*Linear and Multilinear Algebra*, 1982, Vol. 12, pp. 1–20
0308–1087/82/1201–0001 $06.50/0


# Operational Formulas

STEVEN ROMAN
*Department of Mathematics, California State University, Fullerton, California 92634*

*(Received December 17, 1980; in final form August 5, 1981)*

## 1. INTRODUCTION

In this paper we present a technique for deriving operational formulas on the algebra of polynomials. A well-known example of such a formula is

$$(XD)^m = \sum_{k=0}^{m} S(m,k)X^kD^k \tag{1.1}$$

where $D$ is the derivative, $X$ is the operator multiplication by $x$ and $S(m,k)$ are the Stirling numbers of the second kind. Most formulas of this kind are obtained by catch-as-catch-can methods and proved by induction. We shall use the modern umbral calculus to implement a simple technique for deriving such formulas.

In section 2 we give a brief discussion of relevant facts from the umbral calculus. For a more detailed treatment we refer the reader to the references. Section 3 gives a general operational formula and the remaining sections are devoted to applications of this formula.

Let us describe this simple technique by using (1.1) as an example. Suppose we wish to determine the constants $a_{m,k}$ in

$$(XD)^m = \sum_{k=0}^{m} a_{m,k}X^kD^k.$$

We first notice that

$$X^kD^kx^n = (n)_kx^n \tag{1.2}$$

and

$$(XD)^mx^n = n^mx^n. \tag{1.3}$$

It is well known that

$$x^m = \sum_{k=0}^{m} S(m,k)(x)_k \tag{1.4}$$

and so

$$n^m = \sum_{k=0}^{m} S(m,k)(n)_k.$$

Combining these equations gives

$$\begin{aligned}(XD)^m x^n &= n^m x^n \\ &= \sum_{k=0}^{m} S(m,k)(n)_k x^n \\ &= \sum_{k=0}^{m} S(m,k) X^k D^k x^n\end{aligned}$$

from which (1.1) follows.

The idea is quite simple and the real problem here is first to determine which sequence of polynomials to use in (1.2) and (1.3) [in this case it is $x^n$]. Of course the sequence must form a basis for the polynomials. Then it is necessary to find the appropriate formula as in (1.4). As we shall see, both these problems can be handled in some generality by the techniques of the umbral calculus. Actually, applying the umbral calculus to operational formulas is not a new idea and Al-Salam and Ismail have made significant steps in this direction.

We have not aimed at the most general possibilities in our formulas. However, in their present form they are general enough to include many formulas from the literature and should adequately serve to describe the general technique.

The domain of the operators in this paper will be the algebra of polynomials in a single variable. However, Carlitz has pointed out that a polynomial differential operator $T$, that is, a sum of a finite number of monomials of the form

$$p_0(X)Dp_1(X)D\ldots p_{n-1}(X)Dp_n(X)$$

has the property that $T=0$ if and only if $T=0$ when restricted to polynomials. Thus some of the present formulas may be extended to a larger domain.

We would like to point out that the results here can be recast in terms of the algebra of two non-commuting variables $X$ and $Y$ for which

$$XY - YX = f(Y)$$

where $f(Y)$ is a formal power series in $Y$.

We shall use the notation

$$(x)_n = x(x-1)\dots(x-n+1)$$
$$(x)^{(n)} = x(x+1)\dots(x+n-1)$$
$$(x)_{a,n} = x(x-a)\dots(x-(n-1)a)$$
$$(x)_a^{(n)} = x(x+a)\dots(x+(n-1)a)$$
$$\delta_{n,k} = \begin{cases} 1 & \text{if } n = k \\ 0 & \text{if } n \neq k \end{cases}$$

## 2. THE UMBRAL CALCULUS

In this section we give a brief description of relevant topics from the umbral calculus. In the interest of conservation of space we give no proofs.

Let $P^*$ denote the vector space of all linear functionals on $P$. We use the notation $\langle L|p(x)\rangle$ for the action of a linear functional $L$ on a polynomial $p(x)$.

Let $f(A)$ be a formal power series in $A$ of the form

$$f(A) = \sum_{k=0}^{\infty} \frac{a_k}{k!} A^k \tag{2.1}$$

where $a_k$ are constants. Then $f(A)$ can be made into a linear functional by defining

$$\langle f(A)|x^n\rangle = a_n. \tag{2.2}$$

Notice that (2.2) implies that $A^k$ is the $k$th derivative evaluated at 0 and $A^0$ is evaluation at 0.

Conversely, any linear functional $L$ can be represented in the form (2.1). In fact, if

$$f_L(A) = \sum_{k=0}^{\infty} \frac{\langle L|x^k\rangle}{k!} A^k \tag{2.3}$$

Then as linear functionals $L$ and $f_L(A)$ are equal. From now on all linear functionals will be expressed directly as formal power series in $A$, using (2.2) and (2.3).

Notice that from (2.1) and (2.2) we have

$$f(A) = \sum_{k=0}^{\infty} \frac{\langle f(A)|x^k\rangle}{k!} A^k. \tag{2.4}$$

What we are really doing is abusing the vector space isomorphism

$$\psi : L \to \sum_{k=0}^{\infty} \frac{\langle L | x^k \rangle}{k!} A^k$$

from $P^*$ onto the *algebra* of formal power series in $A$. This isomorphism induces a natural product on linear functionals, namely, that of formal power series.

For example, if $a$ is a constant the *evaluation functional* is given by

$$\langle f(A) | p(x) \rangle = p(a).$$

By (2.4) we have

$$\begin{aligned} f(A) &= \sum_{k=0}^{\infty} \frac{a^k}{k!} A^k \\ &= e^{aA}. \end{aligned}$$

If $f(A)$ has the form (2.1) the *degree* of $f(A)$ is the smallest $k$ such that $a_k \neq 0$. If $f_k(A)$ is a sequence of linear functionals for which deg $f_k(A) \geqslant k$ then for any constants $a_k$ the series

$$h(A) = \sum_{k=0}^{\infty} a_k f_k(A)$$

is a well-defined formal power series in $A$, and hence also a linear functional.

Moreover,

$$\langle h(A) | p(x) \rangle = \sum_{k=0}^{\infty} a_k \langle f_k(A) | p(x) \rangle$$

for all polynomials $p(x)$.

By a *sequence* $s_n(x)$ in $P$ we shall always imply that deg $s_n(x) = n$.

THEOREM 2.1 *If deg $f(A) = 1$ and deg $g(A) = 0$ then the identities*

$$\langle g(A) f(A)^k | s_n(x) \rangle = n! \delta_{n,k}$$

*define a unique sequence $s_n(x)$ in $P$.*

We call $s_n(x)$ the *Sheffer sequence* for the pair $(g(A), f(A))$. They are known in the literature (up to multiplicative constants) as sequences of Sheffer $A$-type zero.

THEOREM 2.2 *Let $s_n(x)$ be Sheffer for $(g(A), f(A))$. Then for any $h(A)$ we have*

$$h(A) = \sum_{k=0}^{\infty} \frac{\langle h(A) | s_k(x) \rangle}{k!} g(A) f(A)^k.$$

Next we come to our generalization of Eq. (1.4).

THEOREM 2.3 *Let $s_n(x)$ be Sheffer for $(g(A), f(A))$. Then for any polynomial $p(x)$ we have*

$$p(x) = \sum_{k \geqslant 0} \frac{\langle g(A) f(A)^k | p(x) \rangle}{k!} s_k(x). \tag{2.5}$$

Closely associated with the umbral algebra is the algebra $S$ of formal power series in the derivative operator $D$. If

$$f(D) = \sum_{k=0}^{\infty} \frac{a_k}{k!} D^k \tag{2.6}$$

we have

$$f(D) x^n = \sum_{k=0}^{n} \binom{n}{k} a_k x^{n-k}.$$

The map $\phi$ from $P^*$ onto $S$ defined by

$$\phi f(A) = f(D)$$

is an algebra isomorphism from $P^*$ onto $S$.

THEOREM 2.4 *For all linear functionals $f(A)$ and $g(A)$ we have*

$$\langle f(A) g(A) | x^n \rangle = \langle f(A) | g(D) x^n \rangle.$$

In particular we have

$$\langle f(A) | x^n \rangle = \langle A^0 | f(D) x^n \rangle$$

where as we have remarked $A^0$ is evaluation at 0.

THEOREM 2.5 *The sequence $s_n(x)$ is Sheffer for $(g(A), f(A))$ for some invertible $g(A)$ if and only if*

$$f(D) s_n(x) = n s_{n-1}(x). \tag{2.7}$$

The operator corresponding to evaluation $e^{aA}$ is the translation operator

$$e^{aD} p(x) = p(x+a).$$

The operator corresponding to $e^{aA} - 1$ is the difference operator

$$\Delta_a = e^{aD} - 1$$

where

$$\Delta_a p(x) = p(x+a) - p(x).$$

Thus by (2.7),

$$\Delta_a (x)_{a,n} = n (x)_{a,n-1}.$$

The operator corresponding to $1-e^{-aA}$ is $\nabla_a = 1-e^{-aD}$. We shall write $\Delta_1 = \Delta$ and $\nabla_1 = \nabla$.

Let $s_n(x)$ be a sequence of polynomials. The linear operator $\theta$ defined by

$$\theta s_n(x) = s_{n+1}(x)$$

is called the *shift* for $s_n(x)$.

If $f(A)$ is given by (2.1) then by $f'(A)$ we mean the formal derivative of $f(A)$ with respect to $A$,

$$f'(A) = \sum_{k=1}^{\infty} \frac{a_k}{(k-1)!} A^{k-1}.$$

We define $f'(D)$ analogously.

THEOREM 2.6 *If $s_n(x)$ is Sheffer for $(g(A), f(A))$ and $\theta$ is its shift then*

$$\theta = \left(X - \frac{g'(D)}{g(D)}\right) \frac{1}{f'(D)}$$

*where $X$ is the operator multiplication by $x$*

We remark that any $h(D)$ for which $\deg h(D) = 0$ can be written in the form $g'(D)/g(D)$ and any $l(D)$ for which $\deg l(D) = 0$ can be written in the form $1/f'(D)$. Hence any operator of the form $(X-h(D))l(D)$ is a shift for some Sheffer sequence.

We define $\theta^{-1}$ to be the linear operator

$$\theta^{-1} s_n(x) = s_{n-1}(x)$$

where as always $s_j(x) = 0$ if $j < 0$. Care must be taken here since $\theta \circ \theta^{-1}$ is not the identity $[\theta \circ \theta^{-1} s_0(x) = 0]$. We write $(\theta^{-1})^k = \theta^{-k}$.

THEOREM 2.7 *Let $\theta$ be the shift for $s_n(x)$. Then for integers $i$ and $j$,*

$$\theta^i s_j(x) = \begin{cases} s_{i+j}(x) & \text{if } j \geqslant 0 \text{ or if } i+j < 0 \\ 0 & \text{if } j < 0 \text{ and } i+j \geqslant 0 \end{cases}$$

THEOREM 2.8 *Let $\theta$ be a shift. Then $\theta^i \theta^j = \theta^{i+j}$ if and only if $j \geqslant 0$ or $i \leqslant 0$.*

We finish this section with an often used result.

THEOREM 2.9 *If $s_n(x)$ is Sheffer for $(g(A), f(A))$ then the sequence*

$$r_n(x) = s_n(x+\gamma_0+\gamma_1 n)$$

*is Sheffer for*

$$\left(\frac{[e^{-\gamma_1 A} f(A)]'}{f'(A)} e^{(\gamma_1-\gamma_0)A} g(A), e^{-\gamma_1 A} f(A)\right).$$

## 3. THE GENERAL FORMULA

Let $s_n(x)$ be a sequence of polynomials. Hence $\deg s_n(x) = n$. Suppose $\theta$ is its shift. Thus

$$\theta s_n(x) = s_{n+1}(x)$$
$$\theta^{-1} s_n(x) = s_{n-1}(x).$$

Let $Q_k$ be a sequence of linear operators on $P$ for which

$$Q_k s_n(x) = p_k(n) s_{n+\mu}(x) \tag{3.1}$$

where $\mu$ is an integer independent of $k$ and $n$ and $p_k(x)$ is a polynomial.

Let $T$ be a linear operator on $P$ for which

$$T s_n(x) = t(n) s_{n+\lambda}(x)$$

where $\lambda$ is an integer independent of $n$ and $t(x)$ is a polynomial. Suppose that

$$t(x) = \sum_{k \geqslant 0} d_k p_k(x) \tag{3.2}$$

for some constants $d_k$. Theorem 2.7 implies that if $\mu \geqslant \min\{0, \lambda\}$ then

$$\begin{aligned} T s_n(x) &= t(n) s_{n+\lambda}(x) \\ &= \sum_{k \geqslant 0} d_k p_k(n) s_{n+\lambda}(x) \\ &= \theta^{\lambda-\mu} \sum_{k \geqslant 0} d_k p_k(n) s_{n+\mu}(x) \\ &= \theta^{\lambda-\mu} \sum_{k \geqslant 0} d_k Q_k s_n(x). \end{aligned}$$

Thus

$$T = \theta^{\lambda-\mu} \sum_{k \geqslant 0} d_k Q_k \tag{3.3}$$

whenever $\mu \geqslant \min\{0, \lambda\}$. Equation (3.3) is our operational formula.

Let us turn to a discussion of (3.2). Several cases will present themselves. The simplest is when $p_k(x)$ is a Sheffer sequence, say for $(h(A), l(A))$. Then by Theorem 2.3

$$d_k = \frac{1}{k!} \langle h(A) l(A)^k | t(x) \rangle \tag{3.4}$$

If $p_k(x) = r_{a+k}(x)$ where $r_k(x)$ is Sheffer for $(h(A), l(A))$ then

$$t(x) = \sum_{k \geqslant 0} \frac{1}{k!} \langle h(A) l(A)^k | t(x) \rangle r_k(x)$$

$$= \sum_{k \geqslant -a} \frac{1}{(a+k)!} \langle h(A) l(A)^{a+k} | t(x) \rangle p_k(x)$$

and in order for (3.2) to hold we must have, for $-a \leqslant k < 0$.

$$\langle h(A) l(A)^{a+k} | t(x) \rangle = 0 \tag{3.5}$$

and then for $k \geqslant 0$

$$d_k = (1/(a+k)!) \langle h(A) l(A)^{a+k} | t(x) \rangle. \tag{3.6}$$

Another case which will arise is when $p_k(x) = p(x) r_k(x)$ where $r_k(x)$ is Sheffer for $(h(A), l(A))$ and $p(x)$ is a polynomial independent of $k$. Then if $p(x)$ is a factor of $t(x)$ we have

$$d_k = \frac{1}{k!} \left\langle h(A) l(A)^k \middle| \frac{t(x)}{p(x)} \right\rangle. \tag{3.7}$$

In the subsequent sections it will be assumed that $s_n(x)$ is the Sheffer sequence for $(g(A), f(A))$ and that $\theta$ is its shift. We recall that Theorem 2.6 gives

$$\theta = \left( X - \frac{g'(D)}{g(D)} \right) \frac{1}{f'(D)}.$$

Since $\theta s_n(x) = s_{n+1}(x)$ we see that

$$\theta f(D) s_n(x) = n s_n(x)$$

and so for any polynomial $p(x)$

$$p(\theta f(D)) s_n(x) = p(n) s_n(x).$$

Thus the sequence of operators $Q_k$ satisfying (3.1), where $s_n(x)$ is Sheffer, is

$$Q_k = \theta^\mu p_k(\theta f(D)).$$

As we shall see in the applications, however, the $Q_k$ may take many other forms.

Let us list for future reference some cases which will give the most familiar formulas.

1. $s_n(x) = g^{-1}(D) x^n$ is Sheffer for $(g(A), A)$.

These are known in the literature as Appell polynomials. In this case

$$\theta = (X - g'(D)/g(D)).$$

A popular choice is $g(D) = e^{rD}$. Then $\theta = X - r$.

2. $s_n(x) = (x)_{r,n}$ is Sheffer for $(1, e^{rA}-1)$.

Then

$$\theta = rXe^{-rD}.$$

Notice that $f(D) = \Delta_r$ and $\theta f(D) = rX\nabla_r$.

3. $s_n(x) = (x)_r^{(n)}$ is Sheffer for $(1, 1-e^{-rA})$.

Then

$$\theta = rXe^{rD}.$$

Notice that $f(D) = \nabla_r$ and $\theta f(D) = rX\Delta_r$.

4. $s_n(x) = x(x-rn)^{n-1}$ is Sheffer for $(1, Ae^{rA})$.

Then

$$\theta = X\frac{e^{-rD}}{1+rD}.$$

Here we have

$$\theta f(D) = X\frac{D}{1+rD}$$

5. $s_n(x)$ are the Hermite polynomials which are Sheffer for $(e^{-vA^2/2}, A)$. Then

$$\theta = X + vD$$

and

$$\theta f(D) = (X+vD)D.$$

## 4. APPLICATION 1

In this section we consider the sequence of operators

$$Q_k = \theta^{(1-\alpha)k} f(D)^{a+k}\theta^{\beta+\alpha k}$$

where $a$, $\alpha$ and $\beta$ are integers and $a \geqslant 0$, $\alpha \geqslant 0$.

THEOREM 4.1

$$Q_k s_n(x) = (n+\beta+\alpha k)_{a+k} s_{n+\beta-a}(x) \tag{4.1}$$

*Proof* Using Eq. (2.7) and Theorem 2.7 we have

$$\begin{aligned} Q_k s_n(x) &= \theta^{(1-\alpha)k} f(D)^{a+k} \theta^{\beta+\alpha k} s_n(x) \\ &= \theta^{(1-\alpha)k} f(D)^{a+k} s_{n+\beta+\alpha k}(x) \\ &= (n+\beta+\alpha k)_{a+k} \theta^{(1-\alpha)k} s_{n+\beta+\alpha k-a-k}(x) \\ &= \begin{cases} (n+\beta+\alpha k)_{a+k} s_{n+\beta-a}(x) \\ \quad 0 \qquad \text{when} \quad n+\beta+\alpha k-a-k<0 \\ \qquad\qquad \text{and} \quad n+\beta-a \geqslant 0. \end{cases} \end{aligned}$$

But if $n+\beta-a \geqslant 0$ then $n+\beta+\alpha k \geqslant n+\beta \geqslant n+\beta-a \geqslant 0$ and so $n+\beta+\alpha k -a-k<0$ implies that $(n+\beta+\alpha k)_{a+k}=0$. Thus (4.1) always hold.

By Theorem 4.1 the operators $Q_k$ have the form (3.1) for $\mu=\beta-a$ and $p_k(x) = (x+\beta+\alpha k)_{a+k}$. Thus (3.3) becomes

$$T = \theta^{\lambda-\beta+a} \sum_{k \geqslant 0} d_k \theta^{(1-\alpha)k} f(D)^{a+k} \theta^{\beta+\alpha k}$$

where $a \geqslant 0, \alpha \geqslant 0$ and $\beta-a \geqslant \min\{0,\lambda\}$. In view of the conditions on $a$, $\alpha$ and $\beta$ we may conclude that

$$T = \sum_{k \geqslant 0} d_k \theta^{\lambda-\beta+a+(1-\alpha)k} f(D)^{a+k} \theta^{\beta+\alpha k} \tag{4.2}$$

where $a \geqslant 0$, $\alpha \geqslant 0$ and $\beta-a \geqslant \min\{0,\lambda\}$. To see this if $\lambda \leqslant 0$ then $\lambda-\beta +a \leqslant 0$ and Theorem 2.8 gives (4.2). If $\lambda>0$ and $(1-\alpha)k<0$ then $\beta-a \geqslant 0$ and since both $\beta+\alpha k-a-k \geqslant 0$ and $\beta+\alpha k-a-k-(1-\alpha)k=\beta-a \geqslant 0$ we deduce (4.2) directly as in the proof of Theorem 4.1.

Now we turn to a computation of the constants $d_k$. The form of $p_k(x) = (x+\beta+\alpha k)_{a+k}$ puts us under the case covered by (3.5) and (3.6). From Theorem 2.9 we see that $r_k(x)=(x+\beta-a\alpha+\alpha k)_k$ is Sheffer for

$$((1-\alpha+\alpha e^{-A})e^{(a\alpha-\beta)A}, e^{-\alpha A}(e^A-1)).$$

Therefore we must have

$$\langle (1-\alpha+\alpha e^{-A}) e^{(-\beta-\alpha k)A} (e^A-1)^{a+k} | t(x) \rangle = 0 \tag{4.3}$$

for $-a \leqslant k<0$ and for $k \geqslant 0$

$$d_k = \frac{1}{(a+k)!} \langle (1-\alpha+\alpha e^{-A}) e^{(-\beta-\alpha k)A} (e^A-1)^{a+k} | t(x) \rangle. \tag{4.4}$$

From Theorem 2.4 we may write $d_k$ in the form

$$d_k = \frac{1}{(a+k)!} \langle A^0 | \Delta^{a+k} [(1-\alpha) t(x-\beta-\alpha k) + \alpha t(x-\beta-\alpha k-1)] \rangle. \tag{4.5}$$

Let us give some examples of (4.2).

1. Let $a = 0$ and

$$T = [\theta f(D)]^m$$

for $m > 0$. Then $Ts_n(x) = n^m s_n(x)$ and so

$$\lambda = 0$$
$$t(x) = x^m.$$

Equation (4.3) is automatically satisfied when $a = 0$ and (4.5) gives

$$\begin{aligned} d_k &= \frac{1}{k!}\langle A^0|\Delta^k[(1-\alpha)(x-\beta-\alpha k)^m+\alpha(x-\beta-\alpha k-1)^m]\rangle \\ &= \sum_{j=k}^{m}\binom{j}{k}S(m,j)(-\beta-\alpha j)(-\beta-\alpha k-1)_{j-k-1} \end{aligned}$$

and so (4.2) is

$$[\theta f(D)]^m =$$

$$\sum_{k=0}^{m}\left[\sum_{j=k}^{m}\binom{j}{k}S(m,j)(-\beta-\alpha j)(-\beta-\alpha k-1)_{j-k-1}\right]\theta^{-\beta+(1-\alpha)k}f(D)^k\theta^{\beta+\alpha k}$$

where $\alpha \geqslant 0$ and $\beta \geqslant 0$.

In case $\alpha = 0$ we obtain

$$[\theta f(D)]^m = \sum_{k=0}^{m}\left[\sum_{j=k}^{m}\binom{j}{k}S(m,j)(-\beta)_{j-k}\right]\theta^{-\beta+k}f(D)^k\theta^{\beta}$$

where $\beta \geqslant 0$. If in addition $\beta = 0$ we have

$$[\theta f(D)]^m = \sum_{k=0}^{m} S(m,k)\theta^k f(D)^k.$$

Using the examples at the end of section 3 we obtain

$$[(X-r)D]^m = \sum_{k=0}^{m} S(m,k)(X-r)^k D^k$$

$$(X\nabla_r)^m = \sum_{k=0}^{m} r^{k-m}S(m,k)(Xe^{-rD})^k\Delta_r^k$$

$$(X\Delta_r)^m = \sum_{k=0}^{m} r^{k-m}S(m,k)(Xe^{rD})^k\nabla_r^k$$

$$[XD(1+rD)^{-1}]^m = \sum_{k=0}^{m} S(m,k)[Xe^{-rD}(1+rD)^{-1}]^k D^k e^{rkD}$$

$$[(X+vD)D]^m = \sum_{k=0}^{m} S(m,k)(X+vD)^k D^k.$$

In the interest of space conservation, we shall not repeat this list of examples for each future case.

2. Let $a = 0$ and

$$T = \theta^m f(D)^m$$

for $m > 0$. Then $Ts_n(x) = (n)_m s_n(x)$ and so

$$\lambda = 0$$

$$t(x) = (x)_m.$$

Equation (4.5) gives

$$\begin{aligned} d_k &= \frac{1}{k!} \langle A^0 | \Delta^k [(1-\alpha)(x-\beta-\alpha k)_m + \alpha(x-\beta-\alpha k-1)_m] \rangle \\ &= \binom{m}{k} (-\beta-\alpha m)(-\beta-\alpha k-1)_{m-k-1} \end{aligned}$$

and (4.2) is

$$\theta^m f(D)^m = \sum_{k=0}^{m} \binom{m}{k} (-\beta-\alpha m)(-\beta-\alpha k-1)_{m-k-1} \theta^{-\beta+(1-\alpha)k} f(D)^k \theta^{\beta+\alpha k}$$

where $\alpha \geqslant 0$ and $\beta \geqslant 0$.

When $\alpha = 0$ we get

$$\theta^m f(D)^m = \sum_{k=0}^{m} \binom{m}{k} (-\beta)_{m-k} \theta^{-\beta+k} f(D)^k \theta^{\beta}$$

where $\beta \geqslant 0$.

3. Let

$$T = [f(D)\theta f(D)]^m$$

for $m > 0$. Then $Ts_n(x) = (n)_m^2 s_{n-m}(x)$ and so

$$\lambda = -m$$

$$t(x) = (x)_m^2.$$

Now suppose $a = 0$. Then (4.3) is satisfied and (4.5) yields

$$d_k = (1/k!) \langle A^0 | \Delta^k [(1-\alpha)(x-\beta-\alpha k)_m^2 + \alpha(x-\beta-\alpha k-1)_m^2] \rangle.$$

Expanding $\Delta^k = (e^D - 1)^k$ will give

$$d_k = \sum_{j=0}^{k} \binom{k}{j} \frac{(-1)^{k-j}}{k!} [(1-\alpha)(j-\beta-\alpha k)_m^2 + \alpha(j-\beta-\alpha k-1)_m^2].$$

Thus (4.2) becomes

$$[f(D)\theta f(D)]^m = \sum_{k=0}^{2m}\left[\sum_{j=0}^{k}\binom{k}{j}\frac{(-1)^{k-j}}{k!}[(1-\alpha)(j-\beta-\alpha k)_m^2+\alpha(j-\beta-\alpha k-1)_m^2]\right] \cdot \theta^{-m-\beta+(1-\alpha)k}f(D)^k\theta^{\beta+\alpha k}$$

where $\alpha \geqslant 0$ and $\beta \geqslant -m$. The upper limit is $2m$ here since $d_k = 0$ for $k > 2m$.

In case $\alpha = 0$ we obtain

$$[f(D)\theta f(D)]^m = \sum_{k=0}^{2m}\left[\sum_{j=0}^{k}\binom{k}{j}\frac{(-1)^{k-j}}{k!}(j-\beta)_m^2\right]\theta^{-m-\beta+k}f(D)^k\theta^\beta$$

where $\beta \geqslant -m$. If $\beta = 0$ we have

$$[f(D)\theta f(D)]^m = \sum_{k=0}^{2m}\left[\sum_{j=0}^{k}\binom{k}{j}\frac{(-1)^{k-j}}{k!}(j)_m^2\right]\theta^{-m+k}f(D)^k.$$

Now suppose $\alpha = 0$, $a \geqslant 0$. Rather than turn to (4.3) and (4.5) we can refer back to (3.2) which is

$$(x)_m^2 = \sum_{k\geqslant 0} d_k(x+\beta)_{a+k}.$$

Now $(x+\beta)_{a+k} = (x+\beta)_a(x+\beta-a)_k$ and if $a-m \leqslant \beta \leqslant 0$ then $(x+\beta)_a$ is a factor of $(x)_m^2$. Thus we have

$$\frac{(x)_m^2}{(x+\beta)_a} = \sum_{k\geqslant 0} d_k(x+\beta-a)_k.$$

Translating by $-\beta+a$ gives

$$\frac{(x-\beta+a)_m^2}{(x+a)_a} = \sum_{k\geqslant 0} d_k(x)_k$$

and from Theorem 2.3 we obtain

$$d_k = \frac{1}{k!}\left\langle A^0|\Delta^k\frac{(x-\beta+a)_m^2}{(x+a)_a}\right\rangle.$$

Thus (4.2) becomes

$$[f(D)\theta f(D)]^m = \sum_{k=0}^{2m-a}\frac{1}{k!}\left\langle A^0|\Delta^k\frac{(x-\beta+a)_m^2}{(x+a)_a}\right\rangle\theta^{-m-\beta+a+k}f(D)^{a+k}\theta^\beta$$

where $a \geqslant 0$ and $a-m \leqslant \beta \leqslant 0$.

If $\beta = 0$ this becomes

$$[f(D)\theta f(D)]^m = \sum_{k=0}^{2m-a} \frac{1}{k!} \langle A^0 | \Delta^k (x+a)_m (x)_{m-a} \rangle \theta^{-m+a+k} f(D)^{a+k}$$

where $0 \leqslant a \leqslant m$. When $a = m$ we obtain

$$[f(D)\theta f(D)]^m = \sum_{k=0}^{m} \binom{m}{k} \frac{m!}{k!} \theta^k f(D)^{m+k}.$$

The examples in section 3 give, for example

$$[(X-r)D]^m = \sum_{k=0}^{m} \binom{m}{k} \frac{m!}{k!} (X-r)^k D^{m+k}$$

and

$$(\Delta X \nabla)^m = \sum_{k=0}^{m} \binom{m}{k} \frac{m!}{k!} (Xe^{-D})^k \Delta^{m+k}.$$

4. Let $a = 0$ and

$$T = [\theta^i (r + \theta f(D))]^m$$

where $m > 0$, $i$ is any integer and $r$ is a constant. Then $Ts_n(x) = (n+r)_i^{(m)} s_{n+im}(x)$ and so

$$\lambda = im$$

$$t(x) = (x+r)_i^{(m)}.$$

From (4.5) we get

$$d_k = (1/k!) \langle A^0 | \Delta^k [(1-\alpha)(x+r-\beta-\alpha k)_i^{(m)} + \alpha(x+r-\beta-\alpha k-1)_i^{(m)}] \rangle$$

and so (4.2) is

$$[\theta^i (r+\theta f(D))]^m =$$

$$\sum_{k=0}^{m} \frac{1}{k!} \langle A^0 | \Delta^k [(1-\alpha)(x+r-\beta-\alpha k)_i^{(m)} + \alpha(x+r-\beta-\alpha k-1)_i^{(m)}] \rangle \cdot \theta^{im-\beta+(1-\alpha)k} f(D)^k \theta^{\beta+\alpha}$$

where $\alpha \geqslant 0$ and $\beta \geqslant \min\{0, im\}$.

In case $\alpha = 0$ we have

$$[\theta^i (r+\theta f(D))]^m = \sum_{k=0}^{m} \frac{1}{k!} \langle A^0 | \Delta^k (x+r-\beta)_i^{(m)} \rangle \, \theta^{im-\beta+k} f(D)^k \theta^\beta$$

where $\beta \geqslant \min\{0, im\}$. If $i = 1$, since

$$\Delta^k(x+r-\beta)^{(m)} = (m)_k(x+r-\beta+k)^{(m-k)}$$

we have

$$[\theta(r+\theta f(D))]^m = \sum_{k=0}^{m} \binom{m}{k} (r-\beta+k)^{(m-k)}\theta^{m-\beta+k}f(D)^k\theta^\beta$$

for $\beta \geqslant 0$.

5. Let $a = 0$ and

$$\begin{aligned} T &= (r+\theta f(D))_{i,m} \\ &= (r+\theta f(D))(r+i+\theta f(D))\dots(r+(m-1)i+\theta f(D)) \end{aligned}$$

where $m > 0$ and $r$ and $i$ are constants. Then since $(j+\theta f(D))s_n(x) = (j+n)s_n(x)$ we see that $Ts_n(x) = (r+n)_{i,m}s_n(x)$. So

$$\lambda = 0$$

$$t(x) = (r+x)_{i,m}.$$

Let us go directly to the case $\alpha = 0$. Then

$$d_k = \frac{1}{k!}\langle A^0 | \Delta^k(x+r-\beta)_{i,m}\rangle$$

and (4.2) is

$$(r+\theta f(D))_{i,m} = \sum_{k=0}^{m} \frac{1}{k!}\langle A^0|\Delta^k(x+r-\beta)_{i,m}\rangle\, \theta^{-\beta+k}f(D)^k\theta^\beta$$

where $\beta \geqslant 0$.

In case $i = 1$ we have

$$(r+\theta f(D))_m = \sum_{k=0}^{m} \binom{m}{k} (r-\beta)_{m-k}\theta^{-\beta+k}f(D)^k\theta^\beta$$

and if $\beta = r$ we have

$$(r+\theta f(D))_m = \theta^{-r+m}f(D)^m\theta^r.$$

Our examples give

$$(r+XD)_m = X^{-r+m}D^mX^r$$

$$(r+X\nabla)_m = (Xe^{-D})^{-r+m}\Delta^m(Xe^{-D})^r$$

and so on.

## 5. APPLICATION 2

In this section we consider the sequence of operators

$$Q_k = f(D)^a r_k(S) f(D)^b \theta^\beta$$

where $a \geqslant 0$, $b \geqslant 0$, $S = \theta f(D)$ and $r_k(x)$ is Sheffer for $(h(A), l(A))$.

THEOREM 5.1

$$Q_k s_n(x) = (n+\beta)_{a+b} r_k(n+\beta-b) s_{n+\beta-a-b}(x).$$

*Proof* Using Theorem 2.7 we have

$$\begin{aligned} Q_k s_n(x) &= f(D)^a r_k(S) f(D)^b \theta^\beta s_n(x) \\ &= f(D)^a r_k(S) f(D)^b s_{n+\beta}(x) \\ &= (n+\beta)_b f(D)^a r_k(S) s_{n+\beta-b}(x) \\ &= (n+\beta)_b r_k(n+\beta-b) f(D)^a s_{n+\beta-b}(x) \\ &= (n+\beta)_b (n+\beta-b)_a r_k(n+\beta-b) s_{n+\beta-a-b}(x) \\ &= (n+\beta)_{a+b} r_k(n+\beta-b) s_{n+\beta-a-b}(x). \end{aligned}$$

By Theorem 5.1 the operators $Q_k$ have the form (3.1) with $\mu = \beta-a-b$ and $p_k(x) = (x+\beta)_{a+b} r_k(x+\beta-b)$. Thus (3.3) becomes

$$T = \sum_{k \geqslant 0} d_k \theta^{\lambda-\beta+a+b} f(D)^a r_k(S) f(D)^b \theta^\beta \tag{5.1}$$

where $a \geqslant 0$, $b \geqslant 0$ and $\beta-a-b \geqslant \min\{0, \lambda\}$.

By virtue of the discussion at the end of section 3 we seek operators $T$ for which $Ts_n(x) = t(n) s_{n+\lambda}(x)$ where $(x+\beta)_{a+b}$ is a factor of $t(x)$. In this case (3.7) holds, and if $r_k(x)$ is Sheffer for $(h(A), l(A))$, by Theorem 2.9, $r_k(x+\beta-b)$ is Sheffer for $(e^{-(\beta-b)A} h(A), l(A))$. Thus (3.7) gives

$$d_k = \frac{1}{k!} \left\langle h(A) l(A)^k \,\middle|\, \frac{t(x-\beta+b)}{(x+b)_{a+b}} \right\rangle.$$

A class of operators with this property is $T = R^m$ where

$$R = \theta^\gamma f(D)^c u(S) f(D)^d \theta^\beta$$

where $u(x)$ is a polynomial. However, we shall not need $T$ in this much generality, so we will not prove this fact.

Let us give some examples of (5.1).

1. Let

$$T = f(D)^c u(\theta f(D)) f(D)^d \theta^\sigma$$

where $u(x)$ is a polynomial and $c \geqslant 0$, $d \geqslant 0$, $0 \leqslant \sigma-\beta \leqslant c+d-a-b$. Then

$$Ts_n(x) = (n+\sigma)_{c+d}u(n+\sigma-d)s_{n+\sigma-c-d}(x)$$

and so

$$\lambda = \sigma-c-d$$
$$t(x) = (x+\sigma)_{c+d}u(x+\sigma-d).$$

Noticing that

$$\frac{(x+\sigma-\beta+b)_{c+d}}{(x+b)_{a+b}} = (x+\sigma-\beta+b)_{\sigma-\beta}(x-a)_{\beta-\sigma+c+d-a-b}$$

we have

$$d_k = \frac{1}{k!}\langle h(A)l(A)^k|(x+\sigma-\beta+b)_{\sigma-\beta}(x-a)_{\beta-\sigma+c+d-a-b}u(x+\sigma-d-\beta+b)\rangle.$$

Thus (5.1) becomes

$$f(D)^c u(\theta f(D)) f(D)^d \theta^\sigma =$$

$$\sum_{k\geqslant 0}\frac{1}{k!}\langle h(A)l(A)^k|(x+\sigma-\beta+b)_{\sigma-\beta}(x-a)_{\beta-\sigma+c+d-a-b}u(x+\sigma-d-\beta+b)\rangle$$
$$\cdot\,\theta^{\sigma-\beta-c-d+a+b}f(D)^a r_k(\theta f(D)) f(D)^b\theta^\beta$$

where $a,b,c,d \geqslant 0$, and $0 \leqslant \sigma-\beta \leqslant c+d-a-b$.

In case $a+b = c+d$, we must have $\sigma = \beta$ and so

$$f(D)^c u(\theta f(D)) f(D)^d = \sum_{k\geqslant 0}\frac{1}{k!}\langle h(A)l(A)^k|u(x-d+b)\rangle\, f(D)^a r_k(\theta f(D)) f(D)^b$$

where $a+b = c+d$.

If instead we take $a = b = 0$ then we obtain

$$f(D)^c u(\theta f(D)) f(D)^d \theta^\sigma =$$

$$\sum_{k\geqslant 0}\frac{1}{k!}\langle h(A)l(A)^k|(x+\sigma-\beta)_{c+d}u(x+\sigma-\beta-d)\rangle\,\theta^{\sigma-\beta-c-d}r_k(\theta f(D))\theta^\beta$$

where $c,d \geqslant 0$, and $0 \leqslant \sigma-\beta \leqslant c+d$. For $u(x) = 1$, $r_k(x) = (x-a)^k$ and $d = 0$, since

$$\langle A^k|(x+a+\sigma-\beta)_c\rangle = \sum_{j=k}^{c} s(c,j)(j)_k(a+\sigma-\beta)^{j-k}$$

where $s(c, j)$ is a Stirling number of the first kind, we have

$$f(D)^c\theta^\sigma = \sum_{k=0}^{c}\left[\sum_{j=k}^{c}\binom{j}{k}s(c, j)(a+\sigma-\beta)^{j-k}\right]\theta^{\sigma-\beta-c}(\theta f(D)-a)^k\theta^\beta$$

where $c \geqslant 0$, and $0 \leqslant \sigma-\beta \leqslant c$.

Next we take $r_k(x) = (x)_k$ and $u(x) = 1$. Then we obtain as above

$$f(D)^c\theta^\sigma = \sum_{k=0}^{c}\binom{c}{k}(-\beta+\sigma)_{c-k}\theta^{\sigma-c-\beta}(\theta f(D))_k\theta^\beta$$

where $c \geqslant 0$ and $0 \leqslant \sigma-\beta \leqslant c$. For $\sigma = 0$ we obtain

$$f(D)^c = \sum_{k=0}^{c}\binom{c}{k}(-\beta)_{c-k}\theta^{-c-\beta}(\theta f(D))_k\theta^\beta.$$

2. Let $a = b = 0$ and

$$T = [f(D)^c u(\theta f(D))]^m$$

where $c \geqslant 0$ and $m > 0$. Then

$$Ts_n(x) = (n)_{mc}u(n)u(n-c)\ldots u(n-(m-1)c)s_{n-mc}(x)$$

and so

$$\lambda = -mc$$

$$t(x) = (x)_{mc}u(x)u(x-c)\ldots u(x-(m-1)c).$$

Equation (5.1) is

$$[f(D)^c u(\theta f(D))]^m =$$

$$\sum_{k\geqslant 0}\frac{1}{k!}\langle h(A)l(A)^k|(x-\beta)_{mc}u(x-\beta)\ldots u(x-\beta-(m-1)c)\rangle\theta^{-mc-\beta}\cdot r_k(\theta f(D))\theta^\beta$$

where $c \geqslant 0$ and $\beta \geqslant -mc$.

Let us take $c = 1$, $u(x) = x^i$ and $r_k(x) = x^k$. Then we obtain

$$[f(D)(\theta f(D))^i]^m = \sum_{k\geqslant 0}\frac{1}{k!}\langle A^k|(x-\beta)_m^{i+1}\rangle\,\theta^{-m-\beta}(\theta f(D))^k\theta^\beta$$

where $\beta \geqslant -m$. For $m = 1$ we get

$$f(D)(\theta f(D))^i = \sum_{k\geqslant 0}\binom{i+1}{k}(-\beta)^{i+1-k}\theta^{-1-\beta}(\theta f(D))^k\theta^\beta$$

where $\beta \geqslant -1$.

## 6. APPLICATION 3

In this section we consider the sequence of operators

$$Q_k = \theta^{-\alpha k}[\theta^\alpha(a+S)]^{b+k}$$

where $\alpha > 0$, $b \geqslant 0$ and $S = \theta f(D)$.

THEOREM 6.1

$$Q_k s_n(x) = (a+n)_\alpha^{(b+k)} s_{n+\alpha b}(x).$$

*Proof* First we have

$$\theta^\alpha(a+S)s_n(x) = (a+n)s_{n+\alpha}(x).$$

From this it follows that

$$[\theta^\alpha(a+S)]^m s_n(x) = (a+n)(a+n+\alpha)\dots(a+n+(m-1)\alpha)s_{n+m\alpha}(x).$$

Putting $m = b+k$ and applying $\theta^{-\alpha k}$ gives the result.

Thus the operators $Q_k$ are of the form (3.1) with $\mu = \alpha b$ and $p_k(x) = (x+a)_\alpha^{(b+k)}$. Thus (3.3) becomes

$$T = \theta^{\lambda-\alpha b} \sum_{k \geqslant 0} d_k \theta^{-\alpha k}[\theta^\alpha(a+S)]^{b+k}$$

where $\alpha, b \geqslant 0$. Since $\alpha b \geqslant 0$ we may conclude that

$$T = \sum_{k \geqslant 0} d_k \theta^{\lambda-\alpha b-\alpha k}[\theta^\alpha(a+\theta f(D))]^{b+k} \tag{6.1}$$

where $\alpha > 0$, $b \geqslant 0$.

Now $p_k(x) = r_{b+k}(x)$ where $r_k(x) = (x+a)_\alpha^{(k)}$ is Sheffer for $(e^{-aA}, 1-e^{-\alpha A})$. Thus (3.5) and (3.6) become

$$\langle e^{-aA}(1-e^{-\alpha A})^{b+k} | t(x)\rangle = 0 \tag{6.2}$$

for $-b \leqslant k < 0$ and

$$\begin{aligned} d_k &= \frac{1}{(b+k)!}\langle e^{-aA}(1-e^{-\alpha A})^{b+k} | t(x)\rangle \\ &= \frac{1}{(b+k)!}\langle A^0 | \nabla_\alpha^{b+k} t(x-a)\rangle \end{aligned} \tag{6.3}$$

for $k \geqslant 0$.

We shall content ourselves with a single example of (6.1). Let $\alpha = 1$, $b = 0$ and

$$T = (f(D)\theta)^m.$$

Then $Ts_n(x) = (n+1)^m s_n(x)$ and

$$\lambda = 0$$

$$t(x) = (x+1)^m.$$

Since $b = 0$, (6.2) is satisfied and (6.3) is

$$d_k = \frac{1}{k!}\langle A^0 | \nabla^k (x+1-a)^m \rangle$$

$$= \sum_{j=k}^{m} \binom{j}{k} (-1)^{m+j} S(m, j)(1-a)^{j-k}.$$

Thus (6.1) is

$$(f(D)\theta)^m = \sum_{k=0}^{m} \left[ \sum_{j=k}^{m} \binom{j}{k} (-1)^{m+j} S(m, j)(1-a)^{j-k} \right] \theta^{-k} [\theta(a+\theta f(D))]^k.$$

If $a = 1$ we obtain

$$(f(D)\theta)^m = \sum_{k=0}^{m} (-1)^{m+k} S(m, k) \theta^{-k} [\theta(1+\theta f(D))]^k.$$

## References

There have been many articles on the subject of operational formulas and we shall not attempt to give a list. Rather we mention only those two papers ([1], [3]) which inspired the present work. Reference [5] is always an excellent source for combinatorial identities, including operational formulas. The remaining references are for those interested in a further study of the umbral calculus and [11] and [12] have extensive bibliographies of their own.


[1] W. Al-Salam and M. E. H. Ismail, Some operational formulas, *J. Math. Anal. Appl.* **51** (1975), 208–218.

[2] J. W. Brown and S. Roman, Inverse relations for certain Sheffer sequences, *S.I.A.M. Journal on Mathematical Analysis*, **12** (1981), 186–195.

[3] L. Carlitz, A theorem on differential operators, *American Mathematical Monthly*, **83** (1976), 351–354.

[4] R. Mullin and G.-C. Rota, Theory of binomial enumeration, in "Graph Theory and Its Applications", Academic Press, 1970.

[5] J. Riordan, *Combinatorial Identities*, John Wiley and Sons, 1968.

[6] S. Roman, The algebra of formal series, *Advances in Mathematics*, **31** (1979), 309–329. Erratum **35** (1980), 274.

[7] S. Roman, The algebra of formal series II: Sheffer sequences, *J. Math. Anal. Appl.* **74** (1980), 120–143.

[8] S. Roman, The algebra of formal series III: several variables, *J. Approximation Theory*, **26** (1979), 340–381.

[9] S. Roman, The formula of Faa di Bruno, *American Mathematical Monthly*, **87** (1980), 805–809.

[10] S. Roman, Polynomials, power series and interpolation, *J. Math. Anal. Appl.* **80** (1981), 333–371.

[11] S. Roman and G.-C. Rota, The umbral calculus, *Advances in Mathematics*, **27** (1978), 95–188.

[12] G.-C. Rota, D. Kahaner and A. Odlyzko, Finite operator calculus, *J. Math. Anal. Appl.* **42** (1973), 685–760.