I

MULTIVARIATE STIRLING POLYNOMIALS

1 Introduction

1.1 Background and problem

It is well-known that a close connection exists between iterated differentiation and Stirling numbers (see, e. g., [44, 77, 101]). Let $s_1(n,k)$ denote the signed Stirling numbers of the first kind, $s_2(n,k)$ the Stirling numbers of the second kind, and D the operator d/dx. Then, for all positive integers n, the nth iterate $(xD)^n$ can be expanded into the sum

$$(xD)^n = \sum_{k=1}^{n} s_2(n,k) x^k D^k. \tag{1.1}$$

An expansion in the reverse direction is also known to be valid (see, e. g., [44, p. 197] or [77, p. 45]):

$$D^n = x^{-n} \sum_{k=1}^{n} s_1(n,k)(xD)^k. \tag{1.2}$$

Let us first look at Eq. (1.1). The occurrence of the Stirling numbers can be explained combinatorially as follows. Observing

$$(xD)^n f(x) = D^n (f \circ \exp)(\log x)$$

we can use the classical higher-order chain rule (named after Faà di Bruno; cf. [42, 44], [51, pp. 52, 481]) to calculate the nth derivative of the composite function $f \circ g$:

$$(f \circ g)^{(n)}(x) = \sum_{k=1}^{n} B_{n,k}(g'(x), \ldots, g^{(n-k+1)}(x)) \cdot f^{(k)}(g(x)), \tag{1.3}$$

where $B_{n,k} \in \mathbb{Z}[X_1, \ldots, X_{n-k+1}]$, $1 \le k \le n$, is the (partial) exponential Bell polynomial

$$B_{n,k}(X_1, \ldots, X_{n-k+1}) = \sum_{r_1, r_2, \ldots} \frac{n!}{r_1! r_2! \ldots (1!)^{r_1} (2!)^{r_2} \ldots} X_1^{r_1} X_2^{r_2} \ldots \tag{1.4}$$

the sum to be taken over all non-negative integers $r_1, r_2, \ldots, r_{n-k+1}$ such that $r_1+r_2+\ldots+r_{n-k+1}=k$ and $r_1+2r_2+\ldots+(n-k+1)r_{n-k+1}=n$. The coefficient in $B_{n,k}$ counts the partitions of n distinct objects into k blocks (subsets) with r_j blocks containing exactly j objects ($1 \leq j \leq n-k+1$). Therefore, the sum of these coefficients is equal to the number $s_2(n,k)$ of all such partitions. So we have $B_{n,k}(x,\ldots,x) = s_2(n,k)x^k$. Evaluating $(f \circ \exp)^{(n)}(\log x)$ by Eq. (1.3) then immediately gives the right-hand side of Eq. (1.1).

Question. Can also Eq. (1.2) be interpreted in this way by substituting jth derivatives in place of the indeterminates X_j of some polynomial $S_{n,k} \in \mathbb{Z}[X_1,\ldots,X_{n-k+1}]$, the coefficients of which add up to $s_1(n,k)$?

The main purpose of the present chapter is to give a positive and comprehensive answer to this question including recurrences, a detailed study of the inverse relationship between the polynomial families $B_{n,k}$ and $S_{n,k}$, as well as fully explicit formulas (with some applications to Stirling numbers and Lagrange inversion).

The issue turns out to be closely related to the problem of generalizing Eq. (1.1), that is, finding an expansion for the operator $(\theta D)^n$ ($n \geq 1$, θ a function of x). Note that, in the case of scalar functions, $(\theta D)f$ is the *Lie derivative* of f with respect to θ. Several authors have dealt with this problem. In [21] and [69] a polynomial family $C_{n,k} \in \mathbb{Z}[X_0, X_1, \ldots, X_{n-k}]$ has been defined[1] by differential recurrences and shown to comply with $(\theta D)^n = \sum_{k=1}^{n} C_{n,k}(\theta, \theta', \ldots, \theta^{(n-k)})D^k$. Comtet [21] has tabulated $C_{n,k}$ up to $n = 7$ and stated that $C_{n,k}(x,\ldots,x) = c(n,k)x^n$, where $c(n,k) := |s_1(n,k)|$ denotes the signless Stirling numbers of the first kind ('cycle numbers' according to the terminology in [50]). Since however all coefficients of $C_{n,k}$ are positive, $C_{n,k}$ does not appear to be a suitable companion for $B_{n,k}$ with regard to the desired inversion law.

Todorov [99, 100] has studied the above Lie derivation with respect to a function θ of the special form $\theta(x) = {}^1/_{\varphi'(x)}$, $\varphi'(x) \neq 0$. His main results in [99] ensure the existence of $S_{n,k} \in \mathbb{Z}[X_1,\ldots,X_{n-k+1}]$ such that

$$\big(\varphi'(x)^{-1}D\big)^n f(x) = \sum_{k=1}^{n} A_{n,k}(\varphi'(x),\ldots,\varphi^{(n-k+1)}(x)) \cdot f^{(k)}(x), \qquad (1.5)$$

[1] Here and in Chapter II we write $C_{n,k}$ instead of Comtet's $A_{n,k}$ (cf. [21]) in order to avoid misunderstandings. Note that in both chapters of this book $A_{n,k}$ is exclusively used to denote the 'Lie coefficients' according to Todorov (see Eq. (1.5) below).

where $A_{n,k} := X_1^{-(2n-1)} S_{n,k}$. While differential recurrences for $A_{n,k}$ can readily be derived from Eq. (1.5) (cf. [99, Equation (27)] or a slightly modified version in [100, Theorem 2]), a simple representation for $S_{n,k}$ – as is Eq. (1.4) for $B_{n,k}$ – was still lacking up to now. Todorov [99, p. 224] erroneously believed that the somewhat cumbersome 'explicit' expression in [21] for the coefficients of $C_{n,k}$ would directly yield the coefficients of $S_{n,k}$. Also the determinantal form presented in [99, Theorem 6] for $(D/\varphi')^n$ (and thus also for $S_{n,k}$) may only in a modest sense be regarded as explicit.

Nevertheless, Todorov's choice ($\theta = 1/\varphi'$) eventually proves to be a crucial idea. Among other things, it reveals that $A_{n,k}$ (and thus $S_{n,k}$) is connected with the classical Lagrange problem of computing the compositional inverse $\overline{f}$ of a given series $f(x) = \sum_{n\geq 1}(f_n/n!)x^n$, $f_1 \neq 0$. As we shall see later, the Taylor coefficients $\overline{f}_n$ of $\overline{f}(x)$ can be expressed simply by applying $A_{n,1}$ to the coefficients of f as follows:

$$\overline{f}_n = A_{n,1}(f_1, \ldots, f_n). \tag{1.6}$$

On the other hand, Comtet [22] found an inversion formula that expresses $\overline{f}_n$ in terms of the (partial) exponential Bell polynomials:

$$\overline{f}_n = \sum_{k=0}^{n-1} (-1)^k f_1^{-n-k} B_{n+k-1,k}(0, f_2, \ldots, f_n). \tag{1.7}$$

This result has been shown by Haiman and Schmitt [33, 81] to provide essentially both a combinatorial representation and a cancellation-free computation of the antipode on a Faà di Bruno Hopf algebra (a topic that has received a lot of attention in quantum field theory due to its application to renormalization; cf. [55, 20, 28]). Combining Eq. (1.6) with Eq. (1.7) we obtain an expression for $A_{n,1}$ in terms of the Bell polynomials. This suggests looking for a similar representation for the whole family $A_{n,k}$, $1 \leq k \leq n$. As a main result (Theorem 6.1), we shall prove the following substantially extended version of Eq. (1.6) & Eq. (1.7):

$$A_{n,k} = \sum_{r=k-1}^{n-1} (-1)^{n-1-r} \binom{2n-2-r}{k-1} X_1^{-(2n-1)+r} \widetilde{B}_{2n-1-k-r,n-1-r}. \tag{1.8}$$

The tilde over B indicates that X_1 has been replaced by 0. From Eq. (1.8) we eventually get the desired explicit standard representation for $A_{n,k}$ that corresponds to the one for $B_{n,k}$ given in Eq. (1.4).

Equation Eq. (1.8) states a somewhat intricate relationship between the families $A_{n,k}$ and $B_{n,k}$. A simpler connection of both expressions is the following basic inversion law, which generalizes the orthogonality of the Stirling numbers (cf. Section 5):

$$\sum_{j=k}^{n} A_{n,j} B_{j,k} = \delta_{nk} \quad (1 \leq k \leq n), \tag{1.9}$$

where $\delta_{nn} = 1$, $\delta_{nk} = 0$ if $n \neq k$ (Kronecker symbol).

1.2 Terminology and notation

Considering Eq. (1.9) and the fact that the sum of the coefficients of $A_{n,k}$ and of $B_{n,k}$ are equal to $s_1(n,k)$ and to $s_2(n,k)$, respectively, it may be justified to call $A_{n,k}$ and $B_{n,k}$ *multivariate Stirling polynomials of the first and second kind.* There should be no risk of confusing them with polynomials *in one variable* like those introduced and named after Stirling by Nielsen [73, 74], neither with the closely related 'Stirling polynomials' $f_k(n) := s_2(n+k,n)$ and $g_k(n) := c(n, n-k)$ Gessel and Stanley [31] have investigated as functions of $n \in \mathbb{Z}$.

A sequence $r_1, r_2, r_3, \ldots$ of non-negative integers is said to be an (n,k)-*partition type*, $0 \leq k \leq n$, if $r_1+r_2+r_3+\ldots = k$ and $r_1+2r_2+3r_3+\ldots = n$. The set of all (n,k)-partition types is denoted by $\mathbb{P}(n,k)$; we write $\mathbb{P}$ for the union of all $\mathbb{P}(n,k)$. In the degenerate case ($k=0$) set $\mathbb{P}(n,0) = \emptyset$, if $n > 0$, and $\mathbb{P}(0,0) = \{0\}$ otherwise. Let $k \geq 1$. Since $n-k+1$ is the greatest j such that $r_j > 0$, we often write (n,k)-partition types as ordered $(n-k+1)$-tuples $(r_1, \ldots, r_{n-k+1})$.

The polynomials to be considered in the sequel have the form

$$P_\pi = \sum \pi(r_1, r_2, \ldots) X_1^{r_1} X_2^{r_2} \ldots,$$

where the sum ranges over all elements $(r_1, r_2, \ldots)$ of a full set $\mathbb{P}(n,k)$. As a consequence, P_π is homogeneous of degree k and isobaric of degree n. The coefficients of P_π may be regarded as values of a map $\pi : \mathbb{P} \longrightarrow \mathbb{Z}$ defined by some combinatorially meaningful expression, at least in typical cases like the following:

$$\omega(r_1, r_2, \ldots) := \frac{(r_1 + 2r_2 + \ldots)!}{r_1!\, r_2! \cdot \ldots} \qquad \textit{order function} \text{ (Lah)} \tag{1.10}$$

$$\zeta(r_1, r_2, \ldots) := \frac{\omega(r_1, r_2, \ldots)}{1^{r_1}\, 2^{r_2} \cdot \ldots} \qquad \textit{cycle function} \text{ (Cauchy)} \tag{1.11}$$

$$\beta(r_1, r_2, \ldots) := \frac{\omega(r_1, r_2, \ldots)}{(1!)^{r_1}\, (2!)^{r_2} \cdot \ldots} \qquad \textit{subset function} \text{ (Faà di Bruno)} \tag{1.12}$$

These coefficients count the number of ways a set can be partitioned into non-empty blocks according to a given partition type, that is, r_j denotes the number of blocks containing j elements ($j = 1, 2, \ldots$). The result depends on the meaning of 'block': linearly ordered subset (ω), cyclic order (ζ), or unordered subset (β).

It should be noticed that the corresponding polynomials $P_\omega, P_\zeta, P_\beta (= B_{n,k})$ are closely related to well-known combinatorial number-families:

$$P_\omega(1, \ldots, 1) = l^+(n,k), \quad \text{unsigned Lah numbers [57, 77]}$$
$$P_\zeta(1, \ldots, 1) = c(n,k) = \begin{bmatrix} n \\ k \end{bmatrix}, \quad \text{unsigned Stirling numbers of the 1st kind}$$
$$P_\beta(1, \ldots, 1) = s_2(n,k) = \begin{Bmatrix} n \\ k \end{Bmatrix}, \quad \text{Stirling numbers of the 2nd kind.}$$

1.3 Overview

This chapter is organized as follows: In Section 2 a general setting is sketched that allows functions and derivations to be treated algebraically. Section 3 contains a study of the iterated Lie operator $D(\varphi)^{-1}D$. An expansion formula for $(D(\varphi)^{-1}D)^n$ is established together with a differential recurrence for the resulting Lie coefficients $A_{n,k}$. Doing the same with respect to the inverse function $\overline{\varphi}$ will yield, conversely, D^n expanded and $B_{n,k}$ as the corresponding Lie coefficients. A by-product of Section 3 is Faà di Bruno's formula and its applications to the partial Bell polynomials $B_{n,k}$ to be briefly summarized in Section 4. These basic facts then lead to both inversion and recurrence relations, which we shall demonstrate and discuss in Section 5. The main task in Section 6 is to find an explicit polynomial expression for $S_{n,k}$. This is eventually achieved by means of Eq. (1.8), a proof of which makes up a central part of the section. In Section 7 we give some applications to the Lagrange inversion problem and to exponential generating functions.

2 Function algebra with derivation

2.1 Basic notions

Menger [64] has introduced the notion of a 'tri-operational algebra' of functions, which in the sequel (since 1960) stimulated to a great extent studies of generalized function algebras, e. g., [26, 65, 89, 90, 96]. In what follows I will propose a variant of Menger's original system tailored to our specific purposes of treating functions and their derivatives in a purely algebraic way.

Let $(\mathcal{F}, +, \cdot)$ be a non-trivial commutative ring of characteristic zero, 0 and 1 its identity elements with respect to addition and multiplication. We will think of the elements of $\mathcal{F}$ as 'functions (of one variable)' and therefore assume that $\mathcal{F}$ has a third binary operation $\circ$ (called *composition*) together with an identity element ι such that the following axioms are satisfied:

(F1) $f \circ (g \circ h) = (f \circ g) \circ h$

(F2) $(f + g) \circ h = (f \circ h) + (g \circ h)$

(F3) $(f \cdot g) \circ h = (f \circ h) \cdot (g \circ h)$

(F4) $f \circ \iota = \iota \circ f = f$

(F5) $1 \circ 0 = 1$

(F4) is assumed to be valid for all $f \in \mathcal{F}$; hence ι is unique. Let f be any element of $\mathcal{F}$. From (F2) we conclude $0 \circ f = 0$; so we get $\iota \neq 0$ (by (F4)) and $\iota \neq 1$ (by (F5)). (F2) furthermore implies $(-f) \circ g = -(f \circ g)$.

The least subring of $\mathcal{F}$ containing 1 will in the following conveniently be identified with $\mathbb{Z}$. (F5) then extends to the integers, that is, $n \circ 0 = n$ holds for all $n \in \mathbb{Z}$.

Given a *unit* f in $\mathcal{F}$ (i. e., f is an element invertible with respect to multiplication), we write f^{-1} (or $1/f$) for the inverse (henceforth called *reciprocal*) of f.

Remark 2.1. It must be emphasized that $\circ$ has to be understood as a *partial* operation (of course, $\iota^{-1} \circ 0$ is not defined). We therefore assign truth values to formulas, especially to our postulates (F1–3), whenever the terms involved are meaningful.

Let $f, g \in \mathcal{F}$ be functions such that $f \circ g = g \circ f = \iota$. Then g is called the *compositional inverse* of f, and vice versa. It is unique and will be denoted by $\overline{f}$. The following is obvious: $\overline{\iota} = \iota$, $\overline{\overline{f}} = f$, and $\overline{f \circ g} = \overline{g} \circ \overline{f}$.

Definition 2.1. Suppose $(\mathcal{F}, +, \cdot, \circ)$ satisfies (F1–5). We then call a mapping $D : \mathcal{F} \longrightarrow \mathcal{F}$ *derivation on* $\mathcal{F}$, and $(\mathcal{F}, +, \cdot, \circ, D)$ a *function algebra with derivation*, if D meets the following conditions:

(D1) $D(f + g) = D(f) + D(g)$

(D2) $D(f \cdot g) = D(f) \cdot g + f \cdot D(g)$

(D3) $D(f \circ g) = (D(f) \circ g) \cdot D(g)$

(D4) $D(\iota) = 1$

(D5) $D(f) = 0 \Longrightarrow f \circ 0 = f$

The classical derivation rules (D1), (D2) make $\mathcal{F}$ into a differential ring. Some simple facts are immediate: $D(0) = D(1) = 0$, $D(m \cdot f) = m \cdot D(f)$ for all $m \in \mathbb{Z}$. By an inductive argument the product rule (D2) can be generalized:

$$D(f_1 \cdots f_n) = \sum_{k=1}^{n} f_1 \cdots f_{k-1} \cdot D(f_k) \cdot f_{k+1} \cdots f_n. \tag{2.1}$$

By putting $f_i = f$, $1 \leq i \leq n$, Eq. (2.1) becomes $D(f^n) = nf^{n-1}D(f)$. If f is a unit, this holds also for $n \leq 0$. As usual, f^m for $m < 0$ is defined by $(f^{-1})^{-m}$.

(D4) prevents D from operating trivially. In the case of a field $\mathcal{F}$, (D4) can be weakend to $D(f) \neq 0$ (for some $f \in \mathcal{F}$), since the chain rule (D3) then gives $D(f) = D(f \circ \iota) = D(f) \cdot D(\iota)$.

Applying (D3) and (D4) to $D(f \circ \overline{f})$ we obtain the inversion rule

$$D(\overline{f}) = \frac{1}{D(f) \circ \overline{f}}. \tag{2.2}$$

In a differential ring, it is customary to define the subring $\mathcal{K}$ of constants as the kernel of the additive homomorphism D, that is,

$$\mathcal{K} := \{f \in \mathcal{F} \mid D(f) = 0\}.$$

We have $\mathbb{Z} \subseteq \mathcal{K}$. Constants behave as one would expect.

Proposition 2.1. $c \in \mathcal{K} \iff c \circ f = c$ *for all* $f \in \mathcal{F}$.

Proof. $\Rightarrow$: Suppose $D(c) = 0$. Then, for any $f \in \mathcal{F}$ we have by (F1) and (D5): $c \circ f = (c \circ 0) \circ f = c \circ (0 \circ f) = c \circ 0 = c$. – $\Leftarrow$: Set $f = 0$ and apply the chain rule (D3). ◊