, for the action of the linear functional L on the polynomial $p(x)$. For each nonnegative integer k we define the linear functional A^k by

$$\langle A^k|x^n\rangle = n!\,\delta_{n,k}$$

for all $n \geqslant 0$, where $\delta_{n,k}$ is the Kronecker delta function (that is, $\delta_{n,k}=1$ if $n=k$ and $\delta_{n,k}=0$ if $n\neq k$). Then A^k is extended to any polynomial by linearity. Now any linear functional on P can be expressed as a formal series in A^k. By a formal series in A^k we mean an expression of the form

$$\sum_{k=0}^{\infty} a_k A^k$$

where $a_k \in C$. A series of this form represents a well-defined linear functional if we set

$$\left\langle \sum_{k=0}^{\infty} a_k A^k \middle| p(x)\right\rangle = \sum_{k=0}^{\infty} a_k\langle A^k|p(x)\rangle.$$

This follows because $\langle A^k|p(x)\rangle = 0$ for all but a finite number of integers k and so the sum on the right is a finite one. Now we can prove:

THEOREM 1. *If L is a linear functional on P then L can be written as*

$$L = \sum_{k=0}^{\infty} \frac{\langle L|x^k\rangle}{k!} A^k. \tag{1}$$

Proof. We have seen that the sum above is a well-defined linear functional. Moreover,

$$\left\langle \sum_{k=0}^{\infty} \frac{\langle L|x^k\rangle}{k!} A^k \middle| x^n\right\rangle = \sum_{k=0}^{\infty} \frac{\langle L|x^k\rangle}{k!}\langle A^k|x^n\rangle$$
$$= \langle L|x^n\rangle$$

for all $n \geqslant 0$. Thus (1) holds and the proof is complete.

Theorem 1 implies that the vector space P^* is isomorphic to the vector space F of all formal power series in the variable A. But F is also an algebra (in fact an integral domain). Therefore, so is P^*. To be explicit we make P^* into an algebra by setting

$$A^k A^j = A^{k+j}.$$

Then if L and M are given in the form of equation (1) we set

$$LM = \sum_{k=0}^{\infty} \frac{\langle L|x^k\rangle}{k!} A^k \sum_{j=0}^{\infty} \frac{\langle M|x^j\rangle}{j!} A^j$$
$$= \sum_{n=0}^{\infty} \frac{1}{n!}\left[\sum_{k=0}^{n} \binom{n}{k}\langle L|x^k\rangle\langle M|x^{n-k}\rangle\right] A^n,$$

and so we have the formula

$$\langle LM|x^n\rangle = \sum_{k=0}^{n} \binom{n}{k}\langle L|x^k\rangle\langle M|x^{n-k}\rangle. \tag{2}$$

From this formula an induction argument easily establishes the formula for a multiple product

$$\langle L_1 \cdots L_j | x^n \rangle = \sum_{k_1 + \cdots + k_j = n} \binom{n}{k_1, \ldots, k_j} \langle L_1 | x^{k_1} \rangle \cdots \langle L_j | x^{k_j} \rangle.$$

We call the algebra P^* the *umbral algebra*. Whenever we write $L = f(A)$ we shall mean that $f(A)$ is the series given in equation (1).

We need to establish only one simple fact about the umbral algebra before turning to Faà di Bruno's formula. If $L = f(A)$, then by L' or $f'(A)$ we mean the linear functional obtained by taking the formal derivative of the series $f(A)$ with respect to the variable A. Thus for example $(A^k)' = kA^{k-1}$.

LEMMA 1. *If $L = f(A)$ is a linear functional on P, then*

$$\langle f'(A) | p(x) \rangle = \langle f(A) | xp(x) \rangle$$

for all polynomials $p(x)$.

Proof. By linearity we need only establish this for $f(A) = A^k$ and $p(x) = x^n$. But then we have

$$\begin{aligned}\langle (A^k)' | x^n \rangle &= \langle kA^{k-1} | x^n \rangle \\ &= kn!\,\delta_{n,k-1} \\ &= (n+1)!\,\delta_{n+1,k} \\ &= \langle A^k | x^{n+1} \rangle\end{aligned}$$

and the proof is complete.

3. The Formula of Faà di Bruno. We are now ready for the main result.

THEOREM 2. *If $f(t)$ and $g(t)$ are functions for which all the necessary derivatives are defined, then*

$$D^n f(g(t)) = \sum \frac{n!}{k_1! \cdots k_n!} (D^k f)(g(t)) \left(\frac{Dg(t)}{1!} \right)^{k_1} \cdots \left(\frac{D^n g(t)}{n!} \right)^{k_n},$$

where $k = k_1 + \cdots + k_n$ and the sum is over all $k_1, \ldots, k_n$ for which $k_1 + 2k_2 + \cdots + nk_n = n$.

Proof. We shall follow the general lines of the proof given in [3]. Let us write

$$\begin{aligned}h(t) &= f(g(t)) \\ h_n &= D_t^n h(t) \\ g_n &= D_t^n g(t) \\ f_n &= D_u^n f(u)|_{u=g(t)}.\end{aligned}$$

Then

$$h_1 = D_t h(t) = D_u f(u)|_{u=g(t)} D_t g(t) = f_1 g_1,$$

and similarly

$$\begin{aligned}h_2 &= f_1 g_2 + f_1 g_1^2 \\ h_3 &= f_1 g_3 + f_2 3 g_1 g_2 + f_1 g_1^3.\end{aligned}$$

It is easily established by induction that h_n has the form

$$h_n = \sum_{k=1}^{n} f_k l_{n,k}(g_1, \ldots, g_n) \tag{3}$$

where $l_{n,k}(g_1, \ldots, g_n)$ does not depend on any of the functions f_j. Now, since we wish only to determine $l_{n,k}(g_1, \ldots, g_n)$, we are free to choose $f(t)$ arbitrarily. Let us take $f(t) = e^{at}$ where a is an arbitrary constant. Then

$$f_k = D_u^k f(u)|_{u=g(t)} = a^k e^{ag(t)} \tag{4}$$

and

$$h_n = D_t^n e^{ag(t)}. \tag{5}$$

Substituting (4) and (5) into (3) and multiplying by $e^{-ag(t)}$ gives

$$e^{-ag(t)} D_t^n e^{ag(t)} = \sum_{k=1}^{n} a^k l_{n,k}(g_1, \ldots, g_n).$$

If we set $B_n(t) = e^{-ag(t)} D_t^n e^{ag(t)}$, then for $n \geq 1$ we have

$$\begin{aligned} B_n(t) &= e^{-ag(t)} D_t^{n-1} a g_1(t) e^{ag(t)} \\ &= a e^{-ag(t)} \sum_{k=0}^{n-1} \binom{n-1}{k} g_{k+1}(t) D_t^{n-k-1} e^{ag(t)} \\ &= a \sum_{k=0}^{n-1} \binom{n-1}{k} g_{k+1}(t) B_{n-k-1}(t) \end{aligned} \tag{6}$$

where we have used Leibniz's formula for the second equality. Now we may think of t as being fixed; write $B_n(t) = B_n$ and $g_n(t) = g_n$ and define two linear functionals L and M on P by

$$\langle L | x^n \rangle = B_n$$
$$\langle M | x^n \rangle = g_n.$$

Notice that $\langle L | 1 \rangle = B_0 = 1$, $\langle M | 1 \rangle = g_0 = g$ and

$$L = \sum_{k=0}^{\infty} \frac{B_k}{k!} A^k$$
$$M = \sum_{k=0}^{\infty} \frac{g_k}{k!} A^k.$$

Equation (6) now becomes, by virtue of equation (2),

$$\begin{aligned} \langle L | x^n \rangle &= a \sum_{k=0}^{n-1} \binom{n-1}{k} \langle M | x^{k+1} \rangle \langle L | x^{n-1-k} \rangle \\ &= a \sum_{k=0}^{n-1} \binom{n-1}{k} \langle M' | x^k \rangle \langle L | x^{n-1-k} \rangle \\ &= a \langle M'L | x^{n-1} \rangle \end{aligned}$$

and so

$$\langle L' | x^{n-1} \rangle = a \langle M'L | x^{n-1} \rangle.$$

Since this holds for all $n \geq 1$, we conclude that

$$L' = aM'L. \tag{7}$$

This formal differential equation is easily solved. The linear functional $F(A) = e^{a(M-g_0)}$ clearly satisfies (7) and if $G(A)$ also satisfies (7) then $F(A)/G(A)$ has derivative equal to zero and is therefore a constant. Hence all solutions are of the form

$$L = ce^{a(M-g_0)}$$

where c is a constant. In order to determine c we consider the initial condition

$$1 = B_0 = \langle L | 1 \rangle = \langle ce^{a(M-g_0)} | 1 \rangle = c$$

and so

$$L = e^{a(M-g_0)}.$$

Thus

$$B_n = \langle L | x^n \rangle$$

$$= \langle e^{a(M-g_0)} | x^n \rangle$$
$$= \sum_{k=0}^{\infty} \frac{a_k}{k!} \langle (M-g_0)^k | x^n \rangle$$
$$= \sum_{k=0}^{\infty} \frac{a_k}{k!} \sum_{j_1+\cdots+j_k=n} \binom{n}{j_1,\ldots,j_k} \langle M-g_0 | x^{j_1} \rangle \cdots \langle M-g_0 | x^{j_k} \rangle$$
$$= \sum_{k=0}^{n} \frac{a_k}{k!} \sum_{\substack{j_1+\cdots+j_k=n \\ j_i \geq 1}} \binom{n}{j_1,\ldots,j_k} g_{j_1} \cdots g_{j_k}$$

and so equating coefficients of a^k in the two expressions for B_n gives

$$l_{n,k}(g_1,\ldots,g_n) = \frac{n!}{k!} \sum_{\substack{j_1+\cdots+j_k=n \\ j_i \geq 1}} \left(\frac{g_{j_1}}{j_1!}\right) \cdots \left(\frac{g_{j_k}}{j_k!}\right)$$
$$= \frac{n!}{k!} \sum \binom{k}{k_1,\ldots,k_n} \left(\frac{g_1}{1!}\right)^{k_1} \cdots \left(\frac{g_n}{n!}\right)^{k_n}$$

where the last sum is over all $k_1,\ldots,k_n$ for which $k_1+\cdots+k_n=k$ and $k_1+2k_2+\cdots+nk_n=n$. Finally,

$$h_n(t) = \sum_{k=1}^{n} f_k l_{n,k}(g_1,\ldots,g_n)$$
$$= \sum_{k=1}^{n} f_k \sum \frac{n!}{k_1!\cdots k_n!} \left(\frac{g_1}{1!}\right)^{k_1} \cdots \left(\frac{g_n}{n!}\right)^{k_n}$$
$$= \sum \frac{n!}{k_1!\cdots k_n!} f_k \left(\frac{g_1}{1!}\right)^{k_1} \cdots \left(\frac{g_n}{n!}\right)^{k_n}$$

where $k=k_1+\cdots+k_n$ and the last sum is over all $k_1,\ldots,k_n$ for which $k_1+2k_2+\cdots+nk_n=n$. This is the desired formula and the proof is complete.

The research for this paper was partially supported by National Science Foundation Grant MCS-7900911.

References

1. C. Jordan, Calculus of Finite Differences, Chelsea, New York, 1965, p. 33.
2. L. Comtet, Advanced Combinatorics, Reidel, Boston, 1974, pp. 137–139.
3. J. Riordan, An Introduction to Combinatorial Analysis, Wiley, New York, 1958, pp. 35–37.
4. E. T. Bell, Postulational basis for the umbral calculus, Amer. J. Math., 62 (1940) 717–724.
5. S. M. Roman and G.-C. Rota, The umbral calculus, Advances in Math., 27 (1978) 95–188.
6. R. Mullin and G.-C. Rota, Theory of binomial enumeration, in Graph Theory and Its Applications, Academic Press, 1970.
7. G.-C. Rota, D. Kahaner, and A. Odlyzko, Finite operator calculus, J. Math. Anal. Appl., 42 (1973) 685–760.
8. S. M. Roman, The algebra of formal series, Advances in Math., 31 (1979) 309–329; addendum required—page missing.
9. ______, The algebra of formal series II: Sheffer sequences, J. Math. Anal. Appl. (to appear).
10. ______, The algebra of formal series III: Several variables, J. Approx. Theory, 26 (1979) 340–341.
11. ______, Polynomials, power series and interpolation, J. Math. Anal. Appl. (to appear).
12. J. Cigler, Some remarks on Rota's umbral calculus, Proc. Koninklyke Nederlandse Akad. van Wetenschappen, Amsterdam, Ser. A, 81 (1978).
13. A. Garsia and S. A. Joni, A new expression for umbral operators and power series inversion, Proc. Amer. Math. Soc., 64 (1977) 179–185.
14. A. Garsia, An exposé of the Mullin-Rota theory of polynomials of binomial type, J. Linear and Multi-linear Algebra, 1 (1973) 47–65.
15. S. A. Joni, Lagrange inversion in higher dimensions and umbral operators, J. Linear and Multi-linear Algebra, 6 (1978) 111–121.