A MULTIVARIATE FAA DI BRUNO FORMULA
WITH APPLICATIONS

G. M. CONSTANTINE AND T. H. SAVITS

ABSTRACT. A multivariate Faa di Bruno formula for computing arbitrary partial derivatives of a function composition is presented. It is shown, by way of a general identity, how such derivatives can also be expressed in the form of an infinite series. Applications to stochastic processes and multivariate cumulants are then delineated.

1. INTRODUCTION AND NOTATION

The problem of finding an explicit expression for the nth derivative of a composition of functions has a long and distinguished history. According to Lukacs [7], the need for such a formula was explicitly mentioned as early as 1810 in Lacroix's treatise on calculus. Although several special cases were solved earlier, the first to obtain a general solution was C. F. Faa di Bruno [4].

The formula of Faa di Bruno is as follows. Let $g(x)$ be defined on a neighborhood of x^0 and have derivatives up to order n at x^0; let $f(y)$ be defined on a neighborhood of $y^0 = g(x^0)$ and have derivatives up to order n at y^0. Then the nth derivative of the composition $h(x) = f[g(x)]$ at x^0 is given by the formula

$$(1.1) \qquad h_n = \sum_{k=1}^{n} f_k \sum_{p(n,k)} n! \prod_{i=1}^{n} \frac{g_i^{\lambda_i}}{(\lambda_i!)(i!)^{\lambda_i}} .$$

In the above expression, we set

$$h_n = \frac{d^n}{dx^n} h(x^0), \ f_k = \frac{d^k}{dy^k} f(y^0), \ g_i = \frac{d^i}{dx^i} g(x^0)$$

and

$$p(n,k) = \{(\lambda_1, \dots, \lambda_n) : \lambda_i \in \mathcal{N}_0, \sum_{i=1}^{n} \lambda_i = k, \sum_{i=1}^{n} i\lambda_i = n\}$$

with $\mathcal{N}_0$ the set of nonnegative integers. A member $(\lambda_1, \dots, \lambda_n) \in p(n,k)$ represents a partition of a set with n elements into λ_1 classes of cardinality $1, \dots, \lambda_n$ classes

of cardinality n. The number of such partitions is denoted by S_n^k and is called a *Stirling number of the second kind:*

(1.2)
$$S_n^k = \sum_{p(n,k)} n! \prod_{i=1}^{n} \frac{1}{(\lambda_i!)(i!)^{\lambda_i}} .$$

It plays an important role in combinatorial theory; see Constantine [2].

Several interesting and useful applications of this formula have appeared in the literature. For example, Lukacs [9] uses it to relate the moments and cumulants of a random variable. In that same paper, the formula of Faa di Bruno is used to show that a population is normal if and only if the k-statistic of order p (any $p > 1$) is independent of the sample mean. More recently, Chen and Savits [1] use (1.1) to calculate arbitrary moments of a compound nonhomogeneous Poisson process. These ideas are further extended in Constantine and Savits [3] to obtain generalizations of Dobinski's formula regarding Bell numbers. Hsu [6] uses Faa di Bruno's formula to construct certain classes of identities.

We are interested in multivariate extensions of (1.1). More specifically, let $f(y_1, \ldots, y_m)$ and $g^{(1)}(x_1, \ldots x_d), \ldots, g^{(m)}(x_1, \ldots, x_d)$ be differentiable a sufficient number of times, and set

(1.3)
$$h(x_1, \ldots, x_d) = f[g^{(1)}(x_1, \ldots, x_d), \ldots, g^{(m)}(x_1, \ldots, x_d)] .$$

The task before us is to obtain an explicit expression for an arbitrary partial derivative of h in terms of the various partial derivatives of the functions f and $g^{(1)}, \ldots, g^{(m)}$. The special case when $d = 1$ and arbitrary m appears in Most [11], and in an even more restricted form in Hoppe [5]. To the best of our knowledge, the problem for a general function composition as stated in (1.3) has not been considered. We could not even trace the simpler case with $m = 1$ and arbitrary d.

The paper is organized as follows. Our multivariate Faa di Bruno formula is derived in Section 2. Some special cases are then delineated. The focus of Section 3 is on a general identity concerning series compositions. Noteworthy special cases of the identity are examined, leading to the definition of the multivariate Stirling numbers. Another application offers an infinite sum representation for an arbitrary partial derivative of a series composition. The last section describes some applications to multidimensional compound nonhomogeneous Poisson processes and to relationships between multivariate cumulants and moments.

In order to simplify expressions, it is convenient to recall some multivariate notation. If $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d) \in \mathcal{N}_0^d$ and $\mathbf{z} = (z_1, \ldots, z_d) \in \Re^d$, then

$$|\boldsymbol{\nu}| = \sum_{i=1}^{d} \nu_i ,$$

$$\boldsymbol{\nu}! = \prod_{i=1}^{d} (\nu_i!) ,$$

$$D_{\mathbf{x}}^{\boldsymbol{\nu}} = \frac{\partial^{|\boldsymbol{\nu}|}}{\partial x_1^{\nu_1} \cdots \partial x_d^{\nu_d}} , \text{ for } |\boldsymbol{\nu}| > 0 ,$$

$$D_{\mathbf{x}}^{\mathbf{0}} = \text{identity operator} ,$$

$$\mathbf{z}^{\boldsymbol{\nu}} = \prod_{i=1}^{d} z_i^{\nu_i},$$

and

$$\|\mathbf{z}\| = \max_{1 \le i \le d} |z_i| \; .$$

Also, if $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_d) \in \mathcal{N}_0^d$, we write $\boldsymbol{\ell} \le \boldsymbol{\nu}$ provided $\ell_i \le \nu_i$ for $i = 1, \ldots, d$. In such a case, we set

$$\binom{\boldsymbol{\nu}}{\boldsymbol{\ell}} = \prod_{i=1}^{d} \binom{\nu_i}{\ell_i} = \frac{\boldsymbol{\nu}!}{\boldsymbol{\ell}!(\boldsymbol{\nu} - \boldsymbol{\ell})!} \; .$$

A function h is said to belong to $\mathcal{C}_{\boldsymbol{\nu}}(\mathbf{x}^0)$ if $D_{\mathbf{x}}^{\boldsymbol{\ell}} h$ exists and is continuous in a neighborhood of $\mathbf{x}^0$ for all $\boldsymbol{\ell} \le \boldsymbol{\nu}$; we also write $h \in \mathcal{C}^n(\mathbf{x}^0)$ if $h \in \mathcal{C}_{\boldsymbol{\ell}}(\mathbf{x}^0)$ for all $|\boldsymbol{\ell}| \le n$.

Finally, we introduce a linear order on $\mathcal{N}_0^d$. If $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d)$ are in $\mathcal{N}_0^d$, we write $\boldsymbol{\mu} \prec \boldsymbol{\nu}$ provided one of the following holds:

    (i) $|\boldsymbol{\mu}| < |\boldsymbol{\nu}|$;

    (ii) $|\boldsymbol{\mu}| = |\boldsymbol{\nu}|$ and $\mu_1 < \nu_1$; or

    (iii) $|\boldsymbol{\mu}| = |\boldsymbol{\nu}|, \mu_1 = \nu_1, \ldots, \mu_k = \nu_k$ and $\mu_{k+1} < \nu_{k+1}$ for some $1 \le k < d$.

2. General statement and proof

As mentioned in Section 1, we aim to obtain an explicit expression for an arbitrary partial derivative of a composition of functions

$$h(x_1, \ldots, x_d) = f[g^{(1)}(x_1, \ldots, x_d), \ldots, g^{(m)}(x_1, \ldots, x_d)] \; .$$

Specifically, let $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d) \ne \mathbf{0} = (0, \ldots, 0)$ and $\mathbf{x}^0$ be given. Assume that we have $g^{(1)}, \ldots, g^{(m)} \in \mathcal{C}_{\boldsymbol{\nu}}(\mathbf{x}^0)$ and $f \in \mathcal{C}^n(\mathbf{y}^0)$, where $n = |\boldsymbol{\nu}|$ and $\mathbf{y}^0 = (g^{(1)}(\mathbf{x}^0), \ldots, g^{(m)}(\mathbf{x}^0))$. Under these conditions $D_{\mathbf{x}}^{\boldsymbol{\nu}} h(\mathbf{x}^0)$ exists and can be explicitly expressed as in Theorem 2.1 below.

Setting $h_{\boldsymbol{\nu}} = D_{\mathbf{x}}^{\boldsymbol{\nu}} h(\mathbf{x}^0), f_{\boldsymbol{\lambda}} = D_{\mathbf{y}}^{\boldsymbol{\lambda}} f(\mathbf{y}^0), g_{\boldsymbol{\mu}}^{(i)} = D_{\mathbf{x}}^{\boldsymbol{\mu}} g^{(i)}(\mathbf{x}^0), \mathbf{g}_{\boldsymbol{\mu}} = (g_{\boldsymbol{\mu}}^{(1)}, \ldots, g_{\boldsymbol{\mu}}^{(m)})$, we state our main result.

2.1. Theorem.

$$(2.1) \qquad h_{\boldsymbol{\nu}} = \sum_{1 \le |\boldsymbol{\lambda}| \le n} f_{\boldsymbol{\lambda}} \sum_{s=1}^{n} \sum_{p_s(\boldsymbol{\nu}, \boldsymbol{\lambda})} (\boldsymbol{\nu}!) \prod_{j=1}^{s} \frac{[\mathbf{g}_{\boldsymbol{\ell}_j}]^{\mathbf{k}_j}}{(\mathbf{k}_j!)[\boldsymbol{\ell}_j!]^{|\mathbf{k}_j|}} \; ,$$

where $n = |\boldsymbol{\nu}|$ *and*

$$(2.2) \qquad p_s(\boldsymbol{\nu}, \boldsymbol{\lambda}) = \{(\mathbf{k}_1, \ldots, \mathbf{k}_s; \; \boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_s) : |\mathbf{k}_i| > 0 \; ,$$

$$\mathbf{0} \prec \boldsymbol{\ell}_1 \prec \cdots \prec \boldsymbol{\ell}_s, \; \sum_{i=1}^{s} \mathbf{k}_i = \boldsymbol{\lambda} \; and \; \sum_{i=1}^{s} |\mathbf{k}_i| \boldsymbol{\ell}_i = \boldsymbol{\nu} \} \; .$$

In the above, the vectors $\mathbf{k}$ are m-dimensional, the vectors $\boldsymbol{\ell}$ are d-dimensional and we always set $0^0 = 1$.

2.2. *Remark.* (i) It is perhaps notationally less cumbersome to introduce the set

(2.3)
$$p(\boldsymbol{\nu}, \boldsymbol{\lambda}) = \{(\mathbf{k}_1 \ldots, \mathbf{k}_n; \boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_n) : \text{ for some } 1 \leq s \leq n,$$

$$\mathbf{k}_i = \mathbf{0} \text{ and } \boldsymbol{\ell}_i = \mathbf{0} \text{ for } 1 \leq i \leq n - s; |\mathbf{k}_i| > 0 \text{ for } n - s + 1 \leq i \leq n;$$

$$\text{and } \mathbf{0} \prec \boldsymbol{\ell}_{n-s+1} \prec \cdots \prec \boldsymbol{\ell}_n \text{ are such that}$$

$$\sum_{i=1}^{n} \mathbf{k}_i = \boldsymbol{\lambda}, \sum_{i=1}^{n} |\mathbf{k}_i| \boldsymbol{\ell}_i = \boldsymbol{\nu} \}.$$

In partial analogy to the univariate case, one may view the $\boldsymbol{\ell}_i$'s as the "parts" of $\boldsymbol{\nu}$ and the $|\mathbf{k}_i|$'s as their respective multiplicities. Clearly $p(\boldsymbol{\nu}, \boldsymbol{\lambda})$ can be identified with the union of $p_s(\boldsymbol{\nu}, \boldsymbol{\lambda})$ for $s = 1, \ldots, n$. We may then write formula (2.1) more succinctly as

$$(2.4) \qquad h_{\boldsymbol{\nu}} = \sum_{1 \leq |\boldsymbol{\lambda}| \leq n} f_{\boldsymbol{\lambda}} \sum_{p(\boldsymbol{\nu}, \boldsymbol{\lambda})} (\boldsymbol{\nu}!) \prod_{j=1}^{n} \frac{[\mathbf{g}_{\boldsymbol{\ell}_j}]^{\mathbf{k}_j}}{(\mathbf{k}_j!) \, [\boldsymbol{\ell}_j!]^{|\mathbf{k}_j|}} .$$

(ii) Formula (2.1), or (2.4), is of course valid in a neighborhood of $\mathbf{x}^0$. For clarity of exposition, we shall sometimes distinguish this situation by writing $h_{\boldsymbol{\nu}}(\mathbf{x}) = D_{\mathbf{x}}^{\boldsymbol{\nu}} h(\mathbf{x})$, $f_{\boldsymbol{\lambda}}(\mathbf{y}) = D_{\mathbf{y}}^{\boldsymbol{\lambda}} f(\mathbf{y})$ and $g_{\boldsymbol{\nu}}^{(i)}(\mathbf{x}) = D_{\mathbf{x}}^{\boldsymbol{\nu}} g^{(i)}(\mathbf{x})$.

Our formula is established through a sequence of lemmas by extending the method of proof given in Constantine [2] for the univariate Faa di Bruno's formula (1.1).

2.3. Lemma. *For* $\mathbf{0} \neq \boldsymbol{\omega} \leq \boldsymbol{\nu}$,

$$h_{\boldsymbol{\omega}}(\mathbf{x}) = \sum_{1 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\omega}|} f_{\boldsymbol{\lambda}}[\mathbf{g}(\mathbf{x})] \alpha_{\boldsymbol{\omega}, \boldsymbol{\lambda}}(\mathbf{x})$$

where $\alpha_{\boldsymbol{\omega}, \boldsymbol{\lambda}}(\mathbf{x})$ *is independent of the choice of* $f \in \mathcal{C}^n(\mathbf{y}^0)$ *and is expressible as a sum of terms of the form* $\prod_{k=1}^{|\boldsymbol{\lambda}|} g_{\boldsymbol{\mu}_k}^{(i_k)}(\mathbf{x})$ *with* $1 \leq i_k \leq m$ *and* $\mathbf{0} \neq \boldsymbol{\mu}_k \leq \boldsymbol{\omega}$.

Proof. Let $\mathbf{e}_j$ be a unit vector whose jth component is 1 and the remaining entries are 0, and suppose $\mathbf{e}_j \leq \boldsymbol{\nu}$. Then

$$h_{\mathbf{e}_j}(\mathbf{x}) = \sum_{\alpha=1}^{m} f_{\mathbf{e}_\alpha}[\mathbf{g}(\mathbf{x})] g_{\mathbf{e}_j}^{(\alpha)}(\mathbf{x}) .$$

Since $h_{\boldsymbol{\mu}+\mathbf{e}_j}(\mathbf{x}) = D_{\mathbf{x}}^{\mathbf{e}_j} h_{\boldsymbol{\mu}}(\mathbf{x})$, the result easily follows by induction. $\square$

2.4. *Remark.* It is clear from Lemma 2.3 that in order to compute $h_{\boldsymbol{\nu}} = h_{\boldsymbol{\nu}}(\mathbf{x}^0)$, we only need to determine the values of $\alpha_{\boldsymbol{\nu}, \boldsymbol{\lambda}} = \alpha_{\boldsymbol{\nu}, \boldsymbol{\lambda}}(\mathbf{x}^0)$. Since these values depend only on the values of $g_{\boldsymbol{\mu}}^{(i)} = g_{\boldsymbol{\mu}}^{(i)}(\mathbf{x}^0)$ for $1 \leq i \leq m, 0 \neq \boldsymbol{\mu} \leq \boldsymbol{\nu}$, we obtain the same values for any functions $\tilde{g}^{(i)} \in \mathcal{C}_{\boldsymbol{\nu}}(\mathbf{x}^0)$ provided $\tilde{g}_{\boldsymbol{\mu}}^{(i)}(\mathbf{x}^0) = g_{\boldsymbol{\mu}}^{(i)}(\mathbf{x}^0)$ for $\mathbf{0} \neq \boldsymbol{\mu} \leq \boldsymbol{\nu}$, each $1 \leq i \leq m$. Consequently, without loss of generality, we assume that each $g^{(i)}$ is infinitely differentiable and that $\max_{1 \leq i \leq m} \sup_{|\boldsymbol{\mu}|>0} |g_{\boldsymbol{\mu}}^{(i)}(\mathbf{x})| \leq M < \infty$ in a neighborhood of $\mathbf{x}^0$. Specifically, we replace $g^{(i)}$ with the polynomial

$$\tilde{g}^{(i)}(\mathbf{x}) = \sum_{\mathbf{0} \leq \boldsymbol{\mu} \leq \boldsymbol{\nu}} g_{\boldsymbol{\mu}}^{(i)} \frac{(\mathbf{x} - \mathbf{x}^0)^{\boldsymbol{\mu}}}{\boldsymbol{\mu}!} .$$

Although Theorem 2.1 concerns a specific $\boldsymbol{\nu}$, we now consider all $\boldsymbol{\nu} \geq \mathbf{0}$ and define for $\mathbf{z} \in \Re^m$,

$$B_{\boldsymbol{\nu}}(\mathbf{z}; \mathbf{x}) = \left\{ \begin{array}{ll} \sum_{1 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}|} \alpha_{\boldsymbol{\nu}, \boldsymbol{\lambda}}(\mathbf{x}) \, \mathbf{z}^{\boldsymbol{\lambda}} & \text{if } |\boldsymbol{\nu}| > 0, \\ 1 & \text{if } |\boldsymbol{\nu}| = 0 \ . \end{array} \right.$$

According to our convention, we write $B_{\boldsymbol{\nu}}(\mathbf{z})$ for $B_{\boldsymbol{\nu}}(\mathbf{z}; \mathbf{x}^0)$.

2.5. Lemma. $B_{\boldsymbol{\nu}}(\mathbf{z}; \mathbf{x}) = \exp\{-\sum_{i=1}^{m} z_i \, g^{(i)}(\mathbf{x})\} D_{\mathbf{x}}^{\boldsymbol{\nu}} [\exp\{\sum_{i=1}^{m} z_i \, g^{(i)}(\mathbf{x})\}]$.

Proof. If $f(\mathbf{y}) = f(y_1, \ldots, y_m) = \exp\{\sum_{i=1}^{m} z_i y_i\}$, then $D_{\mathbf{y}}^{\boldsymbol{\lambda}} f(\mathbf{y}) = \mathbf{z}^{\boldsymbol{\lambda}} f(\mathbf{y})$ for all $\boldsymbol{\lambda}$. Hence for $h(\mathbf{x}) = f[g^{(1)}(\mathbf{x}), \ldots, g^{(m)}(\mathbf{x})] = \exp\{\sum_{i=1}^{m} z_i \, g^{(i)}(\mathbf{x})\}$, we obtain $f_{\boldsymbol{\lambda}}[\mathbf{g}(\mathbf{x})] = D_{\mathbf{y}}^{\boldsymbol{\lambda}} f[\mathbf{g}(\mathbf{x})] = \mathbf{z}^{\boldsymbol{\lambda}} h(\mathbf{x})$. Thus, by Lemma 2.3, for $|\boldsymbol{\nu}| > 0$

$$h_{\boldsymbol{\nu}}(\mathbf{x}) = \sum_{1 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}|} f_{\boldsymbol{\lambda}}[\mathbf{g}(\mathbf{x})] \, \alpha_{\boldsymbol{\nu}, \boldsymbol{\lambda}}(\mathbf{x})$$

$$= h(\mathbf{x}) \sum_{1 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}|} \mathbf{z}^{\boldsymbol{\lambda}} \, \alpha_{\boldsymbol{\nu}, \boldsymbol{\lambda}}(\mathbf{x}) = h(\mathbf{x}) \, B_{\boldsymbol{\nu}}(\mathbf{z}; \mathbf{x}) \ .$$

The result now follows. □

2.6. Lemma. $B_{\boldsymbol{\nu}+\mathbf{e}_j}(\mathbf{z}; \mathbf{x}) = \sum_{i=1}^{m} z_i [\sum_{\mathbf{0} \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{\ell}} g_{\boldsymbol{\ell}+\mathbf{e}_j}^{(i)} (\mathbf{x}) \, B_{\boldsymbol{\nu}-\boldsymbol{\ell}}(\mathbf{z}; \mathbf{x})]$.

Proof. We need to recall the multivariate rendition of Leibnitz's formula which provides an expression for the partial derivatives of a product of two functions. Specifically, if $F, G \in \mathcal{C}_{\boldsymbol{\nu}}(\mathbf{x}^0)$, then in a neighborhood of $\mathbf{x}^0$,

$$D_{\mathbf{x}}^{\boldsymbol{\nu}}(FG)(\mathbf{x}) = \sum_{\mathbf{0} \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{\ell}} D_{\mathbf{x}}^{\boldsymbol{\ell}} F(\mathbf{x}) \, D_{\mathbf{x}}^{\boldsymbol{\nu}-\boldsymbol{\ell}} G(\mathbf{x}) \ .$$

This formula is easily obtained by induction. Using Lemma 2.5, we can now write

$$B_{\boldsymbol{\nu}+\mathbf{e}_j}(\mathbf{z}; \mathbf{x}) = \exp\{-\sum_{k=1}^{m} z_k g^{(k)}(\mathbf{x})\} \, D_{\mathbf{x}}^{\boldsymbol{\nu}+\mathbf{e}_j} [\exp\{\sum_{k=1}^{m} z_k g^{(k)}(\mathbf{x})\}]$$

$$= \exp\{-\sum_{k=1}^{m} z_k \, g^{(k)}(\mathbf{x})\} \, D_{\mathbf{x}}^{\boldsymbol{\nu}} [(\sum_{i=1}^{m} z_i g_{\mathbf{e}_j}^{(i)}(\mathbf{x})) \, \exp\{\sum_{k=1}^{m} z_k g^{(k)}(\mathbf{x})\}]$$

$$= \exp\{-\sum_{k=1}^{m} z_k g^{(k)}(\mathbf{x})\}$$

$$\times \left(\sum_{i=1}^{m} z_i \sum_{\mathbf{0} \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{\ell}} g_{\boldsymbol{\ell}+\mathbf{e}_j}^{(i)} (\mathbf{x}) \, D_{\mathbf{x}}^{\boldsymbol{\nu}-\boldsymbol{\ell}} [\exp\{\sum_{k=1}^{m} z_k g^{(k)}(\mathbf{x})\}] \right)$$

$$= \sum_{i=1}^{m} z_i [\sum_{\mathbf{0} \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{\ell}} g_{\boldsymbol{\ell}+\mathbf{e}_j}^{(i)} (\mathbf{x}) \, B_{\boldsymbol{\nu}-\boldsymbol{\ell}}(\mathbf{z}; \mathbf{x})] \ .$$

This ends the proof. □

The next step involves the consideration of two power series in $\mathbf{w} \in \Re^d$:

(2.5) $$\psi(\mathbf{w}) = \sum_{|\boldsymbol{\mu}| > 0} (\sum_{i=1}^{m} z_i \, g_{\boldsymbol{\mu}}^{(i)} \, \frac{\mathbf{w}^{\boldsymbol{\mu}}}{\boldsymbol{\mu}!})$$

and

$$\phi(\mathbf{w}) = \sum_{\boldsymbol{\nu} \geq \mathbf{0}} B_{\boldsymbol{\nu}}(\mathbf{z}) \, \frac{\mathbf{w}^{\boldsymbol{\nu}}}{\boldsymbol{\nu}!} \; . \tag{2.6}$$

Note that $\mathbf{w}$ and $\mathbf{z}$ are variables, but we have substituted $\mathbf{x} = \mathbf{x}^0$. In order to justify the remaining steps we need to show that these series converge in a neighborhood of $(\mathbf{w}, \mathbf{z}) = (\mathbf{0}, \mathbf{0})$.

First consider the series (2.5). By Remark 2.4, we may assume that

$$\max_{1 \leq i \leq m} \; \sup_{|\boldsymbol{\mu}| > 0} |g_{\boldsymbol{\mu}}^{(i)}| \leq M < \infty.$$

Hence, for $||\mathbf{z}|| \leq R, ||\mathbf{w}|| \leq r$,

$$|\psi(\mathbf{w})| \leq (mMR) \sum_{|\boldsymbol{\mu}| > 0} \frac{r^{|\boldsymbol{\mu}|}}{\boldsymbol{\mu}!} \leq (mMR)e^{rd} \; ,$$

any $r, R > 0$.

The convergence of the series (2.6), however, is less straightforward. We need a bound on $B_{\boldsymbol{\nu}}(\mathbf{z}) = \sum_{1 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}|} \alpha_{\boldsymbol{\nu}, \boldsymbol{\lambda}} \mathbf{z}^{\boldsymbol{\lambda}}$. In Lemma 2.3 it was shown that $\alpha_{\boldsymbol{\nu}, \boldsymbol{\lambda}}$ is expressible as a sum of terms of the form $\prod_{j=1}^{|\boldsymbol{\lambda}|} g_{\boldsymbol{\mu}_j}^{(i_j)}$. Hence each such term is bounded by $M^{|\boldsymbol{\lambda}|}$. Thus we need to bound the number of terms in $\alpha_{\boldsymbol{\nu}, \boldsymbol{\lambda}}$. It is easier, however, to bound the number of terms in $\sum_{|\boldsymbol{\lambda}|=k} \alpha_{\boldsymbol{\nu}, \boldsymbol{\lambda}}$ instead. We denote this number by $\xi(\boldsymbol{\nu}, k)$.

2.7. Lemma. *The number of terms $\xi(\boldsymbol{\nu}, k)$ depends on $\boldsymbol{\nu}$ through $|\boldsymbol{\nu}|$ only; we thus write $\xi(\boldsymbol{\nu}, k) = \xi_n(k)$ when $|\boldsymbol{\nu}| = n$. Furthermore, $\xi_n(k) = m^k S_n^k$ for $1 \leq k \leq n$, where S_n^k is a Stirling number of the second kind (see (1.2)).*

Proof. Consider $|\boldsymbol{\nu}| = 1$ so that $\boldsymbol{\nu} = \mathbf{e}_j$ for some $1 \leq j \leq d$. Then

$$h_{\mathbf{e}_j} = \sum_{\alpha=1}^{m} f_{\mathbf{e}_j} \, g_{\mathbf{e}_j}^{(\alpha)} \; ,$$

and therefore $\xi(\mathbf{e}_j, 1) = m$, independent of j; i.e., $\xi_1(1) = m = m S_1^1$. Suppose the result holds for all $\boldsymbol{\mu}$ and k with $1 \leq k \leq |\boldsymbol{\mu}| \leq n$. Let $|\boldsymbol{\nu}| = n + 1$, so that $\boldsymbol{\nu} = \boldsymbol{\mu} + \mathbf{e}_j$ for some $1 \leq j \leq d$ and $\boldsymbol{\mu}$ with $|\boldsymbol{\mu}| = n$. We then write

$$h_{\boldsymbol{\nu}} = D_{\mathbf{x}}^{\mathbf{e}_j} \, h_{\boldsymbol{\mu}}(\mathbf{x}^0) = D_{\mathbf{x}}^{\mathbf{e}_j} \Big[\sum_{1 \leq |\boldsymbol{\lambda}| \leq n} f_{\boldsymbol{\lambda}}[\mathbf{g}(\mathbf{x})] \alpha_{\boldsymbol{\mu}, \boldsymbol{\lambda}}(\mathbf{x}) \Big] \Big|_{\mathbf{x}=\mathbf{x}^0}$$

$$= \sum_{k=1}^{n} \sum_{|\boldsymbol{\lambda}|=k} D_{\mathbf{x}}^{\mathbf{e}_j} \, [f_{\boldsymbol{\lambda}}[\mathbf{g}(\mathbf{x})] \, \alpha_{\boldsymbol{\mu}, \boldsymbol{\lambda}}(\mathbf{x})] \Big|_{\mathbf{x}=\mathbf{x}^0}$$

$$= \sum_{k=1}^{n} \sum_{|\boldsymbol{\lambda}|=k} \Big\{ \alpha_{\boldsymbol{\mu}, \boldsymbol{\lambda}} \, \big(\sum_{\alpha=1}^{m} f_{\boldsymbol{\lambda}+\mathbf{e}_\alpha} \, g_{\mathbf{e}_j}^{(\alpha)} \big) + f_{\boldsymbol{\lambda}} \, [D_{\mathbf{x}}^{\mathbf{e}_j} \, \alpha_{\boldsymbol{\mu}, \boldsymbol{\lambda}}](\mathbf{x}^0) \Big\} \; .$$

But according to Lemma 2.3, each term of $\alpha_{\boldsymbol{\mu}, \boldsymbol{\lambda}}(\mathbf{x})$ for $|\boldsymbol{\lambda}| = k$ is a product of k factors of the form $\prod_{\ell=1}^{k} g_{\boldsymbol{\mu}_\ell}^{(i_\ell)}(\mathbf{x})$ and thus $D_{\mathbf{x}}^{\mathbf{e}_j}$ applied to this product contributes k terms, namely,

$$D_{\mathbf{x}}^{\mathbf{e}_j} \Big[\prod_{\ell=1}^{k} g_{\boldsymbol{\mu}_\ell}^{(i_\ell)}(\mathbf{x}) \Big] \Big|_{\mathbf{x}=\mathbf{x}^0} = \sum_{\ell=1}^{k} g_{\boldsymbol{\mu}_\ell+\mathbf{e}_j}^{(i_\ell)} \prod_{r \neq \ell} g_{\boldsymbol{\mu}_r}^{(i_r)} \; .$$

Hence it follows that $\xi(\boldsymbol{\nu}, k)$ only depends on $\boldsymbol{\nu}$ through $|\boldsymbol{\nu}| = n + 1$. If we set $\xi_n(k) = m^k s_n(k)$, we obtain the following recursion formula :

$$
\begin{aligned}
s_{n+1}(k) &= s_n(k-1) + k \; s_n(k), \qquad 1 < k \leq n \; , \\
s_{n+1}(1) &= s_n(1), \\
s_{n+1}(n+1) &= s_n(n) \; ,
\end{aligned}
$$

with the initial condition $s_1(1) = 1$. The solution to this system of equations is given by $s_n(k) = S_n^k$ (see, e.g., Constantine [2], page 9). Hence $\xi_n(k) = m^k S_n^k$. $\quad\square$

Lemma 2.7 now allows us to bound the series (2.6). Specifically, if $||\mathbf{z}|| \leq R$, then for $|\boldsymbol{\nu}| = n$,

$$
|B_{\boldsymbol{\nu}}(\mathbf{z})| \leq \sum_{k=1}^{n} (RM)^k \xi_n(k) = \sum_{k=1}^{n} (mRM)^k \; S_n^k \; .
$$

Consequently, for $||\mathbf{z}|| \leq R$, $||\mathbf{w}|| \leq r$, and using the identity (see Constantine [2], page 41),

$$
\sum_{k=1}^{n} t^k \; S_n^k = e^{-t} \sum_{\ell=1}^{\infty} \ell^n \; \frac{t^\ell}{\ell!} \; ,
$$

we obtain the bound

$$
\begin{aligned}
|\phi(\mathbf{w})| &\leq 1 \; + \; \sum_{n=1}^{\infty} \sum_{|\boldsymbol{\nu}|=n} \frac{r^n}{\boldsymbol{\nu}!} \; e^{-mRM} \sum_{\ell=1}^{\infty} \ell^n \; \frac{(mMR)^\ell}{\ell!} \\
&= 1 + e^{-mRM} \sum_{\ell=1}^{\infty} \frac{(mMR)^\ell}{\ell!} \sum_{n=1}^{\infty} \sum_{|\boldsymbol{\nu}|=n} \frac{(r\ell)^n}{\boldsymbol{\nu}!} \\
&\leq 1 + e^{-mRM} \sum_{\ell=0}^{\infty} \frac{(mRM)^\ell}{\ell!} \; e^{rd\ell} \\
&= 1 + \exp\{mMR(e^{rd} - 1)\} \; . \quad\square
\end{aligned}
$$

2.8. Lemma. *With the notation of (2.5) and (2.6), $\psi(\mathbf{w}) = \ln \phi(\mathbf{w})$ for $\mathbf{w}$ in a neighborhood of zero.*

Proof. Since $\phi(\mathbf{0}) = 1$, by continuity we can find $r > 0$ such that $||\mathbf{w}|| \leq r$ implies $\phi(\mathbf{w}) \geq \frac{1}{2}$. It then suffices to show that they have the same gradient there since $\psi(\mathbf{0}) = \ln \phi(\mathbf{0}) = 0$.
    Now if $1 \leq j \leq d$,

$$
\begin{aligned}
D_{\mathbf{w}}^{\mathbf{e}_j} \; [\ln \phi(\mathbf{w})] &= \frac{1}{\phi(\mathbf{w})} \sum_{\substack{\boldsymbol{\nu} \geq \mathbf{0} \\ \nu_j > 0}} B_{\boldsymbol{\nu}}(\mathbf{z}) \; \frac{\mathbf{w}^{\boldsymbol{\nu}-\mathbf{e}_j}}{(\boldsymbol{\nu} - \mathbf{e}_j)!} \\
&= \frac{1}{\phi(\mathbf{w})} \sum_{\boldsymbol{\nu} \geq \mathbf{0}} B_{\boldsymbol{\nu}+\mathbf{e}_j} \; (\mathbf{z}) \; \frac{\mathbf{w}^{\boldsymbol{\nu}}}{\boldsymbol{\nu}!}
\end{aligned}
$$

and

$$D_{\mathbf{w}}^{\mathbf{e}_j} \; \psi(\mathbf{w}) = \sum_{\substack{|\boldsymbol{\mu}|>0 \\ \mu_j>0}} (\sum_{i=1}^{m} z_i \; g_{\boldsymbol{\mu}}^{(i)}) \; \frac{\mathbf{w}^{\boldsymbol{\mu}-\mathbf{e}_j}}{(\boldsymbol{\mu}-\mathbf{e}_j)!}$$

$$= \sum_{\boldsymbol{\mu} \geq \mathbf{0}} (\sum_{i=1}^{m} z_i \; g_{\boldsymbol{\mu}+\mathbf{e}_j}^{(i)}) \; \frac{\mathbf{w}^{\boldsymbol{\mu}}}{\boldsymbol{\mu}!} \; .$$

But according to Lemma 2.6,

$$\sum_{\boldsymbol{\nu} \geq \mathbf{0}} B_{\boldsymbol{\nu}+\mathbf{e}_j} \; (\mathbf{z}) \; \frac{\mathbf{w}^{\boldsymbol{\nu}}}{\boldsymbol{\nu}!} = \sum_{\boldsymbol{\nu} \geq \mathbf{0}} (\sum_{i=1}^{m} z_i [\sum_{0 \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{\ell}} g_{\boldsymbol{\ell}+\mathbf{e}_j}^{(i)} \; B_{\boldsymbol{\nu}-\boldsymbol{\ell}} \; (\mathbf{z})]) \; \frac{\mathbf{w}^{\boldsymbol{\nu}}}{\boldsymbol{\nu}!}$$

$$= \sum_{\boldsymbol{\nu} \geq \mathbf{0}} \left(\sum_{0 \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{\ell}} (\sum_{i=1}^{m} z_i \; g_{\boldsymbol{\ell}+\mathbf{e}_j}^{(i)}) \; B_{\boldsymbol{\nu}-\boldsymbol{\ell}}(\mathbf{z}) \right) \; \frac{\mathbf{w}^{\boldsymbol{\nu}}}{\boldsymbol{\nu}!}$$

$$= [\sum_{\boldsymbol{\nu} \geq \mathbf{0}} B_{\boldsymbol{\nu}} \; (\mathbf{z}) \; \frac{\mathbf{w}^{\boldsymbol{\nu}}}{\boldsymbol{\nu}!}] \; [\sum_{\boldsymbol{\mu} \geq \mathbf{0}} (\sum_{i=1}^{m} z_i \; g_{\boldsymbol{\mu}+\mathbf{e}_j}^{(i)}) \; \frac{\mathbf{w}^{\boldsymbol{\mu}}}{\boldsymbol{\mu}!}]$$

$$= \phi(\mathbf{w}) \; D_{\mathbf{w}}^{\mathbf{e}_j} \; \psi(\mathbf{w}) \; .$$

Consequently, $D_{\mathbf{w}}^{\mathbf{e}_j}[\ln \phi(\mathbf{w})] = D_{\mathbf{w}}^{\mathbf{e}_j} \; \psi(\mathbf{w})$ and we are done. $\qquad\square$

We are now ready to prove Theorem 2.1.

Proof of Theorem 2.1. Since $\ln \phi(\mathbf{w}) = \psi(\mathbf{w})$ in a neighborhood of $\mathbf{w} = \mathbf{0}$, the result remains valid if we exponentiate both sides. Hence

$$\sum_{\boldsymbol{\nu} \geq \mathbf{0}} B_{\boldsymbol{\nu}}(\mathbf{z}) \frac{\mathbf{w}^{\boldsymbol{\nu}}}{\boldsymbol{\nu}!} = \exp\{ \sum_{|\boldsymbol{\mu}|>0} (\sum_{i=1}^{m} z_i \; g_{\boldsymbol{\mu}}^{(i)}) \; \frac{\mathbf{w}^{\boldsymbol{\mu}}}{\boldsymbol{\mu}!} \}$$

$$= \prod_{|\boldsymbol{\mu}|>0} \exp\{ (\sum_{i=1}^{m} z_i \; g_{\boldsymbol{\mu}}^{(i)}) \; \frac{\mathbf{w}^{\boldsymbol{\mu}}}{\boldsymbol{\mu}!} \}$$

$$= \prod_{|\boldsymbol{\mu}|>0} \{ \sum_{\mathbf{k} \geq \mathbf{0}} \frac{\mathbf{z}^{\mathbf{k}} [\mathbf{g}_{\boldsymbol{\mu}}]^{\mathbf{k}} \; \mathbf{w}^{|\mathbf{k}|\boldsymbol{\mu}}}{\mathbf{k}!(\boldsymbol{\mu}!)^{|\mathbf{k}|}} \}$$

$$= \sum_{\boldsymbol{\nu} \geq \mathbf{0}} (\sum_{p(\boldsymbol{\nu})} \prod_{i=1}^{q} \frac{\mathbf{z}^{\mathbf{k}_i} \; [\mathbf{g}\boldsymbol{\ell}_i]^{\mathbf{k}_i}}{\mathbf{k}_i!(\boldsymbol{\ell}_i!)^{|\mathbf{k}_i|}}) \; \mathbf{w}^{\boldsymbol{\nu}} \; ,$$

where $q = q(\boldsymbol{\nu}), \mathbf{0} \prec \boldsymbol{\ell}_1 \prec \cdots \prec \boldsymbol{\ell}_q$ is a complete ordered listing of all nonzero vectors $\boldsymbol{\ell} \leq \boldsymbol{\nu}$ and

$$p(\boldsymbol{\nu}) = \{ (\mathbf{k}_1, \dots, \mathbf{k}_q) : \mathbf{k}_i \geq \mathbf{0} \text{ and } \sum_{i=1}^{q} |\mathbf{k}_i| \boldsymbol{\ell}_i = \boldsymbol{\nu} \} \; .$$

The above can be justified by first considering the finite product $\prod_{0<|\boldsymbol{\mu}| \leq n}$ and then passing to a limit as $n \to \infty$.

Since both power series converge in a neighborhood of $\mathbf{w} = \mathbf{0}$, we can equate coefficients to obtain

$$B_{\boldsymbol{\nu}}(\mathbf{z}) = \boldsymbol{\nu}! \sum_{p(\boldsymbol{\nu})} \prod_{i=1}^{q} \frac{\mathbf{z}^{\mathbf{k}_i} \, [\mathbf{g}_{\boldsymbol{\ell}_i}]^{\mathbf{k}_i}}{\mathbf{k}_i! \, (\boldsymbol{\ell}_i!)^{|\mathbf{k}_i|}}$$

$$= \boldsymbol{\nu}! \sum_{p(\boldsymbol{\nu})} \mathbf{z}^{\mathbf{k}_1 + \cdots + \mathbf{k}_q} \, (\prod_{i=1}^{q} \frac{[\mathbf{g}_{\boldsymbol{\ell}_i}]^{\mathbf{k}_i}}{\mathbf{k}_i!(\boldsymbol{\ell}_i!)^{|\mathbf{k}_i|}}) \; .$$

But if $(\mathbf{k}_1, \ldots, \mathbf{k}_q) \in p(\boldsymbol{\nu})$, then setting $\boldsymbol{\lambda} = \mathbf{k}_1 + \cdots + \mathbf{k}_q$ and $n = |\boldsymbol{\nu}|$ we see that

$$n = |\boldsymbol{\nu}| = \sum_{i=1}^{q} |\mathbf{k}_i||\boldsymbol{\ell}_i| \geq \sum_{i=1}^{q} |\mathbf{k}_i| = |\boldsymbol{\lambda}| \; .$$

Consequently, $|\boldsymbol{\lambda}| \leq n$ and there are at most n nonzero $\mathbf{k}_i$. Also, since not all $\mathbf{k}_i$ are zero, $|\boldsymbol{\lambda}| \geq 1$. Thus it is easy to see that we can rewrite the above as

(2.7)

$$\sum_{1 \leq |\boldsymbol{\lambda}| \leq n} \alpha_{\boldsymbol{\nu}, \boldsymbol{\lambda}} \, \mathbf{z}^{\boldsymbol{\lambda}} = B_{\boldsymbol{\nu}}(\mathbf{z}) = \boldsymbol{\nu}! \sum_{1 \leq |\boldsymbol{\lambda}| \leq n} \mathbf{z}^{\boldsymbol{\lambda}} \, (\sum_{s=1}^{n} \sum_{p_s(\boldsymbol{\nu}, \boldsymbol{\lambda})} \prod_{i=1}^{s} \frac{[\mathbf{g}_{\boldsymbol{\ell}_i}]^{\mathbf{k}_i}}{(\mathbf{k}_i!)[\boldsymbol{\ell}_i!]^{|\mathbf{k}_i|}}) \; .$$

Since this is valid in a neighborhood of $\mathbf{z} = \mathbf{0}$, we can equate coefficients and combine with Lemma 2.3 to conclude the proof of Theorem 2.1. $\qquad\square$

2.9. Corollary. *If $1 \leq k \leq n = |\boldsymbol{\nu}|$, then*

(2.8) $$\boldsymbol{\nu}! \sum_{|\boldsymbol{\lambda}|=k} \sum_{p(\boldsymbol{\nu}, \boldsymbol{\lambda})} \prod_{j=1}^{n} \frac{1}{(\mathbf{k}_j!)[\boldsymbol{\ell}_j!]^{|\mathbf{k}_j|}} = m^k S_n^k \; .$$

Proof. Define $g^{(i)}(x_1, \ldots, x_d) = \exp\{\sum_{j=1}^{d} x_j\}$ for $i = 1, \ldots, m$ and set $\mathbf{x}^0 = \mathbf{0}$. Then from (2.7) it follows that

$$B_{\boldsymbol{\nu}}(t\mathbf{e}) = \boldsymbol{\nu}! \sum_{k=1}^{n} t^k \, (\sum_{|\boldsymbol{\lambda}|=k} \sum_{p(\boldsymbol{\nu}, \boldsymbol{\lambda})} \prod_{j=1}^{n} \frac{1}{(\mathbf{k}_j!) \, [\boldsymbol{\ell}_j!]^{|\mathbf{k}_j|}}) \; ,$$

for $t \in \Re$ and $\mathbf{e} = (1, 1, \ldots, 1)$.

On the other hand, according to Lemma 2.5,

$$B_{\boldsymbol{\nu}}(\mathbf{z}; \mathbf{0}) = \exp\{-\sum_{i=1}^{m} z_i\} \, D_{\mathbf{x}}^{\boldsymbol{\nu}} \, [\exp\{\sum_{i=1}^{m} z_i \, g^{(i)}(\mathbf{x})\}]\Big|_{\mathbf{x}=\mathbf{0}} \; .$$

But

$$\exp\{\sum_{i=1}^{m} z_i g^{(i)}(\mathbf{x})\} = h(x_1 + \cdots + x_d),$$

where we set $h(u) = \exp\{(z_1 + \cdots + z_m)e^u\}$. Clearly,

$$D_{\mathbf{x}}^{\boldsymbol{\nu}}[\exp\{\sum_{i=1}^{m} z_i \, g^{(i)}(\mathbf{x})\}]\Big|_{\mathbf{x}=\mathbf{0}} = D_u^n \, h(0) \; ,$$

and, since we can write $h(u) = f[g(u)]$ with $f(v) = \exp\{(z_1 + \cdots + z_m)v\}$ and $g(u) = e^u$, Faa di Bruno's formula (1.1) gives

$$D_u^n h(0) = \sum_{k=1}^{n} (z_1 + \cdots + z_m)^k \sum_{p(n,k)} n! \prod_{i=1}^{n} \frac{1}{(\lambda_i!)(i!)^{\lambda_i}}$$

$$= \sum_{k=1}^{n} (z_1 + \cdots + z_m)^k \ S_n^k \ .$$

Hence

$$B_{\boldsymbol{\nu}}(t\mathbf{e}) = \sum_{k=1}^{n} m^k \ t^k \ S_n^k \ .$$

Equating like powers of t gives the desired conclusion. $\square$

We close this section with two special cases of Theorem 2.1: the case when $m = 1$, arbitrary d, and the case when $d = 1$, arbitrary m. The latter case appeared in Most [11].

2.10. Corollary. *Let* $|\boldsymbol{\nu}| = n \geq 1$ *and* $h(x_1, \ldots, x_d) = f[g(x_1, \ldots, x_d)]$ *with* $g \in \mathcal{C}_{\boldsymbol{\nu}}(\mathbf{x}^0)$ *and* $f \in \mathcal{C}^n(y^0)$, *where* $y^0 = g(\mathbf{x}^0)$. *Then*

$$(2.9) \qquad h_{\boldsymbol{\nu}} = \sum_{r=1}^{n} f_r \sum_{p(\boldsymbol{\nu},r)} (\boldsymbol{\nu}!) \prod_{j=1}^{n} \frac{[g_{\boldsymbol{\ell}_j}]^{k_j}}{(k_j!) \ (\boldsymbol{\ell}_j!)^{k_j}} \ ,$$

where

$$p(\boldsymbol{\nu}, r) = \{(k_1, \ldots, k_n; \ \boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_n) : \ for \ some \ 1 \leq s \leq n,$$
$$k_i = 0 \ and \ \boldsymbol{\ell}_i = \mathbf{0} \ for \ 1 \leq i \leq n - s; k_i > 0 \ for \ n - s + 1 \leq i \leq n;$$
$$and \ \mathbf{0} \prec \boldsymbol{\ell}_{n-s+1} \prec \cdots \prec \boldsymbol{\ell}_n \ are \ such \ that$$

$$\sum_{i=1}^{n} k_i = r, \sum_{i=1}^{n} k_i \boldsymbol{\ell}_i = \boldsymbol{\nu}\}.$$

2.11. Corollary. *Let* $n \geq 1, h(x) = f[g^{(1)}(x), \ldots, g^{(m)}(x)]$ *with* $g^{(i)} \in \mathcal{C}_n(x^0)$ *and* $f \in \mathcal{C}^n(\mathbf{y}^0)$, *where* $\mathbf{y}^0 = (g^{(1)}(x^0), \ldots, g^{(m)}(x^0))$. *Then*

$$(2.10) \qquad h_n = \sum_{1 \leq |\boldsymbol{\lambda}| \leq n} f_{\boldsymbol{\lambda}} \sum_{p(n,\boldsymbol{\lambda})} (\boldsymbol{\nu}!) \prod_{j=1}^{n} \frac{[\mathbf{g}_j]^{\mathbf{k}_j}}{(\mathbf{k}_j!)[j!]^{|\mathbf{k}_j|}} \ ,$$

where

$$p(n, \boldsymbol{\lambda}) = \{(\mathbf{k}_1, \ldots, \mathbf{k}_n) : \mathbf{k}_j \geq \mathbf{0}, \ \sum_{j=1}^{n} \mathbf{k}_j = \boldsymbol{\lambda}, \ \sum_{j=1}^{n} j|\mathbf{k}_j| = n\} \ .$$

Note that since we can take $\ell_j = j$, there is no need to include any ℓ's in the above expression for $p(n, \boldsymbol{\lambda})$.

3. Series composition and a general identity

In this section we prove a general multivariate identity based on a series composition. As a consequence of this identity we obtain another general formula for calculating the derivative of a series composition in any number of variables. Unlike the general Faa di Bruno formula (2.4), proved in the previous section, the formula

which we offer presents itself in the form of an infinite sum. Certain special cases of the general identity lead to a multivariate extension of the Stirling numbers.

We first introduce some further notation. Let $\mathbf{0} \leq \boldsymbol{\nu} = (\nu_1, \ldots, \nu_d)$ and $\mathbf{0} \neq \boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$. Define

(3.1)

$$s(\boldsymbol{\nu}, \boldsymbol{\lambda}) = \{(\boldsymbol{\mu}_1^{(1)}, \ldots, \boldsymbol{\mu}_{\lambda_1}^{(1)}; \ldots; \boldsymbol{\mu}_1^{(m)}, \ldots, \boldsymbol{\mu}_{\lambda_m}^{(m)}) : \ \boldsymbol{\mu}_j^{(i)} \in \mathcal{N}_0^d \text{ and } \sum_{i=1}^m \sum_{j=1}^{\lambda_i} \boldsymbol{\mu}_j^{(i)} = \boldsymbol{\nu}\}$$

and

(3.2)

$$s^+(\boldsymbol{\nu}, \boldsymbol{\lambda}) = \{(\boldsymbol{\mu}_1^{(1)}, \ldots, \boldsymbol{\mu}_{\lambda_1}^{(1)}; \ldots; \boldsymbol{\mu}_1^{(m)}, \ldots, \boldsymbol{\mu}_{\lambda_m}^{(m)}) \in s(\boldsymbol{\nu}, \boldsymbol{\lambda}) : \boldsymbol{\mu}_j^{(i)} \neq \mathbf{0}\} \ .$$

Our main identity concerns series compositions. Let

$$f(y_1, \ldots, y_m) = f(\mathbf{y}) = \sum_{\mathbf{0} \leq \boldsymbol{\lambda}} a_{\boldsymbol{\lambda}} \ \mathbf{y}^{\boldsymbol{\lambda}}$$

and, for $i = 1, \ldots, m$, let

$$g^{(i)}(x_1, \ldots, x_d) = g^{(i)}(\mathbf{x}) = \sum_{\mathbf{0} \leq \boldsymbol{\mu}} b_{\boldsymbol{\mu}}^{(i)} \ \frac{\mathbf{x}^{\boldsymbol{\mu}}}{\boldsymbol{\mu}!}$$

be power series. For technical simplicity, we shall assume that the series are everywhere convergent, i.e., that f and $g^{(i)}$ are entire functions.

3.1. Theorem. *For* $|\boldsymbol{\nu}| > 0$,

(3.3)
$$\sum_{1 \leq |\boldsymbol{\lambda}|} a_{\boldsymbol{\lambda}} \sum_{s(\boldsymbol{\nu}, \boldsymbol{\lambda})} \prod_{i=1}^m \prod_{j=1}^{\lambda_i} \frac{b_{\boldsymbol{\mu}_j^{(i)}}^{(i)}}{\boldsymbol{\mu}_j^{(i)}!}$$
$$= \sum_{1 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}|} f_{\boldsymbol{\lambda}} (\mathbf{b_0}) \sum_{p(\boldsymbol{\nu}, \boldsymbol{\lambda})} \prod_{j=1}^{|\boldsymbol{\nu}|} \frac{[\mathbf{b}_{\boldsymbol{\ell}_j}]^{\mathbf{k}_j}}{(\mathbf{k}_j!)[\boldsymbol{\ell}_j!]^{|\mathbf{k}_j|}} \ ,$$

where

$$\mathbf{b}_{\boldsymbol{\mu}} = (b_{\boldsymbol{\mu}}^{(1)}, \ldots, b_{\boldsymbol{\mu}}^{(m)}) \ .$$

Proof. Consider the series expansion of the series composition $h(\mathbf{x}) = f[g^{(1)}(\mathbf{x}), \ldots, g^{(m)}(\mathbf{x})]$. We have

(3.4)
$$h(\mathbf{x}) = f[g^{(1)}(\mathbf{x}), \ldots, g^{(m)}(\mathbf{x})]$$
$$= \sum_{\mathbf{0} \leq \boldsymbol{\lambda}} a_{\boldsymbol{\lambda}} \left(\sum_{\mathbf{0} \leq \boldsymbol{\mu}^{(1)}} b_{\boldsymbol{\mu}^{(1)}}^{(1)} \frac{\mathbf{x}^{\boldsymbol{\mu}^{(1)}}}{\boldsymbol{\mu}^{(1)}!} \right)^{\lambda_1} \cdots \left(\sum_{\mathbf{0} \leq \boldsymbol{\mu}^{(m)}} b_{\boldsymbol{\mu}^{(m)}}^{(m)} \frac{\mathbf{x}^{\boldsymbol{\mu}^{(m)}}}{\boldsymbol{\mu}^{(m)}!} \right)^{\lambda_m}$$
$$= a_{\mathbf{0}} + \sum_{1 \leq |\boldsymbol{\lambda}|} a_{\boldsymbol{\lambda}} \sum_{\mathbf{0} \leq \boldsymbol{\nu}} \mathbf{x}^{\boldsymbol{\nu}} \left(\sum_{s(\boldsymbol{\nu}, \boldsymbol{\lambda})} \prod_{i=1}^m \prod_{j=1}^{\lambda_i} \frac{b_{\boldsymbol{\mu}_j^{(i)}}^{(i)}}{\boldsymbol{\mu}_j^{(i)}!} \right)$$
$$= a_{\mathbf{0}} + \sum_{\mathbf{0} \leq \boldsymbol{\nu}} \mathbf{x}^{\boldsymbol{\nu}} \left(\sum_{1 \leq |\boldsymbol{\lambda}|} a_{\boldsymbol{\lambda}} \sum_{s(\boldsymbol{\nu}, \boldsymbol{\lambda})} \prod_{i=1}^m \prod_{j=1}^{\lambda_i} \frac{b_{\boldsymbol{\mu}_j^{(i)}}^{(i)}}{\boldsymbol{\mu}_j^{(i)}!} \right) \ .$$

On the other hand, computing the derivatives of h in accordance with Theorem 2.1, and evaluating at zero yields

(3.5)
$$h(\mathbf{x}) = \sum_{\mathbf{0} \leq \boldsymbol{\nu}} h_{\boldsymbol{\nu}}(\mathbf{0}) \, \frac{\mathbf{x}^{\boldsymbol{\nu}}}{\boldsymbol{\nu}!}$$

$$= f(\mathbf{b_0}) + \sum_{1 \leq |\boldsymbol{\nu}|} \mathbf{x}^{\boldsymbol{\nu}} \sum_{1 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}|} f_{\boldsymbol{\lambda}}(\mathbf{b_0}) \sum_{p(\boldsymbol{\nu},\boldsymbol{\lambda})} \prod_{j=1}^{|\boldsymbol{\nu}|} \frac{[\mathbf{b}_{\boldsymbol{\ell}_j}]^{\mathbf{k}_j}}{(\mathbf{k}_j!)[\boldsymbol{\ell}_j!]^{|\mathbf{k}_j|}} \; .$$

Equating the coefficients of $\mathbf{x}^{\boldsymbol{\nu}}$ in (3.4) and (3.5) leads to the conclusion of the theorem.

Many special cases of Theorem 3.1, with $m = 1$, $d = 1$ and $f(y) = e^y$, appear in Constantine and Savits [3]. The best known special case, Dobinski's identity, is obtained by further selecting $g(x) = e^x$. In this case (3.3) yields

(3.6)
$$e^{-1} \sum_{k=1}^{\infty} \frac{k^n}{k!} = \sum_{r=1}^{n} \sum_{p(n,r)} \frac{n!}{\prod_{j=1}^{n} (j!)^{k_j} (k_j!)} = B_n \; ,$$

the familiar Dobinski series expansion of the Bell number B_n which counts the number of partitions of a set with n elements. The merit of this expression lies in the fact that the infinite series on the left of (3.6) is fast converging, and one can thus assess the magnitude or asymptotic behavior of the Bell numbers; see Lovász [8].

A noteworthy special case of Theorem 3.1 takes place when $b_{\mathbf{0}}^{(i)} = g^{(i)}(\mathbf{0}) = 0$ for all $i = 1, \dots, m$. In this case we only need require that f and $g^{(i)}$ be analytic in a neighborhood of $\mathbf{0}$. The infinite sum in (3.3) now reduces to a finite sum. Indeed, for $|\boldsymbol{\lambda}| > |\boldsymbol{\nu}|$, the relation $\sum_{i=1}^{m} \sum_{j=1}^{\lambda_i} \boldsymbol{\mu}_j^{(i)} = \boldsymbol{\nu}$ requires that at least one $\boldsymbol{\mu}_j^{(i)} = \mathbf{0}$. Hence the corresponding product vanishes. Moreover, the assumption $\mathbf{b_0} = \mathbf{0}$ implies that $f_{\boldsymbol{\lambda}}(\mathbf{b_0}) = (\boldsymbol{\lambda}!) \, a_{\boldsymbol{\lambda}}$. With these adjustments, the identity (3.3) becomes

(3.7)
$$\sum_{1 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}|} a_{\boldsymbol{\lambda}} \sum_{s^+(\boldsymbol{\nu},\boldsymbol{\lambda})} \prod_{i=1}^{m} \prod_{j=1}^{\lambda_i} \frac{b_{\boldsymbol{\mu}_j^{(i)}}^{(i)}}{\boldsymbol{\mu}_j^{(i)}!}$$

$$= \sum_{1 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}|} a_{\boldsymbol{\lambda}} \, (\boldsymbol{\lambda}!) \sum_{p(\boldsymbol{\nu},\boldsymbol{\lambda})} \prod_{j=1}^{|\boldsymbol{\nu}|} \frac{[\mathbf{b}_{\boldsymbol{\ell}_j}]^{\mathbf{k}_j}}{(\mathbf{k}_j!)[\boldsymbol{\ell}_j!]^{|\mathbf{k}_j|}} \; . \quad \square$$

3.2. *Remark.* Suppose that $f(\mathbf{y})$ in Theorem 3.1 is an exponential function, that is, $f(\mathbf{y}) = \exp\{\sum_{i=1}^{m} c_i y_i\}$, where $(c_1, \dots, c_m) = \mathbf{c}$. Observe that in this case $h(\mathbf{x}) = f[(\mathbf{g}(\mathbf{x})]$ can also be written as

$$h(\mathbf{x}) = f[\mathbf{g}(\mathbf{0})] \, f[\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{0})] \; .$$

We can therefore combine Theorem 3.1 with (3.7) to obtain

$$\sum_{1 \leq |\boldsymbol{\lambda}|} \left(\frac{\mathbf{c}^{\boldsymbol{\lambda}}}{\boldsymbol{\lambda}!} \right) \sum_{s(\boldsymbol{\nu},\boldsymbol{\lambda})} \prod_{i=1}^{m} \prod_{j=1}^{\lambda_i} \frac{b_{\boldsymbol{\mu}_j^{(i)}}^{(i)}}{\boldsymbol{\mu}_j^{(i)}!}$$

$$= f(\mathbf{b_0}) \sum_{1 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}|} \mathbf{c}^{\boldsymbol{\lambda}} \sum_{p(\boldsymbol{\nu},\boldsymbol{\lambda})} \prod_{j=1}^{|\boldsymbol{\nu}|} \frac{[\mathbf{b}_{\boldsymbol{\ell}_j}]^{\mathbf{k}_j}}{(\mathbf{k}_j!)[\boldsymbol{\ell}_j!]^{|\mathbf{k}_j|}}$$

$$= f(\mathbf{b_0}) \sum_{1 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}|} \left(\frac{\mathbf{c}^{\boldsymbol{\lambda}}}{\boldsymbol{\lambda}!} \right) \sum_{s^+(\boldsymbol{\nu},\boldsymbol{\lambda})} \prod_{i=1}^{m} \prod_{j=1}^{\lambda_i} \frac{b_{\boldsymbol{\mu}_j^{(i)}}^{(i)}}{\boldsymbol{\mu}_j^{(i)}!} \ ,$$

where $f(\mathbf{b_0}) = \exp\{\sum_{i=1}^{m} c_i \, b_{\mathbf{0}}^{(i)}\}$. These equations offer an interesting relationship between summations over s, s^+, and p, by way of the exponential series.

3.3. Corollary. *Let real numbers $b_{\boldsymbol{\mu}}^{(i)}$ be given for $\mathbf{0} \neq \boldsymbol{\mu} \leq \boldsymbol{\nu}, 1 \leq i \leq m$. Then*

$$(3.8) \qquad \sum_{s^+(\boldsymbol{\nu},\boldsymbol{\lambda})} \prod_{i=1}^{m} \prod_{j=1}^{\lambda_i} \frac{b_{\boldsymbol{\mu}_j^{(i)}}^{(i)}}{\boldsymbol{\mu}_j^{(i)}!} = (\boldsymbol{\lambda}!) \sum_{p(\boldsymbol{\nu},\boldsymbol{\lambda})} \prod_{j=1}^{|\boldsymbol{\nu}|} \frac{[\mathbf{b}_{\boldsymbol{\ell}_j}]^{\mathbf{k}_j}}{(\mathbf{k}_j!)[\boldsymbol{\ell}_j!]^{|\mathbf{k}_j|}} \ .$$

Proof. For arbitrary $a_{\boldsymbol{\lambda}} \ (0 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}|)$ define

$$f(\mathbf{y}) = \sum_{0 \leq |\boldsymbol{\lambda}| \leq |\boldsymbol{\nu}|} a_{\boldsymbol{\lambda}} \, \mathbf{y}^{\boldsymbol{\lambda}} \ .$$

Also set, for $i = 1, \dots, m$,

$$g^{(i)}(\mathbf{x}) = \sum_{\mathbf{0} \neq \boldsymbol{\mu} \leq \boldsymbol{\nu}} b_{\boldsymbol{\mu}}^{(i)} \, \frac{\mathbf{x}^{\boldsymbol{\mu}}}{\boldsymbol{\mu}!} \ .$$

Since (3.7) holds for arbitrary $a_{\boldsymbol{\lambda}}$, we obtain (3.8) by equating coefficients of $a_{\boldsymbol{\lambda}}$.

Some special cases prove interesting. For example, by taking $b_{\boldsymbol{\ell}}^{(i)} = \boldsymbol{\ell}!$ for all i, we obtain

$$|s^+(\boldsymbol{\nu},\boldsymbol{\lambda})| = \boldsymbol{\lambda}! \sum_{p(\boldsymbol{\nu},\boldsymbol{\lambda})} \prod_{j=1}^{|\boldsymbol{\nu}|} \frac{1}{(\mathbf{k}_j!)} \ .$$

In the above, $|A|$ denotes the cardinality of the set A. In particular, for the univariate case, i.e., $m = d = 1, \boldsymbol{\nu} = n$ and $\boldsymbol{\lambda} = r$, we conclude that

$$r! \sum_{p(n,r)} \prod_{j=1}^{n} \frac{1}{k_j!} = |s^+(n,r)| = \binom{n-1}{r-1} \ ,$$

a well-known identity. A more interesting case results by selecting $b_{\boldsymbol{\ell}}^{(i)} = 1$ for all $\boldsymbol{\ell}$ and i. In the univariate case, Corollary 3.3 (upon multiplying both sides by $n!$) yields the identity

$$\sum_{s^+(n,r)} \binom{n}{k_1 \ \cdots \ k_r} = (r!) \, S_n^r \ ,$$

expressing the Stirling number of the second kind S_n^r as a sum of multinomial coefficients. It is easy to interpret this identity by way of a constructive bijection: to a partition of a set with n elements into r classes, associate a nondecreasing

path of n steps in the r-dimensional lattice of nonnegative integers by moving a step along coordinate i whenever an element falls within class i. The corresponding multidimensional identity is

$$\sum_{s^+(\boldsymbol{\nu},\boldsymbol{\lambda})} (\boldsymbol{\nu}!) \prod_{i=1}^{m} \prod_{j=1}^{\lambda_i} \frac{1}{\boldsymbol{\mu}_j^{(i)}!} = (\boldsymbol{\lambda}!) \sum_{p(\boldsymbol{\nu},\boldsymbol{\lambda})} (\boldsymbol{\nu}!) \prod_{j=1}^{|\boldsymbol{\nu}|} \frac{1}{(\mathbf{k}_j!)[\boldsymbol{\ell}_j!]^{|\mathbf{k}_j|}} \ .$$

In anology with the univariate case, we call the numbers

$$(3.9) \qquad\qquad S_{\boldsymbol{\nu}}^{\boldsymbol{\lambda}} = \sum_{p(\boldsymbol{\nu},\boldsymbol{\lambda})} (\boldsymbol{\nu}!) \prod_{j=1}^{|\boldsymbol{\nu}|} \frac{1}{(\mathbf{k}_j!)[\boldsymbol{\ell}_j!]^{|\mathbf{k}_j|}}$$

multivariate Stirling numbers. Their properties will be studied in a future paper.

We close this section with another theorem. Since the right-hand side of the identity (3.3) is in essence Faa di Bruno's formula, we can develop an infinite series expansion for the derivative $h_{\boldsymbol{\nu}}$ of a series composition. □

3.4. Theorem. *Let $f(\mathbf{y}) = f(y_1, \ldots, y_m)$ be an entire function and $g^{(i)} \in C_{\boldsymbol{\nu}}(\mathbf{t}^0)$ for $i = 1, \ldots, m$. Then in a neighborhood of $\mathbf{t}^0$, the composition $h(\mathbf{t}) = f[g^{(1)}(\mathbf{t}), \ldots, g^{(m)}(\mathbf{t})]$ has $\boldsymbol{\nu}^{th}$ derivative given by*

$$(3.10) \qquad\qquad h_{\boldsymbol{\nu}} = \boldsymbol{\nu}! \sum_{1 \le |\boldsymbol{\lambda}|} \frac{f_{\boldsymbol{\lambda}}(\mathbf{0})}{\boldsymbol{\lambda}!} \sum_{s(\boldsymbol{\nu},\boldsymbol{\lambda})} \prod_{i=1}^{m} \prod_{j=1}^{\lambda_i} \frac{g_{\boldsymbol{\mu}_j^{(i)}}^{(i)}}{\boldsymbol{\mu}_j^{(i)}!} \ .$$

Proof. In Theorem 3.1, replace $g^{(i)}$ with

$$G^{(i)}(\mathbf{x};\mathbf{t}) = \sum_{\mathbf{0} \le \boldsymbol{\mu} \le \boldsymbol{\nu}} g_{\boldsymbol{\mu}}^{(i)}(\mathbf{t}) \frac{\mathbf{x}^{\boldsymbol{\mu}}}{\boldsymbol{\mu}!}$$

and consider the composition $H(\mathbf{x};\mathbf{t}) = f[G^{(1)}(\mathbf{x};\mathbf{t}), \ldots, G^{(m)}(\mathbf{x};\mathbf{t})]$. Noting that $G^{(i)}(\mathbf{0};\mathbf{t}) = g^{(i)}(\mathbf{t})$ and $H(\mathbf{0};\mathbf{t}) = h(\mathbf{t})$, we fix $\mathbf{t}$ and apply Theorem 3.1. The right-hand side of (3.3) becomes

$$\sum_{1 \le |\boldsymbol{\lambda}| \le |\boldsymbol{\nu}|} f_{\boldsymbol{\lambda}}[\mathbf{g}(\mathbf{t})] \sum_{p(\boldsymbol{\nu},\boldsymbol{\lambda})} \prod_{j=1}^{|\boldsymbol{\nu}|} \frac{[\mathbf{g}_{\boldsymbol{\ell}_j}(\mathbf{t})]^{\mathbf{k}_j}}{(\mathbf{k}_j!)[\boldsymbol{\ell}_j!]^{|\mathbf{k}_j|}} = \frac{1}{\boldsymbol{\nu}!} h_{\boldsymbol{\nu}}(\mathbf{t})$$

by equation (2.4). The left-hand side of (3.3), except for the $(\boldsymbol{\nu}!)$ factor, coincides with what appears on the right-hand side of (3.10), since $a_{\boldsymbol{\lambda}} = f_{\boldsymbol{\lambda}}(\mathbf{0})/\boldsymbol{\lambda}!$. This ends the proof. □

The expression for $h_{\boldsymbol{\nu}}$ written in Theorem 3.4 is generally not the most efficient way to explicitly calculate the derivatives of a function composition and in fact imposes stronger requirements on the function f than are necessary for existence. Its usefulness lies primarily in establishing bounds or rates of growth for such derivatives in much the same way that Dobinski's formula (3.6) allows for the estimation of the Bell numbers B_n.

4. Applications

A. Compound nonhomogeneous Poisson processes. Many physical situations can be modeled by a compound nonhomogeneous Poisson process. In particular, in maintenance theory, a minimal repair process is modeled by a nonhomogeneous Poisson process $N = \{N(t), t \geq 0\}$ with a continuous mean function $\Lambda(t) = E[N(t)]$. The repair times correspond to the arrival times $S_1, S_2, \ldots$ of the process N. If we associate costs $Y_1, Y_2, \ldots$ to each repair time, then

$$(4.1) \qquad X(t) = \sum_{i=1}^{N(t)} Y_i$$

is a *compound nonhomogeneous Poisson process.* We assume that the costs are conditionally independent given the repair times, i.e.,

$$P\{Y_1 \in B_1, \ldots, Y_n \in B_n | S_1, \ldots, S_n\} = \prod_{i=1}^{n} L(S_i; B_i)$$

where $L(s; B)$ is a stochastic kernel on $[0, \infty) \times \mathcal{B}(\Re^1)$. A detailed study of this process was initiated in Chen and Savits [1]. They showed that X is characterized as an independent increment process with characteristic function

$$(4.2) \qquad E[e^{iuX(t)}] = \exp\left\{ \int_0^t \int_\Re (e^{iuy} - 1) \, L(s; dy) \, d\Lambda(s) \right\} .$$

Thus it is easy to calculate moments from Faa di Bruno formula (1.1).

If we now replace the cost Y_k by a random vector $\mathbf{Y}_k = (Y_{k1}, \ldots, Y_{kd})$ of costs, the corresponding process (4.1) is d-dimensional, i.e., $\mathbf{X}(t) = (X_1(t), \ldots, X_d(t))$. Analogously, its characteristic function for $\mathbf{u} = (u_1, \ldots, u_d)$ is given by the expression

$$(4.3) \qquad E[e^{i(\mathbf{u}, \mathbf{X}(t))}] = \exp\left\{ \int_0^t \int_{\Re^d} (e^{i(\mathbf{u},\mathbf{y})} - 1) \, L(s; d\mathbf{y}) \, d\Lambda(s) \right\} .$$

Hence, to calculate mixed moments, we can use our generalized Faa di Bruno formula (2.9).

4.1. Theorem. *Let $t \geq 0$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d)$ be given. Suppose that*

$$\int_0^t \int_{\Re^d} \prod_{i=1}^{d} |y_i^{\nu_i}| \, L(s; d\mathbf{y}) d\Lambda(s) < \infty .$$

Then $E[\mathbf{X}^{\boldsymbol{\nu}}(t)] = E[\prod_{i=1}^d X_i^{\nu_i}(t)]$ exists and is given by

$$(4.4) \qquad E[\mathbf{X}^{\boldsymbol{\nu}}(t)] = \boldsymbol{\nu}! \sum_{r=1}^{|\boldsymbol{\nu}|} \sum_{p(\boldsymbol{\nu},r)} \prod_{j=1}^{|\boldsymbol{\nu}|} \frac{\mu_{\boldsymbol{\ell}_j}^{k_j}(t)}{(k_j!)(\boldsymbol{\ell}_j!)^{k_j}}$$

where

$$(4.5) \qquad \mu_{\boldsymbol{\ell}}(t) = \int_0^t \int_{\Re^d} \mathbf{y}^{\boldsymbol{\ell}} \, L(s; d\mathbf{y}) d\Lambda(s) .$$

Proof. The existence can be shown in a manner analogous to the one-dimensional case ($d = 1$) which appears in Chen and Savits [3]. We now apply Corollary 2.10

with $f(y) = e^y, g(\mathbf{u}) = \int_0^t \int_{\Re^d} \{e^{i(\mathbf{u},\mathbf{y})} - 1\} L(s; d\mathbf{y}) d\Lambda(s)$ and $h(\mathbf{u}) = f[g(\mathbf{u})]$, noting that

$$(4.6) \qquad E[\mathbf{X}^{\boldsymbol{\nu}}(t)] = \frac{1}{(i)^{|\boldsymbol{\nu}|}} \, h_{\boldsymbol{\nu}}(\mathbf{0}) \, . \quad \square$$

B. Filtered nonhomogeneous Poisson processes. Keeping the same notation as in Section 4A, we call $X = \{X(t), t \geq 0\}$ a *filtered nonhomogeneous Poisson process* if it can be represented as

$$(4.7) \qquad X(t) = \sum_{k=1}^{N(t)} w(t, S_k, Y_k)$$

where $w(t, s, y)$ is a Borel measurable function. One could also allow Y_k and w to be vector-valued, but we shall be content to only consider the scalar version here.

Since X no longer has independent increments, one might be interested in obtaining moments of the form $E[\prod_{j=1}^d X^{\nu_j}(t_j)]$ for $0 \leq t_1 < \cdots < t_d$. To handle this situation, we first obtain the characteristic function. Therefore, given $0 \leq t_1 < t_2 < \cdots < t_d = \tau$, we define for $1 \leq j \leq d$,

$$w_j(s, y) = w(t_j, s, y) \, I_{[0, t_j]}(s) \, .$$

If $\mathbf{u} = (u_1, \ldots, u_d)$, consider

$$\sum_{j=1}^d u_j \, X(t_j) = \sum_{j=1}^d u_j \sum_{\ell=1}^{N(t_j)} w(t_j, S_\ell, Y_\ell) = \sum_{j=1}^d u_j \sum_{\ell=1}^{N(\tau)} w_j(S_\ell, Y_\ell)$$

$$= \sum_{\ell=1}^{N(\tau)} f(S_\ell, Y_\ell)$$

where $f(s, y) = \sum_{j=1}^d u_j w_j(s, y)$. Thus if we define

$$X_f(t) = \sum_{\ell=1}^{N(t)} f(S_\ell, Y_\ell) \, ,$$

then

$$E[\prod_{j=1}^d e^{iu_j X(t_j)}] = E[e^{ivX_f(t)}]\Big|_{v=1, t=\tau} \, .$$

But $X_f = \{X_f(t), t \geq 0\}$ is an ordinary compound nonhomogeneous Poisson process (cf. Chen and Savits [1], Remark 3.2. (ii)) with characteristic function

$$E[e^{ivX_f(t)}] = \exp\{\int_0^t \int_\Re [e^{ivf(s,y)} - 1] \, L(s; dy) \, d\Lambda(s)\} \, .$$

Consequently,

(4.8)

$$E[\prod_{j=1}^d e^{iu_j X(t_j)}] = \exp\{\int_0^\tau \int_\Re (\prod_{j=1}^d e^{iu_j w_j(s,y)} - 1) \, L(s; dy) \, d\Lambda(s)\} \, .$$

We can now obtain the moments by using Faa di Bruno's formula (2.9).

4.2. Theorem. *Let $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d)$ and $t_1, \ldots, t_d \geq 0$. Set $\tau = \max(t_1, \ldots, t_d)$. Assuming that*

$$\int_0^\tau \int_\Re \prod_{j=1}^d |w_j^{\nu_j}(s, y)|\ L(s; dy)\ d\Lambda(s) < \infty\ ,$$

then $E[\prod_{j=1}^d X^{\nu_j}(t_j)]$ exists and is given by

$$(4.9) \qquad E[\prod_{j=1}^d X^{\nu_j}(t_j)] = \boldsymbol{\nu}! \sum_{r=1}^{|\boldsymbol{\nu}|} \sum_{p(\boldsymbol{\nu}, r)} \prod_{j=1}^{|\boldsymbol{\nu}|} \frac{\mu_{\boldsymbol{\ell}_j}^{k_j}}{(k_j!)[\boldsymbol{\ell}_j!]^{k_j}}\ .$$

Here, for $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_d)$,

$$(4.10) \qquad \mu_{\boldsymbol{\ell}} = \int_0^\tau \int_\Re [\prod_{j=1}^d w_j^{\ell_j}(s, y)]\ L(s; dy)\ d\Lambda(s)\ .$$

4.3. *Remark.* Since $w_j(s, y) = 0$ for $s > t_j$, we can rewrite (4.10) as

$$\mu_{\boldsymbol{\ell}} = \int_0^{\tau(\boldsymbol{\ell})} \int_\Re [\prod_{j \in \sigma(\boldsymbol{\ell})} w^{\ell_j}(t_j, s, y)]\ L(s; dy)\ d\Lambda(s)\ ,$$

where

$$\sigma(\boldsymbol{\ell}) = \{\ell_j : \ell_j \neq 0\}$$

and

$$\tau(\boldsymbol{\ell}) = \min\{t_j : j \epsilon\ \sigma(\boldsymbol{\ell})\}\ .$$

4.4. Corollary. *Assuming the appropriate integrals exist, we have*

$$E[X(t)] = \int_0^t \int_\Re w(t, s, y)\ L(s; dy)\ d\Lambda(s)\ ,$$

$$Var[X(t)] = \int_0^t \int_\Re w^2(t, s, y)\ L(s; dy)\ d\Lambda(s)\ ,\ and$$

$$Cov[X(t_1), X(t_2)] = \int_0^{\min(t_1, t_2)} \int_\Re w(t_1, s, y)\ w(t_2, s, y)\ L(s; dy)\ d\Lambda(s)\ .$$

We may view Theorem 4.2 as a generalization of Cambell's Theorem (e.g., see Parzen [12], page 149). Cambell's Theorem usually only refers to expressions for the mean, variance and covariance, as in the above Corollary 4.4.

C. Multivariate moments and cumulants. Let $\mathbf{X} = (X_1, \ldots, X_d)$ be a d-dimensional random vector. The characteristic function $\phi(\mathbf{u}) = \phi(u_1, \ldots, u_d)$ is defined by

$$\phi(\mathbf{u}) = E[e^{i(\mathbf{u}, \mathbf{X})}] = E[\prod_{j=1}^d e^{iu_j X_j}]\ .$$

Assuming that the mixed moment $E[\mathbf{X}^{\boldsymbol{\nu}}] = E[\prod_{j=1}^d X_j^{\nu_i}]$ exists, it is then given by

$$(4.11) \qquad \alpha_{\boldsymbol{\nu}} = E[\mathbf{X}^{\boldsymbol{\nu}}] = \frac{1}{(i)^{|\boldsymbol{\nu}|}}\ D_{\mathbf{u}}^{\boldsymbol{\nu}}\ \phi(\mathbf{0})\ .$$

The corresponding $\boldsymbol{\nu}$-th cumulant $\kappa_{\boldsymbol{\nu}}$ is

$$(4.12) \qquad \kappa_{\boldsymbol{\nu}} = \frac{1}{(i)^{|\boldsymbol{\nu}|}} \, D_{\mathbf{u}}^{\boldsymbol{\nu}} \, \psi(\mathbf{0}) \ ,$$

where $\psi(\mathbf{u}) = \ln \phi(\mathbf{u})$. This is well-defined in a neighborhood of $\mathbf{u} = \mathbf{0}$.

Since $\psi(\mathbf{u}) = \ln \phi(\mathbf{u})$ and $\phi(\mathbf{u}) = \exp\{\psi(\mathbf{u})\}$, we can use Faa di Bruno's formula (2.9) to relate these quantities.

4.5. Theorem. *Assume that $E[|\mathbf{X}^{\boldsymbol{\nu}}|] < \infty$. Then*

$$(4.13) \qquad \kappa_{\boldsymbol{\nu}} = \boldsymbol{\nu}! \sum_{r=1}^{|\boldsymbol{\nu}|} (-1)^{r-1}(r-1)! \sum_{p(\boldsymbol{\nu},r)} \prod_{j=1}^{|\boldsymbol{\nu}|} \frac{\alpha_{\boldsymbol{\ell}_j}^{k_j}}{(k_j!)[\boldsymbol{\ell}_j!]^{k_j}}$$

and

$$(4.14) \qquad \alpha_{\boldsymbol{\nu}} = \boldsymbol{\nu}! \sum_{r=1}^{|\boldsymbol{\nu}|} \sum_{p(\boldsymbol{\nu},r)} \prod_{j=1}^{|\boldsymbol{\nu}|} \frac{\kappa_{\boldsymbol{\ell}_j}^{k_j}}{(k_j!)[\boldsymbol{\ell}_j!]^{k_j}} \ .$$

Lukacs [9] uses Faa di Bruno's formula (1.1) to obtain these results in the one-dimensional ($d = 1$) case. The multivariate generalizations were considered in McCullagh [10]. His expressions and method of proof are different from ours.

References

1. Chen, C. S. and Savits, T. H. (1993). Some remarks on compound nonhomogenous Poisson processes. *Statistics and Probability Letters,* **17**, 179-187. MR **94b:**60056
2. Constantine, G. M. (1987). *Combinatorial Theory and Statistical Design.* Wiley, New York. MR **88k:**05001
3. Constantine, G. M. and Savits, T. H. (1994). A stochastic process interpretation of partition identities. *SIAM J. Discrete Mathematics*, **2**, 194-202. MR **95d:**05010
4. Faa di Bruno, C. F. (1855). Note sur une nouvelle formule du calcul differentiel. *Quart. J. Math.*, **1**, 359-360.
5. Hoppe, R. (1871). Ueber independente Darstellung der höheren differentialquotienten. *Mathematische Annalen,* **4**, 85-87.
6. Hsu, L. C. (1993). Finding some strange identities via Faa di Bruno's formula. *Math. Research and Exposition*, **13**, 159-165. MR **94f:**05007
7. Lacroix, S. F. (1810). *Traité du calcul integral.* Tome Premier, Paris.
8. Lovász, L. (1979). *Combinatorial problems and exercises.* North-Holland, Amsterdam. MR **80m:**05001
9. Lukacs, E. (1955). Applications of Faa di Bruno's formula in mathematical statistics. *Am. Math. Monthly*, **62**, 340-348. MR **16:**1037g
10. McCullagh, P. (1987). *Tensor Methods in Statistics.* Chapman and Hall, New York. MR **88k:**62004
11. Most, R. (1871). Ueber die höheren differentialquotienten. *Mathematische Annalen*, **4**, 499-504.
12. Parzen, E. (1962). *Stochastic Processes.* Holden-Day, San Francisco. MR **25:**2628

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF PITTSBURGH, PITTSBURGH, PENNSYLVANIA 15260

E-mail address: `gmc@vms.cis.pitt.edu`

E-mail address: `ths@stat.pitt.edu`