Universal Constructions in Umbral Calculus

Nigel Ray

1 Introduction

Modern umbral calculus is steadily approaching maturity, as applications develop in several areas of mathematics. To maximize this utility it is important to work in the most general (as opposed to the most abstract) setting.

The origins of the 19th century theory lie in analysis. In a beautiful recent article [13] Rota and Taylor have returned to these roots, and their bibliography details many of the appropriate works. Given this context the subject naturally developed over the real and complex numbers, and versions with a more algebraic flavour often maintained the requirement of working over a field. Even in Roman's book [12] of 1984, for example, the author invites us to suppose that the coefficients of the generic delta operator are invertible, although the hypothesis is never fully used.

Once we accept the need for a general theory, we may jettison the field of scalars and work over a commutative ring R (which we assume to have an identity). It is then a short step to describing generic examples in the categorical language of universality, and insisting on invertibility only when necessary. This viewpoint offers many new insights and challenges, and informs our belief that the most flexible basis for the study of umbral calculus lies in the category $\mathcal{C}_R$ of coassociative coalgebras over R, together with the category $\mathcal{A}_R$ of dual algebras. Additional features such as gradings and Hopf algebra structures may naturally be present in certain circumstances. This viewpoint was conceived by Joni and Rota [5] and developed by Nichols and Sweedler [8], although both works continued to suggest that R should usually be a field.

Our purpose here is to popularize elements of umbral calculus which have already been translated into the language of universal algebra, and to introduce new and related constructions which are motivated by emerging applications. For readers who are unfamiliar with the standard definitions and notations of coalgebra theory, we refer to [8] as a convenient source.

Over the last decade the author has received a great deal of encouragement from Gian-Carlo in developing this material, and offers his grateful

thanks; currently he is engaged with Brian Taylor in reconciling his approach with that of the classics. It is a pleasure to thank Cristian Lenart for many hours of fruitful discussion, and also Francis Clarke and both referees for improving the accuracy of the final draft.

2 Basics

We define an *umbral calculus* (C, r) to consist of a coalgebra C in $\mathcal{C}_{\mathcal{R}}$ and an *evaluation* functional r in C^* (the dual $\mathrm{Hom}_R(C, R)$). For added convenience we shall insist that C be *supplemented*, in the sense that it is equipped with a summand of *scalars* whose projection is a counit $\varepsilon: C \to R$, and that r acts on this summand as the identity. Since C is a left C-comodule under the coproduct δ, it is also a right C^*-module by duality; in this guise, r acts on C as a right-invariant (or *shift-invariant*) R-linear endomorphism, whose value on c in C is usually denoted by $c \prec r$, or simply cr. In cocommutative situations we write this action on the left as rc. We define positive powers $r^{\star m}$ of r by convolution, and may then identify the action of $r^{\star m}$ with the m-fold iterate of the action of r.

A fundamental example is provided by taking C to be the polynomial algebra $R[x]$, with binomial coproduct $\delta(x^n) = \sum_k \binom{n}{k} x^k \otimes x^{n-k}$ for all $n \geq 0$, and counit $\varepsilon(p(x)) = p(0)$. Such an r is determined by its *umbra*, or sequence of values $r(x^n) = r_n$, and acts on the left of $R[x]$ by cocommutativity. Of course $R[x]$ is actually a Hopf algebra with respect to the polynomial product and the antipode $x \mapsto -x$, but we do not yet need this additional structure. The dual algebra consists of the divided power algebra (or *Hurwitz algebra* [1]) $R\{\{D\}\}$, where the functional D satisfies $D(x^n) = \delta_{1,n}$ and therefore acts on $R[x]$ as d/dx. In this context r may be expressed as the formal differential operator

$$r_1 D + r_2 D^2/2! + \cdots + r_n D^n/n! + \cdots, \tag{2.1}$$

and composition of operators is given by the Cauchy product of formal power series.

The original umbral calculus acquired its shady reputation because of seemingly arbitrary manipulations with symbolic notation. Following Rota's lead, these manipulations are demythologized by the use of the functional r, expressing its value on an arbitrary polynomial $p(x)$ by the substitution

$$r(p(x)) = p(r) \qquad r^n \equiv r_n.$$

Similarly, we may rewrite the expression (2.1) for r as e^{rD} in $R\{\{D\}\}$, so long as we insist that r^n denotes $r(x^n)$ whenever we formally expand in powers of rD; we may again describe the appropriate substitution in the form $r^n \equiv r_n$. With this convention, the action of r as an endomorphism on $R[x]$ may be abbreviated to

$$e^{rD}p(x) = p(x+r) \qquad r^n \equiv r_n,$$

thereby revealing e^{rD} as the *umbral shift*. If we have two functionals r and s then their convolution product $r \star s$ acts as $e^{rD}e^{sD}$, which we may contract to $e^{(r+s)D}$ so long as we insist that $(r+s)^n$ denotes $r \star s(x^n)$ whenever we expand. Since

$$r \star s(x^n) = \mu(r \otimes s(x \otimes 1 + 1 \otimes x)^n) = \sum_k \binom{n}{k} r_k s_{n-k} \tag{2.2}$$

(where μ is the product map in $R[x]$), we may write $r \star s(x^n)$ as

$$(r+s)^n \qquad r^k \equiv r_k, \quad s^k \equiv s_k \tag{2.3}$$

by applying the binomial theorem and the appropriate substitutions.

If r_1 is 1 we refer to the endomorphism $e^{rD} - 1$ as a *delta operator*, and label it Δ^r; it corresponds to the functional $r - \varepsilon$. An alternative pseudobasis for $R\{\{D\}\}$ is then given by the divided powers $\Delta^r_{(n)}$, and dualizing back (continuously, to recover $R[x]$) yields a new basis $B^r_n(x)$ of monic polynomials, whose elements form the *associated sequence* for Δ^r. This sequence is characterized by the binomial property

$$\delta(B^r_n(x)) = \sum_k \binom{n}{k} B^r_k(x) \otimes B^r_{n-k}(x),$$

together with the fact that $\varepsilon(B^r_n(x)) = 0$ for all $n > 0$. When working with Δ^r it is extremely convenient to reindex the umbra by $r(x^n) = r_{n-1}$ (so that $r_0 = 1$), and we adopt this convention henceforth.

We construct the universal delta operator Δ^ϕ by taking R to be the polynomial algebra $\mathbb{Z}[\phi_1, \phi_2, \ldots]$, abbreviated to Φ, and r to be the functional ϕ satisfying $\phi(x^n) = \phi_{n-1}$ for $n \geq 1$. Thus the endomorphism ϕ is the universal umbral shift. The associated sequence $B^\phi_n(x)$ consists of the *conjugate Bell polynomials*, for which (as we shall explain below) there exists an alternative description in terms of posets of partitions. The first three such polynomials are

$$x, \qquad x^2 - \phi_1 x, \qquad \text{and} \qquad x^3 - 3\phi_1 x^2 + (3\phi_1^2 - \phi_2)x. \tag{2.4}$$

We often refer to the m-fold convolution $\phi^{\star m}$ as the mth *umbral integer*, and note that elementary computation along the lines of (2.2) reveals

$$\begin{gathered}\phi^{\star m}(x) = m, \quad \phi^{\star m}(x^2) = m^2 + m(\phi_1 - 1), \\ \phi^{\star m}(x^3) = m^3 + 3m^2(\phi_1 - 1) + m(\phi_2 - 3\phi_1 + 2), \ldots\end{gathered} \tag{2.5}$$

Thus we may interpret $(\phi^{\star x})^n$ as a polynomial of degree n in $\Phi[x]$, for all $n \geq 1$. In fact the formulae (2.5) are equivalent to

$$\phi^{\star m}(B_n^{\phi}(x)) = [m]_n \tag{2.6}$$

(where $[m]_n$ denotes the falling factorial $m(m-1)\cdots(m-n+1)$ for all $m, n \geq 0$), which follows directly from the definitions.

Our arbitrary delta operator Δ^r defines a homomorphism $\Phi \to R$ by $\phi_n \mapsto r_n$, through which its properties may be studied in terms of the universal example. The conjugate Bell polynomials are the universal binomial sequence, and the same homomorphism maps them to the associated sequence for Δ^r; again, the properties of the $B_n^r(x)$ may be investigated in terms of those of the universal example. One particular instance motivates much of our terminology, namely when R is $\mathbb{Z}$ and each r_n is 1. Then the functional $r^{\star m}$ is the substitution $x = m$ and the endomorphism r is the forward shift e^D, whilst Δ^r is the forward difference operator Δ and the associated sequence consists of the falling factorial polynomials $[x]_n$.

An important variation is provided by the subalgebra $R[[\Delta^r]]$ of the Hurwitz algebra. When we form its continuous dual we obtain an R-coalgebra, freely generated by elements b_n^r (where $n \geq 0$ and $b_0^r = 1$) which are characterized by the *divided power* property

$$\delta(b_n^r) = \sum_k b_k^r \otimes b_{n-k}^r,$$

together with the fact that $\varepsilon(b_n^r) = 0$ for all $n > 0$. We label this coalgebra as the *penumbral coalgebra* $\Pi(\Delta^r)$, and remark that the homomorphism $R[x] \to \Pi(\Delta^r)$ (dual to the inclusion) acts such that $B_n^r(x)$ maps to $n!\, b_n^r$. We therefore refer to the b_n^r as the *divided sequence* of Δ^r, although they may only be expressed as rational polynomials in x when R is free of additive torsion. The functional r, its associated umbral shift, and the delta operator Δ^r all extend naturally to the penumbral coalgebra; for example, Δ^r maps b_n^r to b_{n-1}^r for each $n \geq 1$, and so is an isomorphism modulo scalars. In order to ensure that $\Pi(\Delta^r)$ is closed under multiplication (and is therefore a Hopf algebra), we may have to extend the ring of scalars. By duality, this is tantamount to insisting that the differential operator Δ^r admits a Leibniz formula for its action on products, so we label the appropriate extension as R^L and the resulting Hopf algebra as $R^L\langle b_n^r\rangle$. Determining the explicit structure of R^L in any given case remains an intriguing unsolved problem.

In the universal example Φ is torsion free, so the elements of the universal divided sequence b_n^{ϕ} may be written as normalized conjugate Bell polynomials $B_n^{\phi}(x)/n!$ in the rationalization $\Phi\mathbb{Q}[x]$. In this case Φ^L

is a genuine extension of Φ, satisfying $\Phi < \Phi^L < \Phi\mathbb{Q}$; it is known as the *Lazard ring*, and as explained in [3] it features prominently in the theories of formal group laws and stable homotopy. The Hopf algebra $\Phi^L\langle b_n^\phi\rangle$ is the covariant bialgebra of the universal formal group law, and masquerades as the complex bordism module of infinite dimensional complex projective space in algebraic topology.

The homomorphism $\Phi \to R$ extends naturally to $\Phi^L \to R^L$, and again allows the properties of an arbitrary case to be studied in terms of the universal example.

3 Sheffer Sequences

Associated sequences are special cases of Sheffer sequences, which arise when we invest our umbral calculus (C, r) with a right-invariant R-linear isomorphism u of C, which fixes the scalars. Thus u is a map of right C-comodules with respect to δ.

When C is $R[x]$ we may describe the isomorphism by a power series

$$u(\Delta^r) = 1 + u_1\Delta^r + u_2(\Delta^r)^2 + \cdots + u_n(\Delta^r)^n + \cdots, \tag{3.1}$$

where the elements r_n and u_n of R may (or may not) be algebraically independent. Thus u is a unit in $R[[\Delta^r]]$, although it is more usually expressed in terms of D (as in [12], for example); however, the form (3.1) is equivalent whenever $\mathbb{Q}$ is a subring of R, and is imperative for the universal viewpoint. The corresponding *Sheffer sequence* $S_n^{u,r}(x)$ consists of the monic, degree n polynomials $u(\Delta^r)B_n^r(x)$ in $R[x]$, where $n \geq 0$. Thus the $S_n^{u,r}(x)$ are given by $\sum_j [n]_j u_j B_{n-j}^r(x)$, and are characterized by the properties

$$\delta(S_n^{u,r}(x)) = \sum_k \binom{n}{k} S_k^{u,r}(x) \otimes B_{n-k}^r(x) \tag{3.2}$$

and $\varepsilon(S_n^{u,r}(x)) = n!\, u_n$. Alternatively, the $S_n^{u,r}(x)$ may be defined to satisfy the orthogonality relations

$$\langle(\Delta^r)^m u^{-1}, S_n^{u,r}(x)\rangle = m!\,\delta_{m,n} \quad \text{for all} \quad m, n \geq 0,$$

with respect to the sequence of operators $(\Delta^r)^m u^{-1}$.

The universal Sheffer sequence lies over $\Psi \otimes \Phi$, where Ψ denotes the polynomial algebra $\mathbb{Z}[\psi_1, \psi_2, \ldots]$. Writing ψ_n for $\psi_n \otimes 1$ and ϕ_n for $1 \otimes \phi_n$, we may describe the universal isomorphism by means of the power series

$$\psi(\Delta^\phi) = 1 + \psi_1\Delta^\phi + \psi_2(\Delta^\phi)^2 + \cdots + \psi_n(\Delta^\phi)^n + \cdots.$$

The first three universal polynomials $S_n^{\psi,\phi}(x)$ are

$$x+\psi_1, \qquad x^2+(2\psi_1-\phi_1)x+2\psi_2, \qquad \text{and}$$
$$x^3+3(\psi_1-\phi_1)x^2+(6\psi_2-3\psi_1\phi_1+3\phi_1^2-\phi_2)x+6\psi_3,$$

although there are other, equivalent forms.

Our arbitrary unit u defines a homomorphism $\Psi\otimes\Phi\to R$ by means of $\psi_n\mapsto u_n$ and $\phi_n\mapsto r_n$, which acts on polynomials by reducing each $S_n^{\psi,\phi}(x)$ to $S_n^{u,r}(x)$. The properties of the latter may therefore be investigated in terms of the universal example.

By way of illustration, we again consider the forward difference operator Δ over $\mathbb{Z}$; if we then let u be $\Delta/\log(1+\Delta)$, we obtain the Bernoulli polynomials of the second kind as the corresponding Sheffer sequence over $\mathbb{Q}$ (see [12] for further details).

Returning to the penumbral coalgebra $\Pi(\Delta^r)$ and the divided sequence b_n^r, we define the *divided Sheffer sequence* $s_n^{u,r}$ for the pair (u,Δ^r) by $u(\Delta^r)b_n^r$ for each $n\geq 0$. Thus $s_n^{u,r}=\sum_j u_jb_{n-j}^r$, and the sequence is characterized by the property

$$\delta(s_n^{u,r})=\sum_k s_k^{u,r}\otimes b_{n-k}^r,$$

together with the values $\varepsilon(s_n^{u,r})=u_n$ for all $n\geq 0$. Alternatively, the basis $s_n^{u,r}$ is dual to the pseudobasis $(\Delta^r)^nu^{-1}$ for $R[[\Delta^r]]$. In the universal example $\Psi\otimes\Phi$ is torsion free, so the elements $s_n^{\psi,\phi}$ may be written as $S_n^{\psi,\phi}(x)/n!$ in $\Psi\otimes\Phi\mathbb{Q}[x]$.

There is more structure to the universal example than meets the eye! This involves the polynomial algebra $\Psi\otimes\Psi$, in which we write ψ_n for $\psi_n\otimes 1$ and ψ'_n for $1\otimes\psi_n$, as convenient.

We define a second delta operator $\psi_+(\Delta^\phi)=\Delta^\phi\psi(\Delta^\phi)$, and employ its compositional inverse to construct a second unit $\overline{\psi}(\Delta^\phi)$ satisfying $\psi_+(\Delta^\phi\overline{\psi}(\Delta^\phi))=\Delta^\phi$ over $\Psi\otimes\Phi$. Thus the units $\psi'(\psi_+(\Delta^\phi))$ and $\pi(\Delta^\phi)=\psi(\Delta^\phi)\psi'(\psi_+(\Delta^\phi))$ over $\Psi\otimes\Psi\otimes\Phi$ are such that

$$\pi_+=\psi'_+\circ\psi_+ \qquad \text{and} \qquad \psi_+\circ\overline{\psi}_+=1$$

respectively. Simple computation (and Lagrange inversion) reveals that

$$\pi_n=\sum_{k=1}^{n}\psi'_k\psi_{n-k}^{k+1} \qquad \text{and} \qquad \overline{\psi}_n=\psi_n^{-(n+1)}/(n+1),$$

where $\psi=1+\psi_1+\psi_2+\cdots$ and ψ_{n-k}^k denotes the component of degree $n-k$ in the formal expansion ψ^k (assuming that each ψ_m has degree m).

We consider the homomorphism $\delta\colon \Psi \to \Psi \otimes \Psi$ induced by $\psi_n \mapsto \pi_n$ for all $n \geq 0$; this defines a coproduct map

$$\delta(\psi_n) = \sum_{k=0}^{n} \psi_{n-k}^{k+1} \otimes \psi_k,$$

for which there is an alternative construction in terms of the *Sheffer operator* $s_n^{\psi,\phi} \mapsto s_n^{\pi,\phi}$ (in the sense of [12]). We may verify directly that δ is coassociative, respects the product map μ, and has counit the projection onto $\mathbb{Z}$. The endomorphism χ of Ψ, induced similarly by $\psi_n \mapsto \overline{\psi}_n$ for all $n \geq 0$, is defined so that the composition $\mu \circ (1 \otimes \chi) \circ \delta$ coincides with the counit, and we therefore conclude that Ψ is a Hopf algebra with coproduct δ and antipode χ.

This Hopf algebra is better known to topologists as the dual of the Landweber-Novikov algebra, and to algebraic geometers as representing the affine group scheme which assigns to R the group of formal power series $t + t^2R[[t]]$ under composition. We conclude by outlining one of its more important properties.

Proposition 3.3 *The Hopf algebra Ψ coacts on the integral subcoalgebras of $\Phi[x]$ and $\Pi(\Delta^\phi)$ spanned by the $B_n^\phi(x)$ and b_n^ϕ respectively; furthermore, the coactions are compatible with respect to inclusion.*

Proof For convenience we work rationally and consider the units $b(\Delta^\phi) = \sum_{n\geq 0} b_n^\phi(\Delta^\phi)^n$ over $\Phi\mathbb{Q}[x]$, and $b(\psi_+(\Delta^\phi))$ over $\Psi \otimes \Phi\mathbb{Q}[w]$. We consider the Φ-linear homomorphism $c\colon \Phi\mathbb{Q}[w] \to \Psi \otimes \Phi\mathbb{Q}[w]$ induced by $b_n^\phi \mapsto b(\psi_+)_n = \sum_{k=1}^{n} \psi_{n-k}^k \otimes b_k^\phi$, for each $n \geq 1$. This clearly restricts compatibly to both $\Phi[x]$ and $\Pi(\Delta^\phi)$, and it remains only to confirm that it satisfies the coaction condition $(1 \otimes c) \circ c = (\delta \otimes 1) \circ c$ as homomorphisms $\Phi\mathbb{Q}[w] \to \Psi \otimes \Psi \otimes \Phi\mathbb{Q}[w]$. Since both possibilities are induced by the same map $b_n^\phi \mapsto b(\psi'_+(\psi_+))_n$, the result follows. □

We remark that $b_n^\phi \mapsto \psi_{n-1}$ defines a homomorphism $\sigma\colon \mathbb{Z}\langle b_n^\phi\rangle \to \Psi$ of left Ψ-comodules. Moreover, we may extend the structures above so that $\Phi \otimes \Psi$ becomes a Hopf algebroid, which coacts on the Φ-modules $\Phi[w]$ and $\Pi(\Delta^\phi)$ in their entirety. Correspondingly, σ extends to a comodule homomorphism $\Pi(\Delta^\phi) \to \Phi \otimes \Psi$; but thereby hangs another tale.

4 Number theory

To define the classical Bernoulli numbers B_n, we work in the dual of the coalgebra $\mathbb{Q}[x]$ and consider the right-invariant operator

$$D/(e^D - 1) = \sum_{n\geq 0} B_n D^n/n! \qquad (4.1)$$

(abbreviated to e^{BD}) in $\mathbb{Q}[[D]]$; of course $e^D - 1$ is the forward difference operator Δ. Thus e^{BD} arises from the functional B on $\mathbb{Q}[x]$, specified by $B(x^n) = B_n$ for all $n \geq 0$. The B_n are rational numbers, and are clearly zero when n is odd and > 1.

A convenient method of computation (which is recursive, and therefore subject to the usual limitations) may be succinctly described by the symbolic manipulations of (2.3). We rearrange (4.1) as $D + e^{BD} = e^{(B+1)D}$ in $\mathbb{Q}[[D]]$, and deduce that

$$B_n = (B+1)^n \qquad B^k \equiv B_k, \qquad \text{for all } n > 1,$$

where 1 denotes the functional defined by $1(x^n) = 1$ for all $n \geq 0$. Observing that $B_1 = -\frac{1}{2}$, we then calculate the next three Bernoulli numbers to be

$$B_2 = \frac{1}{6}, \qquad B_3 = 0, \quad \text{and} \quad B_4 = -\frac{1}{30}.$$

A famous theorem of von Staudt (from 1840) asserts that

$$B_n \equiv -\sum \frac{1}{p} \bmod \mathbb{Z} \tag{4.2}$$

for even n, where the summation ranges over all primes p such that $p-1$ divides n.

The operator e^{BD} is an isomorphism on $\mathbb{Q}[x]$, and its inverse J is the operator $\int_x^{x+1}$. Since J is Δ/D by (4.1), we may write $1 = \Delta e^{BD}/D$ in $\mathbb{Q}[[D]]$ and immediately obtain the classical Euler-MacLaurin summation formula

$$p(x) = Jp(x) + \sum_{n \geq 1} \frac{B_n}{n!} \Delta p^{(n-1)}(x)$$

for any polynomial $p(x)$ in $\mathbb{Q}[x]$.

To define the *universal Bernoulli numbers* B_n^ϕ, we work in the dual of the coalgebra $\Phi\mathbb{Q}[x]$ and consider the right-invariant operator

$$D/\Delta^\phi = \sum_{n \geq 0} B_n^\phi D^n / n!$$

(abbreviated to $e^{B^\phi D}$) in $\Phi\mathbb{Q}[[D]]$. The B_n^ϕ are therefore homogeneous rational polynomials in the ϕ_n. By analogy with the classical case, we may compute them recursively after noting that $B_1^\phi = -\frac{1}{2}\phi_1$ and applying

$$B_n^\phi = (B^\phi + \phi)^n \qquad (B^\phi)^k \equiv B_k^\phi, \quad \phi^k \equiv \phi_{k-1}$$

for all $n > 1$. We obtain

$$B_2^\phi = \tfrac{1}{6}(3\phi_1^2 - 2\phi_2), \qquad B_3^\phi = \tfrac{1}{4}(-3\phi_1^3 + 4\phi_1\phi_2 - \phi_3),$$

$$\text{and} \qquad B_4^\phi = \tfrac{1}{30}(45\phi_1^4 - 90\phi_1^2\phi_2 + 30\phi_1\phi_3 + 20\phi_2^2 - 6\phi_4).$$

These computations may readily be automated using standard symbolic algebra packages, and the first ten numbers are displayed in [2]. We remark that B_n^ϕ reduces to B_n under the homomorphism $\Phi \to \mathbb{Z}$ representing the forward difference operator; this sets each ϕ_n to 1.

The universal Bernoulli numbers were first defined explicitly by Miller in [6], but special cases had already appeared in the literature many years earlier. For a useful bibliography, see [2].

It is often more convenient to express the B_n^ϕ in terms of the power series which is compositionally inverse to Δ^ϕ. This is traditionally written as

$$D + c_1 D^2/2 + \cdots + c_n D^n/n + \cdots$$

in $\Phi\mathbb{Q}[[D]]$ by analogy with the standard exponential and logarithmic series, so that each c_n is a homogeneous rational polynomial in the ϕ_k, where $k \le n$. Moreover, each c_n maps to $(-1)^n$ whenever we reduce Φ to $\mathbb{Z}$, and lies in the Lazard ring $\Phi^L < \Phi\mathbb{Q}$ of §2. We may now state the *universal von Staudt Theorem* as

$$B_n^\phi \equiv -\sum \frac{1}{p} c_{p-1}^{n/(p-1)} \bmod \Phi^L$$

for even n, where the summation is over all primes p such that $p-1$ divides n, and as $B_n^\phi \equiv 0 \bmod \Phi^L$ for odd $n > 1$. Different proofs may be found in [2] and [11], neither of which appeals to (4.2) and both of which therefore reprove the classical result.

We may provide a *universal Euler-Maclaurin formula*, by writing J^ϕ for the symbolic operator $\int_x^{x+\phi}$, which is inverse to $e^{B^\phi D}$ on $\Phi\mathbb{Q}[x]$. Since $1 = \Delta^\phi e^{B^\phi D}/D$ in $\Phi\mathbb{Q}[[D]]$, we immediately obtain

$$p(x) = J^\phi p(x) + \sum_{n \ge 1} \frac{B_n^\phi}{n!} \Delta^\phi p^{(n-1)}(x)$$

for any polynomial $p(x)$ in $\Phi\mathbb{Q}[x]$.

As in [11], we may also construct universal Stirling numbers $S^\phi(n,k)$ and $s^\phi(n,k)$ from the Bell polynomials and their conjugates. These numbers lie in Φ, and have interesting properties which generalize their classical counterparts; in particular, they are closely related to the B_n^ϕ.

5 Negative integers

We have utilized the convolution product $\phi^{\star m}$ to define the positive umbral integers, thereby raising the question of finding a consistent interpretation when m is negative. Since the umbral shift is an isomorphism on $\Phi[x]$ we may consider its inverse, which is known as the *backward* *umbral shift*. We write the associated functional as ϕ^{-1}, and its m-fold convolution as $\phi^{\star(-m)}$ for all positive integers m; these functionals are our candidates for the negative umbral integers. We note immediately that $\phi^{\star(-m)}$ acts on $\Phi[x]$ as $(1+\Delta^\phi)^{-m}$, and so deduce that

$$\phi^{\star(-m)}(B_n^\phi(x)) = (-1)^n[m]^n \tag{5.1}$$

(where $[m]^n$ is the rising factorial) directly from the definitions.

Comparison with (2.6) demonstrates that the formulae (2.5) are equally valid for negative values of m, confirming that our candidates pass the simplest available test. By way of corroboration, we shall now consider a more subtle application in the theory of chromatic polynomials.

We recall that any partition π of an n-element set V has *type* $\tau(\pi)$, namely the monomial $\prod \phi_j$ in Φ to which each block of cardinality $j+1$ contributes a factor ϕ_j. Thus any function $f\colon V \to [m]$ also has a type $\tau(f)$, defined by its kernel. As described in [10], we may utilize τ to enrich the standard theory of zeta and Möbius functions of posets $\mathcal{P}$ of partitions, and in particular to define the characteristic *type polynomial* $c^\phi(\mathcal{P};x)$ in $\Phi[x]$. For example, we obtain the conjugate Bell polynomial $B_n^\phi(x)$ when $\mathcal{P}$ is the complete partition lattice on V, and x^n when it consists solely of the partition of V into singletons. If we apply the homomorphism $\Phi \to \mathbb{Z}$ which sets each ϕ_n to 1, then $c^\phi(\mathcal{P};x)$ reduces to the classical characteristic polynomial.

For any simple graph G there is an *umbral chromatic polynomial* $\chi^\phi(G;x)$ in $\Phi[x]$. This may either be defined as the characteristic type polynomial of a certain poset $\mathcal{A}$ of admissible partitions of the vertices V of G, or else in interpolated form as $\sum_\pi \tau(\pi)B^\phi_{|\pi|}(x)$, where the summation ranges over all proper colour partitions of V. In either event the evaluation $\phi^{\star m}(\chi^\phi(G;x))$ enumerates the colourings f of G by type, as a sum of monomials in Φ. If we apply the homomorphism $\Phi \to \mathbb{Z}$, then $\chi^\phi(G;x)$ reduces to the classical chromatic polynomial $\chi(G;x)$, and $\phi^{\star m}$ reduces to the substitution $x=m$. We remark that $\chi^\phi(G;x)$ encodes the same information as Stanley's symmetric function $X(G)$ [15]. In the case of the complete graph K_n the poset $\mathcal{A}$ consists of the complete partition lattice on $[n]$, so that $\chi^\phi(K_n;x)$ is $B_n^\phi(x)$ (as also follows from the interpolated form); for the null graph N_n the poset $\mathcal{A}$ is trivial, so that $\chi^\phi(N_n;x)$ is x^n.

We stress that the study of umbral chromatic polynomials and partition types is best construed as a combinatorial realization of the theory of the universal formal group law, or alternatively of the universal delta operator Δ^ϕ.

Motivated by the fact that the classical chromatic polynomial yields important combinatorial information when evaluated at negative integers, we aim for a generalization in terms of the elements $\phi^{\star(-m)}(\chi^\phi(G;x))$. To this end we introduce the set $D(m)$ of proper colourings $g\colon V \to [s+m-1]$ whose image has cardinality $s = s(g)$ for some $1 \le s \le n$; we declare that such colourings have *deficiency* m, and refer to $s(g)$ as the *span* of g.

Proposition 5.2 *For all positive integers m, we have that*

$$\phi^{\star(-m)}(\chi^\phi(G;x)) = \sum_{D(m)} (-1)^{s(g)}\tau(g)$$

in Φ.

Proof We apply $\phi^{\star(-m)}$ to the interpolated form of $\chi^\phi(G;x)$, and immediately obtain from (5.1) that

$$\phi^{\star(-m)}(\chi^\phi(G;x)) = \sum (-1)^{|\pi|}\tau(\pi)[m]^{|\pi|},$$

summed over all proper colour partitions. Since the rising factorial $[m]^{|\pi|}$ enumerates those colourings with kernel π and deficiency m, the formula follows at once. □

Even when reduced to the classical case by the homomorphism $\Phi \to \mathbb{Z}$, this result does not appear to be well documented in the literature.

Every colouring g in $D(m)$ gives rise to an acyclic orientation $\mathcal{O}$ of G by insisting that each edge be oriented in the direction of increasing colour. Furthermore, we may decompose the image of g into maximal subintervals $I(n_1)$, $I(n_2)$, $\ldots$ within $[s+m-1]$; there are at most m of these, and the indices n_j are chosen to exceed by 1 the number of omitted elements less than the smallest member of $I(n_j)$. We may then define a function $q\colon V \to [m]$ by requiring that $g(v)$ lie in $I(q(v))$, and observe that q is compatible with $\mathcal{O}$ insofar as $q(u) \le q(v)$ whenever $u < v$. We thereby obtain a *Stanley m-pair* $(\mathcal{O}, q)$, consisting of an acyclic orientation $\mathcal{O}$ and a compatible function q with codomain $[m]$. We say that g *covers* the pair $(\mathcal{O}, q)$, and write $C(\mathcal{O}, q)$ for the subset of $D(m)$ consisting of all such colourings.

Proposition 5.3 *For any Stanley m-pair $(\mathcal{O}, q)$, we have that*

$$\sum (-1)^{s(g)} = (-1)^n,$$

where the summation ranges over all $g \in C(\mathcal{O}, q)$.

Proof Suppose that the result holds for all graphs with n vertices and $> q$ edges, and with $< n$ vertices. For induction, choose G with n vertices and q edges, and let $(\mathcal{O}, q)$ be a Stanley m-pair. If G is K_n (including the possibility that $n = 1$) then there is a unique g which covers $(\mathcal{O}, q)$, and q dictates precisely which n of the $n+m-1$ colours are used. Thus $\sum(-1)^{s(g)} = (-1)^n$ immediately, and our induction may begin.

We select a non-edge $d = uv$ of G, and suppose either $q(u) \neq q(v)$, or that there exists a directed path between u and v. Then $\mathcal{O}$ extends uniquely (and compatibly with q) to $\mathcal{O}'$ on $G \cup d$; moreover, there is a bijection of $(\mathcal{O}, q)$ with $(\mathcal{O}', q)$ which preserves spans. So

$$\sum_{(\mathcal{O},q)} (-1)^{s(g)} = \sum_{(\mathcal{O}',q)} (-1)^{s(g)} = (-1)^n,$$

where the second equality follows by induction.

On the other hand, if $q(u) = q(v)$ and there is no directed path between u and v, then $\mathcal{O}$ extends to $\mathcal{O}_1$ on $G \cup (u,v)$, to $\mathcal{O}_2$ on $G \cup (v,u)$, and to $\mathcal{O}_3$ on G/d. Each of the first two extensions is compatible with q, and the third is compatible with q', induced on G/d. So we may partition $C(\mathcal{O}, q)$ into three blocks, namely $C(\mathcal{O}_1, q)$, $C(\mathcal{O}_2, q)$, and a block which corresponds bijectively to $C(\mathcal{O}_3, q')$, preserving spans. Thus

$$\sum_{(\mathcal{O},q)} (-1)^{s(g)} = \sum_{(\mathcal{O}_1,q)} (-1)^{s(g)} + \sum_{(\mathcal{O}_2,q)} (-1)^{s(g)} + \sum_{(\mathcal{O}_3,q')} (-1)^{s(g')},$$

yielding $(-1)^n + (-1)^n + (-1)^{n-1}$ by induction and hence completing the proof. □

We may now recover Stanley's original result [14].

Corollary 5.4 *For any simple graph G and positive integer m, the Stanley m-pairs are enumerated by the expression $(-1)^{|V|}\chi(G; -m)$.*

Proof We set each ϕ_n to 1 in 5.2, and apply 5.3. Since each $\tau(g)$ reduces to 1, the result follows. □

Our candidates have passed another test!

6 Hopf rings

A ring object in the category $\mathcal{C}_R$ is known as a Hopf ring (although the epithet *coalgebraic ring* would be strictly more appropriate). Such objects arise naturally in the study of ring schemes as described by Mumford [7], but were first fully exploited in algebraic topology, where they have served to organize a mass of complicated algebraic and geometric information into a coherent framework. The definitive treatment was given by Ravenel and Wilson in 1977 [9].

After discussion with Gian-Carlo, we suggest that Hopf rings will eventually find similar employment in combinatorics. We therefore take the opportunity to abstract from [4] certain applications of umbral calculus to computations in algebraic topology, and place them in the more sytematic context of the previous sections. The appropriate Hopf rings are actually $\mathbb{Z}$-graded, and therefore belong to an enriched category $\mathcal{GC}_R$. We could easily incorporate these into our description, and so pay additional tribute to Gian-Carlo; we resist this temptation, and reserve such discussions for the future.

The most basic example of a Hopf ring may be constructed from two commutative rings R and T with identity. We first restrict attention to the additive group structure $+$ of T, and form the classical group ring $R[T]$ of finite linear combinations $\sum_\alpha r_\alpha[s_\alpha]$, generated over R by the elements of T. The multiplicative structure (which we write by juxtaposition) arises from $+$. Once we take account of the multiplication $\cdot$ in T, we induce a second binary operation $\circ$ by means of $[t_1] \circ [t_2] = [t_1 \cdot t_2]$; in this sense we refer to $R[T]$ as the *ring ring* of T over R. In fact $R[T]$ is a Hopf algebra with respect to juxtaposition, having diagonal defined by $\delta[t] = [t] \otimes [t]$ and antipode induced by $t \mapsto -t$ for all t in T. The distributive law in T ensures that multiplication and $\circ$ are linked via the formula

$$x \circ (yz) = \sum (x' \circ y)(x'' \circ z) \tag{6.1}$$

for arbitrary x, y, and z in $R[T]$, where $\delta(x) = \sum x' \otimes x''$. There are other interrelations between the structure maps, all of which are encapsulated in the fact that $R[T]$ is a ring object in the category $\mathcal{C}_R$.

By analogy with polynomials in a single variable, we may form the *free Hopf ring* $R[T]\langle x\rangle$ over $R[T]$ on the single primitive generator x (which we assume augments to zero). In Ravenel and Wilson's terminology, this is the free $R[T]$-Hopf ring on the binomial coalgebra $\mathbb{Z}[x]$, and consists of all possible products and $\circ$ combinations of x with itself and elements of $R[T]$, subject to two types of relation. The first equates the multiplicative identity $x^0 = 1$ with the element $[0]$ (where 0 is the additive identity in T), whilst the second consists of all relations imposed by the requirements of being a ring object; rather than make the latter explicit, we refer readers to [9] for a comprehensive account. By way of clarification, we note that the monomials incorporating only x have the form $x^{\circ k_1} \cdots x^{\circ k_n}$, since any expression $x^{k_1} \circ \cdots \circ x^{k_n}$ may be simplified by repeated application of the distributivity law (6.1). Primitivity demands that $\delta(x)$ be $x \otimes 1 + 1 \otimes x$, so there is a canonical inclusion $R[x] \to R[T]\langle x\rangle$ of R-coalgebras; this identifies each scalar r with $r[0]$. Since the relation $x \circ [0] = 0$ follows from the ring structure, every element $x^{\circ k}$ is also primitive. Readers should not confuse $R[T]\langle x\rangle$ with the Hopf ring $R[T[x]]$.

We now consider the $R[T]$-linear functional ∂ on $R[T]\langle x\rangle$, defined by $\partial(x^n) = \delta_{1,n}$ and $\partial(y \circ z) = 0$ for all y and z; this restricts to D on the terms rx^n. Straightforward computation shows that the corresponding right-invariant endomorphism ∂ acts as $\partial/\partial x$ (with respect to juxtaposition), and restricts to D on $R[x]$. Given a delta operator Δ^r on $R[x]$ it therefore extends over $R[T]\langle x\rangle$ to an operator ∂^r, defined by $\sum_{n\geq 1} r_{n-1}\partial^n$.

We refer to the R-coalgebra map $R[x] \to R[T]\langle x\rangle$ as *internal differentiation* D_*, and define $D_*^{\circ n}(x^m)$ to be $(x^m)^{\circ n}$; it is an instructive exercise to deduce from the axioms that $D_*^{\circ n}(x^m)$ is zero unless $m = nd$, in which case it yields $m!(x^{\circ n})^d/d!$. We then define the *internal delta operator* Δ_*^t to be the R-linear extension of

$$x^n \mapsto \sum \binom{n}{q_1, 2q_2, \ldots, nq_n} \prod_{i=1}^{n} \frac{(iq_i)!}{i!q_i!}[t_{i-1}] \circ (x^{\circ i})^{q_i}$$

as a map $R[x] \to R[T]\langle x\rangle$. This is tantamount to interpreting Δ_*^t as $\prod_{n\geq 1} \frac{1}{n!}[t_{n-1}] \circ D_*^{\circ n}$. Because Δ_*^t respects coproducts (it is certainly not multiplicative, in general), the polynomials $\Delta_*^t B^r(x)$ continue to exhibit the binomial property (3.2). We write them as $B_n^{r,t}(x)$, and label them the *mixed associated sequence* for the pair (Δ^r, Δ_*^t). We record that $\partial^r B_n^{r,t}(x) = nB_{n-1}^{r,t}(x)$ for all $n \geq 1$, by construction.

The universal example is provided by selecting both r and t to be ϕ over Φ. In this case, we compute the first three *mixed conjugate Bell polynomials* $B^{\phi,\phi}(x)$ to be

$$x, \qquad x^2 - \phi_1 x + [\phi_1] \circ x^{\circ 2}, \qquad \text{and}$$

$$x^3 - 3\phi_1 x^2 + (3\phi_1^2 - \phi_2)x + 3x([\phi_1] \circ x^{\circ 2}) - 3\phi_1[\phi_1] \circ x^{\circ 2} + [\phi_2] \circ x^{\circ 3}.$$

in $\Phi[\Phi]\langle x\rangle$. These reduce to (2.4) on setting each $[\phi_n]$ to zero; it is again an instructive exercise to check from the axioms that the binomial property holds, and that $\partial^\phi B_n^{\phi,\phi}(x) = nB_{n-1}^{\phi,\phi}(x)$ for all $n \geq 1$.

Our delta operators Δ^r and Δ^t define a homomorphism $\Phi[\Phi] \to R[T]$, permitting investigation of the polynomials $B^{r,t}(x)$ in terms of the universal example. There are also analogues for the divided polynomials and penumbral coalgebra, which are closer to the topological applications.

References

[1] L Carlitz, *Some properties of Hurwitz series*, Duke Mathematical Journal **16** (1949), 285–295

[2] Francis Clarke, *The Universal von Staudt theorems*, Transactions of the American Mathematical Society **315** (1989), 591–603

[3] Michiel Hazewinkel, *Formal Groups and Applications*, Academic Press, London (1978)

[4] John Hunton and Nigel Ray, *A rational approach to Hopf rings*, Journal of Pure and Applied Algebra **101** (1995), 313–333

[5] S A Joni and G-C Rota, *Coalgebras and bialgebras in combinatorics*, Studies in Applied Mathematics **61** (1979), 93–139

[6] Haynes Miller, *Universal Bernoulli numbers and the S^1-transfer*, in *Current Trends in Algebraic Topology, part 2 (London, Ontario 1981)*, CMS Conference Proceedings **2**, American Mathematical Society (1982), 437–449

[7] David Mumford, *Lectures on Curves on an Algebraic Surface*, Annals of Mathematics Studies **59**, Princeton University Press (1966)

[8] Warren Nichols and Moss Sweedler, *Hopf Algebras and combinatorics*, in Umbral Calculus and Hopf Algebras, Contemporary Mathematics **6**, American Mathematical Society (1982), 49–84

[9] Douglas C Ravenel and W Stephen Wilson, *The Hopf ring for complex cobordism*, Journal of Pure and Applied Algebra **9** (1977), 241–280

[10] Nigel Ray, *Umbral calculus, binomial enumeration and chromatic polynomials*, Transactions of the American Mathematical Society **309** (1988), 191–213

[11] Nigel Ray, *Stirling and Bernoulli numbers for complex oriented homology theories*, in *Algebraic Topology: Proceedings, Arcata 1986*, Lecture Notes in Mathematics **1370**, Springer-Verlag (1989), 362–373

[12] Steven Roman, *The Umbral Calculus*, Academic Press, Orlando FL (1984)

[13] G-C Rota and B D Taylor, *The classical umbral calculus*, SIAM Journal of Mathematical Analysis **25** (1994), 694–711

[14] Richard P Stanley, *Acyclic orientations of graphs*, Discrete Mathematics **5** (1973), 171–178

[15] Richard P Stanley, *A symmetric function generalization of the chromatic polynomial of a graph*, Advances in Mathematics **111** (1995), 166–194

University of Manchester