ANZIAM J. 48(2007), 327–341

THE MULTIVARIATE FAÀ DI BRUNO FORMULA AND MULTIVARIATE TAYLOR EXPANSIONS WITH EXPLICIT INTEGRAL REMAINDER TERM

ROY B. LEIPNIK[1] and CHARLES E. M. PEARCE[✉2]

(Received 7 August, 2006)

Abstract

The Faà di Bruno formulæ for higher-order derivatives of a composite function are important in analysis for a variety of applications. There is a substantial literature on the univariate case, but despite significant applications the multivariate case has until recently received limited study. We present a succinct result which is a natural generalization of the univariate version. The derivation makes use of an explicit integral form of the remainder term for multivariate Taylor expansions.

2000 *Mathematics subject classification*: primary 46G05, 05A15.
Keywords and phrases: multivariate composite functions, differentiation theory, integral remainder term, multivariate Taylor series, Faà di Bruno formula.

1. Introduction and history

Francesco Faà di Bruno is remembered nowadays for his formulæ for the pth derivative $G^{(p)}(z)$ of a composite function $G(z) = F(u(z))$ ([7, 8]). His little-known determinantal form seems to have been new, but the alternative form of this generalized chain rule, whose extensions are treated in this article, appears earlier in the work of several researchers. Craik [5] has traced the result back to Arbogast [2]. Further accounts of the early history and comments on alternative forms are given by Flanders [9], Gould [11] and Johnson [12].

Most applications are for $p = 2, 3, 4$, though exceptionally $p = 5, 6$ occur in statistical or plasma physics. As detailed by Johnson, the Faà di Bruno formula is mentioned in books on partitions, mathematical statistics, matrix theory, calculus

[1](Deceased October 10, 2006) Department of Mathematics, UCSB, CA 93106–3080, USA.
[2]School of Mathematical Sciences, University of Adelaide, Adelaide, SA 5005, Australia; e-mail: charles.pearce@adelaide.edu.au.

of finite differences, computer science and symmetric functions, to which we add stochastic processes [21].

Bruno's main formula for general $p \geq 1$ involves a $(p+1)$-dimensional summation over indices $m, k_1, k_2, \dots, k_p$ with $1 \leq m \leq p$ and $0 \leq k_1, k_2, \dots, k_p \leq p$, subject to coupling conditions involving

$$\sigma_p(\boldsymbol{k}) := \sum_{\ell=1}^{p} k_\ell \quad \text{and} \quad \tau_p(\boldsymbol{k}) := \sum_{\ell=1}^{p} \ell k_\ell .$$

Here and subsequently we employ boldface for a vector, in this case $\boldsymbol{k} = (k_1, \dots, k_p)$. Bruno's formula is characterized by what we may call the Bruno products

$$B_{\boldsymbol{k},p}(u(z)) = \prod_{\nu=1}^{p} \left\{ \frac{1}{k_\nu !} \left(\frac{d^\nu u(z)}{dz^\nu} \frac{1}{\nu !} \right)^{k_\nu} \right\}. \tag{1.1}$$

These appear in the Bruno formulæ for one intermediate variable and also in our multivariate analogues. The basic formula takes the compressed form

$$G^{(p)}(z) = p! \sum_{m=1}^{p} F^{(m)}(u(z)) \sum_{\substack{\boldsymbol{k}:\sigma_p(\boldsymbol{k})=m;\\ \tau_p(\boldsymbol{k})=p}} B_{\boldsymbol{k},p}(u(z)). \tag{1.2}$$

In multivariate application of the chain rule, the proliferation of additive terms becomes awkward and the derivation of the terms somewhat tiresome even for $p = 3$. This is evident already in the 2×2 case

$$\boldsymbol{u}(\boldsymbol{z}) = (u_1(z_1, z_2), u_2(z_1, z_2)).$$

The possibility of a "compressed" Bruno formula becomes more attractive with increase in the order p or the multivariate dimensions M for $\boldsymbol{u}$ and N for $\boldsymbol{z}$. Such a formula can also be used for symbolic computation instead of recursive application of the standard first-order chain rule. Certain features of the multivariate general chain rule, such as the type of cross-coupling between different intermediate variables in the τ conditions, are possibly more illuminating in a Bruno-type formula than in the full expression. Pure derivatives, such as $\partial^4 G(z_1, z_2)/\partial z_1^4$, are easier to deal with than mixed derivatives like $\partial^4 G/\partial z_1^2 \partial z_2^2$, though these cannot always be avoided, as in the calculation of $(\nabla^2)^2$ terms for elasticity and fluid mechanics. In multivariate probability theory, the verification of sign alternation in derivatives of Laplace transforms of densities is a major application of multivariable derivatives for which multivariate Faà di Bruno formulæ of all orders are desirable.

The use of a symbolic manipulator can easily produce higher-order differential results. However, it achieves these results by recursively applying a chain rule. The outputs are often quite messy and further manipulation is needed to simplify the results. If only a select number of terms are needed, one then has to go back and try to weed out such terms. By utilizing the Bruno formulæ, one can easily isolate the needed terms, which are already simplified.

Bruno gave only a hazy proof of (1.2), neither rigorous nor algorithmically convincing. Königsberger derived a more difficult formula for the general problem of calculating higher-order differentials $d^p f(z_1, \dots, z_N)$, using a symbolic calculus and induction. In principle this should yield multivariable chain rules. However, higher differentials package together a variety of different orders of derivatives, which yield a Bruno formula conveniently only in the case $M = N = 1$. In that case, Königsberger's proof also is an inductive proof of (1.2), as pointed out by Bieberbach [3]. Somewhat later, de la Vallée Poussin [17] produced a concise proof of (1.2), based on a weak form of the Taylor expansion with remainder and a weak uniqueness theorem for "almost" power series of the form

$$a_0 + a_1 h + \cdots + a_{q-1} h^{q-1} + M_q(h) h^q,$$

where $M_q(h) \to a_q$ as $h \to 0$ and $M_q(h)$ is bounded in h for small h. That proof is less elementary than a longer one based on the integral form of the Taylor remainder. The latter can be made quite explicit, at the cost of assuming slightly more regularity than needed. The integral remainder version generalizes well in the multivariable context.

We note that Schwatt [20] contains a good collection of higher-derivative formulæ, including many infinite series, but does not include the di Bruno formula, but rather various substitutes oriented to special cases of interest, and has little on the important topic of asymptotic series for higher-order derivatives, of interest in statistical mechanics (Fowler [10]). Lukács [14] has discussed the problem using formal power series.

Symbolic manipulation by computer via Macsyma, Maple, Mathematica, *etc.* can produce any required order of Bruno or Schwatt formulæ. A multivariate symbolic program produces multivariable versions of such formulæ. In another direction, similar formulæ appear in the Whitney [22] and Dieudonné [6] theories of extensions of differentiable functions. Abraham and Robbin [1] provide a detailed account.

The general problem has received attention recently from several researchers. Constantine and Savits [4] make use of a combinatorial identity, Mishkov [15] employs differential operators and Diophantine equations, while Noschese and Ricci [16] use a connection with a generalization of Bell polynomials. A number of papers in the literature come from the standpoint that derivatives are essentially integer partitions — see, for example [18, 19, 24]. See also [23].

We present a simple treatment based on Taylor series. Following a leisurely exploration of the remainder term idea in a one-dimensional rehearsal in Section 2, we provide, in Section 3, an efficient and explicit Bruno-type formula for two-variable chains

$$G(z_1, z_2) = F(u_1(z_1, z_2), u_2(z_1, z_2)).$$

Section 4 presents briefly the corresponding results for the general multivariate case of the problem.

An earlier version of this paper was presented from the standpoint of symbolic computation at the Ninth International Conference on Technology in Collegiate Mathematics [13]. In the present version we have clarified the arguments, simplified the notation and taken the opportunity to attend to some errors and inconsistencies in [13].

2. Proof of the one-dimensional Bruno formula

In this section we give an integral-remainder proof of the Bruno formula. This is directly generalizable to higher dimensions.

If $u(z)$ and $F(u)$ are $(p+1)$-times differentiable in suitable domains, the p-th-order Taylor expansions with integral remainder are given by

$$F(u+j) = \sum_{m=0}^{p} \frac{j^m}{m!} F^{(m)}(u) + R_p\left(F^{(p+1)}(u), u, j\right), \tag{2.1}$$

$$u(z+h) = \sum_{n=0}^{p} \frac{h^n}{n!} u^{(n)}(z) + R_p\left(u^{(p+1)}(z), z, h\right), \tag{2.2}$$

where the remainder terms are defined by

$$R_p(Y, v, k) = \int_v^{v+k} \frac{(v+k-y)^p}{p!} Y(y) dy. \qquad p = 0, 1, 2, \ldots .$$

The following lemma covers the relevant behaviour of the remainder terms.

LEMMA 2.1. *Suppose X is continuously differentiable and u is p-times continuously differentiable on the requisite interval and put*

$$\phi_{X,u,q} := \int_{u(z)}^{u(z+h)} \frac{[u(z+h)-y]^q}{q!} X'(y)\, dy,$$

where q is a nonnegative integer. Then for $0 \le s \le p$ and all $q \ge s$ we have

$$\left(\frac{\partial}{\partial h}\right)^s \phi_{X,u,q} \to 0 \quad \text{as } h \to 0.$$

PROOF. First take $q = 0$. We have $\phi_{X,u,0} = X(u(z+h)) - X(u(z))$, so

$$\phi_{X,u,0} \to 0 \quad \text{as} \quad h \to 0,$$

giving the result for $q = 0$. The result is also immediate for $s = 0$ and $q > 0$, providing a basis for induction on s. Suppose the desired result holds for $s = 0, \ldots, s_0 < p$ and all $q \geq s$. We have for $q > 0$ that $\partial\phi_{X,u,q}/\partial h = \phi_{X,u,q-1}u'(z+h)$. Thus by Leibnitz' theorem, we have for $s = s_0 + 1$ and $q \geq s$ that

$$\left(\frac{\partial}{\partial h}\right)^s \phi_{X,u,q} = \sum_{\ell=0}^{s-1} \binom{s-1}{\ell} \left[\left(\frac{\partial}{\partial h}\right)^\ell \phi_{X,u,q-1}\right] \left(\frac{\partial}{\partial h}\right)^{s-\ell} u'(z+h) \xrightarrow{h\to 0} 0$$

by the inductive hypothesis. This establishes the induction. $\square$

COROLLARY 2.2. *Suppose that the assumptions of Lemma 2.1 apply with $0 \leq s \leq p$. If $r > 0$ and $\tau \geq 0$, with $L_{s,\tau,r} := (\partial/\partial h)^s\left[h^\tau(\phi_{X,u,q})^r\right]$, then $L_{s,\tau,r} \to 0$ as $h \to 0$.*

PROOF. We have by Leibnitz' theorem that

$$\begin{aligned} L_{s,\tau,r} &= \sum_{\ell=0}^{s} \binom{s}{\ell} \left[\left(\frac{\partial}{\partial h}\right)^{s-\ell} h^\tau\right] \left(\frac{\partial}{\partial h}\right)^\ell \phi^r \\ &= \sum_{\ell=0}^{s} \binom{s}{\ell} \frac{\tau!}{(\tau-s+\ell)!} h^{\tau-s+\ell} \left(\frac{\partial}{\partial h}\right)^\ell \phi^r. \end{aligned}$$

If $s \leq \tau$, the term in h disappears as $h \to 0$ unless $s = \tau$ and $\ell = 0$, in which case the term in ϕ vanishes in the limit, by Lemma 2.1. So suppose $s > \tau$. Then the term in h vanishes in the limit except if $\ell = s - \tau$, and we have

$$L_{s,\tau,r} \to \frac{s!}{(s-\tau)!} \lim_{h\to 0} \left(\frac{\partial}{\partial h}\right)^{s-\tau} \phi^r \quad \text{as} \quad h \to 0.$$

A further use of Leibnitz' theorem with Lemma 2.1 gives the desired result. $\square$

THEOREM 2.3. *If $F(u)$ and $u(z)$ are $(p+1)$-times continuously differentiable $(p > 0)$ on the appropriate domains, then (1.2) holds.*

PROOF. We may set $j = j(z,h) = u(z+h) - u(z)$ in (2.1) and substitute for u from (2.2) to obtain

$$F(u(z+h)) = \sum_{m=0}^{p} \frac{F^{(m)}(u(z))}{m!} \left[\sum_{n=1}^{q} \frac{h^n}{n!} u^{(n)}(z) + {}_uR_p\right]^m + {}_FR_p, \tag{2.3}$$

where the abbreviated remainder expressions are given by

$$
\begin{aligned}
{}_uR_p &= R_p\left(u^{(p+1)}(z), z, h\right),\\
{}_FR_p &= R_p\left(F^{(p+1)}(u), u, j\right)\Big|_{u=u(z)}\ .
\end{aligned}
$$

The powered bracket expression in (2.3) has a multinomial expansion

$$
\begin{aligned}
&\sum_{r=0}^{m} \sum_{\boldsymbol{k}:\sigma_p(\boldsymbol{k})=m-r} \frac{m!}{k_1!\cdots k_p!r!}\left\{\prod_{\nu=1}^{p}\left(\frac{h^\nu}{\nu!}\frac{d^\nu u(z)}{dz^\nu}\right)^{k_\nu}\right\}\left({}_uR_p\right)^r\\
&\qquad = \sum_{r=0}^{m} \sum_{\boldsymbol{k}:\sigma_p(\boldsymbol{k})=m-r} \frac{m!}{r!} h^{\tau_p(\boldsymbol{k})} B_{\boldsymbol{k},p}(u(z))\left({}_uR_p\right)^r\ ,
\end{aligned}
$$

by use of (1.1). This leads to

$$
\begin{aligned}
G(z+h) &= F(u(z+h))\\
&= \sum_{m=0}^{p} F^{(m)}(u) \sum_{\boldsymbol{k}:\sigma_p(\boldsymbol{k})=m} h^{\tau_p(\boldsymbol{k})} B_{\boldsymbol{k},p}(u(z)) + {}_FR_p\\
&\quad + \sum_{m=0}^{p} F^{(m)}(u) \sum_{r=1}^{m} \frac{1}{r!} \sum_{\boldsymbol{k}:\sigma_p(\boldsymbol{k})=m-r} h^{\tau_p(\boldsymbol{k})} B_{\boldsymbol{k},p}(u(z))\left({}_uR_p\right)^r,
\end{aligned}
$$

where we have separated out the contribution for $r = 0$.

Thus for $p > 0$ we have

$$
\frac{\partial^p}{\partial h^p} G(z+h) = \sum_{m=0}^{p} F^{(m)}(u) \sum_{\boldsymbol{k}:\sigma_p(\boldsymbol{k})=m} B_{\boldsymbol{k},p}(u(z)) \frac{\partial^p}{\partial h^p} h^{\tau_p(\boldsymbol{k})} + \frac{\partial^p}{\partial h^p} {}_FR_p \tag{2.4}
$$

$$
+ \sum_{m=0}^{p} F^{(m)}(u) \sum_{r=1}^{m} \frac{1}{r!} \sum_{\boldsymbol{k}:\sigma_p(\boldsymbol{k})=m-r} B_{\boldsymbol{k},p}(u(z)) \frac{\partial^p}{\partial h^p}\left[h^{\tau_p(\boldsymbol{k})}\left({}_uR_p\right)^r\right]. \tag{2.5}
$$

For $h \to 0$, the term involving the derivative of a power of h on the right in (2.4) vanishes unless $p = \tau_p(\boldsymbol{k})$. Since ${}_FR_p = \phi_{X,u,p}$ for $X = F^{(p)}$, we have by Lemma 2.1 that

$$
\left(\frac{\partial}{\partial h}\right)^p {}_FR_p \to 0 \quad \text{as } h \to 0.
$$

Hence the right-hand side of (2.4) has limit

$$
p! \sum_{m=0}^{p} F^{(m)}(u(z)) \sum_{\boldsymbol{k}:\sigma_p(\boldsymbol{k})=m;\ \tau_p(\boldsymbol{k})=p} B_{\boldsymbol{k},p}(u(z))
$$

for $h \to 0$. Similarly ${}_uR_p = \phi_{u^{(p)},z,p}$, so by Corollary 2.2 the right-hand side of (2.5) has limit zero as $h \to 0$.

On combining these results we obtain

$$G^{(p)}(z) = p! \sum_{m=0}^{p} F^{(m)}(u(z)) \sum_{\boldsymbol{k}:\sigma_p(\boldsymbol{k})=m;\ \tau_p(\boldsymbol{k})=p} B_{\boldsymbol{k},p}(u(z)). \tag{2.6}$$

which is close to (1.2). Finally $\tau_p(\boldsymbol{k}) = p > 0$, which implies that at least one of $k_1, \ldots, k_p$ is positive and so $m = \sum_{\ell=1}^{p} k_\ell > 0$, which converts (2.6) into (1.2). $\square$

3. Bruno formulæ with two variables and two functions

We proceed to a multivariate chain rule for $G(z) = F\big(\boldsymbol{u}(z)\big)$ with scalar F, where $\boldsymbol{u}(z) = \big(u_1(z), \ldots, u_M(z)\big)$ and $\boldsymbol{z} = (z_1, \ldots, z_N)$. So as not to become too encumbered with algebraic detail, we begin in this section with the basic multivariate case $M = N = 2$. This has obvious application for derivative orders $p = 2, 3, 4$ to physically meaningful two-dimensional Laplacians, and to curls and repeated curls. For orthogonal coordinates, such as polar, confocal and elliptic, the results are well-known and more easily found by variational integral methods.

First we establish a double Taylor series expansion for a bivariate function, with double-integral remainder form.

THEOREM 3.1. *Suppose*

$$\left(\frac{\partial}{\partial z_1}\right)^{q_1} \left(\frac{\partial}{\partial z_2}\right)^{q_2} f(z_1, z_2)$$

exists and is jointly continuous in z_1 and z_2. Then for $0 \le s_1 \le q_1$ and $0 \le s_2 \le q_2$, we have

$$f(z_1 + h_1, z_2 + h_2) = \sum_{m_1=0}^{s_1} \sum_{m_2=0}^{s_2} \frac{h_1^{m_1} h_2^{m_2}}{m_1! m_2!} D_{\boldsymbol{z}}^{\boldsymbol{m}} f(\boldsymbol{z}) + R_{s_1,s_2}, \tag{3.1}$$

where

$$R_{s_1,s_2} = \sum_{m_1=0}^{s_1} \int_{z_2}^{z_2+h_2} \frac{h_1^{m_1}}{m_1!} \frac{(z_2 + h_2 - y_2)^{s_2}}{s_2!} D_{z_1,y_2}^{m_1,s_2+1} f(z_1, y_2)\, dy_2 \tag{3.2}$$

$$+ \sum_{m_2=0}^{s_2} \int_{z_1}^{z_1+h_1} \frac{h_2^{m_2}}{m_2!} \frac{(z_1 + h_1 - y_1)^{s_1}}{s_1!} D_{y_1,z_2}^{s_1+1,m_2} f(y_1, z_2)\, dy_1$$

$$+ \int_{z_1}^{z_1+h_1} \int_{z_2}^{z_2+h_2} \frac{(z_1+h_1-y_1)^{s_1}}{s_1!} \frac{(z_2+h_2-y_2)^{s_2}}{s_2!} D_{y_1,y_2}^{s_1+1,s_2+1} f(y_1, y_2)\, dy_1\, dy_2. \tag{3.3}$$

In particular, we have for $s_1 = 0 = s_2$ that

$$R_{0,0}(f) = \int_{z_1}^{z_1+h_1} \frac{\partial f(y_1, z_2)}{\partial y_1}\, dy_1 + \int_{z_2}^{z_2+h_2} \frac{\partial f(z_1, y_2)}{\partial y_2}\, dy_2 + \int_{z_1}^{z_1+h_1} \int_{z_2}^{z_2+h_2} \frac{\partial^2 f(y_1, y_2)}{\partial y_1 \partial y_2}\, dy_1\, dy_2. \tag{3.4}$$

PROOF. We provide a proof by induction in s_1 and s_2. To begin, the first integral in (3.4) may be evaluated as $f(z_1 + h_1, z_2) - f(z_1, z_2)$ and by symmetry, the second equals $f(z_1, z_2 + h_2) - f(z_1, z_2)$. The third integral is

$$\int_{z_1}^{z_1+h_1} \left[\frac{\partial f(y_1, z_2 + h_2)}{\partial y_1} - \frac{\partial f(y_1, z_2)}{\partial y_1} \right] dy_1 = [f(z_1 + h_1, z_2 + h_2) - f(z_1, z_2 + h_2)] - [f(z_1 + h_1, z_2) - f(z_1, z_2)].$$

Addition yields

$$R_{0,0}(f) = f(z_1 + h_1, z_2 + h_2) - f(z_1, z_2),$$

that is, (3.1) holds for $s_1 = 0 = s_2$, which is a basis for our induction.

For the inductive step, suppose (3.1) holds for $s_1 = t_1 \leq q_1$ and $s_2 = t_2 < q_2$. With these choices for s_1, s_2, the right-hand side of (3.2) may be recast by integration by parts as

$$\sum_{m_1=0}^{t_1} \frac{h_1^{m_1}}{m_1!} \frac{h_2^{t_2+1}}{(t_2+1)!} D_{z_1,z_2}^{m_1,t_2+1} f(z_1, z_2) + \sum_{m_1=0}^{t_1} \frac{h_1^{m_1}}{m_1!} \int_{z_2}^{z_2+h_2} \frac{(z_2 + h_2 - y_2)^{t_2+1}}{(s_2+1)!} D_{z_1,y_2}^{m_1,s_2+2} f(z_1, y_2)\, dy_2.$$

Similarly the right-hand side of (3.3) may be re-expressed for the same choices of s_1, s_2 as

$$\int_{z_1}^{z_1+h_1} \frac{(z_1 + h_1 - y_1)^{t_1}}{t_1!} \frac{h_2^{t_2+1}}{(t_2+1)!} D_{y_1,y_2}^{t_1+1,t_2+1} f(y_1, z_2)\, dy_1 + \int_{z_1}^{z_1+h_1} \int_{z_2}^{z_2+h_2} \frac{(z_1 + h_1 - y_1)^{t_1}}{t_1!} \frac{(z_2 + h_2 - y_2)^{t_2+1}}{(t_2+1)!} D_{y_1,y_2}^{t_1+1,t_2+2} f(y_1, y_2)\, dy_1\, dy_2.$$

Substitution of these values in (3.2) and (3.3) yields

$$R_{t_1,t_2} = \sum_{m_1=0}^{t_1} \frac{h_1^{m_1}}{m_1!} \frac{h_2^{t_2+1}}{(t_2+1)!} D_{z_1,z_2}^{m_1,t_2+1} f(z_1, z_2) + R_{t_1,t_2+1}.$$

Substitution into (3.1) with $s_1 = t_1$, $s_2 = t_2$ shows that (3.1) holds also for $s = t_1$, $s_2 = t_2 + 1$. By symmetry we have that if (3.1) holds for $s_1 = t_1 < q_1$ and $s_2 = t_2 \leq q_2$, then it applies also for $s = t_1 + 1$, $s_2 = t_2$. This proves the inductive step and so the theorem. □

This establishes the desired double induction on f, and with it a Taylor series integral remainder formula for functions of two variables which the writers have not seen elsewhere.

Before we proceed, it is convenient to introduce some notation. We take p_1 and p_2 as fixed nonnegative integers and put $p = p_1 + p_2$ and

$$A(\boldsymbol{p}) = \big(\{0, 1, \dots, p_1\} \times \{0, 1, \dots, p_2\}\big)\backslash\big(\{0\} \times \{0\}\big),$$
$$C(\boldsymbol{p}) = \big\{(m, m') \in A(p, p) : m + m' \leq p\big\}.$$

Let $U(m)$ denote the family of maps $\phi : A(\boldsymbol{p}) \to \{0, 1, \dots, m\}$. We define

$$T(m) = \left\{\phi \in U(m) : \sum_{\boldsymbol{n} \in A(\boldsymbol{p})} \phi(\boldsymbol{n}) = m\right\}$$

and for $\phi \in T(m)$

$$\binom{m}{\phi} := \frac{m!}{\prod_{\boldsymbol{n} \in A(\boldsymbol{p})} \phi(\boldsymbol{n})!} .$$

If $\phi \in U(m)$ and $\phi' \in U(m')$ for some nonnegative integers m, m', we put

$$\tau_i(\phi, \phi') = \sum_{\boldsymbol{n} \in A(\boldsymbol{p})} n_i \left[\phi(\boldsymbol{n}) + \phi'(\boldsymbol{n})\right] \quad (i = 1, 2).$$

Finally, we define

$$V(m, m') = \big\{(\phi, \phi') : \ \phi \in T(m),\ \phi' \in T(m'),\ \tau_i(\phi, \phi') = p_i \ \text{ for } \ i = 1, 2\big\}.$$

THEOREM 3.2. *Suppose $F(u_1, u_2)$ has continuous derivatives up to order $(p + 1, p + 1)$ and $u_i(z_1, z_2)$ $(i = 1, 2)$ have continuous derivatives to order $(p_1 + 1, p_2 + 1)$ on appropriate domains. Define*

$$B_\phi(u_i(\boldsymbol{z})) = \prod_{\boldsymbol{n} \in A(\boldsymbol{p})} \left\{\frac{1}{\phi(\boldsymbol{n})!} \left(\frac{D_{\boldsymbol{z}}^{\boldsymbol{n}} u_i(\boldsymbol{z})}{n_1! n_2!}\right)^{\phi(\boldsymbol{n})}\right\} \quad \textit{for } \ i = 1, 2.$$

If $G(z_1, z_2), = F(u_1(z_1, z_2), u_2(z_1, z_2))$ and $\boldsymbol{p} \neq \boldsymbol{0}$, then

$$\frac{\partial^p G(\boldsymbol{z})}{\partial z_1^{p_1} \partial z_2^{p_2}} = p_1! p_2! \sum_{\boldsymbol{m} \in C(\boldsymbol{p})} D_{\boldsymbol{u}}^{\boldsymbol{m}} F \sum_{(\phi, \phi') \in V(\boldsymbol{m})} B_\phi(u_1(\boldsymbol{z})) B_{\phi'}(u_2(\boldsymbol{z})). \tag{3.5}$$

PROOF. If $j_s = u_s(z + \boldsymbol{h}) - u_s(\boldsymbol{z})$ for $s = 1, 2$, Theorem 3.1 provides

$$\begin{aligned} F(\boldsymbol{u}+\boldsymbol{j}) &= F\left(u_1(z_1+h_1, z_2+h_2), u_2(z_1+h_1, z_2+h_2)\right) \\ &= \sum_{m_1=0}^{p}\sum_{m_2=0}^{p} \frac{j_1^{m_1} j_2^{m_2}}{m_1!m_2!} D_{\boldsymbol{u}}^{\boldsymbol{m}} F + {}_F R_{p,p}(\boldsymbol{u}, \boldsymbol{j}). \end{aligned}$$

We have readily for $0 \le i \le p_1$ and $0 \le j \le p_2$ that

$$\frac{\partial^{i+j}}{\partial h_1^i \partial h_2^j}\ {}_F R_{p,p}(\boldsymbol{u}, \boldsymbol{j})\bigg|_{\boldsymbol{h}=\boldsymbol{0}} = 0.$$

Theorem 3.1 yields also that

$$j_\ell = \sum_{\boldsymbol{n}\in A(\boldsymbol{p})} \frac{h_1^{n_1} h_2^{n_2}}{n_1!n_2!} D_{\boldsymbol{z}}^{\boldsymbol{n}} u_\ell + {}_{u_\ell} R_{p_1,p_2}(\boldsymbol{z}, \boldsymbol{h}) \quad (\ell = 1, 2)$$

and it is immediate that for $0 \le i \le p_1$ and $0 \le j \le p_2$

$$\frac{\partial^{i+j}}{\partial h_1^i \partial h_2^j}\ {}_{u_\ell} R_{p_1,p_2}(\boldsymbol{z}, \boldsymbol{h})\bigg|_{\boldsymbol{h}=\boldsymbol{0}} = 0 \quad (\ell = 1, 2).$$

It follows that the three integral remainder terms all make zero contribution to

$$\frac{\partial^p}{\partial h_1^{p_1} \partial h_2^{p_2}} F(\boldsymbol{u}+\boldsymbol{j})\bigg|_{\boldsymbol{h}=\boldsymbol{0}}.$$

Thus

$$\begin{aligned} &\frac{\partial^p}{\partial h_1^{p_1} \partial h_2^{p_2}} F(\boldsymbol{u}+\boldsymbol{j})\bigg|_{\boldsymbol{h}=\boldsymbol{0}} \\ &\quad = \frac{\partial^p}{\partial h_1^{p_1} \partial h_2^{p_2}} \sum_{m_1=0}^{p}\sum_{m_2=0}^{p} \frac{D_{\boldsymbol{u}}^{\boldsymbol{m}} F}{m_1!m_2!} \left(\sum_{\boldsymbol{n}\in A(\boldsymbol{p})} \frac{h_1^{n_1} h_2^{n_2}}{n_1!n_2!} D_{\boldsymbol{z}}^{\boldsymbol{n}} u_1 \right)^{m_1} \\ &\quad\quad \times \left(\sum_{\boldsymbol{n}'\in A(\boldsymbol{p})} \frac{h_1^{n_1'} h_2^{n_2'}}{n_1'!n_2'!} D_{\boldsymbol{z}}^{\boldsymbol{n}'} u_2 \right)^{m_2} \Bigg|_{\boldsymbol{h}=\boldsymbol{0}} \\ &\quad = \frac{\partial^p}{\partial h_1^{p_1} \partial h_2^{p_2}}\ H\bigg|_{\boldsymbol{h}=\boldsymbol{0}}, \end{aligned}$$

say, which is $p_1!p_2!$ times the coefficient of $h_1^{p_1} h_2^{p_2}$ in H.

In terms of the notation immediately preceding the theorem, we have that

$$\left(\sum_{\boldsymbol{n}\in A(\boldsymbol{p})} \frac{h_1^{n_1} h_2^{n_2}}{n_1!n_2!} D_{\boldsymbol{z}}^{\boldsymbol{n}} u_i \right)^{m_i} = \sum_{\phi\in T(m_i)} \binom{m_i}{\phi} \prod_{\boldsymbol{n}\in A(\boldsymbol{p})} \left(\frac{h_1^{n_1} h_2^{n_2}}{n_1!n_2!} D_{\boldsymbol{z}}^{\boldsymbol{n}} u_i \right)^{\phi(\boldsymbol{n})} \quad \text{for } i = 1, 2.$$

Hence

$$H=\sum_{m_1=0}^{p}\sum_{m_2=0}^{p} D_{\boldsymbol{u}}^{\boldsymbol{m}} F \sum_{\phi\in T(m_1)}\sum_{\phi'\in T(m_2)} B_\phi(u_1(z))B_{\phi'}(u_2(z))h_1^{\tau_1(\phi,\phi')}h_2^{\tau_2(\phi,\phi')}.$$

The coefficient of $h_1^{p_1}h_2^{p_2}$ is obtained when $\tau_i(\phi,\phi')=p_i$ for $i=1,2$ respectively, that is, when $(\phi,\phi')\in V(m_1,m_2)$. This yields

$$\begin{aligned}\frac{\partial^{p_1+p_2}G(z)}{\partial z_1^{p_1}\partial z_2^{p_2}} &= \frac{\partial^p}{\partial h_1^{p_1}\partial h_2^{p_2}}F(\boldsymbol{u}+\boldsymbol{j})\bigg|_{\boldsymbol{h}=\boldsymbol{0}}\\ &= p_1!p_2!\sum_{m_1=0}^{p}\sum_{m_2=0}^{p} D_{\boldsymbol{u}}^{\boldsymbol{m}} F\sum_{(\phi,\phi')\in V(m_1,m_2)} B_\phi(u_1(z))B_{\phi'}(u_2(z)).\end{aligned}$$

If $\boldsymbol{m}=\boldsymbol{0}$, then $\phi\in T(m_1)$ implies $\phi(\boldsymbol{n})=0$ for all $\boldsymbol{n}\in A(\boldsymbol{p})$ and similarly $\phi'\in T(m_1)$ implies $\phi'(\boldsymbol{n})=0$ for all $\boldsymbol{n}\in A(\boldsymbol{p})$. Since $\boldsymbol{p}\neq\boldsymbol{0}$, this is incompatible with $\tau_i(\phi,\phi')=p_i$ for $i=1,2$. Thus we may remove $\boldsymbol{0}$ from the domain of $\boldsymbol{m}$. Also $\boldsymbol{n}\in A(\boldsymbol{p})$ entails $n_1+n_2\geq 1$, so that

$$\tau_1(\phi,\phi')=\sum_{\boldsymbol{n}\in A(\boldsymbol{p})}(n_1+n_2)\left[\phi(\boldsymbol{n})+\phi'(\boldsymbol{n})\right]\geq\sum_{\boldsymbol{n}\in A(\boldsymbol{p})}\left[\phi(\boldsymbol{n})+\phi'(\boldsymbol{n})\right].$$

For $(\phi,\phi')\in V(m_1,m_2)$, we thus have $p\geq m_1+m_2$. On combining these constraints, we have that the domain of $\boldsymbol{m}$ can be restricted to $\boldsymbol{m}\in C(\boldsymbol{p})$, establishing the result of the enunciation. □

While we have assumed for simplicity of exposition that F has continuous derivatives up to order $p+1$ in each variable, the final result (3.5) involves derivatives of F only to total order $p+1$. A mollification argument can be used to strengthen Theorem 3.2 as follows.

THEOREM 3.3. *The conclusion of Theorem 3.2 holds when the regularity assumptions on F are weakened to F being of class $C^{(p+1)}$.*

4. Bruno formulæ in the general case

Formulæ appropriate for general M and N follow from a development similar to that of the previous section. A multivariate analogue to Theorem 3.1 provides the necessary underpinnings. To this end, we first provide some notation. Put $S=\{1,2,\ldots,L\}$ and define $\Omega=2^S\backslash\{\emptyset\}$. Suppose $Q=\{n_1,n_2,\ldots,n_t\}$, with $n_1<n_2<\cdots<n_t$. For $0\leq m_i\leq s_i$ given $(1\leq i\leq L)$, put

$$w_i=\begin{cases}s_i+1 & \text{if } i\in Q,\\ m_i & \text{if } i\in\Omega\backslash Q,\end{cases}\qquad \alpha_i=\begin{cases}y_i & \text{if } i\in Q,\\ z_i & \text{if } i\in\Omega\backslash Q.\end{cases}$$

For $f = f(z_1, \dots, z_L)$, we define

$$D(Q, \boldsymbol{s}, \boldsymbol{m})\ f(\boldsymbol{\alpha}) = \left\{ \prod_{i=1}^{L} \left(\frac{\partial}{\partial \alpha_i} \right)^{w_i} \right\} f(\boldsymbol{\alpha}).$$

For $\boldsymbol{h} = (h_1, \dots h_L)$ given,

$$\Theta(Q, \boldsymbol{s}, \boldsymbol{h}) := \prod_{i \in \Omega \setminus Q} \left\{ \sum_{m_i=0}^{s_i} \frac{h_i^{m_i}}{m_1!} \right\} \int_{z_{n_1}}^{z_{n_1}+h_{n_1}} \cdots \int_{z_{n_t}}^{z_{n_t}+h_{n_t}} \left\{ \prod_{j \in Q} \frac{(z_j + h_j - y_j)^{s_j}}{s_j!} \right\}$$
$$\times D(Q, \boldsymbol{s}, \boldsymbol{m}) f(\boldsymbol{\alpha})\, dy_{n_1} \cdots dy_{n_t}.$$

Finally, ${}_f R_{\boldsymbol{s}} := \sum_{Q \in \Omega} \Theta(Q, \boldsymbol{s}, \boldsymbol{h})$.

THEOREM 4.1. *Suppose $\prod_{i=1}^{L} \left(\partial/\partial z_i\right)^{q_i} f(z_1, \dots, z_L)$ exists and is jointly continuous in $z_1, \dots, z_L$. Then for $0 \le s_i \le q_i$ $(1 \le i \le L)$, we have*

$$f(\boldsymbol{z} + \boldsymbol{h}) = \sum_{m_1=0}^{s_1} \cdots \sum_{m_L=0}^{s_L} \frac{h_1^{m_1}}{m_1!} \cdots \frac{h_L^{m_L}}{m_L!} D_{\boldsymbol{z}}^{\boldsymbol{m}} f(\boldsymbol{z}) + {}_f R_{\boldsymbol{s}}.$$

PROOF. Like Theorem 3.1, this is established inductively. The only novelty is in obtaining the basis result for $\boldsymbol{s} = \boldsymbol{0}$. Put

$$\Omega_j = \{Q \in \Omega :\ \max\{i \in Q\} = j\} \quad (1 \le j \le L).$$

We employ an inner induction to show for $1 \le j \le L$ that

$$\sum_{Q \in \Omega_j} \Theta(Q, \boldsymbol{0}, \boldsymbol{h}) = f(z_1 + h_1, \dots, z_j + h_j, z_{j+1}, \dots, z_L) - f(z_1, \dots, z_L). \tag{4.1}$$

The result for $j = L$ then provides the basis for the outer induction.

We have

$$\Theta(Q, \boldsymbol{0}, \boldsymbol{h}) = \int_{z_{n_1}}^{z_{n_1}+h_{n_1}} \cdots \int_{z_{n_t}}^{z_{n_t}+h_{n_t}} \prod_{i \in Q} \left(\frac{\partial}{\partial y_i} \right) f(\boldsymbol{\alpha}) dy_{n_1} \dots dy_{n_t}. \tag{4.2}$$

Since $\Omega_1 = \{1\}$,

$$\sum_{Q \in \Omega_1} \Theta(Q, \boldsymbol{0}, \boldsymbol{h}) = \int_{z_1}^{z_1+h_1} \left(\frac{\partial}{\partial y_1} \right) f(y_1, z_2, \dots, z_L) dy_1$$
$$= f(z_1 + h_1, z_2, \dots, z_L) - f(z_1, \dots, z_L),$$

yielding a basis for the inner induction.

For the inductive step, suppose (4.1) applies for some j with $1 \le j < L$. The sets Q contributing to the $(j+1)$-th sum $\sum_{Q \in \Omega_{j+1}} \Theta(Q, \boldsymbol{0}, \boldsymbol{h})$ belong to three disjoint classes:

(a) Ω_j;
(b) sets of the form $Q \cup \{j+1\}$ with $Q \in \Omega_j$;
(c) the singleton class $\{j+1\}$.

By inductive assumption, the contribution from (a) is given by the right-hand side of (4.1). For any set in (b), for which the Q makes contribution $g(z_1, \ldots, z_L; \boldsymbol{h})$, say, the augmention $Q \cup \{j+1\}$ may be seen from (4.2) to make a contribution

$$g(z_1, \ldots, z_j, z_{j+1} + h_{j+1}, z_{j+2}, \ldots, z_L; \boldsymbol{h}) - g(z_1, \ldots, z_L; \boldsymbol{h}).$$

By the inductive assumption, the total contribution from (b) is thus

$$\begin{aligned}\big[f(z_1+h_1, \ldots, z_{j+1}+h_{j+1}, z_{j+2}, \ldots, z_L) - f(z_1+h_1, \ldots, z_j+h_j, z_{j+1}, \ldots, z_L)\big]\\ - \big[f(z_1, \ldots, z_j, z_{j+1}+h_{j+1}, z_{j+2}, \ldots, z_L) - f(z_1, \ldots, z_L)\big].\end{aligned} \tag{4.3}$$

The contribution from (c) is trivially

$$f(z_1, \ldots, z_j, z_{j+1}+h_{j+1}, z_{j+2}, \ldots, z_L) - f(z_1, \ldots, z_L). \tag{4.4}$$

Addition of the right-hand side of (4.1) to the expressions in (4.3) and (4.4) yield that the $(j+1)$-th sum is

$$\sum_{Q \in \Omega_{j+1}} \Theta(Q, \mathbf{0}, \boldsymbol{h}) = f(z_1+h_1, \ldots, z_{j+1}+h_{j+1}, z_{j+2}, \ldots, z_L) - f(z_1, \ldots, z_L).$$

This completes the inductive step for the inner induction and so the proof of the theorem. □

The notation of Section 3 extends as follows. We have $\boldsymbol{m} = (m_1, \ldots, m_M)$, $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_M)$, $\boldsymbol{p} = (p_1, \ldots, p_N)$ with $\sum_{i=1}^N p_i = p$,

$$A(\boldsymbol{p}) = \big(\{0, \ldots, p_1\} \times \ldots \times \{0, \ldots, p_N\}\big) \backslash \big(\{0\} \times \ldots \times \{0\}\big),$$

$$C(\boldsymbol{p}) = \left\{\boldsymbol{m} \in A(p, \ldots, p) : \sum_{\ell=1}^{M} m_\ell \le p\right\}.$$

As before, $\phi_i : A(\boldsymbol{p}) \to \{0, 1, \ldots, m_i\}$. Finally

$$V(\boldsymbol{m}) = \left\{\boldsymbol{\phi} \;\middle|\; \begin{array}{ll} \sum_{\boldsymbol{n} \in A(\boldsymbol{p})} \phi_i(\boldsymbol{n}) = m_i & (1 \le i \le M); \\ \sum_{\boldsymbol{n} \in A(\boldsymbol{p})} n_\ell \sum_{j=1}^{M} \phi_j(\boldsymbol{n}) = p_\ell & (1 \le \ell \le N) \end{array}\right\}.$$

The following theorem may now be established with an argument parallel to that of Theorem 3.3.

THEOREM 4.2. *Suppose $F(\boldsymbol{u})$ is $C^{(p+1)}$ and $u_i(z)$ $(i = 1, 2)$ have continuous derivatives to order $(p_1 + 1, \ldots, p_N + 1)$ on appropriate domains. Define*

$$B_{\phi_i}(u_i(z)) = \prod_{\boldsymbol{n}\in A(\boldsymbol{p})} \left\{ \frac{1}{\phi_i(\boldsymbol{n})!} \left(\frac{D_z^{\boldsymbol{n}} u_i(z)}{\prod_{\ell=1}^{M} n_\ell!} \right)^{\phi_i(\boldsymbol{n})} \right\} \quad (1 \leq i \leq M).$$

If $G(z) = F(\boldsymbol{u}(z))$ and $\boldsymbol{p} \neq \boldsymbol{0}$, then

$$\frac{\partial^p G(z)}{\partial z_1^{p_1} \ldots \partial z_N^{p_N}} = \left(\prod_{j=1}^{N} p_j! \right) \sum_{\boldsymbol{m}\in C(\boldsymbol{p})} D_{\boldsymbol{u}}^{\boldsymbol{m}} F \sum_{\boldsymbol{\phi}\in V(\boldsymbol{m})} \prod_{\ell=1}^{M} B_{\phi_\ell}(u_\ell(z)).$$

References

[1] R. Abraham and J. W. Robbin, *Transverse Mappings and Flows* (W. A. Benjamin, New York, 1967).

[2] L. F. A. Arbogast, *Du Calcul des Dérivations* (Levrault, Strasbourg, 1800).

[3] L. Bieberbach, *Differential Geometry* (Johnson Reprint Co., New York, 1968).

[4] G. M. Constantine and T. H. Savits, "A multivariate Faà di Bruno formula with applications", *Trans. Amer. Math. Soc.* **348** (1996) 503–520.

[5] A. D. D. Craik, "Prehistory of Faà di Bruno's formula", *Amer. Math. Month.* **112** (2005) 119–130.

[6] J. A. Dieudonne, *Infinitesimal Calculus* (Houghton Miffin, New York, 1971).

[7] F. Faà di Bruno, "Sullo sviluppo delle funzione", *Annali di Scienze Matematiche e Fisiche* **6** (1855) 479–480.

[8] F. Faà di Bruno, "Note sur une nouvelle formule de calcul différentiel", *Quart. J. Math.* **1** (1857) 359–360.

[9] H. Flanders, "From Ford to Faà", *Amer. Math. Month.* **108** (2001) 559–561.

[10] R. H. Fowler, *Statistical Mechanics* (Cambridge University Press, Cambridge, 1936).

[11] H. W. Gould, "The generalized chain rule of differentiation with historical notes", *Utilitas Mathematica* **61** (2002) 97–106.

[12] W. P. Johnson, "The curious history of Faà di Bruno's formula", *Amer. Math. Month.* **109** (2002) 217–234.

[13] R. Leipnik and T. Reid, "Multivariable Faà di Bruno formulas", *Elec. Proc. 9th Ann. Intern. Conf. Tech. Colleg. Math.* (1996), available at http://archives.math.utk.edu/ICTCM/EP-9.html#C23.

[14] E. Lukács, "Application of Faà di Bruno's formula in mathematical statistics", *Amer. Math. Month.* **62** (1955) 340–348.

[15] R. Mishkov, "Generalization of the formula of Faà di Bruno for a composite function with a vector argument", *Intern. J. Math. Sci.* **24** (2000) 481–491.

[16] S. Noschese and P. E. Ricci, "Differentiation of multivariable composite functions and Bell polynomials", *J. Comput. Anal. Appl.* **5** (2003) 333–340.

[17] Ch. J. de la Vallée Poussin, *Cours d'Analyse Infinitésimal* (Dover, New York, 1946).

[18] G.-C. Rota, "The number of partitions of a set", *Amer. Math. Month.* **71** (1964) 498–504.

[19] G.-C. Rota, "Geometric probability", *Math. Intell.* **20** (1998) 11–16.

[20] I. J. Schwatt, *Introduction to Operations with Series* (Chelsea, New York, 1962).

[21] L. Takács, *Introduction to the Theory of Queues* (Oxford University Press, Oxford, 1962).
[22] H. Whitney, *Geometric Integration Theory* (Princeton University Press, Princeton, 1957).
[23] C. S. Withers, "A chain rule for differentiation with application to multivariate hermite polynomials", *Bull. Austral. Math. Soc.* **30** (1984) 247–250.
[24] W. C. Yang, "Derivatives are essentially integer partitions", *Discr. Math.* **222** (2000) 235–245.