

Formal calculus and umbral calculus

Thomas J. Robinson

Department of Mathematics
Rutgers University, New Brunswick/Piscataway, USA

`thomasro@math.rutgers.edu`

Submitted: Mar 12, 2010; Accepted: Jun 28, 2010; Published: Jul 10, 2010

Mathematics Subject Classification: 05A40, 17B69

Abstract

We use the viewpoint of the formal calculus underlying vertex operator algebra theory to study certain aspects of the classical umbral calculus. We begin by calculating the exponential generating function of the higher derivatives of a composite function, following a very short proof which naturally arose as a motivating computation related to a certain crucial “associativity” property of an important class of vertex operator algebras. Very similar (somewhat forgotten) proofs had appeared by the 19-th century, of course without any motivation related to vertex operator algebras. Using this formula, we derive certain results, including especially the calculation of certain adjoint operators, of the classical umbral calculus. This is, roughly speaking, a reversal of the logical development of some standard treatments, which have obtained formulas for the higher derivatives of a composite function, most notably Faà di Bruno’s formula, as a consequence of umbral calculus. We also show a connection between the Virasoro algebra and the classical umbral shifts. This leads naturally to a more general class of operators, which we introduce, and which include the classical umbral shifts as a special case. We prove a few basic facts about these operators.

1 Introduction

We present from first principles certain aspects of the classical umbral calculus, concluding with a connection to the Virasoro algebra. One of our main purposes is to show connections between the classical umbral calculus and certain central considerations in vertex operator algebra theory. The first major connection is an analogue, noted in [FLM], of those authors’ original argument showing that lattice vertex operators satisfy a certain fundamental associativity property. Those authors observed that this analogue amounts to a simple calculation of the higher derivatives of a composite function, often formulated as Faà di Bruno’s formula. The philosophy of vertex operator algebra theory led those

authors to emphasize the exponential generating function of the higher derivatives rather than the coefficients (which are easily extracted). That generating function was the analogue of a certain vertex operator. We shall show how taking this as a starting point, one may easily (and rigorously) recover significant portions of the classical umbral calculus of Sheffer sequences. The main aim, part of ongoing research, is to further develop the analogy between vertex operator algebra theory and classical umbral calculus. In addition, a direct connection between the classical umbral shifts and the Virasoro algebra (which plays a central role in vertex operator algebra theory) is established in the second half of this paper. Further analogies between vertex algebra formulas and classical umbral calculus formulas are noted in connection with this result and these motivate a generalization of the classical umbral shifts, which we briefly develop at the conclusion of this paper.

The classical umbral calculus has been treated rigorously in many works following the pioneering research of Gian-Carlo Rota, such as e.g. [MR], [RKO], [Ga], [Rt], [RR], [Rm1], [Fr], [T] and [Ch]. For an extensive bibliography through 2000 we refer the reader to [BL]. The general principle of umbral techniques reaches far beyond the classical umbral calculus and continues to be a subject of research (see e.g. [DS], [N] and [Z2]). Our treatment involves only certain portions of the classical umbral calculus of Sheffer sequences as developed in [Rm1].

There are many proofs of Faà di Bruno's formula for the higher derivatives of a composite function as well as related formulas dating back to at least the early 19th century (see [Jo] for a brief history, as well as [A], [B], [Bli], [F1], [F2], [Lu], [Me], and [Sc]). Moreover, it is a result that seems basic enough to be prone to showing up in numerous unexpected places, such as in connection with vertex operator algebra theory and also, as I recently learned from Professor Robert Wilson, in the theory of divided power algebras, to give just one more example. Here, for instance, a special case of Faà di Bruno's formula implies that certain coefficients are combinatorial and therefore integral, which is the point of interest since one wants a certain construction to work over fields of finite characteristic (see e.g. Lemma 1.3 of [Wi]). Faà di Bruno's formula is purely algebraic or combinatorial. For a couple of combinatorial proofs we refer the reader to [Z1] and [Ch], however we shall only be concerned with algebraic aspects of the result in this paper.

Our interest in Faà di Bruno's formula is due to its appearance in two completely separate subjects. First, it has long well-known connections with umbral calculus and second, perhaps more subtly, it shows up in the theory of vertex operator algebras. There are several umbral style proofs of Faà di Bruno's formula. According to [Jo], an early one of these is due to Riordan [Ri1] using an argument later completely rigorized in [Rm2] and [Ch]. Perhaps even more important, though, is the point of view taken in Section 4.1.8 of [Rm1], where the author discusses what he calls the "generic associated sequence," which he relates to the Bell polynomials, which themselves are closely related to Faà di Bruno's formula. The first part of this paper may, very roughly, be regarded as showing a way to develop some of the classical umbral calculus beginning from such "generic" sequences. We also bring attention more fully to [Ch] in which the formalism of "grammars" and some of the techniques quite closely resemble our approach at this stage, as I recently became aware.

Faà di Bruno's formula (in generating function form) appears in the theory of vertex operator algebras as originally observed in [FLM]. Briefly, Faà di Bruno's formula appeared in generating function form as an analogue, noted in [FLM], of those authors' original argument showing that lattice vertex operators satisfy a certain fundamental associativity property. The work [FLM] deals with many topics, but the parts which are of interest to us have to do with vertex operator algebra theory as well as, in particular, the Virasoro algebra, which is a very important ingredient in vertex operator algebra theory. We note that although certain crucial material from the theory of vertex operator algebras plays an essential role in the motivation of this paper, it turns out that we do not need explicit material directly about vertex operator algebras for the present work. By way of the literature, we briefly mention that the mathematical notion of vertex algebras was introduced in [B] and the variant notion of vertex operator algebra was introduced in [FLM]. An axiomatic treatment of vertex operator algebras was given in [FHL] and a more recent treatment was presented in [LL]. The interested reader may consult [L2] for an exposition of the history of the area.

This work began, unexpectedly, with certain considerations of the formal calculus developed to handle some of the algebraic, and ultimately, analytic aspects of vertex operator algebra theory. Those considerations were related to elementary results in the logarithmic formal calculus as developed in [Mi] and [HLZ]. However, we shall not discuss the connection to the logarithmic formal calculus here (for this see [R1] and [R2]) since another more classical result stemming from vertex algebra theory turns out to be more central to this material, namely that calculation which amounted to a calculation of the higher derivatives of a composite function, which was mentioned above. For the details of this calculation, see the introduction to Chapter 8 as well as Sections 8.3 and 8.4 of [FLM] and in particular Proposition 8.3.4, formula (8.4.32) and the comment following it.

The Virasoro algebra was studied in the characteristic 0 case in [GF] and the characteristic p analogue was introduced by R. Block in [Bl]. Over \mathbb{C} it may be realized as a central extension of the complexified Lie algebra of polynomial vector fields on the circle, which is itself called the Witt algebra. A certain crucial operator representation was introduced by Virasoro in [V] with unpublished contributions made by J.H. Weis, and the operators of this representation play a well known and essential role in string theory and vertex operator algebra theory (cf. [FLM]). Our connection with umbral calculus is made via one of these operators.

Since this paper is interdisciplinary, relating ideas in vertex operator algebra theory and umbral calculus, we have made certain choices regarding terminology and exposition in an effort to make it more accessible to readers who are not specialists in both of these fields. Out of convenience we have chosen [Rm1] as a reference for standard well-known results of umbral calculus. A well known feature of umbral calculus is that it is amenable to many different recastings. For instance, as the referee has pointed out, many of the main classical results, recovered from our point of view in Section 4, concerning adjoint relationships also appeared in [Fr], where what Roman [Rm1] refers to as "adjoints" are very nicely handled by a certain type of "transform." The change in point of view, among other things, gives a very interesting alternative perspective on the results and we

encourage the interested reader to compare the treatments. However, in the interests of space, when we wish to show the equivalence of certain of our results with the literature we will restrict ourselves to using the notation and framework in [Rm1].

In this paper we attempt to avoid specialized vocabulary as much as possible, although we shall try to indicate in remarks at least some of the important vocabulary from classical works. We shall use the name umbral calculus or classical umbral calculus since this seems to enjoy widespread name-recognition, but as the referee pointed out “finite operator calculus” might be a more appropriate name for much of the material such as the method in Proposition 3.1 and relevant material beginning in Section 4 of this work. We have also attempted to keep specialized notation to a minimum. However, because the notation which seems natural to begin with differs from that used in [Rm1] we do include calculations bridging the notational gap in Section 4 for the convenience of the reader. We note that the proofs of the results in Section 4 are much more roundabout than necessary if indeed those results in and of themselves were what was sought. The point is to show that from natural considerations based on the generating function of the higher derivatives of a composite function, one does indeed recover certain results of classical umbral calculus.

We shall now outline the present work section-by-section. In Section 2, along with some basic preliminary material, we begin by presenting a special case of the concise calculation of the exponential generating function of the higher derivatives of a composite function which appeared in the proof of Proposition 8.3.4 in [FLM]. Using this as our starting point, in Section 3 we then abstract this calculation and use the resulting abstract version to derive various results of the classical umbral calculus related to what Roman [Rm1] called associated Sheffer sequences. The umbral results we derive in this section essentially calculate certain adjoint operators, though in a somewhat disguised form. In Section 4, we then translate these “disguised” results into more familiar language using essentially the formalism of [Rm1]. We shall also note in this section how umbral shifts are defined as those operators satisfying what may be regarded as an umbral analogue of the $L(-1)$ -bracket-derivative property (cf. formula (8.7.30) in [FLM]). The observation that such analogues might be playing a role was suggested by Professor James Lepowsky after looking at a preliminary version of this paper.

In Section 5 we make an observation about umbral shifts which will be useful in the last phase of the paper.

In Section 6 we begin the final phase of this paper, in which we relate the classical umbral calculus to the Virasoro algebra of central charge 1. Here we recall the definition of the Virasoro algebra along with one special case of a standard “quadratic” representation; cf. Section 1.9 of [FLM] for an exposition of this well-known quadratic representation. We then show how an operator which was central to our development of the classical umbral calculus is precisely the $L(-1)$ operator of this particular representation of the Virasoro algebra of central charge 1. Using a result which we obtain in Section 5, we show a relationship between the classical umbral shifts and the operator now identified as $L(-1)$ and we then introduce those operators which in a parallel sense correspond to $L(n)$ for $n \geq 0$. (Strictly speaking, by focusing on only those operators $L(n)$ with $n \geq -1$, which them-

selves span a Lie algebra, the full Virasoro algebra along with its central extension remain effectively invisible.) We conclude by showing a couple of characterizations of these new operators in parallel to characterizations we already had of the umbral shifts. In particular we also note how the second of these characterizations, formulated as Proposition 7.2, may be regarded as an umbral analogue of (8.7.37) in [FLM], extending an analogue already noted concerning the $L(-1)$ -bracket-derivative property.

We note also that Bernoulli polynomials have long had connections to umbral calculus (see e.g. [Mel]) and have recently appeared in vertex algebra theory (see e.g. [L1] and [DLM]). It might be interesting to investigate further connections between the two subjects that involve Bernoulli polynomials explicitly.

This paper is an abbreviated version of part of [R3] (cf. also [R4]). The additional material in the longer versions is largely expository, for the convenience of readers who are not specialists.

I wish to thank my advisor, Professor James Lepowsky, as well as the attendees (regular and irregular) of the Lie Groups/Quantum Mathematics Seminar at Rutgers University for all of their helpful comments concerning certain portions of the material which I presented to them there. I also want to thank Professors Louis Shapiro, Robert Wilson and Doron Zeilberger for their useful remarks. Additionally, I would like to thank the referee for many helpful comments.

Finally, I am grateful for partial support from NSF grant PHY0901237.

2 Preliminaries

We set up some notation and recall some well-known and easy preliminary propositions in this section. For a more complete treatment, we refer the reader to the first three sections of Chapter 8 of [FLM] (cf. Chapter 2 of [LL]), while noting that in this paper we shall not need any of the material on “expansions of zero,” the heart of the formal calculus treated in those works.

We shall write $t, u, v, w, x, y, z, x_n, y_m, z_n$ for commuting formal variables, where $n \geq 0$ and $m \in \mathbb{Z}$. All vector spaces will be over \mathbb{C} . Let V be a vector space. We use the following:

$$\mathbb{C}[[x]] = \left\{ \sum_{n \geq 0} c_n x^n \mid c_n \in \mathbb{C} \right\}$$

(formal power series), and

$$\mathbb{C}[x] = \left\{ \sum_{n \geq 0} c_n x^n \mid c_n \in \mathbb{C}, c_n = 0 \text{ for all but finitely many } n \right\}$$

(formal polynomials).

We denote by $\frac{d}{dx}$ the formal derivative acting on either $\mathbb{C}[x]$ or $\mathbb{C}[[x]]$. Further, we shall frequently use the notation e^\square to refer to the formal exponential expansion, where

\square is any formal object for which such expansion makes sense. By “makes sense” we mean that the coefficients of the monomials of the expansion are finite objects. For instance, we have the linear operator $e^{w\frac{d}{dx}} : \mathbb{C}[[x, x^{-1}]] \rightarrow \mathbb{C}[[x, x^{-1}]][[w]]$:

$$e^{w\frac{d}{dx}} = \sum_{n \geq 0} \frac{w^n}{n!} \left(\frac{d}{dx} \right)^n.$$

We recall that a linear map D on an algebra A which satisfies

$$D(ab) = (Da)b + a(Db) \quad \text{for all } a, b \in A$$

is called a *derivation*. Of course, the linear operator $\frac{d}{dx}$ when acting on either $\mathbb{C}[x]$ or $\mathbb{C}[[x]]$ is an example of a derivation.

It is a simple matter to verify, by induction for instance, the following version of the elementary binomial theorem. Let A be an algebra with derivation D . Then for all $a, b \in A$, we have:

$$D^n ab = \sum_{k+l=n} (k+l)! \frac{D^k a}{k!} \frac{D^l b}{l!} \tag{2.1}$$

$$e^{wD} ab = (e^{wD} a) (e^{wD} b). \quad (\text{the automorphism property}) \tag{2.2}$$

Further, we separately state the following important special case of the automorphism property. For $f(x), g(x) \in \mathbb{C}[[x]]$,

$$e^{w\frac{d}{dx}} f(x)g(x) = \left(e^{w\frac{d}{dx}} f(x) \right) \left(e^{w\frac{d}{dx}} g(x) \right).$$

The automorphism property shows, among other things, how the operator $e^{w\frac{d}{dx}}$ may be regarded as a formal substitution, since, for $n \geq 0$, we have:

$$e^{w\frac{d}{dx}} x^n = \left(e^{w\frac{d}{dx}} x \right)^n = (x+w)^n.$$

Therefore, by linearity, we get the following polynomial formal Taylor formula. For $p(x) \in \mathbb{C}[x]$,

$$e^{w\frac{d}{dx}} p(x) = p(x+w).$$

Since the total degree of every term in $(x+w)^n$ is n , we see that $e^{w\frac{d}{dx}}$ preserves total degree. By equating terms with the same total degree we can therefore extend the previous proposition to get the following. For $f(x) \in \mathbb{C}[[x]]$,

$$e^{w\frac{d}{dx}} f(x) = f(x+w). \tag{2.3}$$

Remark 2.1. We note that the identity (2.3) can be derived immediately by direct expansion as the reader may easily check. However, in the formal calculus used in vertex operator algebra theory it is often better to think of this minor result within a context like that provided above. For instance, it is often useful to regard such formal Taylor theorems concerning formal translation operators as representations of the automorphism property (see [R1], [R2] and Remark 2.2).

We have calculated the higher derivatives of a product of two polynomials using the automorphism property. We next reproduce (in a very special case, for the derivation $\frac{d}{dx}$), the quick argument given in Proposition 8.3.4 of [FLM] to calculate the higher derivatives of the composition of two formal power series. Let $f(x), g(x) \in \mathbb{C}[[x]]$. We further require that $g(x)$ have zero constant term, so that, for instance, the composition $f(g(x))$ is always well defined. We shall approach the problem by calculating the exponential generating function of the higher derivatives of $f(g(x))$. We get

$$\begin{aligned}
e^{w\frac{d}{dx}}f(g(x)) &= f(g(x+w)) \\
&= f(g(x) + (g(x+w) - g(x))) \\
&= \left(e^{(g(x+w)-g(x))\frac{d}{dz}}f(z) \right) \Big|_{z=g(x)} \\
&= \sum_{n \geq 0} \frac{f^{(n)}(g(x))}{n!} \left(e^{w\frac{d}{dx}}g(x) - g(x) \right)^n \\
&= \sum_{n \geq 0} \frac{f^{(n)}(g(x))}{n!} \left(\sum_{m \geq 1} \frac{g^{(m)}(x)}{m!} w^m \right)^n. \tag{2.4}
\end{aligned}$$

While our calculation of the higher derivatives is not, strictly speaking, complete at this stage (although all that remains is a little work to extract the coefficients in powers of w), it is in fact this formula which will be of importance to us, since, roughly speaking, many results of the classical umbral calculus follow because of it, and so we shall record it as a proposition.

Proposition 2.1. *Let $f(x)$ and $g(x) \in \mathbb{C}[[x]]$. Let $g(x)$ have zero constant term. Then we have*

$$e^{w\frac{d}{dx}}f(g(x)) = \sum_{n \geq 0} \frac{f^{(n)}(g(x))}{n!} \left(\sum_{m \geq 1} \frac{g^{(m)}(x)}{m!} w^m \right)^n. \tag{2.5}$$

□

A derivation of Faà di Bruno’s classical formula may be found in Section 12.3 of [An]. We shall not need the fully expanded formula.

Remark 2.2. The more general version of this calculation (based on a use of the automorphism property instead of the formal Taylor theorem) appeared in [FLM] because it was related to a much more subtle and elaborate argument showing that vertex operators associated to lattices satisfied a certain associativity property (see [FLM], Sections 8.3 and 8.4 and in particular, formula (8.4.32) and the comment following it). The connection is due in part to the rough resemblance between the exponential generating function of the higher derivatives of a composite function in the special case $f(x) = e^x$ (see (2.6) below) and “half of” a vertex operator.

Noting that in (2.4) we treated $g(x+w) - g(x)$ as one atomic object suggests a reorganization. Indeed by calling $g(x+w) - g(x) = v$ and $g(x) = u$, the second, third and fourth lines of (2.4) become

$$f(u+v) = e^{v \frac{d}{du}} f(u) = \sum_{n \geq 0} \frac{f^{(n)}(u)}{n!} v^n.$$

This is just the formal Taylor theorem, of course, and so we could have begun here and then re-substituted for u and v to get the result. This, according to [Jo], is how the proof of U. Meyer [Me] runs.

It is also interesting to specialize to the case where $f(x) = e^x$, as is often done, and indeed was the case which interested the authors of [FLM] and will interest us in later sections. We have simply

$$e^{w \frac{d}{dx}} e^{g(x)} = e^{g(x+w)} = e^{g(x)} e^{g(x+w)-g(x)} = e^{g(x)} e^{\sum_{m \geq 1} \frac{g^{(m)}(x)}{m!} w^m}. \quad (2.6)$$

Remark 2.3. The generating function for what are called the Bell polynomials (cf. Chapter 12.3 and in particular (12.3.6) in [An]) easily follows from (2.6) using a sort of slightly unrigorous old-fashioned umbral argument replacing $g^{(m)}$ with g_m . See the proof of Proposition 3.1 for one way of handling such arguments. (The referee has pointed out that one may also rigorize this argument with certain evaluations of umbral elements in the umbral calculus whereas our argument in Proposition 3.1 is closer to the related finite operator calculus.) Of course, if we also set $g(x) = e^x - 1$, we get $e^{w \frac{d}{dx}} e^{e^x - 1} = e^{e^{x+w} - 1}$ and setting $x = 0$ is easily seen to give the well-known result that $e^{e^w - 1}$ is the generating function of the Bell numbers, which are themselves the Bell polynomials with all variables evaluated at 1.

For convenience we shall globally name three generic (up to the indicated restrictions) elements of $\mathbb{C}[[t]]$:

$$A(t) = \sum_{n \geq 0} A_n \frac{t^n}{n!}, \quad B(t) = \sum_{n \geq 1} B_n \frac{t^n}{n!}, \quad \text{and} \quad C(t) = \sum_{n \geq 0} C_n \frac{t^n}{n!}, \quad (2.7)$$

where both $B_1 \neq 0$ and $C_0 \neq 0$ (and note the ranges of summation). We recall, and it is easy for the reader to check, that $B(t)$ has a compositional inverse, which we denote by $\overline{B}(t)$, and that $C(t)$ has a multiplicative inverse, $C(t)^{-1}$. We note further that since $\overline{B}(t)$ has zero constant term, $B'(\overline{B}(t))$ is well defined, and we shall denote it by $B^*(t)$. In addition, $p(x)$ will always be a formal polynomial and sometimes we shall feel free to use a different variable such as z in the argument of one of our generic series, so that $A(z)$ is the same type of series as $A(t)$, only with the name of the variable changed.

Remark 2.4. We defined $B^*(t) = B'(\overline{B}(t))$. As the referee pointed out, it is also true that $B^*(t) = \frac{1}{\overline{B}(t)}$, which follows from the chain rule by taking the derivative of both sides of $B(\overline{B}(t)) = t$.

Remark 2.5. Series of the form $B(t)$ are sometimes called “delta series” in umbral calculus, or finite operator calculus (cf. [Rm1]).

We shall also use the notation $A^{(n)}(t)$ for the derivatives of, in this case, $A(t)$, and it will be convenient to define this for all $n \in \mathbb{Z}$ to include anti-derivatives. Of course, to make that well-defined we need to choose particular integration constants and only one choice is useful for us, as it turns out.

Notation 2.1. For all $n \in \mathbb{Z}$, given a fixed sequence $A_m \in \mathbb{C}$ for all $m \in \mathbb{Z}$, we shall define

$$A^{(n)}(t) = \sum_{m \geq n} \frac{A_m t^{m-n}}{(m-n)!}.$$

3 A restatement of the problem and further developments

In the last section we considered the problem of calculating the higher formal derivatives of a composite function of two formal power series, $f(g(x))$, where we obtained an answer involving only expressions of the form $f^{(n)}(g(x))$ and $g^{(m)}(x)$. Because of the restricted form of the answer it is convenient to translate the result into a more abstract notation which retains only those properties needed for arriving at Proposition 2.1. This essential structure depends only on the observation that $\frac{d}{dx} f^{(n)}(g(x)) = f^{(n+1)}(g(x))(g^{(1)}(x))$ for $n \geq 0$ and that $\frac{d}{dx} g^{(m)}(x) = g^{(m+1)}(x)$ for $m \geq 1$.

Motivated by the above paragraph, we consider the algebra

$$\mathbb{C}[\dots, y_{-2}, y_{-1}, y_0, y_1, \dots, x_1, x_2, \dots].$$

Then let D be the unique derivation on $\mathbb{C}[\dots, y_{-2}, y_{-1}, y_0, y_1, y_2, \dots, x_1, x_2, \dots]$ satisfying

$$\begin{aligned} Dy_i &= y_{i+1}x_1 & i \in \mathbb{Z} \\ Dx_j &= x_{j+1} & j \geq 1. \end{aligned}$$

Then the question of calculating $e^{w \frac{d}{dx}} f(g(x))$ as in the last section is seen to be essentially equivalent to calculating

$$e^{wD} y_0,$$

where we “secretly” identify D with $\frac{d}{dx}$, $f^{(n)}(g(x))$ with y_n and $g^{(m)}(x)$ with x_m . We shall make this identification rigorous in the proof of the following proposition, while noting that the statement of said following proposition is already (unrigorously) clear, by the “secret” identification in conjunction with Proposition 2.1.

Proposition 3.1. *We have*

$$e^{wD} y_0 = \sum_{n \geq 0} \frac{y_n \left(\sum_{m \geq 1} \frac{w^m x_m}{m!} \right)^n}{n!}. \tag{3.1}$$

Proof. Let $f(x), g(x) \in \mathbb{C}[[x]]$ such that $g(x)$ has zero constant term as in Proposition 2.1. Consider the unique algebra homomorphism

$$\phi_{f,g} : \mathbb{C}[\dots, y_{-2}, y_{-1}, y_0, y_1, \dots, x_1, x_2, \dots] \rightarrow \mathbb{C}[[x]]$$

satisfying

$$\begin{aligned} \phi_{f,g} y_i &= f^{(i)}(g(x)) \quad i \in \mathbb{Z} \quad \text{and} \\ \phi_{f,g} x_i &= g^{(i)}(x) \quad i \geq 1. \end{aligned}$$

Then we claim that we have

$$\phi_{f,g} \circ D = \frac{d}{dx} \circ \phi_{f,g}.$$

Since $\phi_{f,g}$ is a homomorphism and D is a derivation, it is clear that we need only check that these operators agree when acting on y_i for $i \in \mathbb{Z}$ and x_j for $j \geq 1$. We get

$$(\phi_{f,g} \circ D) y_i = \phi_{f,g}(y_{i+1} x_1) = f^{(i+1)}(g(x)) g'(x) = \frac{d}{dx} f^{(i)}(g(x)) = \left(\frac{d}{dx} \circ \phi_{f,g} \right) y_i$$

and

$$(\phi_{f,g} \circ D) x_i = \phi_{f,g} x_{i+1} = g^{(i+1)}(x) = \frac{d}{dx} g^{(i)}(x) = \left(\frac{d}{dx} \circ \phi_{f,g} \right) x_i,$$

which gives us the claim. Then, using the obvious extension of $\phi_{f,g}$, by (2.5) we have

$$\phi_{f,g} (e^{wD} y_0) = e^{w \frac{d}{dx}} \phi_{f,g} y_0 = e^{w \frac{d}{dx}} f(g(x)) = \phi_{f,g} \left(\sum_{n \geq 0} \frac{y_n \left(\sum_{m \geq 1} \frac{w^m x_m}{m!} \right)^n}{n!} \right), \quad (3.2)$$

for all $f(x)$ and $g(x)$.

Next take the formal limit as $x \rightarrow 0$ of the first and last terms of (3.2). These identities clearly show that we get identities when we substitute $f^{(n)}(0)$ for y_n and $g^{(n)}(0)$ for x_n in (3.1). But $f^{(n)}(0)$ and $g^{(n)}(0)$ are arbitrary and since (3.1) amounts to a sequence of multinomial polynomial identities when equating the coefficients of w^n , we are done. \square

We observe that it would have been convenient in the previous proof if the maps $\phi_{f,g}$ had been invertible. We provide a second proof of Proposition 3.1 using such a set-up. This proof is closely based on a proof appearing in [Ch]. We hope the reader won't mind a little repetition.

Proof. (second proof of Proposition 3.1)

Let $F(x) = \sum_{n \geq 0} \frac{y_n x^n}{n!}$ and $G(x) = \sum_{n \geq 1} \frac{x_n x^n}{n!}$. Consider the unique algebra homomorphism

$$\psi : \mathbb{C}[\dots, y_{-2}, y_{-1}, y_0, y_1, \dots, x_1, x_2, \dots] \rightarrow \mathbb{C}[\dots, y_{-2}, y_{-1}, y_0, y_1, \dots, x_1, x_2, \dots][[x]]$$

satisfying

$$\begin{aligned}\psi(y_i) &= F^{(i)}(G(x)) \quad i \in \mathbb{Z} \quad \text{and} \\ \psi(x_i) &= G^{(i)}(x) \quad i \geq 1.\end{aligned}$$

Then we claim that we have

$$\psi \circ D = \frac{d}{dx} \circ \psi.$$

Since ψ is a homomorphism and D is a derivation, it is clear that we need only check that these operators agree when acting on y_i for $i \in \mathbb{Z}$ and x_j for $j \geq 1$. We get

$$(\psi \circ D) y_i = \psi(y_{i+1}x_1) = F^{(i+1)}(G(x))G'(x) = \frac{d}{dx}F^{(i)}(G(x)) = \left(\frac{d}{dx} \circ \psi\right) y_i$$

and

$$(\psi \circ D) x_i = \psi(x_{i+1}) = G^{(i+1)}(x) = \frac{d}{dx}G^{(i)}(x) = \left(\frac{d}{dx} \circ \psi\right) x_i,$$

which gives us the claim. Then, using the obvious extension of ψ , we have

$$\psi(e^{wD}y_0) = e^{w\frac{d}{dx}}\psi(y_0) = e^{w\frac{d}{dx}}F(G(x)). \tag{3.3}$$

But now we get to note that ψ has a left inverse, namely setting $x = 0$, because $F^{(i)}(G(0)) = y_i$ for $i \in \mathbb{Z}$ and $G^{(i)}(0) = x_i$ for $i \geq 1$. Thus we get

$$e^{wD}y_0 = \left(e^{w\frac{d}{dx}}F(G(x))\right)|_{x=0} = F(G(x+w))|_{x=0} = F(G(w)), \tag{3.4}$$

which is exactly what we want. □

We note that our second proof of Proposition 3.1 did not depend on Proposition 2.1. Completing a natural circle of reasoning, by using the first proof of Proposition 3.1, before invoking Proposition 2.1, we had from (3.2)

$$\phi_{f,g}(e^{wD}y_0) = e^{w\frac{d}{dx}}\phi_{f,g}y_0 = e^{w\frac{d}{dx}}f(g(x)),$$

which by (3.4) gives

$$e^{w\frac{d}{dx}}f(g(x)) = \phi_{f,g}(F(G(w))),$$

which gives us back Proposition 2.1. Thus we have shown in a natural way how Propositions 2.1 and 3.1 are equivalent.

Remark 3.1. One nice aspect of our second proof of Proposition 3.1, based closely on a proof in [Ch], is that its key brings to the fore of the argument perhaps the most striking feature of the result, which is that the exponential generating function of higher derivatives of a composite function is itself in the form of a composite function. This, of course, is an old-fashioned umbral feature. Furthermore, it was the form of the answer, that it roughly resembled “half of a vertex operator,” which was what interested the authors of [FLM]. This feature is also central to what follows.

We may now clearly state the trick on which (from our point of view) much of the classical umbral calculus is based. It is clear that if we substitute A_n for y_n and B_n for x_n in (3.1) then the right-hand side will become $A(B(w))$. Actually, it will be more interesting to substitute xB_n for x_n . With this as motivation, we formally define two (for flexibility) substitution maps. Let $\chi_{B(t)}$ and $\psi_{A(t)}$ be the algebra homomorphisms uniquely defined by the following:

$$\chi_{B(t)} : \mathbb{C}[\dots, y_{-1}, y_0, y_1, \dots, x_1, x_2, \dots] \rightarrow \mathbb{C}[\dots, y_{-1}, y_0, y_1, \dots, x]$$

with

$$\begin{aligned} \chi_{B(t)}(y_i) &= y_i & i \in \mathbb{Z} \\ \chi_{B(t)}(x_j) &= B_j x & j \geq 1. \end{aligned}$$

and

$$\psi_{A(t)} : \mathbb{C}[\dots, y_{-1}, y_0, y_1, \dots, x] \rightarrow \mathbb{C}[x]$$

with

$$\begin{aligned} \psi_{A(t)}(y_i) &= A_i & i \in \mathbb{Z} \\ \psi_{A(t)}(x) &= x. \end{aligned}$$

Then we have

$$\psi_{A(t)} \circ \chi_{B(t)} (e^{wD} y_0) = A(xB(w)). \quad (3.5)$$

To keep the notation from becoming cluttered, we shall sometimes abbreviate $A(t)$ by simply A and make other similar abbreviations when there should be no confusion.

We next note that it is not difficult to explicitly calculate the action of $\psi_A \circ \chi_B \circ e^{wD}$ on $\mathbb{C}[\dots, y_{-1}, y_0, y_1, \dots, x_1, x_2, \dots]$. Indeed it is easy to see that we have

$$\psi_A \circ \chi_B \circ e^{wD} y_n = A^{(n)}(xB(w)) \quad n \in \mathbb{Z} \quad (3.6)$$

$$\text{and } \psi_A \circ \chi_B \circ e^{wD} x_n = xB^{(n)}(w) \quad n \geq 1. \quad (3.7)$$

These identities determine the action completely because of the automorphism property satisfied by e^{wD} .

The following series of identities (one of which is (3.5) repeated) is immediate from

what we have shown:

$$\psi_A \circ \chi_B \circ e^{wD} y_1 = A'(xB(w)) \quad (3.8)$$

$$\psi_{A'} \circ \chi_B \circ e^{wD} y_0 = A'(xB(w)) \quad (3.9)$$

$$\frac{\partial}{\partial x} \circ \psi_A \circ \chi_B \circ e^{wD} y_{-1} = A(xB(w))B(w) \quad (3.10)$$

$$\psi_{tA(t)} \circ \chi_B \circ e^{wD} y_0 = xB(w)A(xB(w)) \quad (3.11)$$

$$\psi_A \circ \chi_B \circ e^{wD} y_0 = A(xB(w)) \quad (3.12)$$

$$\psi_{A \circ B} \circ \chi_t \circ e^{wD} y_0 = A(B(xw)) \quad (3.13)$$

$$\frac{\partial}{\partial w} \circ \psi_A \circ \chi_B \circ e^{wD} y_0 = A'(x(B(w)))xB'(w) \quad (3.14)$$

$$\psi_{B^*(t)A'(t)} \circ \chi_B \circ e^{wD} y_0 = B^*(xB(w))A'(xB(w)). \quad (3.15)$$

We can now easily get the following proposition.

Proposition 3.2. *We have*

1. $A'(B(w)) = (\psi_A \circ \chi_B \circ e^{wD} y_1)|_{x=1} = (\psi_{A'} \circ \chi_B \circ e^{wD} y_0)|_{x=1}$,
2. $A(B(w))B(w) = (\frac{\partial}{\partial x} \circ \psi_A \circ \chi_B \circ e^{wD} y_{-1})|_{x=1} = (\psi_{tA(t)} \circ \chi_B \circ e^{wD} y_0)|_{x=1}$,
3. $A(B(w)) = (\psi_A \circ \chi_B \circ e^{wD} y_0)|_{x=1} = (\psi_{A \circ B} \circ \chi_t \circ e^{wD} y_0)|_{x=1}$, and
4. $A'(B(w))B'(w) = \frac{\partial}{\partial w} ((\psi_A \circ \chi_B \circ e^{wD} y_0)|_{x=1}) = (\psi_{B^*(t)A'(t)} \circ \chi_B \circ e^{wD} y_0)|_{x=1}$.

Proof. All the identities are proved by setting $x = 1$ in (3.8), (3.9), (3.10), (3.11), (3.12), (3.13), (3.14) and (3.15) and equating the results pairwise as follows. Equations (3.8) and (3.9) give (1); equations (3.10) and (3.11) give (2); equations (3.12) and (3.13) give (3); and equations (3.14) and (3.15) give (4). \square

Each of the identities in Proposition 3.2 turns out to be equivalent to the fact that a certain pair of operators are adjoints. In order to see this, our next task will be to put the procedure of setting $x = 1$, used in Proposition 3.2, into a context of linear functionals. We shall do this in the next section.

4 Umbral connection

We set up a bra-ket notation following [Rm1] so that we may display some of the results obtained in the formalism there presented.

Notation 4.1. Let $f(x) = \sum_{n \geq 0} f_n x^n \in \mathbb{C}[x]$. Then we define

$$\langle A(v) | f(x) \rangle = \sum_{n \geq 0} f_n A_n,$$

where the symbol $\langle \cdot | \cdot \rangle$ is linear in each entry. In particular, $\langle v^k/k! | x^n \rangle = \delta_{k,n}$, where $\delta_{k,n}$ is the Kronecker delta.

So we are now viewing $A(v)$ as a linear functional on $\mathbb{C}[x]$. This leads us to the notion of adjoint operators, a key notion in the umbral calculus as presented in [Rm1]. We shall soon show how to recover certain of the same results about adjoints from our point of view.

Definition 4.1. We say that a linear operator ϕ on $\mathbb{C}[x]$ and a linear operator ϕ^* on $\mathbb{C}[[v]]$ are adjoints exactly when, for all $A(v)$ and for all $p(x)$, the following identity is satisfied:

$$\langle \phi^*(A(v)) | p(x) \rangle = \langle A(v) | \phi(p(x)) \rangle.$$

Of course, by linearity, it is equivalent to require that the identity in Definition 4.1 be satisfied for $p(x)$ ranging over a basis of $\mathbb{C}[x]$. In addition, we extend the bra-ket notation in the obvious way to handle elements of $\mathbb{C}[x][[w]]$ “coefficient-wise.”

Proposition 4.1. *If ϕ is a linear operator on $\mathbb{C}[x]$ and ϕ^* is a linear operator on $\mathbb{C}[[v]]$ such that*

$$\langle \phi^*(A(v)) | e^{xB(w)} \rangle = \langle A(v) | \phi(e^{xB(w)}) \rangle,$$

for all $A(v)$ and $B(w)$, then ϕ and ϕ^ are adjoints.*

Proof. Equating coefficients of w^n gives us the adjoint equation for a sequence of polynomials $B_n(x)$ of degree exactly n and arbitrary $A(v)$. Since the degree of $B_n(x)$ is n , these polynomials form a basis and so the result follows by linearity. \square

Remark 4.1. The sequence of polynomials $B_n(x)$ which appeared in the proof of Proposition 4.1 have been called “basic sequences” or sequences of “binomial type” (see [MR]). We shall call them “attached umbral sequences” (see Definition 4.4 and Remark 4.11).

The next theorem allows us to translate our “set $x = 1$ ” procedure from Proposition 3.2 into the bra-ket notation.

Theorem 4.2. *Let $u \in \mathbb{C}[y_0, y_1, \dots, x]$ be of the form $u = \sum_{n \geq 0} u_n y_n x^n$ where $u_n \in \mathbb{C}$. Then we have:*

$$\langle A(v) | \psi_{e^t}(u) \rangle = (\psi_A(u))|_{x=1}.$$

Proof. We calculate to get:

$$\langle A(v) | \psi_{e^t}(u) \rangle = \langle A(v) | \sum_{n \geq 0} u_n x^n \rangle = \sum_{n \geq 0} u_n A_n = (\psi_A(u))|_{x=1}.$$

\square

We may now easily observe that parts (1) and (2) of Proposition 3.2 yield adjoint relationships as the following theorem formalizes. Since the proofs of each statement are similar to reasoning described below (regarding parts (3) and (4) of Proposition 3.2) and are routine, we omit them, referring the reader to [R3] and [R4] for the details.

Theorem 4.3. *We have*

1. $p(x) \in \mathbb{C}[x]$, viewed as a multiplication operator on $\mathbb{C}[x]$ and $p(\frac{d}{dv})$ are adjoint operators.
2. $F(\frac{d}{dx}) \in \mathbb{C}[[\frac{d}{dx}]]$ and $F(v)$ viewed as a multiplication operator on $\mathbb{C}[[v]]$ are adjoint operators.

□

Remark 4.2. Part (1) of Theorem 4.3 appeared as Theorem 2.1.10 in [Rm1] and Part (2) of Theorem 4.3 appeared as Theorem 2.2.5 in [Rm1].

Parts (3) and (4) of Proposition 3.2 also amount to adjoint relationships. By (3.1) and Theorem 4.2, we have that part (3) of Proposition 3.2 is essentially equivalent to

$$A(B(w)) = \langle A(v) | \psi_{e^t} \circ \chi_B \circ e^{wD} y_0 \rangle = \langle A(B(v)) | \psi_{e^t} \circ \chi_t \circ e^{wD} y_0 \rangle, \quad (4.1)$$

which in turn, by (3.5), gives

$$\langle A(v) | e^{xB(w)} \rangle = \langle A(B(v)) | e^{xw} \rangle. \quad (4.2)$$

We have therefore effectively calculated the adjoint to the substitution map S_B which acts by $S_B(g(v)) = g(B(v))$ for all $g(v) \in \mathbb{C}[[v]]$. We simply need to make a couple of definitions.

Remark 4.3. As mentioned in the introduction, some proofs in this section are “inefficient” if the results are desired merely in and of themselves. As an example of this, we may observe that equation (4.2), which essentially records a classical result as mentioned below, is obvious once one notes that $\langle A(v) | e^{xw} \rangle = A(w)$.

Remark 4.4. We shall be defining certain linear operators on $\mathbb{C}[x]$ by specifying, for instance, how they act on e^{xw} , which, recall, stands for the formal exponential expansion. Of course, by this we mean that the operator acts only on the coefficients of w^n $n \geq 0$. We have already employed similar abuses of notation with the action of $\phi_{f,g}$ in the proof of Proposition 3.1 and with the bra-ket notation as mentioned in the comment preceding Proposition 4.1.

We now recall the definition of certain “umbral operators”; cf. Section 3.4 in [Rm1] More particularly, the umbral operator attached to a series $B(w)$ in this work is the same as the umbral operator for $\bar{B}(w)$ in [Rm1].

Remark 4.5. We shall attempt to always use the word “attached” in this context to indicate the slight discrepancy of notation from Roman’s [Rm1] usage regarding the switch to the compositional inverse.

Definition 4.2. We define the umbral operator attached to $B(w)$ to be the unique linear map $\theta_B : \mathbb{C}[x] \rightarrow \mathbb{C}[x]$ satisfying:

$$\theta_B e^{xw} = e^{xB(w)}.$$

Theorem 4.4. *We have that S_B and θ_B are adjoint operators.*

Proof. The result follows from Proposition 4.1 and (4.2). □

Remark 4.6. Theorem 4.4 essentially appeared as Theorem 3.4.1 in [Rm1], although in this work we have chosen some different characterizations of certain objects as definitions as discussed in the introduction. It is not difficult to tie up all the relevant information. For more details see [R3].

By (3.1) and Theorem 4.2 we have that part (4) of Proposition 3.2 is essentially equivalent to

$$A'(B(w))B'(w) = \frac{\partial}{\partial w} \langle A(v) | \psi_{e^t} \circ \chi_B \circ e^{wD} y_0 \rangle = \langle B^*(v)A'(v) | \psi_{e^t} \circ \chi_B \circ e^{wD} y_0 \rangle,$$

which in turn, by (3.5), gives

$$\begin{aligned} \frac{\partial}{\partial w} \langle A(v) | e^{xB(w)} \rangle &= \langle B^*(v)A'(v) | e^{xB(w)} \rangle \quad \Leftrightarrow \\ \langle A(v) | \frac{\partial}{\partial w} e^{xB(w)} \rangle &= \langle B^*(v)A'(v) | e^{xB(w)} \rangle. \end{aligned} \tag{4.3}$$

We now recall the definition of a certain important special class of “Sheffer shifts”; cf. Section 3.6 in [Rm1]. More particularly, the umbral shift attached to a series $B(w)$ in this work is the same as the umbral shift for $\overline{B}(w)$ in [Rm1].

Remark 4.7. Regarding the word “attached,” in this context see Remark 4.5.

Definition 4.3. For each $B(w)$, let $D_B : \mathbb{C}[x] \rightarrow \mathbb{C}[x]$ be the unique linear map satisfying

$$D_B e^{xB(w)} = \frac{\partial}{\partial w} e^{xB(w)}. \tag{4.4}$$

We call D_B the *umbral shift* attached to $B(w)$.

Remark 4.8. As discussed in the Introduction and Remark 2.2, the authors of [FLM] were concerned with the exponential generating function of the higher derivatives of a composite function, because it roughly resembled “half of” a vertex operator. Following this analogy, we might say that $\psi_{e^t} \circ \chi_B \circ e^{wD} y_0$ is an umbral analogue of (“half of”) a vertex operator. Having made this analogy, one can see how using (4.4), we have defined “attached” umbral shifts as those operators satisfying an analogue of the $L(-1)$ -bracket-derivative property, which is stated as the equality of the first and third expressions of formula (8.7.30) in [FLM]. See also Remark 7.5.

Theorem 4.5. *We have that D_B and $B^*(v) \circ \frac{d}{dv}$ are adjoints.*

Proof. The result follows from Proposition 4.1 and (4.3). □

Remark 4.9. Theorem 4.5 essentially appeared as part of Theorem 3.6.1 in [Rm1], where the author of that work had already, in addition, shown that the operators $B^*(t) \circ \frac{d}{dt}$ are exactly the surjective derivations on $\mathbb{C}[[t]]$, a routine matter once we note that $B^*(t)$ is an arbitrary element of $\mathbb{C}[[t]]$ having a multiplicative inverse. We also mention a similar caveat for the reader regarding different choices of definitions between the present work and [Rm1] just as discussed in Remark 4.6 and in the Introduction.

In closing this section we note obvious characterizations of the attached umbral operators and attached umbral shifts in terms of the coefficients of their generating function definitions. For this it is convenient for us to recall the definition of attached umbral sequences; cf. Section 2.3 and Theorem 2.3.4 in particular in [Rm1]. We note that the umbral sequence “attached” to a series $B(w)$ in this work is the same as the Sheffer sequence associated to $\overline{B}(w)$ in [Rm1].

Remark 4.10. Regarding the use of the word “attached” in conjunction with umbral sequence, see Remark 4.5.

Definition 4.4. We define the sequence of polynomials $B_n(x)$, the umbral sequence attached to $B(w)$, to be the unique sequence satisfying the following:

$$e^{xB(w)} = \sum_{n \geq 0} \frac{x^n B(w)^n}{n!} = \sum_{n \geq 0} \frac{B_n(x) w^n}{n!}.$$

Remark 4.11. We recall that the attached umbral sequences already appeared explicitly (though, of course not by name) in the proof of Proposition 4.1 (see Remark 4.1).

Proposition 4.2. *We have that $\theta_B : \mathbb{C}[x] \rightarrow \mathbb{C}[x]$, the umbral operator attached to $B(w)$, is characterized as the unique linear map satisfying:*

$$\theta_B x^n = B_n(x).$$

□

Proposition 4.3. *We have that $D_B : \mathbb{C}[x] \rightarrow \mathbb{C}[x]$, the umbral shift attached to $B(w)$, is characterized as the unique linear map satisfying:*

$$D_B B_n(x) = B_{n+1}(x).$$

□

5 Umbral shifts revisited

In this section we shall show a characterization of the attached umbral shifts which will be useful in Section 7. We begin by (for temporary convenience) generalizing Definition 4.3.

Definition 5.1. For each $A(t)$ and $B(t)$, let $D_B^A : \mathbb{C}[x] \rightarrow \mathbb{C}[x]$ be the unique linear map satisfying

$$D_B^A e^{xB(t)} = \frac{\partial}{\partial t} A(xB(t)).$$

Recalling the identities (3.6) and (3.7), we note that

$$\begin{aligned} \psi_A \circ \chi_B \circ D e^{wD} y_0 &= \psi_A \circ \chi_B \circ e^{wD} y_1 x_1 \\ &= \psi_A \circ \chi_B \circ (e^{wD} y_1) (e^{wD} x_1) \\ &= A'(xB(w)) x B'(w), \end{aligned}$$

so that

$$\begin{aligned} A'(xB(w)) x B'(w) &= \psi_A \circ \chi_B \circ D e^{wD} y_0 \\ &= \frac{\partial}{\partial w} (\psi_A \circ \chi_B \circ e^{wD} y_0) \\ &= D_B^A \circ \psi_{e^t} \circ \chi_B \circ e^{wD} y_0. \end{aligned}$$

Recalling that e^{wD} stands for the formal exponential Taylor series, and extracting the coefficients in $w^n/n!$ for $n \geq 0$ from the second and fourth terms from the above identity yields:

$$D_B^A \circ \psi_{e^t} \circ \chi_B \circ D^n y_0 = \psi_A \circ \chi_B \circ D^{n+1} y_0.$$

Furthermore, because $\psi_{e^t} \circ \chi_B \circ D^n y_0$ is a polynomial of degree exactly n , this formula characterizes the maps D_B^A .

Although we have briefly generalized the definition for the attached umbral shifts (in order to fit more closely with our calculations from Section 3), the previous identity shows how it is natural to restrict our attention to the attached umbral shifts, and it is this case that will later interest us anyway. We may now state the characterization of the attached umbral shifts mentioned in the introduction to this section.

Remark 5.1. While we only temporarily generalized the definition for attached umbral shifts, as the referee has pointed out, it might be nice to investigate whether extensions of standard umbral calculus calculations could be developed for the operators D_B^A .

Proposition 5.1. *The attached umbral shift, $D_B : \mathbb{C}[x] \rightarrow \mathbb{C}[x]$ is the unique linear map satisfying*

$$D_B \circ \psi_{e^t} \circ \chi_B \circ D^n y_0 = \psi_{e^t} \circ \chi_B \circ D^{n+1} y_0,$$

for all $n \geq 0$.

□

Remark 5.2. Proposition 5.1 was announced, together with a more direct proof, as Proposition 6.1 in [R2].

Remark 5.3. The classical umbral calculus can be considered to be the study of Sheffer sequences through “umbral” techniques. So far we have only considered a special case, attached umbral sequences. The general case may be obtained easily by using the results of the special case and this is probably the shortest route given the efforts we have already made. Alternatively one could give a complete parallel development. Recall that our approach began by calculating the higher derivatives of a formal composite function. That is, letting $f(x), g(x) \in \mathbb{C}[[x]]$ such that the constant term of $g(x)$ is zero, we began calculating the higher derivatives of $f(g(x))$. Let $h(x) \in \mathbb{C}[[x]]$. Then the general theory follows by the parallel argument with the starting point of calculating the higher derivatives of the product $h(x)f(g(x))$. The interested reader may see this worked out in detail in both [R3] and [R4].

6 Connection with the Virasoro algebra

Our goal in this section will be to show that an operator closely related to the derivation D (which appeared in Section 3) is one of the standard quadratic representations of the $L(-1)$ operator of the Virasoro algebra.

Recall that we began our main investigation by calculating the higher derivatives of the composition of two formal power series $f(x)$ and $g(x)$, where the constant term of $g(x)$ was required to be 0, following a proof given in [FLM]. In fact, the case that interested the authors in [FLM] was when $f(x) = e^x$. It is not difficult to specialize our arguments to this case. When we abstract, we get the following set-up: Consider the vector space $y\mathbb{C}[x_1, x_2, x_3, \dots]$ where x_j for $j \geq 1$ are commuting formal variables. Then let \mathcal{D} be the unique derivation on $y\mathbb{C}[x_1, x_2, x_3, \dots]$ satisfying

$$\begin{aligned} \mathcal{D}y &= yx_1 \\ \mathcal{D}x_j &= x_{j+1} \quad j \geq 1. \end{aligned} \tag{6.1}$$

The question of calculating $e^{w \frac{d}{dx}} e^{g(x)}$ is seen to be essentially equivalent to calculating

$$e^{w\mathcal{D}}y,$$

where we “secretly” identify \mathcal{D} with $\frac{d}{dx}$, $e^{g(x)}$ with y and $g^{(m)}(x)$ with x_m . It is clear by our identification, and rigorously as an easy corollary of Proposition 3.1, that:

$$e^{w\mathcal{D}}y = \sum_{n \geq 0} \frac{y \left(\sum_{m \geq 1} \frac{w^m x_m}{m!} \right)^n}{n!} = ye^{\sum_{m \geq 1} \frac{w^m x_m}{m!}}. \tag{6.2}$$

We note that

$$\begin{aligned} \mathcal{D} &= x_1 y \frac{\partial}{\partial y} + x_2 \frac{\partial}{\partial x_1} + x_3 \frac{\partial}{\partial x_2} + \cdots \\ &= x_1 y \frac{\partial}{\partial y} + \sum_{k \geq 1} x_{k+1} \frac{\partial}{\partial x_k}. \end{aligned}$$

We next shall switch gears in order to recall certain basics about the Virasoro algebra using operators arising from certain Heisenberg Lie algebras. We follow (a variant of) the exposition of this well-known material in [FLM]. Let \mathfrak{h} be the one-dimensional abelian (complex) Lie algebra with basis element h . We define a nonsingular symmetric bilinear form on \mathfrak{h} by $(ah, bh) = ab$ for all $a, b \in \mathbb{C}$. We recall the (particular) affine Heisenberg Lie algebra $\widehat{\mathfrak{h}}$ which is the vector space

$$\widehat{\mathfrak{h}} = \mathfrak{h} \otimes \mathbb{C}[t, t^{-1}] \oplus \mathbb{C}c,$$

with Lie brackets determined by

$$[ah \otimes t^m, bh \otimes t^n] = (ah, bh)m\delta_{m+n,0}c = abm\delta_{m+n,0}c,$$

where c is central and δ is the Kronecker delta.

We may realize $\widehat{\mathfrak{h}}$ as differential and multiplication operators on a space with infinitely many variables as follows. We consider the space $y\mathbb{C}[x_1, x_2, x_3, \dots]$ and make the following identification:

$$h \otimes t^n = \begin{cases} \alpha(-n)x_{-n} & n < 0 \\ \beta(n)\frac{\partial}{\partial x_n} & n > 0 \\ y\frac{\partial}{\partial y} & n = 0, \end{cases}$$

where $\alpha(n), \beta(n) \in \mathbb{C}$ for $n \neq 0$ and we identify c , the central element, with the multiplication by identity operator. Of course, in this setting $y\frac{\partial}{\partial y}$ is a fancy name for the identity operator, but we wrote it this way so that it appears explicitly as a derivation. It is easy to see that

$$\left[\alpha(n)x_n, \beta(n)\frac{\partial}{\partial x_n} \right] = -\alpha(n)\beta(n),$$

with all other pairs of operators commuting. Thus our identification gives a representation of the Heisenberg Lie algebra exactly when we require that for all $n \neq 0$

$$\alpha(n)\beta(n) = n.$$

For this representation we shall sometimes use the notation $h(n)$ to denote the image of $h \otimes t^n$.

Definition 6.1. The Virasoro algebra of central charge 1 is the Lie algebra generated by basis elements 1 , a central element, and $L(n)$ for $n \in \mathbb{Z}$ which satisfy the following relations for all $m, n \in \mathbb{Z}$.

$$[L(m), L(n)] = (m - n)L(m + n) + (1/12)(m^3 - m)\delta_{m+n,0}. \quad (6.3)$$

Remark 6.1. The Virasoro algebra is a central extension of the Witt algebra, the Lie algebra of the derivations of Laurent polynomials in a single formal variable. The correspondence can be seen by identifying $L(n)$ with $-t^{n+1}\frac{d}{dt}$. In fact, (cf. Proposition 1.9.4 in [FLM]) the Virasoro algebra with a general central element is the unique, up to isomorphism, one dimensional central extension of the Witt algebra.

We state the following theorem without proof. It is a special case, for instance of Theorem 1.9.6 in [FLM] where a complete proof is provided.

Theorem 6.1. *The operators*

$$L(n) = \frac{1}{2} \sum_{k \in \mathbb{Z}} h(n-k)h(k) \quad n \neq 0 \quad \text{and} \quad (6.4)$$

$$L(0) = \frac{1}{2} \sum_{k \in \mathbb{Z}} h(-|k|)h(|k|) \quad (6.5)$$

give a representation of the Virasoro algebra of central charge 1,

The space $y\mathbb{C}[x_1, x_2, x_3, \dots]$ is obviously a module for the Virasoro algebra. It is graded by $L(0)$ eigenvalues, which are called *weights*. In the literature, such a module is often called a lowest weight module; this module has y as a lowest weight vector.

We have

$$\begin{aligned} L(0) &= \frac{1}{2}h(0)^2 + h(-1)h(1) + h(-2)h(2) + \dots \\ &= \frac{1}{2}y \frac{\partial}{\partial y} \circ y \frac{\partial}{\partial y} + \alpha(1)\beta(1)x_1 \frac{\partial}{\partial x_1} + \alpha(2)\beta(2)x_2 \frac{\partial}{\partial x_2} + \dots \end{aligned}$$

so that $L(0)y = \frac{1}{2}y$. Thus the lowest weight of the module is $\frac{1}{2}$.

We may now show that by an appropriate (unique) choice of $\alpha(n)$ and $\beta(n)$ we get $\mathcal{D} = L(-1)$. We have

$$\begin{aligned} L(-1) &= h(-1)h(0) + h(-2)h(1) + h(-3)h(2) + \dots \\ &= \alpha(1)x_1y \frac{\partial}{\partial y} + \alpha(2)\beta(1)x_2 \frac{\partial}{\partial x_1} + \alpha(3)\beta(2)x_3 \frac{\partial}{\partial x_2} + \dots \end{aligned}$$

Therefore it is clear that in order to have $\mathcal{D} = L(-1)$, we need exactly that

$$\begin{aligned} \alpha(1) &= 1 \quad \text{and} \\ \alpha(n+1)\beta(n) &= 1 \quad n \geq 1, \end{aligned}$$

where we recall that we already have the restriction that $\alpha(n)\beta(n) = n$ for all $n \geq 1$. These two sets of restrictions imply that

$$\begin{aligned} (n+1)\beta(n) &= \beta(n+1) \quad n \geq 1 \\ \beta(1) &= 1, \end{aligned}$$

so that

$$\begin{aligned}\beta(n) &= n! \\ \alpha(n) &= \frac{1}{(n-1)!},\end{aligned}$$

for all $n \geq 1$, is the unique solution. We record this as a proposition.

Proposition 6.1. *The operator $L(-1)$, given by (6.4), with (and only with) both $\alpha(n) = \frac{1}{(n-1)!}$ and $\beta(n) = n!$, is identical to the operator \mathcal{D} , given by (6.1).*

□

For the remainder of this paper we shall assume that $\alpha(n) = \frac{1}{(n-1)!}$ and $\beta(n) = n!$.

7 Umbral shifts revisited and generalized

We shall continue to consider the space $y\mathbb{C}[x_1, x_2, x_3, \dots]$ as in the previous section, and similarly to some of our previous work, such as in Section 3, we shall consider certain substitution maps. Let $\phi_{B(t)}$ denote the following algebra homomorphism.

$$\phi_B : y\mathbb{C}[x_1, x_2, x_3, \dots] \rightarrow \mathbb{C}[x]$$

with

$$\begin{aligned}\phi_{B(t)}x_j &= B_jx & j \geq 1 \\ \text{and } \phi_{B(t)}y &= 1.\end{aligned}$$

Then we have

$$\phi_B \circ e^{w\mathcal{D}}y = \phi_B \circ e^{wL(-1)}y = e^{xB(w)}.$$

In light of Proposition 5.1, it is routine to show the following.

Theorem 7.1. *The attached umbral shift, $D_B : \mathbb{C}[x] \rightarrow \mathbb{C}[x]$ is the unique linear map satisfying*

$$D_B \circ \phi_B \circ L(-1)^ny = \phi_B \circ L(-1)^{n+1}y,$$

for all $n \geq 0$.

□

With Theorem 7.1 as motivation, we make the following definition.

Definition 7.1. For $m \geq -1$ we define the operators $\mathcal{D}_B(m) : \mathbb{C}[x] \rightarrow \mathbb{C}[x]$ to be the unique linear maps satisfying

$$\mathcal{D}_B(m) \circ \phi_B \circ L(-1)^n y = \phi_B \circ L(m)L(-1)^n y,$$

for all $n \geq 0$.

Of course, $\mathcal{D}_B(-1) = D_B$. These operators are well-defined because $\phi_B \circ L(-1)^n y$ has degree exactly n . In fact, $\phi_B \circ L(-1)^n y = B_n(x)$, the umbral polynomial attached to $B(t)$. We call the operators $\mathcal{D}_B(n)$ generalized attached umbral shifts.

We would like to use the Virasoro relations to help compute the generalized attached umbral shifts. We begin with the following lemma.

Lemma 7.1. *There exist rational numbers $f_m(n)$ such that*

$$L(m)L(-1)^n y = f_m(n)L(-1)^{n-m} y,$$

for all $m \geq -1, n \geq 0$ such that $n \geq m$.

Proof. We will need that $L(m)L(-1)^n y = 0$ when $m > n$. This is really due to the weights of the vectors, but in this paper we shall proceed, in just this special case, with an elementary induction argument. We induct on n . For $n = 0$ this follows essentially because y is a lowest weight vector, but even without considering weights it is easy to directly see given the definition of the operators. By induction (used twice) we have

$$\begin{aligned} L(m)L(-1)^n y &= L(-1)L(m)L(-1)^{n-1} y + [L(m), L(-1)]L(-1)^{n-1} y \\ &= (m+1)L(m-1)L(-1)^{n-1} y \\ &= 0. \end{aligned}$$

We may now focus on the main argument. We establish the boundary cases. Letting $m = -1$ we easily check that $f_{-1}(n) = 1$. The other boundary is $m = n$. We shall use another intermediate induction to establish this case. Our base case then is $m = n = 0$ for which it is easy to check that $f_0(0) = 1/2$. We also have

$$\begin{aligned} L(n)L(-1)^n y &= L(-1)L(n)L(-1)^{n-1} y + [L(n), L(-1)]L(-1)^{n-1} y \\ &= (n+1)L(n-1)L(-1)^{n-1} y, \end{aligned}$$

so that inducting on n we get our result. Moreover we now have the recurrence

$$f_n(n) = (n+1)f_{n-1}(n-1) \quad n \geq 1, \tag{7.1}$$

with, as we have seen, the boundary $f_0(0) = 1/2$.

For our main argument we induct on $m+n$. We have already checked the base case. We then have by induction and using the Virasoro relations that for the remaining cases $m \geq 0$ and $n > m$, we have

$$\begin{aligned} L(m)L(-1)^n y &= L(-1)L(m)L(-1)^{n-1} y + [L(m), L(-1)]L(-1)^{n-1} y \\ &= f_m(n-1)L(-1)^{n-m} y + (m+1)L(m-1)L(-1)^{n-1} y \\ &= (f_m(n-1) + (m+1)f_{m-1}(n-1)) L(-1)^{n-m} y. \end{aligned}$$

Therefore, not only do the values $f_n(m)$ exist but we have a recurrence for them

$$f_m(n) = f_m(n-1) + (m+1)f_{m-1}(n-1). \quad (7.2)$$

□

In the last proposition we found a recurrence for certain values $f_m(n)$. We could extend the range of m and n to include all $m \geq -1$ and $n \in \mathbb{Z}$ and define $f_m(n)$ to be the solution to the recurrence equation found above which coincides when $n = 0$ and $n > m$ with the values already defined. In fact, it is easy to find a simpler boundary condition yielding the desired solution other than using the boundary with $m = n$. It is easy to see that instead we may specify that $f_m(0) = 0$ for $m \geq 1$, by considering (7.1), which shows that we may specify 0's below the diagonal. Further, it is easy to see from this recurrence equation with given boundary, that $f_m(n)$ is an integer for $m \neq 0$ and that $f_0(n)$ are half integers. It is also easy to see from this recurrence, by induction on n , that we have for $n \geq 0$ that

$$\begin{aligned} f_m(n) &= f_m(0) + (m+1) \sum_{i=1}^n f_{m-1}(n-i) \\ &= f_m(0) + (m+1) \sum_{i=0}^{n-1} f_{m-1}(i). \end{aligned} \quad (7.3)$$

We now give the natural generalization to Proposition 4.3.

Proposition 7.1. *We have that $\mathcal{D}_B(m) : \mathbb{C}[x] \rightarrow \mathbb{C}[x]$, the generalized umbral shift attached to $B(t)$, is characterized as the unique linear map satisfying:*

$$\mathcal{D}_B(m)B_n(x) = f_m(n)B_{n-m}(x), \quad (7.4)$$

where by convention $B_n(x) = 0$ for $n \leq -1$.

Proof. By definition 7.1 and Lemma 7.1 we have

$$\begin{aligned} \mathcal{D}_B(m)B_n(x) &= \phi_B L(m) L(-1)^n y \\ &= f_m(n) \phi_B L(-1)^{n-m} y \\ &= f_m(n) B_{n-m}(x). \end{aligned}$$

□

Remark 7.1. The convention in Proposition 7.1 that $B_n(x) = 0$ for $n \leq -1$ is only used to ensure that in all cases the right hand side of (7.4) is well defined. This condition could have allowed for much more flexibility since we already have that $f_m(n) = 0$ whenever $n - m \leq -1$.

We shall next solve for simple closed (polynomial) expressions for $f_m(n)$ for fixed m . If we are willing to sum over squares, cubes etc. we could compute the answer for nonnegative n for each (fixed) m in turn, using (7.3). It is easy to verify that $f_0(n) = n + 1/2$ and $f_1(n) = n^2$ solve the first two cases. To solve for the remaining cases however, for variety, we shall use a heuristic argument making use of the Virasoro algebra relations to derive the answer. We have, for $m + l \leq n$,

$$\begin{aligned}
(l - m)f_{l+m}(n)L(-1)^{n-l-m}y &= (l - m)L(l + m)L(-1)^ny \\
&= [L(l), L(m)]L(-1)^ny \\
&= L(l)L(m)L(-1)^ny - L(m)L(l)L(-1)^ny \\
&= L(l)f_m(n)L(-1)^{n-m}y - L(m)f_l(n)L(-1)^{n-l}y \\
&= f_l(n - m)f_m(n)L(-1)^{n-m-l}y \\
&\quad - f_m(n - l)f_l(n)L(-1)^{n-l-m}y \\
&= (f_l(n - m)f_m(n) - f_m(n - l)f_l(n))L(-1)^{n-l-m}y.
\end{aligned}$$

We shall for the time being (unmathematically) ignore the restriction on the indices and get, for whenever all terms are well defined, the identity

$$(l - m)f_{l+m}(n) = f_l(n - m)f_m(n) - f_m(n - l)f_l(n).$$

It is easy to see that the case $l = -1$ recovers 7.2. Further, one can check that the case $l = 0$ does not yield any new information. For $l = 1$ we get

$$\begin{aligned}
(1 - m)f_{m+1}(n) &= f_1(n - m)f_m(n) - f_m(n - 1)f_1(n) \\
&= (n - m)^2f_m(n) - n^2f_m(n - 1),
\end{aligned}$$

so that it is easy to calculate, by simple substitution, each higher case (in m) starting with $m = 2$. The calculations are not difficult, of course, but I myself “cheated” and used Maple to find and factor the first few answers, which yield an easy and obvious pattern as follows:

$$\begin{aligned}
f_{-1}(n) &= 1 \\
f_0(n) &= n + 1/2 \\
f_1(n) &= n^2 \\
f_2(n) &= (1/2)n(n - 1)(2n - 1) \\
f_3(n) &= n(n - 1)^2(n - 2) \\
f_4(n) &= (1/2)n(n - 1)(n - 2)(n - 3)(2n - 3) \\
f_5(n) &= n(n - 1)(n - 2)^2(n - 3)(n - 4) \\
f_6(n) &= (1/2)n(n - 1)(n - 2)(n - 3)(n - 4)(n - 5)(2n - 5) \\
f_7(n) &= n(n - 1)(n - 2)(n - 3)^2(n - 4)(n - 5)(n - 6) \\
&\vdots
\end{aligned}$$

With that as a guide, we may now return to doing rigorous math and state and prove the following theorem.

Theorem 7.2. *The unique solution to the recurrence equation (7.2) with $m \geq -1$ and $n \in \mathbb{Z}$ and with boundary given by $f_{-1}(n) = 1$, $f_0(0) = 1/2$ and $f_m(0) = 0$ for $m \geq 1$ is given by*

$$\begin{aligned} f_{-1}(n) &= 1 \\ f_m(n) &= (1/2)(n(n-1)(n-2)\cdots(n-m+1))(2n-m+1) \quad \text{for } m \geq 0. \end{aligned}$$

Proof. The proof is a straightforward calculation. Let $m \geq 0$ (although admittedly the low m cases are a bit degenerate in this notation). Then we have

$$\begin{aligned} f_m(n) - f_m(n-1) &= (1/2)(n(n-1)(n-2)\cdots(n-m+1))(2n-m+1) \\ &\quad - (1/2)(n-1)(n-2)\cdots(n-m)(2(n-1)-m+1) \\ &= (1/2)(n-1)(n-2)\cdots(n-m+1) \cdot \\ &\quad \cdot (n(2n-m+1) - (n-m)(2(n-1)-m+1)) \\ &= (1/2)(n-1)(n-2)\cdots(n-m+1)(2mn+2n-m^2-m) \\ &= (n-1)(n-2)\cdots(n-m+1)(m+1)(n-m/2) \\ &= (m+1)f_{m-1}(n-1). \end{aligned}$$

□

Remark 7.2. Recalling (7.3), it is easy to see that we could use the last result, perhaps somewhat awkwardly, to solve for the sum of squares and cubes etc., which happens to be related to the Bernoulli numbers, one of the motivating subjects for Blissard [Bli] and is one of the classic problems solved via umbral methods (cf. Chapter 11 [Do] for a nice, succinct old-fashioned umbral style proof and also Chapter 3.11 [Mel]).

Remark 7.3. We note that the umbral calculus has long been known to have connections to the Bernoulli numbers and polynomials (see e.g. [Mel]). Bernoulli polynomials have also appeared in the literature of vertex algebra theory (see e.g. [L1] and [DLM]). Just as we have been establishing some analogues and connections between umbral calculus and vertex algebra theory, it might be interesting in future work to investigate further possible connections explicitly related to Bernoulli numbers and polynomials.

Remark 7.4. We used a heuristic argument emphasizing the connection between umbral calculus and the Virasoro algebra to help guess a solution to $f_m(n)$ leading up to Theorem 7.2. However, as the referee has pointed out, one can connect the result with classical umbral calculus as well. The calculations below essentially reproduce results obtained and shown to me by the referee and we shall follow their reasoning. Any errors or inelegance etc. in the particular exposition presented here are entirely due to me. Observing in all that follows the typical caveats about the degeneracy of cases for “low values of n ”, for $n \geq 0$ let

$$t_n(x) = f_{n-1}(x+n). \tag{7.5}$$

Assuming we know that $t_n(x)$ are polynomials, then using (7.2) for $n - 1$ we get

$$t_n(x) = t_n(x - 1) + nt_{n-1}(x),$$

since the result holds for all positive integral values of x . Using the formal Taylor theorem (2.3), we get for $n - 1$ that

$$\begin{aligned} \left(1 - e^{-\frac{d}{dx}}\right) t_n(x) &= t_n(x) - t_n(x - 1) \\ &= t_n(x - 1) + nt_{n-1}(x) - t_n(x - 1) \\ &= nt_{n-1}(x). \end{aligned}$$

Therefore $t_n(x)$ is a Sheffer sequence (cf. Theorem 2.3.7 [Rm1]). It is easy to see (cf. Theorem 2.4.3 in [Rm1]) that the relevant associated sequence $p_k(y)$ satisfies

$$\sum_{k=0}^{\infty} \frac{p_k(y)}{k!} t^k = e^{y \log(1/(1-t))} = (1 - t)^{-y},$$

which yields

$$p_k(y) = y(y + 1) \cdots (y + k - 1).$$

We further have (cf. Theorem 2.3.9 [Rm1])

$$t_n(x + y) = \sum_{k=0}^n \binom{n}{k} p_k(y) t_{n-k}(x).$$

It is easy to see that $t_0(-2) = 1$, $t_1(-2) = -1/2$ and that $t_n(-2) = 0$ for $n > 1$ so that for $n \geq 0$

$$\begin{aligned} t_n(x) &= \sum_{k=0}^n \binom{n}{k} p_k(x + 2) t_{n-k}(-2) \\ &= p_n(x + 2) - \frac{n}{2} p_{n-1}(x + 2) \\ &= (x + 2)(x + 3) \cdots (x + n + 1) - \frac{n}{2} (x + 2)(x + 3) \cdots (x + n) \\ &= \frac{1}{2} (x + 2)(x + 3) \cdots (x + n)(2x + n + 2). \end{aligned} \tag{7.6}$$

It is easy to check this formula using (7.5) and Theorem 7.2. The referee also pointed out that letting

$$s_n(x) = \frac{t_n(x)}{n!} \left(= \frac{f_{n-1}(x + n)}{n!} \right),$$

it is easy to see that the following particularly simple recursion holds

$$s_n(x) = s_{n-1}(x) + s_n(x - 1),$$

and, continuing to follow the referee, using (7.6) it is easy to see that

$$s_n(x) = \binom{x+n+1}{n} - \frac{1}{2} \binom{x+n}{n-1}.$$

We shall conclude this paper by stating and proving the natural generalization to the original formula defining the attached umbral shifts in Definition 4.3.

Proposition 7.2. *For each $m \geq -1$, the map $\mathcal{D}_B(m) : \mathbb{C}[x] \rightarrow \mathbb{C}[x]$ is the unique linear map satisfying:*

$$\mathcal{D}_B(m)e^{xB(w)} = \left(w^{m+1} \frac{\partial}{\partial w} + \frac{m+1}{2} w^m \right) e^{xB(w)}.$$

Proof. We calculate, for $m \geq -1$ (although once again the low m cases are a bit degenerate) to get

$$\begin{aligned} w^{m+1} \frac{\partial}{\partial w} e^{xB(w)} &= w^{m+1} \frac{\partial}{\partial w} \sum_{n \geq 0} \frac{B_n(x) w^n}{n!} \\ &= \sum_{n \geq 0} \frac{n B_n(x) w^{n+m}}{n!} \\ &= \sum_{n \geq m} \frac{(n-m) B_{n-m}(x) w^n}{(n-m)!} \\ &= \sum_{n \geq m} \frac{n(n-1) \cdots (n-m) B_{n-m}(x) w^n}{n!} \end{aligned}$$

and

$$\begin{aligned} w^m e^{xB(w)} &= \sum_{n \geq 0} \frac{B_n(x) w^{n+m}}{n!} \\ &= \sum_{n \geq m} \frac{B_{n-m}(x) w^n}{(n-m)!} \\ &= \sum_{n \geq m} \frac{n(n-1) \cdots (n-m+1) B_{n-m}(x) w^n}{n!}, \end{aligned}$$

so that it is easy to check that

$$\begin{aligned} \left(w^{m+1} \frac{\partial}{\partial w} + \frac{m+1}{2} w^m \right) e^{xB(w)} &= \sum_{n \geq m} \frac{f_m(n) B_{n-m}(x) w^n}{n!} \\ &= \mathcal{D}_B(m) e^{xB(w)}. \end{aligned}$$

□

Remark 7.5. Building on Remarks 2.2 and 4.8 we may regard Proposition 7.2 as an analogue of formula (8.7.37) in [FLM] in the cases where the n in (8.7.37) in [FLM] is restricted so that $n \geq -1$. The h in formula (8.7.37) in [FLM] should be replaced by the weight of the relevant lowest weight vector, which in our setting seems perhaps to correspond with the lowest weight of the module of the Virasoro algebra which we have been considering, which as we have noted is indeed $1/2$.

References

- [A] T.A., Sur la différentiation des fonctions de fonctions, *Nouvelles Annales de Mathématiques* **9** (1850), 119–125.
- [An] G. Andrews, The Theory of Partitions, in *Encyclopedia of Mathematics and its Applications*, ed. by G.-C. Rota, Vol. 2, Addison-Wesley, Reading, Massachusetts, (1976), Cambridge University Press, Cambridge (1984), Cambridge Mathematical Library (1998).
- [B] E.T. Bell, The history of Blissard’s symbolic calculus, with a sketch of the inventor’s life. *Amer. Math. Monthly*, **45** (1938) 414–421.
- [Bli] J. Blissard, Theory of generic functions, *Quarterly J. Pure Appl. Math.* **4** (1861), 279–305; **5** (1862), 58–75, 185–208.
- [Bl] R. Block, On the Mills-Seligman axioms for Lie algebras of classical type, *Trans. Amer. Math. Soc.* **121** (1966), 378–392.
- [BL] A. Di Bucchiano and D.E. Loeb, A selected survey of umbral calculus, *Elec. J. Combin.*, 3:Dynamical Surveys Section, 1995.
- [Ch] W.Y.C. Chen, Context-free grammars, differential operators and formal power series, Conference on Formal Power Series and Algebraic Combinatorics (Bordeaux, 1991); *Theoret. Comput. Sci.* **117** (1993), 113–129.
- [DS] E. DiNardo and D. Senato, Umbral nature of the Poisson random variables, *Algebraic combinatorics and computer science*, Springer Italia, Milan (2001), 245–266.
- [Do] H. Dörrie (transl. D. Antin), *100 Great Problems of Elementary Mathematics*, Dover Publications, Inc., New York, 1965 (orig. German version appeared in 1932).
- [DLM] B. Doyon, J. Lepowsky and A. Milas, Twisted vertex operators and Bernoulli polynomials, *Contemp. Math.* **8** (2006) no. 2, 247–307.
- [F1] C.F. Faà di Bruno, Sullo sviluppo delle funzioni, *Annali di Scienze Matematiche e Fisiche* **6** (1855), 479–480.
- [F2] C.F. Faà di Bruno, Note sur une nouvelle formule de calcul différentiel, *Quarterly J. Pure Appl. Math.* **1** (1857), 359–360.
- [FHL] I. Frenkel, Y.-Z. Huang and J. Lepowsky, On axiomatic approaches to vertex operators and modules, *Memoirs Amer. Math. Soc.* **104**, 1993.

- [FLM] I. Frenkel, J. Lepowsky and A. Meurman, *Vertex Operator Algebras and the Monster*, Pure and Appl. Math., Vol. 134, Academic Press, New York, 1988.
- [Fr] J.M. Freeman, Transforms of Operators on $k[x][[t]]$, *Congressus Numerantium* **48** (1985), 115–132.
- [Ga] A. Garsia, An exposé of the Mullin-Rota theory of polynomials of binomial type, *Lin. and Multilin. Alg.* **1** (1973), 47–65.
- [GF] I.M. Gelfand and D.B. Fuchs, The cohomology of the Lie algebra of vector fields on a circle, *Funk. Anal. i Prilozhen.* **2** (1968), 342–343.
- [H] Y.-Z. Huang, *Two-dimensional Conformal Field Theory and Vertex Operator Algebras* Progress in Math. **148**, Birkhäuser, Boston, 1998.
- [HLZ] Y.-Z. Huang, J. Lepowsky and L. Zhang, Logarithmic tensor product theory for generalized modules for a conformal vertex algebra, arXiv:0710.2687v3 [math.QA]
- [Jo] W.P. Johnson, The curious history of Faà di Bruno’s formula, *Amer. Math. Monthly* **109** (2002), 217–234.
- [L1] J. Lepowsky, Application of a “Jacobi identity” for vertex operator algebras to zeta values and differential operators, *Lett. Math. Phys.* **53** (2000), no. 2, 551–579.
- [L2] J. Lepowsky, Some developments in vertex operator algebra theory, old and new, *Lie algebras, vertex operator algebras and their applications* Contemp. Math. **442**, Amer. Math. Soc., Providence, RI (2007), 355–387.
- [LL] J. Lepowsky and H. Li, *Introduction to Vertex Operator Algebras and Their Representations*, Progress in Math., Vol. 227, Birkhäuser, Boston, 2003.
- [Lu] E. Lucas, Théorie nouvelle des nombres de Bernoulli et d’Euler, *Comptes rendus de l’Academie des Sciences (Paris)*, **83** (1876), 539–541; *Annali di Matematica pura ed applicata, Serie 2* **8** (1877), 56–79.
- [Mel] Z.A. Melzak, *Companion to Concrete Mathematics: Mathematical techniques and various applications*, Pure and Applied Mathematics, John Wiley and Sons, New York-London-Sydney-Toronto, 1973.
- [Me] U.H. Meyer, Sur les dérivées d’une fonction de fonction, *Archiv der Mathematik und Physik* **9** (1847), 96–100.
- [Mi] A. Milas, Weak modules and logarithmic intertwining operators for vertex operator algebras, in: *Recent Developments in Infinite-Dimensional Lie Algebras and Conformal Field Theory*, ed. S. Berman, P. Fendley, Y.-Z. Huang, K. Misra, and B. Parshall, Contemp. Math., Vol. 297, American Mathematical Society, Providence, RI, 2002, 201–225.
- [MR] R. Mullin, and G.-C. Rota, On the foundations of combinatorial theory, in *Graph Theory and its Applications* (B. Harris, ed.), Academic Press, New York-London (1970).

- [N] H. Niederhausen, Rota’s umbral calculus and recursions: Dedicated to the memory of Gian-Carlo Rota, *Algebra Universalis* **49** (2003) no. 4, 435–457.
- [Ri1] J. Riordan, Derivatives of composite functions, *Bull. Amer. Math. Soc.* **52** (1946), 664–667.
- [Ri2] J. Riordan, *Combinatorial Identities*, Wiley, New York, 1968.
- [R1] T.J. Robinson, The formal Taylor theorem revisited, to appear.
- [R2] T.J. Robinson, New perspectives on exponentiated derivations, the formal Taylor theorem, and Faà di Bruno’s formula in: Proceedings of the Conference on Vertex Operator Algebras, Illinois State University, (2008) ed. by Maarten Bergvelt, Gaywalee Yamskulna, and Wenhua Zhao, Contemporary Math. 497, Amer. Math. Soc., Providence, (2009), 185–198.
- [R3] Formal calculus, umbral calculus, and basic axiomatics of vertex algebras, Ph.D. thesis, Rutgers University, 2009.
- [R4] T.J. Robinson, Formal calculus and umbral calculus, arXiv:0912.0961 v2 [math.QA].
- [Rm1] S. Roman, *The Umbral Calculus*, Pure and Appl. Math., **111**, Academic Press, New York, 1984.
- [Rm2] S. Roman, The formula of Faà di Bruno, *Amer. Math. Monthly*, **87** (1980), 805–809.
- [Rt] G.-C. Rota, *Finite Operator Calculus*, Academic Press, New York, 1975.
- [RKO] G.-C. Rota, D. Kahaner and A. Odlyzko, On the foundations of combinatorial theory VIII: finite operator calculus, *J. Math. Anal. Appl.* **42** (1973), 684–760.
- [RR] G.-C. Rota and S. Roman, The umbral calculus, *Adv. in Math.* **27** (1978), 95–188.
- [Sc] G. Scott, Formulae of successive differentiation, *Quarterly J. Pure Appl. Math.* **4** (1861), 77–92.
- [T] B. Taylor, Difference equations via the classical umbral calculus in: Mathematical essays in honor of Gian-Carlo Rota (Cambridge, MA, 1996), ed. B. Sagan and R. Stanley, *Progr. Math.* **161**, Birkhäuser, Boston (1998) 397–411.
- [V] M.A. Virasoro, Subsidiary conditions and ghosts in dual-resonance models, *Phys. Rev.* **D1** (1970), 2933–2936.
- [Wi] R.L. Wilson, Classification of generalized Witt algebras over algebraically closed fields, *Trans. Amer. Math. Soc.* **153** (1971), 191–210.
- [Z1] D. Zeilberger, Toward a combinatorial proof of the Jacobian conjecture?, in: Proc. of the “Colloque de combinatoire enumerative,” ed. by G. Labelle and P. Leroux; *Lecture Notes in Math.* **1234**, Springer-Verlag, Berlin (1987).
- [Z2] D. Zeilberger, Using Rota’s umbral calculus to enumerate Stanley’s P - Partitions, *Adv. Applied Math.* **41** (2008), 206–217.