JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **87**, 58–115 (1982)

The Theory of the Umbral Calculus. I

STEVEN ROMAN

Department of Mathematics, California State University, Fullerton, California 92634

Submitted by G.-C. Rota

Section 1. Introduction. Section 2. Formal Power Series. Section 3. Linear Functionals. Section 4. Linear Operators. Section 5. Polynomial Sequences. Section 6. Recurrence Formulas. Section 7. Transfer Formulas. Section 8. Umbral Composition and Transfer Operators. Section 9. Examples: Gegenbauer, Chebyshev and Others. Section 10. Examples: Jacobi and Others. Section 11. Examples: The q-case. Section 12. Future Directions. Section 13. References.

1. INTRODUCTION

This is the first in a series of papers intended to develop the modern theory of the umbral calculus. No previous knowledge of the subject is required for this series.

Let us give a brief explanation of the term modern umbral calculus. A large part of applied analysis is concerned with the study of certain sequences of special polynomials. Some of the most important of these sequences are associated with the names of Jacobi, Gegenbauer, Legendre, Chebyshev, Bessel, Laguerre, Hermite and Bernoulli. All of these sequences, and many more, fall into a special class. Boas and Buck, in their work on polynomial expansions of analytic functions, used the term sequences of *generalized Appell type* for members of this class. A sequence $p_n(x)$ of polynomials is of generalized Appell type if it has a generating function of the form

$$A(t)\,\Psi(xh(t)) = \sum_{k=0}^{\infty} p_k(x)\,t^k$$

where

$$A(t) = \sum_{k=0}^{\infty} a_k t^k, \qquad a_0 \neq 0,$$

$$\Psi(t) = \sum_{k=0}^{\infty} \Psi_k t^k, \qquad \Psi_k \neq 0 \text{ for all } k,$$

0022-247X/82/050058-58$02.00/0

and

$$h(t)=\sum_{k=1}^{\infty} h_k t^k, \qquad h_1 \neq 0.$$

The modern umbral calculus grew out of an attempt to develop a unified theory for this class of polynomial sequences. In an earlier paper this author and G.-C. Rota developed the theory to deal effectively with an important subclass known as the sequences of Sheffer A-type zero, whose generating functions are of the form

$$A(t)\,e^{xh(t)}=\sum_{k=0}^{\infty} \frac{p_k(x)}{k!}\,t^k. \tag{1.1}$$

This subclass includes the important sequences of Hermite, Laguerre and Bernoulli. We remark that the sequence $p_n(x)$ is normalized by the presence of $k!$ on the right side of (1.1). All attempts at that time to extend the theory to the entire class of generalized Appell sequences failed. It is the purpose of this paper to make that extension.

Let us give a breif description of the contents of this paper. Section 2 contains a review of needed facts about formal power series. Section 3 discusses the dual vector space P^* of all linear functionals on the algebra P of polynomials. For it is in the umbral calculus that one studies the algebra P via its dual space P^*. In this section the structure of an algebra is put on the vector space P^*. Then since we may multiply linear functionals, the notion of a geometric sequence ML^k for $k=0, 1, 2,\ldots$ and M and L in P^* makes sense. Section 4 defines a certain algebra of linear operators on P which is isomorphic to the algebra P^*. Thus technically we introduce no new mathematical concepts; however, the notational convenience of the linear operator proves indispensible to the theory. In Section 5 we define the main object of study—*the Sheffer sequence*. Briefly, a sequence $s_n(x)$ of polynomials is the Sheffer sequence for a pair of linear functionals (M, L) if it is orthogonal to the geometric sequence ML^k, that is, if

$$\langle ML^k \mid s_n(x)\rangle = c_n \delta_{n,k}$$

for all $n, k \geqslant 0$ where c_n is a fixed sequence of non-zero constants and the notation $\langle N \mid p(x)\rangle$ is used for the action of N in P^* on $p(x)$ in P. Of course $\delta_{n,k}$ is the Kronecker delta function, $\delta_{n,k}=0$ if $n\neq k$ and $\delta_{n,n}=1$. In this same section we give several characterizations of Sheffer sequences, including the generating function

$$A(t)\,\varepsilon_x(h(t))=\sum_{k=0}^{\infty} \frac{s_k(x)}{c_k}\,t^k. \tag{1.2}$$

where $A(t)$ and $h(t)$ are as before and

$$\varepsilon_x(t) = \sum_{k=0}^{\infty} \frac{x^k}{c_k} t^k$$

is a generalization of the exponential series ($\varepsilon_x(t) = e^{xt}$ if $c_n = n!$). Thus we see that if $\Psi_n = 1/c_n$, then the sequence $s_n(x)/c_n$ is of generalized Appell type as defined by Boas and Buck. Also included in Section 5 is an algebraic characterization of Sheffer sequences which may be thought of as a generalization of the binomial formula. In Section 6 we derive recurrence formulas for Sheffer sequences and in Section 7 we give a powerful formula for the direct computation of Sheffer sequences. Section 8 is devoted in part to the *connection-constants* problem of determining the constants $a_{n,k}$ in

$$s_n(x) = \sum_{k=0}^{\infty} a_{n,k} p_k(x),$$

where $s_n(x)$ and $p_n(x)$ are given Sheffer sequences. The remainder of the paper is devoted to examples.

The author realizes all too well the motivational difficulties encountered in reading a large amount of theory without the benefit of example. Accordingly, one may resonably omit Sections 7 and 8 at first reading.

Let us point out one of the most innovative aspects of the present theory. Suppose $s_n(x)$ is the Sheffer sequence for the pair of linear functionals (M, L) where M is *not the* multiplicative identity in the algebra P^*. If we denote this identity by ε, then the Sheffer sequence $p_n(x)$ for the pair (ε, L) bears a strong association to $s_n(x)$. Many of the properties of $s_n(x)$ are possessed by $p_n(x)$ and yet in some sense $p_n(x)$ is a simpler sequence. Now the point is that almost all of the well-known classical sequences are of the type $s_n(x)$. In the Hermite and Bernoulli cases the associated sequence $p_n(x)$ is the simple sequence x^n. In the Laguerre case $L_n^{(\alpha)}(x)$ of order α, the simpler associated sequence is the Laguerre sequence $L_n^{(-1)}(x)$. But up to now there had been no clue to the existence of such sequences $p_n(x)$ associated to, for example, the Jacobi, Gegenbauer or Chebyshev sequences. A major portion of the examples is devoted to the study of the properties of these new sequences.

It is painfully evident from even a superficial screening of the literature that one man's Hermite polynomial, say, is not another man's Hermite polynomial. The difference is mainly due to normalization factors. We are no exception in this regard. For the Gegenbauer polynomials $G_n(x)$ Rainville sets

$$(1 - 2xt + t^2)^{-\lambda} = \sum_{k=0}^{\infty} G_n(x)\, t^k.$$

In Section 9 we shall study the Sheffer sequence $s_n(x)$ characterized by

$$(1 - 2xt + t^2)^{-\lambda} = \sum_{k=0}^{\infty} \binom{-\lambda}{k} s_k(x)\, t^k.$$

Thus $\binom{-\lambda}{k} s_n(x)$ is the Gegenbauer sequence. In the hope of minimizing the confusion we shall reserve the name Gegenbauer for the sequence $G_n(x)$. referring to Section 9 as the Gegenbauer case. For the Chebyshev polynomials of the first kind we have

$$(1 - x^2)(1 - 2xt + t^2)^{-1} = 1 + 2 \sum_{k=0}^{\infty} T_k(x)\, t^k.$$

In Section 9 we study the Sheffer sequence characterized by

$$(1 - x^2)(1 - 2xt + t^2)^{-1} = \sum_{k=0}^{\infty} (-1)^k s_k(x)\, t^k.$$

A similar normalization factor is required for the Chebyshev polynomials of the second kind. Finally, for the Jacobi polynomials $P_n(x)$ Rainville gives the generating function

$$(1-t)^{-1-\alpha-\beta}\, {}_2F_1 \left[\begin{matrix} \dfrac{1+\alpha+\beta}{2}, \dfrac{2+\alpha+\beta}{2}; \\ 1+\alpha; \end{matrix} \; \frac{2t(x-1)}{(1-t)^2} \right]$$
$$= \sum_{k=0}^{\infty} \frac{(1+\alpha+\beta)^{(k)}}{(1+\alpha)^{(k)}} P_k(x)\, t^k.$$

In Section 10 we study the Sheffer sequence $J_n(x)$ satisfying

$$(1-t)^{-1-\alpha-\beta}\, {}_2F_1 \left[\begin{matrix} \dfrac{1+\alpha+\beta}{2}, \dfrac{2+\alpha+\beta}{2}; \\ 1+\alpha; \end{matrix} \; \frac{2xt}{(1-t)^2} \right]$$
$$= \sum_{k=0}^{\infty} \frac{((1+\alpha+\beta)/2)^{(k)}((2+\alpha+\beta)/2)^{(k)}}{(1+\alpha)^{(k)}} J_k(x)\, t^k.$$

From this we may easily obtain $J_n(x)$ in terms of the classical Jacobi polynomials (see Section 10).

We have decided to postpone any discussion of applications of the umbral calculus to future works. The calculus may be applied successfully to the study of orthogonality, inverse relations, formal power series, solutions to recurrence relations and counting techniques to mention only a few.

However, these applications can be broadened and await further generalizations of the theory to appear as sequels to this paper. We briefly discuss two such generalizations in the last section.

2. Formal Power Series

In this section we give a few basic facts about formal power series. Let $\mathscr{F}$ be the algebra of all formal power series in the variable t over the field K (of characteristic zero). Addition and multiplication in $\mathscr{F}$ is purely formal and $\mathscr{F}$ is well known to be an integral domain. If

$$f(t) = \sum_{k=0}^{\infty} a_k t^k, \tag{2.1}$$

then the *degree* of $f(t)$ is the smallest k such that $a_k \neq 0$. It is easy to see that $\deg f(t)\, g(t) = \deg f(t) + \deg g(t)$.

The series $f(t)$ has a multiplicative inverse in $\mathscr{F}$, denoted by $f^{-1}(t)$ or $1/f(t)$, if and only if $\deg f(t) = 0$. We call such a series *invertible*.

Suppose $g_k(t)$ is a sequence in $\mathscr{F}$ for which $\deg g_k(t) \geqslant k$. Then if a_k is a sequence of constants, the sum

$$\sum_{k=0}^{\infty} a_k g_k(t)$$

is a well-defined series in $\mathscr{F}$, found by simply collecting coefficients of like powers of t. In particular, we may take $g_k(t) = g(t)^k$ where $\deg g(t) \geqslant 1$.

If $f(t)$ is given by (2.1), we may form the composition

$$f(g(t)) = \sum_{k=0}^{\infty} a_k g(t)^k$$

which is a well-defined element of $\mathscr{F}$ provided $\deg g(t) \geqslant 1$. It is clear that $\deg f(g(t)) = \deg f(t) \cdot \deg g(t)$.

The series $f(t)$ has a compositional inverse, denoted by $\bar{f}(t)$ and satisfying $\bar{f}(f(t)) = f(\bar{f}(t)) = t$, if and only if $\deg f(t) = 1$. We call any series $f(t)$ with $\deg f(t) = 1$ a *delta series*.

A sequence $g_k(t)$ for which $\deg g_k(t) = k$ forms a pseudobasis for $\mathscr{F}$. In other words, for each series $f(t)$ there is a unique sequence of constants a_k for which

$$f(t) = \sum_{k=0}^{\infty} a_k g_k(t).$$

In particular, the powers of a delta series form a pseudobasis for $\mathscr{F}$.

3. LINEAR FUNCTIONALS

Let P be the algebra of polynomials in a single variable over K and let P^* be the dual vector space of all linear functionals on P. We use the notation $\langle L \mid p(x)\rangle$ for the action of L in P^* on $p(x)$ in P. Any linear functional L in P^* is uniquely defined by specifying the values $\langle L \mid x^n\rangle$ for $n \geqslant 0$.

Let c_n be a fixed sequence of non-zero constants. We use this sequence to define, for each $f(t)$ in $\mathscr{F}$, a linear functional in P^* as follows. If $f(t) = \sum_{k=0}^{\infty} a_k t^k$, then the linear functional $f(t)$ satisfies

$$\langle f(t) \mid x^n\rangle = c_n a_n \tag{3.1}$$

for all $n \geqslant 0$. Notice that we have used the same notation $f(t)$ for the power series and the linear functional. This should cause no confusion since if $f(t)$ and $g(t)$ are in $\mathscr{F}$, then $f(t) = g(t)$ if and only if $\langle f(t) \mid x^n\rangle = \langle g(t) \mid x^n\rangle$ for all $n \geqslant 0$. In other words, $f(t)$ and $g(t)$ are equal as formal series if and only if they are equal as linear functionals.

The action defined in (3.1) depends on the particular choice of the sequence c_n, although the notation does not reflect this. We will generally think of c_n as fixed and no confusion should arise.

As a consequence of (3.1) we have

$$\langle t^k \mid x^n\rangle = c_n \delta_{n,k}$$

and

$$\left\langle \sum_{k=0}^{\infty} a_k t^k \,\middle|\, x^n \right\rangle = \sum_{k=0}^{\infty} a_k \langle t^k \mid x^n\rangle;$$

and so for any $p(x)$ in P,

$$\left\langle \sum_{k=0}^{\infty} a_k t^k \,\middle|\, p(x) \right\rangle = \sum_{k=0}^{\infty} a_k \langle t^k \mid p(x)\rangle.$$

Now any linear functional L in P^* can be represented as a series in $\mathscr{F}$. In fact, if

$$f_L(t) = \sum_{k=0}^{\infty} \frac{\langle L \mid x^k\rangle}{c_k} t^k, \tag{3.2}$$

then

$$\begin{aligned}\langle f_L(t) \mid x^n\rangle &= \sum_{k=0}^{\infty} \frac{\langle L \mid x^k\rangle}{c_k} \langle t^k \mid x^n\rangle \\ &= \langle L \mid x^n\rangle\end{aligned}$$

and so as linear functionals $f_L(t) = L$.

It is easily verified that the map $L \to f_L(t)$ is a vector space isomorphism from P^* onto $\mathscr{F}$. We shall obscure this map by identifying P^* as the vector space $\mathscr{F}$ of all formal power series in t. Thus from now on we shall write our linear functionals in the form of power series in t.

The isomorphism $L \to f_L(t)$ has induced a natural product on linear functionals—namely, that of formal power series. In symbols $f_{LM}(t) = f_L(t)\, f_M(t)$.

Let us give some simple consequences of these results.

PROPOSITION 3.1. *If* $\deg f(t) > \deg p(x)$, *then*

$$\langle f(t) \mid p(x) \rangle = 0.$$

PROPOSITION 3.2. *If* $f(t)$ *is in* $\mathscr{F}$, *then*

$$f(t) = \sum_{k=0}^{\infty} \frac{\langle f(t) \mid c^k \rangle}{c_k} t^k. \tag{3.3}$$

Proof. Applying the right side to x^n gives

$$\left\langle \sum_{k=0}^{\infty} \frac{\langle f(t) \mid x^k \rangle}{c_k} t^k \;\middle|\; x^n \right\rangle = \sum_{k=0}^{\infty} \frac{\langle f(t) \mid x^k \rangle}{c_k} \langle t^k \mid x^n \rangle$$
$$= \langle f(t) \mid x^n \rangle.$$

Thus the two sides of (3.3) are equal as linear functionals, and so also as formal power series.

In view of (3.3) we have for any $f(t)$, $g(t)$ in $\mathscr{F}$,

$$\langle f(t)\, g(t) \mid x^n \rangle = \sum_{k=0}^{n} \frac{c_n}{c_k c_{n-k}} \langle f(t) \mid x^k \rangle \langle g(t) \mid x^{n-k} \rangle.$$

COROLLARY 1. *If* $\deg p_n(x) = n$ *and* $\langle f(t) \mid p_n(x) \rangle = 0$ *for all* $n \geqslant 0$, *then* $f(t) = 0$.

PROPOSITION 3.3. *If* $p(x)$ *is in* P, *then*

$$p(x) = \sum_{k \geqslant 0} \frac{\langle t^k \mid p(x) \rangle}{c_k} x^k. \tag{3.4}$$

Proof. Applying t^n to both sides of (3.4) shows that corresponding coefficients of like powers of x are equal. This proves the result.

COROLLARY 2. *If* $\deg f_k(t) = k$ *and* $\langle f_k(t) \mid p(x) \rangle = 0$ *for all* $k \geqslant 0$, *then* $p(x) = 0$.

Proof. This follows from the fact that $f_k(t)$ forms a pseudobasis for $\mathscr{F}$ and so $\langle t^j \mid p(x) \rangle = 0$ for all $j \geqslant 0$.

PROPOSITION 3.4. *If $f(t)$ is a delta series and*

$$g(t) = \sum_{k=0}^{\infty} a_k f(t)^k,$$

then

$$\langle g(t) \mid x^n \rangle = \sum_{k=0}^{\infty} a_k \langle f(t)^k \mid x^n \rangle.$$

Proof. We have

$$\begin{aligned} \langle g(t) \mid x^n \rangle &= \left\langle \sum_{k=0}^{\infty} a_k f(t)^k \,\middle|\, x^n \right\rangle \\ &= \left\langle \sum_{k=0}^{n} a_k f(t)^k \,\middle|\, x^n \right\rangle \\ &= \sum_{k=0}^{n} a_k \langle f(t)^k \mid x^n \rangle \\ &= \sum_{k=0}^{\infty} a_k \langle f(t)^k \mid x^n \rangle. \end{aligned}$$

One of the most important linear functionals on P is the *evaluation functional* denoted, for y in K, by $\varepsilon_y(t)$ and defined by

$$\langle \varepsilon_y(t) \mid p(x) \rangle = p(y).$$

In view of (3.3) we have

$$\varepsilon_y(t) = \sum_{k=0}^{\infty} \frac{y^k}{c_k} t^k.$$

It is interesting to note the form of $\varepsilon_y(t)$ for various choices of the sequence c_n. For example, if $c_n = n!$, then

$$\varepsilon_y(t) = e^{yt},$$

and if $c_n = 1/\binom{\lambda}{n}$, then

$$\varepsilon_y(t) = (1 + yt)^{\lambda}.$$

It will be convenient to introduce a linear operator ∂_t on $\mathscr{F}$ by setting

$$\begin{aligned} \partial_t t^k &= \frac{c_k}{c_{k-1}} t^{k-1}, \qquad & k \geqslant 1 \\ &= 0, & k = 0 \end{aligned}$$

and extending by infinite linearity. Then we have for all $n \geqslant 0$ and $k \geqslant 0$,

$$\begin{aligned} \langle t^k \mid x \cdot x^n \rangle &= \langle t^k \mid x^{n+1} \rangle \\ &= c_{n+1} \delta_{n+1,k} \\ &= \frac{c_k}{c_{k-1}} c_n \delta_{n,k-1} \\ &= \left\langle \frac{c_k}{c_{k-1}} t^{k-1} \,\middle|\, x^n \right\rangle \\ &= \langle \partial_t t^k \mid x^n \rangle. \end{aligned}$$

This proves the following proposition.

PROPOSITION 3.5. *If $f(t)$ is in $\mathscr{F}$, then*

$$\langle \partial_t f(t) \mid p(x) \rangle = \langle f(t) \mid x p(x) \rangle$$

for all $p(x)$ in P.

For practice, we compute

$$\begin{aligned} \partial_t \varepsilon_y(t) &= \partial_t \sum_{k=0}^{\infty} \frac{y^k}{c_k} t^k \\ &= \sum_{k=1}^{\infty} \frac{y^k}{c_{k-1}} t^{k-1} \\ &= y \varepsilon_y(t). \end{aligned} \tag{3.5}$$

As a final remark, when we are thinking of a delta (or invertible) series $f(t)$ as a linear functional we shall refer to it as a *delta* (or *invertible*) *functional.*

4. Linear Operators

If $f(t)$ in $\mathscr{F}$ has the form $f(t)=\sum_{k=0}^{\infty} a_k t^k$, then we shall define the linear operator $f(t)$ on $\mathscr{F}$ by

$$f(t)\,x^n=\sum_{k=0}^{n}\frac{c_n}{c_{n-k}}\,a_k x^{n-k}. \tag{4.1}$$

Again we have used the same notation $f(t)$ for a formal power series and a linear operator. No problems will arise (except perhaps for the temporary confusion of the reader) since $f(t)$ and $g(t)$ are equal as formal power series if and only if they are equal as linear operators. [To see the "if" part take successively $n=0, 1, 2,\ldots$ in (4.1)].

Notice that we are using juxtaposition to denote the action of an operator on a polynomial. A little practice will remove the discomfort involved in thinking of an element $f(t)$ in $\mathscr{F}$ as either a formal power series, a linear functional or a linear operator, and the notational difference between

$$\langle f(t) \mid p(x)\rangle$$

and

$$f(t)\;p(x)$$

will make the particular type of action of $f(t)$ on $p(x)$ clear.

The action $f(t)\;p(x)$ depends on the sequence c_n. However, we shall think of this sequence as being fixed and so no confusion should arise.

It follows from (4.1) that

$$\begin{aligned} t^k x^n &= \frac{c_n}{c_{n-k}}\,x^{n-k}, & n\geqslant k\\ &=0, & n<k \end{aligned}$$

and so

$$\begin{aligned} (t^k t^j)\,x^n = t^{k+j}x^n &= \frac{c_n}{c_{n-k-j}}\,x^{n-k-j}\\ &=\frac{c_n}{c_{n-j}}\,\frac{c_{n-j}}{c_{n-k-j}}\,x^{n-k-j}\\ &=\frac{c_n}{c_{n-j}}\,t^k x^{n-j}\\ &=t^k(t^j x^n). \end{aligned}$$

Therefore, if $f(t)$, $g(t)$ are in $\mathscr{F}$ we have

$$[f(t)\, g(t)]\, p(x) = f(t)[g(t)\, p(x)] \tag{4.2}$$

and we may write $f(t)\, g(t)\, p(x)$ without ambiguity.

Notice also that

$$f(t)\, g(t)\, p(x) = g(t)\, f(t)\, p(x).$$

Actually, (4.2) shows that the product in $\mathscr{F}$ is composition of operators.

When we are thinking of a delta (or invertible) series $f(t)$ as an operator we may refer to it as a *delta* (or *invertible*) *operator.*

A key relationship between the linear functional $f(t)$ and the linear operator $f(t)$ is given in the next theorem.

THEOREM 4.1. *If $f(t)$, $g(t)$ are in $\mathscr{F}$, then*

$$\langle g(t)\, f(t) \mid p(x)\rangle = \langle g(t) \mid f(t)\, p(x)\rangle$$

for all $p(x)$ in P.

Proof. By linearity we need only show this for $g(t) = t^k$, $f(t) = t^j$ and $p(x) = x^n$. But then we have

$$\begin{aligned}
\langle t^k t^j \mid x^n\rangle &= \langle t^{k+j} \mid x^n\rangle \\
&= c_n \delta_{n,k+j} \\
&= \frac{c_n}{c_{n-j}}\, c_{n-j} \delta_{n-j,k} \\
&= \left\langle t^k \;\middle|\; \frac{c_n}{c_{n-j}}\, x^{n-j} \right\rangle \\
&= \langle t^k \mid t^j x^n\rangle.
\end{aligned}$$

For many choices of the sequence c_n, the operator t may be expressed in terms of some more familiar operators. To fix the notation we use $(x)_n = x(x-1) \cdots (x-n+1)$, $x^{(n)} = x(x+1) \cdots (x+n-1)$

$$Dx^n = nx^{n-1},$$

$$D^{-1}x^n = \frac{1}{n+1}\, x^{n+1},$$

$$x^{-1}x^n = x^{n-1}, \qquad x^{-1}1 = 0.$$

Then

(1) when $c_n = n!$, $t = D$;

(2) when $c_n = 1$, $t = x^{-1}$;

(3) when $c_n = (n!)^{m+1}$, $t = (Dx)^m D$;

(4) when $c_n = 1/(-\lambda)_n$, $t = -(\lambda + xD)^{-1} x^{-1}$;

(5) when $c_n = 1/n$, $t = x^{-1}D^{-1}x^{-1}$;

(6) when $c_n = 1/\binom{-\lambda}{n}$, $t = -(\lambda + xD)^{-1} D$;

(7) when $c_n = 2^{2n}(1+\alpha)^{(n)}/(1+\alpha+\beta)^{(2n)}$,
$t = 4(1+\alpha+\beta+2xD)^{-1}(2+\alpha+\beta+2xD)^{-1}x^{-1}(\alpha+xD)$;

(8) when $c_n = (1-q)\cdots(1-q^n)/(1-q)^n$,
$tp(x) = (p(qx) - p(x))/(qx - x)$.

As we shall see, from the point of view of the present theory the operator t is the natural operator for studying various polynomial sequences. Case 1 is related to sequences of Sheffer A-type zero, such as the Hermite and Laguerre polynomials. This case has been studied by the present author and G.-C. Rota. Case 6 concerns itself with such polynomials as those of Gegenbauer and Chebyshev. Case 7 relates to the Jacobi polynomials and Case 8 is the so called q-case.

Let us make some remarks concerning the series $\varepsilon_y(t)$. This series acts as the same linear functional, namely, evaluation at y, regardless of the particular sequence c_n. However, this is not the case for the operator $\varepsilon_y(t)$. We have

$$\varepsilon_y(t)\, x^n = \sum_{k=0}^{n} \frac{c_n}{c_k c_{n-k}} y^k x^{n-k}.$$

Now if for example $c_n = n!$, we obtain

$$\varepsilon_y(t)\, x^n = (x+y)^n,$$

and if $c_n = 1$, we obtain

$$\varepsilon_y(t)\, x^n = \frac{x^{n+1} - y^{n+1}}{x - y}.$$

It is not hard to see that not all linear operators on P are of the form $f(t)$ in $\mathscr{F}$. We have the following characterization of such operators.

PROPOSITION 4.1. *Let U be a linear operator on P. There exists a series $f(t)$ in $\mathscr{F}$ such that $Up(x) = f(t)\, p(x)$ for all $p(x)$ in P if and only if U commutes with the operator t, that is, $Utp(x) = tUp(x)$ for all $p(x)$ in P.*

Proof. The necessity is clear. For the converse, suppose U commutes with t. We define the series $f(t)$ by

$$f(t) = \sum_{k=0}^{\infty} \frac{\langle t^0 \mid Ux^k \rangle}{c_k} t^k.$$

Then

$$\begin{aligned}
f(t)\,x^n &= \sum_{k=0}^{n} \frac{\langle t^0 \mid Ux^k\rangle}{c_k}\, t^k x^n \\
&= \sum_{k=0}^{n} \frac{c_n}{c_k c_{n-k}} \langle t^0 \mid Ux^k\rangle\, x^{n-k} \\
&= \sum_{k=0}^{n} \frac{\langle t^0 \mid Ut^{n-k}x^n\rangle}{c_{n-k}}\, x^{n-k} \\
&= \sum_{k=0}^{n} \frac{\langle t^{n-k} \mid Ux^n\rangle}{c_{n-k}}\, x^{n-k} \\
&= \sum_{k=0}^{n} \frac{\langle t^k \mid Ux^n\rangle}{c_k}\, x^k \\
&= Ux^n.
\end{aligned}$$

The last equality follows from the fact (easily proved by induction) that deg $Ux^n \leqslant n$ and from Eq. (3.4).

COROLLARY 1. *A linear operator on P has the form $f(t)$ in $\mathscr{F}$ if and only if it commutes with any delta operator.*

Proof. This follows from the fact that the sequence of powers of a delta operator form a pseudobasis for $\mathscr{F}$. We may then apply Proposition 4.1.

COROLLARY 2. *A linear operator on P has the form $f(t)$ in $\mathscr{F}$ if and only if it commutes with any evaluation operator $\varepsilon_y(t)$.*

Proof. This follows from Corollary 1 since $\varepsilon_y(t) - c_0^{-1} t^0$ is a delta series.

5. POLYNOMIAL SEQUENCES

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

THEOREM 5.1. *Let $f(t)$ be a delta series and let $g(t)$ be an invertible series. Then the identity*

$$\langle g(t)\, f(t)^k \mid s_n(x)\rangle = c_n \delta_{n,k} \tag{5.1}$$

for all n, $k \geqslant 0$ determines a unique sequence $s_n(x)$ in P.

Proof. The uniqueness follows from Corollary 1 of Proposition 3.3. For the existence, suppose $s_n(x)=\sum_{j=0}^{n} a_{n,j}x^j$ where $a_{n,n}\neq 0$, and $g(t)\,f(t)^k=\sum_{i=k}^{\infty} b_{k,i}t^i$ where $b_{k,k}\neq 0$. Then (5.1) becomes

$$\begin{aligned} c_n\delta_{n,k} &= \left\langle \sum_{i=k}^{\infty} b_{k,i}t^i \,\middle|\, \sum_{j=0}^{n} a_{n,j}x^j \right\rangle \\ &= \sum_{i=k}^{\infty}\sum_{j=0}^{n} b_{k,i}a_{n,j}\langle t^i \mid x^j\rangle \\ &= \sum_{j=k}^{n} b_{k,j}a_{n,j}c_j. \end{aligned} \tag{5.2}$$

By taking $k=n$ one obtains

$$a_{n,n}=\frac{1}{b_{n,n}}\cdot$$

By taking successively $k=n-1, n-2,..., 0$ one obtains a triangular system of equations which can be solved for $a_{n,k}$.

We will say that the sequence $s_n(x)$ is the *Sheffer sequence* for the pair $(g(t), f(t))$, or $s_n(x)$ is *Sheffer for* $(g(t), f(t))$. Notice that $g(t)$ must be invertible and $f(t)$ must be a delta series. The case $g(t)=t^0$ calls for special attention. The Sheffer sequence for $(t^0, f(t))$ will be called the *associated sequence* for $f(t)$, and we say $s_n(x)$ is *associated to* $f(t)$ or $f(t)$ is *associated to* $s_n(x)$.

THEOREM 5.2 (The Expansion Theorem). *Let $s_n(x)$ be Sheffer for $(g(t), f(t))$. Then for any $h(t)$ in $\mathscr{F}$*

$$h(t)=\sum_{k=0}^{\infty}\frac{\langle h(t)\mid s_k(x)\rangle}{c_k}\,g(t)\,f(t)^k.$$

Proof. We simply apply the left side to $s_n(x)$ to obtain $\langle h(t)\mid s_n(x)\rangle$. The fact that $\deg s_n(x)=n$ completes the proof.

COROLLARY 1. *If $p_n(x)$ is the associated sequence for $f(t)$, then for any $h(t)$ in $\mathscr{F}$*

$$h(t)=\sum_{k=0}^{\infty}\frac{\langle h(t)\mid p_k(x)\rangle}{c_k}\,f(t)^k.$$

The next results show how to expand an arbitrary polynomial as a linear combination of polynomials from a Sheffer sequence. They follow from Theorem 5.2 by taking $h(t)=\varepsilon_y(t)$.

COROLLARY 2. *Let $s_n(x)$ be Sheffer for $(g(t), f(t))$. Then for any $p(x)$ in P*

$$p(x) = \sum_{k \geqslant 0} \frac{\langle g(t) f(t)^k \mid p(x) \rangle}{c_k} s_k(x)$$

COROLLARY 3. *Let $p_n(x)$ be associated to $f(t)$. Then for any $p(x)$ in P*

$$p(x) = \sum_{k \geqslant 0} \frac{\langle f(t)^k \mid p(x) \rangle}{c_k} p_k(x).$$

It is our intention now to characterize Sheffer sequences in several ways. We begin with the generating function.

THEOREM 5.3 (Generating Function). *The sequence $s_n(x)$ is Sheffer for $(g(t), f(t))$ if and only if*

$$\frac{1}{g(\bar{f}(t))} \varepsilon_y(\bar{f}(t)) = \sum_{k=0}^{\infty} \frac{s_k(y)}{c_k} t^k \tag{5.3}$$

for all y in K.

Proof. If $s_n(x)$ is Sheffer for $(g(t), f(t))$, then by the Expansion Theorem

$$\varepsilon_y(t) = \sum_{k=0}^{\infty} \frac{s_k(y)}{c_k} g(t) f(t)^k$$

and so

$$\frac{1}{g(t)} \varepsilon_y(t) = \sum_{k=0}^{\infty} \frac{s_k(y)}{c_k} f(t)^k$$

and

$$\frac{1}{g(\bar{f}(t))} \varepsilon_y(\bar{f}(t)) = \sum_{k=0}^{\infty} \frac{s_k(y)}{c_k} t^k.$$

For the converse, suppose (5.3) holds. Then if $r_n(x)$ is the Sheffer sequence for $(g(t), f(t))$, we have

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{r_k(y)}{c_k} t^k &= \frac{1}{g(\bar{f}(t))} \varepsilon_y(\bar{f}(t)) \\ &= \sum_{k=0}^{\infty} \frac{s_k(y)}{c_k} t^k \end{aligned}$$

and so $r_k(x) = s_k(x)$ for all $k \geqslant 0$.

COROLLARY 1. *The sequence $p_n(x)$ is associated to $f(t)$ if and only if*

$$\varepsilon_y(\bar{f}(t)) = \sum_{k=0}^{\infty} \frac{p_k(x)}{c_k} t^k.$$

Equation (5.3) defines the sequence $c_n^{-1} s_n(x)$ as a so-called generalized Appell sequence by Boas and Buck. Thus we see that the present theory applies to a rather broad class of polynomial sequences.

The generating function leads us to a representation for Sheffer sequence.

THEOREM 4.5. *The sequence $s_n(x)$ is Sheffer for $(g(t), f(t))$ if and only if*

$$s_n(x) = \sum_{k=0}^{n} \frac{\langle g(\bar{f}(t))^{-1} \bar{f}(t)^k \mid x^n \rangle}{c_k} x^k. \tag{5.4}$$

Proof. Applying the right-hand side of (5.3) to x^n gives

$$\left\langle \sum_{k=0}^{\infty} \frac{s_k(y)}{c_k} t^k \;\middle|\; x^n \right\rangle = s_n(y)$$

and applying the left-hand side of (5.3) to x^n gives

$$\langle g(\bar{f}(t))^{-1} \varepsilon_y(\bar{f}(t)) \mid x^n \rangle = \sum_{k=0}^{n} \frac{y_k}{c_k} \langle g(\bar{f}(t))^{-1} \bar{f}(t)^k \mid x^n \rangle$$

Since these equations hold for all y in K, the result follows.

Equation (5.4) is called the *conjugate representation* for $s_n(x)$.

THEOREM 5.5. *The sequence $s_n(x)$ is Sheffer for $(g(t), f(t))$ if and only if $g(t)\, s_n(x)$ is the associated sequence for $f(t)$.*

Proof. This follows directly from the definitions and Theorem 4.1.

Theorem 5.5 says that each associated sequence generates a class of Sheffer sequences, one for each invertible operator $g(t)$ in $\mathscr{F}$.

Next we give an operator characterization of Sheffer sequences.

THEOREM 5.6. *A sequence $p_n(x)$ is the associated sequence for $f(t)$ if and only if*

(i) $\langle t^0 \mid p_n(x) \rangle = c_0 \delta_{n,0}$,

(ii) $f(t)\, p_n(x) = \dfrac{c_n}{c_{n-1}}\, p_{n-1}(x)$

Proof. Suppose $\langle f(t)^k \mid p_n(x)\rangle = c_n \delta_{n,k}$. Then $k=0$ gives (i). To get (ii) we have

$$\begin{aligned}\langle f(t)^k \mid f(t)\, p_n(x)\rangle &= \langle f(t)^{k+1} \mid p_n(x)\rangle \\ &= c_n \delta_{n,k+1} \\ &= \left\langle f(t)^k \,\middle|\, \frac{c_n}{c_{n-1}}\, p_{n-1}(x)\right\rangle\end{aligned}$$

and so (ii) follows since $f(t)$ is a delta series. Conversely, if (i) and (ii) hold, then

$$\begin{aligned}\langle f(t)^k \mid p_n(x)\rangle &= \left\langle t^0 \,\middle|\, \frac{c_n}{c_{n-k}}\, p_{n-k}(x)\right\rangle \\ &= \frac{c_n}{c_{n-k}}\, c_{n-k} \delta_{n-k,0} \\ &= c_n \delta_{n,k}.\end{aligned}$$

THEOREM 5.7. *A sequence $s_n(x)$ is Sheffer for $(g(t), f(t))$ for some invertible $g(t)$ if and only if*

$$f(t)\, s_n(x) = \frac{c_n}{c_{n-1}}\, s_{n-1}(x). \tag{5.5}$$

Proof. If $s_n(x)$ is Sheffer for $(g(t), f(t))$, then $p_n(x) = g(t)\, s_n(x)$ is associated to $f(t)$. Hence

$$\begin{aligned}g(t)\, s_n(x) &= f(t)\, g^{-1}(t)\, p_n(x) \\ &= g^{-1}(t)\, f(t)\, p_n(x) \\ &= \frac{c_n}{c_{n-1}}\, g^{-1}(t)\, p_{n-1}(x) \\ &= \frac{c_n}{c_{n-1}}\, s_{n-1}(x).\end{aligned}$$

Conversely, if (5.5) holds, then we define a linear operator U on P by

$$U s_n(x) = p_n(x),$$

where $p_n(x)$ is associated to $f(t)$. Then

$$
\begin{aligned}
Uf(t)\, s_n(x) &= \frac{c_n}{c_{n-1}}\, Us_{n-1}(x) \\
&= \frac{c_n}{c_{n-1}}\, p_{n-1}(x) \\
&= f(t)\, p_n(x) \\
&= f(t)\, Us_n(x)
\end{aligned}
$$

and so $Uf(t) = f(t)U$. By Corollary 1 of Proposition 4.1 we deduce the existence of an invertible $g(t)$ in $\mathscr{F}$ for which $g(t)\, s_n(x) = p_n(x)$. The result follows from Theorem 5.5.

We now turn to a characterization of Sheffer sequences which generalizes the binomial formula.

THEOREM 5.8 (The Sheffer Identity). *A sequence $s_n(x)$ is Sheffer for the pair $(g(t), f(t))$ for some $g(t)$ if and only if*

$$\varepsilon_y(t)\, s_n(x) = \sum_{k=0}^{n} \frac{c_n}{c_k c_{n-k}}\, p_k(y)\, s_{n-k}(x) \tag{5.6}$$

for all y in K where $p_n(x)$ is associated to $f(t)$.

Proof. Suppose $s_n(x)$ is Sheffer for $(g(t), f(t))$. Then by the Expansion Theorem

$$\varepsilon_y(t) = \sum_{k=0}^{\infty} \frac{p_k(y)}{c_k}\, f(t)^k.$$

Applying both sides to $s_n(x)$ and using Theorem 5.7 gives Eq. (5.6). For the converse, let U be the linear operator on P defined by $Us_n(x) = p_n(x)$. Then it is sufficient to show that $U = h(t)$ for some $h(t)$ in $\mathscr{F}$. Now

$$
\begin{aligned}
\varepsilon_y(t)\, Us_n(x) &= \varepsilon_y(t)\, p_n(x) \\
&= \sum_{k=0}^{n} \frac{c_n}{c_k c_{n-k}}\, p_k(y)\, p_{n-k}(x) \\
&= U \sum_{k=0}^{n} \frac{c_n}{c_k c_{n-k}}\, p_k(y)\, s_{n-k}(x) \\
&= U\varepsilon_y(t)\, s_n(x)
\end{aligned}
$$

and so there exists an invertible $h(t)$ for which $U = h(t)$. This completes the proof.

One may observe that for $c_n = n!$ the Sheffer identity is

$$s_n(x+y) = \sum_{k=0}^{n} \binom{n}{k} s_k(x)\, p_{n-k}(x).$$

When $s_n(x) = p_n(x) = x^n$ we get the binomial formula.

An important property of Sheffer sequences is their performance with respect to multiplication in $\mathscr{F}$.

THEOREM 5.9. *Let $s_n(x)$ be Sheffer for the pair $(g(t), f(t))$ and let $p_n(x)$ be associated to $f(t)$. Then for all $h(t)$ and $l(t)$ in $\mathscr{F}$ we have*

$$\langle h(t)\, l(t) \mid s_n(x)\rangle = \sum_{k=0}^{n} \frac{c_n}{c_k c_{n-k}} \langle h(t) \mid s_k(x)\rangle \langle l(t) \mid p_{n-k}(x)\rangle$$

Proof. An instructive proof of this result is based on the algebra $P[x, y]$ of polynomials in two variables x and y. For $h(t)$ in $\mathscr{F}$ by $h_x(t)$ we mean the linear operator on $P[x, y]$ defined by

$$h_x(t)\, x^k y^j = \langle h(t) \mid x^k > y^j$$

and by $h_y(t)$ we mean

$$h_y(t)\, x^k y^j = x^k \langle h(t) \mid x^j\rangle.$$

Then the identity

$$\langle h(t)\, l(t) \mid x^n\rangle = \sum_{k=0}^{n} \frac{c_n}{c_k c_{n-k}} \langle h(t) \mid x^{n-k}\rangle \langle l(t) \mid x^k\rangle$$

can be written as

$$\begin{aligned}\langle h(t)\, l(t) \mid x^n\rangle &= h_x(t)\, l_y(t) \sum_{k=0}^{n} \frac{c_n}{c_k c_{n-k}} x^{n-k} y^k \\ &= h_x(t)\, l_y(t)\, \varepsilon_y(t)\, x^n.\end{aligned}$$

By linearity we may replace x^n by $s_n(x)$ giving

$$\begin{aligned}\langle h(t)\, l(t) \mid s_n(x)\rangle &= h_x(t)\, l_y(t)\, \varepsilon_y(t)\, s_n(x) \\ &= h_x(t)\, l_y(t) \sum_{k=0}^{n} \frac{c_n}{c_k c_{n-k}} s_k(y)\, p_{n-k}(x) \\ &= \sum_{k=0}^{n} \frac{c_n}{c_k c_{n-k}} \langle l(t) \mid s_k(x)\rangle \langle h(t) \mid p_{n-k}(x)\rangle.\end{aligned}$$

The result follows by interchanging $h(t)$ and $l(t)$.

6. Recurrence Formulas

If μ is a linear operator on P, then the *adjoint* μ^* is the linear operator on $\mathscr{F}$ defined by

$$\mu^* f(t) = \sum_{k=0}^{\infty} \frac{\langle f(t) \mid \mu x^k \rangle}{c_k} t^k.$$

PROPOSITION 6.1. *If* $\deg f_j(t) \geqslant j$ *for* $j \geqslant 0$ *and*

$$f(t) = \sum_{j=0}^{\infty} a_j f_j(t),$$

then

$$\mu^* f(t) = \sum_{j=0}^{\infty} a_j \mu^* f_j(t). \tag{6.1}$$

Proof. We have

$$\begin{aligned} \mu^* f(t) &= \sum_{k=0}^{\infty} \sum_{j=0}^{\deg \mu x^k} a_j \langle f_j(t) \mid \mu x^k \rangle \frac{t^k}{c_k} \\ &= \sum_{j=0}^{\infty} a_j \sum_{k=0}^{\infty} \frac{\langle f_j(t) \mid \mu x^k \rangle}{c_k} t^k \\ &= \sum_{j=0}^{\infty} a_j \mu^* f_j(t). \end{aligned}$$

PROPOSITION 6.2. *If μ is a linear operator on P, then*

$$\langle \mu^* f(t) \mid p(x) \rangle = \langle f(t) \mid \mu p(x) \rangle$$

for all $f(t)$ in $\mathscr{F}$ and $p(x)$ in P.

Proof. By Proposition 6.1 we need only show this for $f(t) = t^j$ and $p(x) = x^n$. In this case

$$\begin{aligned} \langle \mu^* t^j \mid x^n \rangle &= \sum_{k=0}^{\infty} \frac{\langle t^j \mid \mu x^k \rangle}{c_k} \langle t^k \mid x^n \rangle \\ &= \langle t^j \mid \mu x^n \rangle. \end{aligned}$$

If $p_n(x)$ is associated to $f(t)$, then the *umbral shift* θ_f *associated* to $f(t)$ [or $p_n(x)$] is the linear operator on P is defined by

$$\theta_f p_n(x) = \frac{(n+1) c_n}{c_{n+1}} p_{n+1}(x)$$

for all $n \geqslant 0$.

Recall that a derivation ∂ on an algebra A is a linear operator on A satisfying

$$\partial(ab) = (\partial a)b + a\,\partial b$$

for all a, b in A.

We may characterize umbral shifts by their adjoints.

THEOREM 6.1. *An operator θ on P is the umbral shift for $f(t)$ if and only if its adjoint θ^* is a derivation on $\mathscr{F}$ satisfying* Eq. (6.1) *and*

$$\theta^* f(t)^k = kf(t)^{k-1}$$

for all $k \geqslant 0$.

Proof. Suppose θ_f is the umbral shift for $f(t)$, with associated sequence $p_n(x)$. Then

$$\begin{aligned}\langle \theta_f^* f(t)^k \mid p_n(x)\rangle &= \langle f(t)^k \mid \theta p_n(x)\rangle \\ &= \frac{(n+1)\,c_n}{c_{n+1}} \langle f(t)^k \mid p_{n+1}(x)\rangle \\ &= kc_n \delta_{n+1,k} \\ &= \langle kf(t)^{k-1} \mid p_n(x)\rangle\end{aligned}$$

and so $\theta_f^* f(t)^k = kf(t)^{k-1}$. From Proposition 6.1 we conclude that θ_f^* is a derivation on $\mathscr{F}$. For the converse, let ω be a derivation on $\mathscr{F}$ for which $\omega f(t)^k = kf(t)^{k-1}$. Then if $p_n(x)$ is associated to $f(t)$, we have

$$\begin{aligned}\langle \omega f(t)^k \mid p_n(x)\rangle &= \langle kf(t)^{k-1} \mid p_n(x)\rangle \\ &= kc_n \delta_{n,k-1} \\ &= \langle f(t)^k \mid \theta_f p_n(x)\rangle \\ &= \langle \theta_f^* f(t)^k \mid p_n(x)\rangle.\end{aligned}$$

Thus since ω satisfies (6.1), we conclude $\omega = \theta_f^*$.

Notice that $\theta_t^* t^k = kt^{k-1}$ and so θ_t^* is the derivative with respect to t. That is,

$$\theta_t^* g(t) = g'(t),$$

and

$$\langle g(t) \mid \theta_t x^n\rangle = \langle g'(t) \mid x^n\rangle.$$

We also observe that

$$\theta_t t x^n = \frac{c_n}{c_{n-1}} \theta_t x^{n-1}$$
$$= n x^n$$

and so

$$\theta_t t = xD,$$

where D is the derivative with respect to x on P.

Next we derive the chain rule for derivations.

PROPOSITION 6.3. *If $f(t)$ and $g(t)$ are delta series, then*

$$\theta_f^* = (\theta_f^* g(t))\, \theta_g^*$$

Proof. Since θ_f^* is a derivation,

$$\theta_f^* g(t)^k = k g(t)^{k-1}\, \theta_f^* g(t)$$
$$= (\theta_f^* g(t))\, \theta_g^* g(t)^k$$

and Proposition 6.1 completes the proof.

Now we can relate two umbral shifts.

THEOREM 6.2. *If θ_f and θ_g are umbral shifts, then*

$$\theta_f = \theta_g \circ (\theta_g^* f(t))^{-1}.$$

Proof. For any $p(x)$ in P,

$$\langle t^k \mid \theta_f p(x)\rangle = \langle \theta_f^* t^k \mid p(x)\rangle$$
$$= \langle (\theta_g^* f(t))^{-1} (\theta_g^* t^k) \mid p(x)\rangle$$
$$= \langle \theta_g^* t^k \mid (\theta_g^* f(t))^{-1}\, p(x)\rangle$$
$$= \langle t^k \mid \theta_g \circ (\theta_g^* f(t))^{-1}\, p(x)\rangle$$

from which the result follows.

From this theorem we obtain our first recurrence formula.

THEOREM 6.3. *If $p_n(x)$ is associated to $f(t)$, then*

$$p_{n+1}(x) = \frac{c_{n+1}}{(n+1)\, c_n}\, \theta_t (f'(t))^{-1}\, p_n(x), \tag{6.2}$$

where $f'(t)$ is the ordinary derivative of $f(t)$ with respect to t and $\theta_t: x^n \to ((n+1)c_n/c_{n+1})\, x^{n+1}$.

Proof. The result follows from Theorem 6.2 by taking $g(t) = t$ and applying to $p_n(x)$.

We wish to derive a recurrence formula for Sheffer sequences. To this end we derive a formula for the adjoint θ_f^* of an umbıal shift.

PROPOSITION 6.4. *Let θ_f be the umbral shift for $f(t)$. Then*

$$\theta_f^* h(t) = h(t)\,\theta_f - \theta_f h(t)$$

for all $h(t)$ in $\mathscr{F}$.

Proof. For any $g(t)$ in $\mathscr{F}$ and $p(x)$ in P we have

$$\begin{aligned}\langle g(t) \mid \theta_f^*(h(t))\, p(x)\rangle &= \langle \theta_f^*(h(t))\, g(t) \mid p(x)\rangle \\ &= \langle \theta_f^*(h(t)\, g(t)) - h(t)\,\theta_f^*(g(t)) \mid p(x)\rangle \\ &= \langle g(t) \mid (h(t)\,\theta_f - \theta_f h(t))\, p(x)\rangle.\end{aligned}$$

PROPOSITION 6.5. *Let $s_n(x)$ be Sheffer for $(g(t), f(t))$. Then if θ_f is the umbral shift for $f(t)$,*

$$s_{n+1}(x) = \frac{c_{n+1}}{(n+1)\,c_n}\,(g(t)\,\theta_f^*\, g^{-1}(t) + \theta_f)\, s_n(x).$$

Proof. Let $p_n(x)$ be the associated sequence for $f(t)$. Then

$$\begin{aligned}s_{n+1}(x) &= g^{-1}(t)\, p_{n+1}(x) \\ &= \frac{c_{n+1}}{(n+1)\,c_n}\, g^{-1}(t)\,\theta_f\, s_n(x) \\ &= \frac{c_{n+1}}{(n+1)\,c_n}\, g^{-1}(t)\,\theta_f\, g(t)\, s_n(x).\end{aligned}$$

But

$$\begin{aligned}g^{-1}(t)\,\theta_f\, g(t) &= (g^{-1}(t)\,\theta_f - \theta_f\, g^{-1}(t))\, g(t) + \theta_f \\ &= g(t)\,\theta_f^*(g^{-1}(t)) + \theta_f.\end{aligned}$$

This completes the proof.

Now we have our recurrence formula.

THEOREM 6.4. *Let $s_n(x)$ be the Sheffer sequence for $(g(t), f(t))$. Then*

$$s_{n+1}(x) = \frac{c_{n+1}}{(n+1)\,c_n}\left(\theta_t - \frac{g'(t)}{g(t)}\right)\frac{1}{f'(t)}\, s_n(x). \tag{6.3}$$

Proof. This follows from Proposition 6.5 by noticing that

$$\begin{aligned}\theta_f^* g^{-1}(t) &= (\theta_f^* t)\, \theta_t^* g^{-1}(t)\\ &= (f'(t))^{-1}(-g^{-2}(t)\, g'(t))\end{aligned}$$

and

$$\begin{aligned}\theta_f &= \theta_t(\theta_t^* f(t))^{-1}\\ &= \theta_t(f'(t))^{-1}.\end{aligned}$$

THEOREM 6.5. *Let $s_n(x)$ be Sheffer for $(g(t), f(t))$. If*

$$\begin{aligned}T &= \left(\theta_t - \frac{g'(t)}{g(t)}\right)\frac{f(t)}{f'(t)}\\ &= \left(xD - \frac{tg'(t)}{g(t)}\right)\frac{f(t)}{tf'(t)},\end{aligned}$$

then

$$Ts_n(x) = ns_n(x). \tag{6.4}$$

In other words, $s_n(x)$ is an eigenfunction for T with eigenvalue n.

Proof. This follows from Theorem 6.4 by noticing that

$$\frac{c_{n+1}}{(n+1)\, c_n} s_n(x) = f(t)\, s_{n+1}(x).$$

The two forms of T are equivalent since $\theta_t t = xD$.

We remark that since $s_n(x)$ forms a basis for P, any polynomial solution to $Tp(x) = np(x)$ is a constant multiple of $s_n(x)$.

Next we require two lemmas.

LEMMA 1. *Let $h(t)$ be invertible, with leading coefficient equal to 1. Then the equation*

$$h(t) = \frac{g'(t)}{g(t)}$$

has a unique (up to multiplicative constant) solution given by

$$g(t) = \exp\left(\int h(t)\, dt\right).$$

LEMMA 2. *Let $l(t)$ be a delta series, with leading coefficient equal to 1. Then the equation*

$$l(t)=\frac{f(t)}{f'(t)}$$

has a unique (up to multiplicative constant) solution given by

$$f(t)=t\exp\left(\int\frac{(l(t)/t)^{-1}-1}{t}\,dt\right).$$

THEOREM 6.6. *Let T be a linear operator of the form*

$$\begin{aligned}T&=(\theta_t-h(t))\,l(t)\\&=(xD-th(t))\frac{l(t)}{t}\end{aligned}$$

where $h(t)$ is invertible and $l(t)$ is a delta series both having leading coefficient equal to 1. *Then a solution to the equation*

$$Ts_n(x)=ns_n(x)$$

is given by nth polynomial in the Sheffer sequence for the pair

$$\left(\exp\left(\int h(t)\,dt\right), t\exp\left(\int\frac{(l(t)/t)^{-1}-1}{t}\,dt\right)\right).$$

Proof. A solution to $Ts_n(x)=ns_n(x)$ is given by the Sheffer sequence for the pair $(g(t), f(t))$ where

$$\frac{g'(t)}{g(t)}=h(t),$$

$$\frac{f(t)}{f'(t)}=l(t).$$

The result follows from the lemmas.

We shall now derive another set of recurrence formulas.

THEOREM 6.7. *Let $s_n(x)$ be Sheffer for $(g(t), f(t))$. Suppose*

$$\bar{f}(t)=\sum_{k=1}^{\infty}\frac{a_k}{kc_k}t^k,$$

$$\left(\frac{\bar{f}(t)}{t}\right)^{-1}\bar{f}'(t)=\sum_{k=0}^{\infty}\frac{b_k}{c_k}t^k,$$

$$\frac{t[g(\bar{f}(t))]'}{g(\bar{f}(t))}=\sum_{k=0}^{\infty}\frac{d_k}{c_k}t^k \qquad [d_0=0].$$

Then

$$ns_n(x) = \sum_{k=1}^{n} \frac{c_n}{c_k c_{n-k}} (a_k \theta_t - d_k) s_{n-k}(x) \tag{6.5}$$

and

$$ns_n(x) = \sum_{k=0}^{n} \frac{c_n}{c_k c_{n-k}} (b_k xD - d_k) s_{n-k}(x) \qquad [d_0 = 0] \tag{6.6}$$

Proof. To prove (6.5) we have

$$ns_n(x) = \left(\theta_t \frac{f(t)}{f'(t)} - \frac{g'(t) f(t)}{g(t) f'(t)}\right) s_n(x).$$

But

$$\begin{aligned} \frac{f(t)}{f'(t)} &= f(t) \bar{f}'(f(t)) \\ &= t\bar{f}'(t)\,|_{t=f(t)} \\ &= \sum_{k=1}^{\infty} \frac{a_k}{c_k} f(t)^k \end{aligned}$$

and

$$\begin{aligned} \frac{g'(t) f(t)}{g(t) f'(t)} &= \frac{g'(\bar{f}(t))\, t\bar{f}'(t)}{g(\bar{f}(t))} \Big|_{t=f(t)} \\ &= \frac{[g(\bar{f}(t))]}{g(\bar{f}(t))} \Big|_{t=f(t)} \\ &= \sum_{k=0}^{\infty} \frac{d_k}{c_k} f(t)^k. \end{aligned}$$

The result follows by substitution and $f(t)^k s_n(x) = (c_n/c_{n-k}) s_{n-k}(x)$. To prove (6.6) we proceed in the same manner using

$$\begin{aligned} \frac{f(t)}{tf'(t)} &= \frac{f(t)}{t} \bar{f}'(f(t)) \\ &= \frac{f(t)}{\bar{f}(f(t))} \bar{f}'(f(t)) \end{aligned}$$

$$= \left(\frac{\bar{f}(t)}{t}\right)^{-1} \bar{f}'(t)\Big|_{t=f(t)}$$

$$= \sum_{k=0}^{\infty} \frac{b_k}{c_k} f(t)^k.$$

Theorem 6.7 has a sort of converse.

THEOREM 6.8. *Let $r_n(x)$ be a sequence of polynomials satisfying* (6.5). *Let*

$$l(t) = \sum_{k=1}^{\infty} \frac{a_k}{kc_k} t^k$$

and

$$th(t) = \sum_{k=0}^{\infty} \frac{d_k}{c_k} t^k.$$

Then there is a constant a for which $ar_n(x)$ is Sheffer for the pair

$$\left(\exp\left(\int h(\bar{l}(t))\, d\bar{l}(t)\right), \bar{l}(t)\right).$$

Proof. First we notice that solutions to (6.5) are unique up to multiplicative constant. Now suppose $s_n(x)$ is the Sheffer sequence for the pair $(g(t), f(t))$, where

$$\bar{f}(t) = l(t)$$

and

$$\frac{t[g(\bar{f}(t))]'}{g(\bar{f}(t))} = th(t).$$

Then in view of Theorem 6.7 the sequence $s_n(x)$ satisfies (6.5). Hence $s_n(x) = \alpha r_n(x)$ for some constant α. It remains only to solve the above equations for $f(t)$ and $g(t)$ using the previous lemmas.

In order to complete the converse to Theorem 6.7 we need one more lemma.

LEMMA 3. *Let $h(t)$ be invertible. Then there is a unique (up to multiplicative constant) solution to the equation*

$$l(t) = \left(\frac{\bar{f}(t)}{t}\right)^{-1} \bar{f}'(t)$$

given by

$$\bar{f}(t) = t \exp \left(\int \frac{l(t) - 1}{t} dt \right).$$

THEOREM 6.9. *Let $r_n(x)$ be a sequence of polynomials satisfying* (6.6). *Let*

$$l(t) = \sum_{k=0}^{\infty} \frac{b_k}{c_k} t^k$$

and

$$th(t) = \sum_{k=0}^{\infty} \frac{d_k}{c_k} t^k.$$

Then there is a constant a for which $ar_n(x)$ is Sheffer for the pair $(g(t), f(t))$ where

$$\bar{f}(t) = t \exp \left(\int \frac{l(t) - 1}{t} dt \right),$$

$$g(t) = \exp \left(\int h(\bar{l}(t)) \, d\bar{l}(t) \right).$$

Proof. The solution to (6.6) is unique up to the multiplicative constant. Now let $s_n(x)$ be the Sheffer sequence for $(g(t), f(t))$ where

$$\left(\frac{\bar{f}(t)}{t} \right)^{-1} \bar{f}'(t) = l(t)$$

and

$$\frac{t[g(\bar{f}(t))]'}{g(\bar{f}(t))} = th(t).$$

Then by Theorem 6.7 the sequence $s_n(x)$ satisfies (6.6). Hence $s_n(x) = \alpha r_n(x)$ for some constant α. It remains only to solve for $f(t)$ and $g(t)$ using the previous lemmas.

7. TRANSFER FORMULAS

In this section we develop formulas for the direct computation of associated sequences. Using these formulas and Theorem 5.5 we can compute Sheffer sequences.

THEOREM 7.1 (The Transfer Formula). *If $p_n(x)$ is the associated sequence for $f(t)$, then*

$$p_n(x) = f'(t)\left(\frac{f(t)}{t}\right)^{-n-1} x^n \tag{7.1}$$

for all $n \geqslant 0$. As usual $f'(t)$ is the derivative of $f(t)$ with respect to t.

Proof. We check the conditions of Theorem 5.6 for the sequence $p_n(x)$ in (7.1). For condition (i) we have

$$\begin{aligned}
&\left\langle t^0 \,\middle|\, f'(t)\left(\frac{f(t)}{t}\right)^{-n-1} x^n\right\rangle \\
&\quad = \left\langle \left(\frac{f(t)}{t}\,t\right)'\left(\frac{f(t)}{t}\right)^{-n-1} \,\middle|\, x^n\right\rangle \\
&\quad = \left\langle \left(\frac{f(t)}{t}\right)^{-n} + t\left(\frac{f(t)}{t}\right)'\left(\frac{f(t)}{t}\right)^{-n-1} \,\middle|\, x^n\right\rangle.
\end{aligned}$$

For $n = 0$, we get

$$\begin{aligned}
\left\langle t^0 + t\left(\frac{f(t)}{t}\right)'\left(\frac{f(t)}{t}\right)^{-1} \,\middle|\, 1\right\rangle &= \langle t^0 \mid 1\rangle \\
&= c_0.
\end{aligned}$$

For $n > 0$, we get

$$\begin{aligned}
&\left\langle \left(\frac{f(t)}{t}\right)^{-n} \,\middle|\, x^n\right\rangle + \left\langle t\left(\frac{f(t)}{t}\right)'\left(\frac{f(t)}{t}\right)^{-n-1} \,\middle|\, x^n\right\rangle \\
&\quad = \left\langle \left(\frac{f(t)}{t}\right)^{-n} \,\middle|\, x^n\right\rangle - \frac{1}{n}\left\langle \left[\left(\frac{f(t)}{t}\right)^{-n}\right]' \,\middle|\, tx^n\right\rangle \\
&\quad = \left\langle \left(\frac{f(t)}{t}\right)^{-n} \,\middle|\, x^n\right\rangle - \frac{1}{n}\frac{c_n}{c_{n-1}}\left\langle \left[\left(\frac{f(t)}{t}\right)^{-n}\right]' \,\middle|\, x^{n-1}\right\rangle \\
&\quad = \left\langle \left(\frac{f(t)}{t}\right)^{-n} \,\middle|\, x^n\right\rangle - \frac{1}{n}\frac{c_n}{c_{n-1}}\left\langle \theta_t^*\left(\frac{f(t)}{t}\right)^{-n} \,\middle|\, x^{n-1}\right\rangle \\
&\quad = \left\langle \left(\frac{f(t)}{t}\right)^{-n} \,\middle|\, x^n\right\rangle - \frac{1}{n}\frac{c_n}{c_{n-1}}\left\langle \left(\frac{f(t)}{t}\right)^{-n} \,\middle|\, \theta_t x^{n-1}\right\rangle \\
&\quad = \left\langle \left(\frac{f(t)}{t}\right)^{-n} \,\middle|\, x^n\right\rangle - \frac{1}{n}\frac{c_n}{c_{n-1}}\left\langle \left(\frac{f(t)}{t}\right)^{-n} \,\middle|\, \frac{nc_{n-1}}{c_n}x^n\right\rangle \\
&\quad = 0.
\end{aligned}$$

As for (ii) we have

$$f(t)\, p_n(x) = t\,\frac{f(t)}{t}\, f'(t) \left(\frac{f(t)}{t}\right)^{-n-1} x^n$$

$$= \frac{c_n}{c_{n-1}}\, f'(t) \left(\frac{f(t)}{t}\right)^{-n} x^{n-1}$$

$$= \frac{c_n}{c_{n-1}}\, p_{n-1}(x).$$

This completes the proof.

An alternative form of the Transfer Formula can be derived.

THEOREM 7.2 (The Transer Formula). *If $p_n(x)$ is the associated sequence for $f(t)$, then*

$$p_n(x) = \frac{c_n}{nc_{n-1}}\, \theta_t \left(\frac{f(t)}{t}\right)^{-n} x^{n-1} \tag{7.2}$$

for all $n \geqslant 1$.

Proof. In view of Proposition 6.4 we have

$$\begin{aligned}
p_n(x) &= f'(t) \left(\frac{f(t)}{t}\right)^{-n-1} x^n \\
&= \left(t\frac{f(t)}{t}\right)' \left(\frac{f(t)}{t}\right)^{-n-1} x^n \\
&= \left(\frac{f(t)}{t} + t\left(\frac{f(t)}{t}\right)'\right)\left(\frac{f(t)}{t}\right)^{-n-1} x^n \\
&= \left(\frac{f(t)}{t}\right)^{-n} x^n + \frac{c_n}{c_{n-1}} \left(\frac{f(t)}{t}\right)^{-n-1} \left(\frac{f(t)}{t}\right)' x^{n-1} \\
&= \left(\frac{f(t)}{t}\right)^{-n} x^n - \frac{c_n}{nc_{n-1}} \left(\left(\frac{f(t)}{t}\right)^{-n}\right)' x^{n-1} \\
&= \left(\frac{f(t)}{t}\right)^{-n} x^n - \frac{c_n}{nc_{n-1}} \left[\left(\frac{f(t)}{t}\right)^{-n} \theta_t - \theta_t \left(\frac{f(t)}{t}\right)^{-n}\right] x^n \\
&= \frac{c_n}{nc_{n-1}}\, \theta_t \left(\frac{f(t)}{t}\right)^{-n} x^{n-1}.
\end{aligned}$$

We can use the Transfer Formula to relate any two associated sequences.

COROLLARY 1. *Let $p_n(x)$ be associated to $f(t)$ and $q_n(x)$ be associated to $g(t)\,f(t)$, where $g(t)$ is any invertible series. Then*

$$q_n(x) = \theta_t\, g(t)^{-n}\, \theta_t^{-1} p_n(x),$$

where $\theta_t^{-1} x^{n+1} = (c_{n+1}/(n+1)\,c_n)\,x^n$ and $\theta_t^{-1} 1 = 0$.

The Transfer Formula readily gives Lagrange's formula for the compositional inverse of a delta series. To see this we take $y = 0$ in the generating function of a Sheffer sequence,

$$\frac{1}{g(\bar{f}(t))} = \sum_{k=0}^{\infty} \frac{c_0 s_k(0)}{c_k}\, t^k,$$

since $\varepsilon_0(\bar{f}(t)) = c_0^{-1}$. Now by the Transfer Formula

$$\begin{aligned} c_0 s_k(0) &= \left\langle t^0 \,\middle|\, g^{-1}(t)\, f'(t) \left(\frac{f(t)}{t}\right)^{-k-1} x^k \right\rangle \\ &= \left\langle g^{-1}(t)\, f'(t) \left(\frac{f(t)}{t}\right)^{-k-1} \,\middle|\, x^k \right\rangle. \end{aligned}$$

Thus the coefficient of t^k/c_k in $g^{-1}(\bar{f}(t))$ equals c_k times the coefficient of t^k in $g^{-1}(t)\, f'(t) (f(t)/t)^{-k-1}$. Other versions of Lagrange's formula are similarly derived.

The next proposition is an application of the Transfer Formula.

PROPOSITION 7.1. *Let $s_n(x)$ be Sheffer for $(g(t), f(t))$. Let $h(t)$ and $l(t)$ be invertible. Then the sequence*

$$r_n(x) = h(t)\, l(t)^n\, s_n(x)$$

is Sheffer for

$$\left(\frac{[l^{-1}(t)\, f(t)]'}{f'(t)\, h(t)}\, l(t)\, g(t),\, l^{-1}(t)\, f(t)\right).$$

Proof. First we have

$$\begin{aligned} l^{-1}(t)\, f(t)\, r_n(x) &= f(t)\, h(t)\, l(t)^{n-1}\, s_n(x) \\ &= \frac{c_n}{c_{n-1}}\, h(t)\, l(t)^{n-1}\, s_{n-1}(x) \\ &= \frac{c_n}{c_{n-1}}\, r_{n-1}(x). \end{aligned}$$

Also,

$$\begin{aligned}
&\frac{[l^{-1}(t)\,f(t)]'}{f'(t)\,h(t)}\,l(t)\,g(t)\,r_n(x)\\
&\quad=\frac{[l^{-1}(t)\,f(t)]'}{f'(t)}\,l(t)^{n+1}\,g(t)\,s_n(x)\\
&\quad=\frac{[l^{-1}(t)\,f(t)]'}{f'(t)}\,l(t)^{n+1}\,f'(t)\left(\frac{f(t)}{t}\right)^{-n-1}x^n\\
&\quad=[l^{-1}(t)\,f(t)]'\left(\frac{l^{-1}(t)\,f(t)}{t}\right)^{-n-1}x^n
\end{aligned}$$

is the associated sequence for $l^{-1}(t)\,f(t)$, and the proof is complete.

8. Umbral Composition and Transfer Operators

Let $p_n(x)$ be associated to $f(t)$. The *transfer operator for* $p_n(x)$ or $f(t)$ is the linear operator λ_f on P defined by

$$\lambda_f x^n = p_n(x).$$

We have

$$\lambda_f^* g(t) = \sum_{k=0}^{\infty} \frac{\langle g(t) \mid p_k(x)\rangle}{c_k}\, t^k. \tag{8.1}$$

We can characterize transfer operators by their adjoints.

THEOREM 8.1. *A linear operator λ on P is the transfer operator for $f(t)$ if and only if its adjoint λ^* is an automorphism of $\mathcal{F}$ satisfying* (6.1) *and for which*

$$\lambda^* f(t) = t.$$

Proof. Suppose λ_f is the transfer operator for $f(t)$, with associated sequence $p_n(x)$. By Theorem 5.9 we have

$$\begin{aligned}
\langle \lambda_f^* h(t)\, g(t) \mid x^n\rangle &= \langle h(t)\, g(t) \mid p_n(x)\rangle\\
&= \sum_{k=0}^{n} \frac{c_n}{c_k c_{n-k}} \langle h(t) \mid p_k(x)\rangle\langle g(t) \mid p_{n-k}(x)\rangle\\
&= \sum_{k=0}^{n} \frac{c_n}{c_k c_{n-k}} \langle \lambda_f^* h(t) \mid x^k\rangle\langle \lambda_f^* g(t) \mid x^{n-k}\rangle\\
&= \langle \lambda_f^* h(t)\, \lambda_f^* g(t) \mid x_n\rangle
\end{aligned}$$

and so $\lambda_f^* h(t)\, g(t) = \lambda_f^* h(t)\, \lambda_f^* g(t)$. Also, from (8.1) we see that $\lambda_f^* f(t) = t$. For the converse, suppose ω is an automorphism of $\mathscr{F}$ satisfying (6.1) and for which $\omega f(t) = t$. Then if λ_f is the transfer operator for $f(t)$, with associated sequence $p_n(x)$ we have

$$\begin{aligned}
\langle \lambda_f^* f(t)^k \mid x^n \rangle &= \langle f(t)^k \mid \lambda x^n \rangle \\
&= \langle f(t)^k \mid p_n(x) \rangle \\
&= c_n \delta_{n,k} \\
&= \langle t^k \mid x^n \rangle \\
&= \langle \omega f(t)^k \mid x^n \rangle
\end{aligned}$$

and so $\omega = \lambda_f^*$.

The most important properties of transfer operators are contained in the next result.

THEOREM 8.2. (a) *A transfer operator maps associated sequences to associated sequences.*

(b) *If λ: $p_n(x) \to q_n(x)$ is a linear operator where $p_n(x)$ is associated to $f(t)$ and $q_n(x)$ is associated to $g(t)$, then $\lambda^* g(t) = f(t)$.*

(c) *If $p_n(x)$ and $q_n(x)$ are associated sequences and $\lambda p_n(x) = q_n(x)$, then λ is a transfer operator.*

Proof. (a) Let $\lambda: x^n \to p_n(x)$ be a transfer operator and let $q_n(x)$ be associated to $g(t)$. Then

$$\begin{aligned}
\langle \lambda^{*-1} g(t)^k \mid \lambda q_n(x) \rangle &= \langle g(t) \mid q_n(x) \rangle \\
&= c_n \delta_{n,k}
\end{aligned}$$

and so $\lambda q_n(x)$ is the associated sequence for $\lambda^{*-1} g(t)$.

(b) We have $\langle \lambda^* g(t) \mid p_n(x) \rangle = \langle g(t) \mid \lambda p_n(x) \rangle = \langle g(t) \mid q_n(x) \rangle = c_n \delta_{n,1} = \langle f(t) \mid p_n(x) \rangle$ and so $\lambda^* g(t) = f(t)$.

(c) Suppose $p_n(x)$ is associated to $f(t)$, and $q_n(x)$ is associated to $g(t)$. Then by (b), $\lambda^* g(t) = f(t)$ and $\lambda^* g(\bar{f}(t))^k = t^k$. Hence $\langle g(\bar{f}(t))^k \mid \lambda x^n \rangle = \langle t^k \mid x^n \rangle = c_n \delta_{n,k}$ and so λx^n is associated to $g(\bar{f}(t))$.

Suppose $p_n(x)$ and $q_n(x)$ are two sequences of polynomials with $q_n(x) = \sum_{k=0}^n q_{n,k} x^k$. Then the *umbral composition* of $q_n(x)$ with $p_n(x)$ is the sequence

$$q_n(\underline{p}(x)) = \sum_{k=0}^{n} q_{n,k} p_k(x).$$

Notice that if $\lambda: x^n \to p_n(x)$, then

$$q_n(\underline{p}(x)) = \lambda q_n(x).$$

THEOREM 8.3. *Let $p_n(x)$ be associated to $f(t)$ and $q_n(x)$ be associated to $g(t)$. Then $q_n(\underline{p}(x))$ is associated to $g(f(t))$.*

Proof. Let $\lambda: x^n \to p_n(x)$ be the transfer operator for $f(t)$. Then by the proof of part (a) of Theorem 8.2, $\lambda q_n(x)$ is the associated sequence for $\lambda^{*-1} g(t)$. But $\lambda q_n(x) = q_n(\underline{p}(x))$ and, by part (b) of Theorem 8.2, $\lambda^{*-1} g(t) = g(\lambda^{*-1}(t)) = g(f(t))$.

We would like to extend this result to Sheffer sequences.

THEOREM 8.4. *Let $s_n(x)$ be Sheffer for $(g(t), f(t))$ and let $r_n(x)$ be Sheffer for $(h(t), l(t))$. Then $r_n(\underline{s}(x))$ is Sheffer for the pair*

$$(g(t)\, h(f(t)), l(f(t))).$$

Proof. Let $\lambda_f: x_n \to p_n(x)$ be the transfer operator associated to $f(t)$ and let $\mu: x^n \to s_n(x)$. Thus we have $\mu = g^{-1}(t)\, \lambda_f$. Now

$$\begin{aligned}
&\langle g(t)\, h(f(t))\, l(f(t))^k \mid r_n(\underline{s}(x))\rangle \\
&\quad = \langle g(t)\, h(f(t))\, l(f(t))^k \mid \mu r_n(x)\rangle \\
&\quad = \langle g(t)\, h(f(t))\, l(f(t))^k \mid g^{-1}(t)\, \lambda_f r_n(x)\rangle \\
&\quad = \langle \lambda_f^*(h(f(t))\, l(f(t))^k) \mid r_n(x)\rangle \\
&\quad = \langle h(t)\, l(t)^k \mid r_n(x)\rangle \\
&\quad = c_n \delta_{n,k}.
\end{aligned}$$

This completes the proof.

Suppose $s_n(x)$ and $r_n(x)$ are two sequences of polynomials related by

$$r_n(x) = \sum_{k=0}^{n} a_{n,k} s_k(x).$$

The *connection-constants problem is to determine the constants $a_{n,k}$*. In case $s_n(x)$ and $r_n(x)$ are Sheffer sequences we can give a solution to this problem.

THEOREM 8.5. *Let $s_n(x)$ be Sheffer for $(g(t), f(t))$ and let $r_n(x)$ be Sheffer for $(h(t), l(t))$. Suppose*

$$r_n(x) = \sum_{k=0}^{n} a_{n,k} s_k(x). \tag{8.2}$$

Then the sequence

$$t_n(x) = \sum_{k=0}^{n} a_{n,k} x^k$$

is the Sheffer sequence for the pair

$$\left(\frac{h(\bar{f}(t))}{g(\bar{f}(t))}, l(\bar{f}(t))\right).$$

Proof. Equation (8.2) can be written as

$$r_n(x) = t_n(s(x)).$$

If $t_n(x)$ is Sheffer for the pair $(X(t), Y(t))$, then by Theorem 8.4 we have

$$h(t) = g(t)\, X(f(t)),$$
$$l(t) = Y(f(t))$$

and so

$$Y(t) = l(\bar{f}(t)),$$
$$X(t) = \frac{h(\bar{f}(t))}{g(\bar{f}(t))}.$$

COROLLARY 1. *If $p_n(x)$ is associated to $f(t)$ and $q_n(x)$ is associated to $l(t)$ and*

$$q_n(x) = \sum_{k=0}^{n} a_{n,k} p_k(x),$$

then $t_n(x) = \sum_{k=0}^{n} a_{n,k} x^k$ is associated to $l(\bar{f}(t))$.

9. Examples: Gegenbauer, Chebyshev and Others

In this section we study the delta series

$$f(t) = \frac{\sqrt{1-t^2} - 1}{t}$$

This will lead us to the Gegenbauer and Chebyshev polynomials.

Associated Sequence

We begin by computing

$$f(t)=\frac{-t}{1+\sqrt{1-t^2}},$$

$$\bar{f}(t)=\frac{-2t}{1+t^2},$$

$$f'(t)=\frac{-1}{\sqrt{1-t^2}(1+\sqrt{1-t^2})}$$

$$=\frac{f(t)}{t\sqrt{1-t^2}}.$$

We shall denote the associated sequence for $f(t)$ by $p_n(x)$. The generating function for $p_n(x)$ is

$$\varepsilon_y\left(\frac{-2t}{1+t^2}\right)=\sum_{k=0}^{\infty}\frac{p_k(y)}{c_k}t^k.$$

The conjugate representation for $p_n(x)$ is

$$p_n(x)=\sum_{k=0}^{n}\frac{\langle(-2t/(1+t^2))^k \mid x^n\rangle}{c_k}x^k.$$

Now

$$\left\langle\left(\frac{-2t}{1+t^2}\right)^k \middle| x^n\right\rangle=(-2)^k\sum_{j=0}^{k}\binom{-k}{j}\langle t^{2j+k} \mid x^n\rangle$$

$$=(-2)^k\sum_{j=0}^{k}\binom{-k}{j}c_n\delta_{n,2j+k}$$

$$=\begin{cases}(-2)^{n-2j}\dbinom{2j-n}{j}c_n & \text{if } k=n-2j\\ 0 & \text{if } k\neq n-2j\end{cases}$$

and so

$$p_n(x)=\sum_{j=0}^{[n/2]}\frac{c_n}{c_{n-2j}}\binom{2j-n}{j}(-2x)^{n-2j}.$$

Let us give some consequences of the Expansion Theorem. We have

$$t^n=\sum_{k=n}^{\infty}\frac{\langle t^n \mid p_k(x)\rangle}{c_k}f(t)^k.$$

But

$$\langle t^n \mid p_k(x)\rangle = 0 \qquad \text{if} \quad n+k \text{ odd}$$

and

$$\langle t^n \mid p_{n+2j}(x)\rangle = (-2)^n \binom{-n}{j} c_{n+2j}$$

and so

$$t^n = \sum_{j=0}^{\infty} (-2)^n \binom{-n}{j} \left(\frac{-t}{1+\sqrt{1-t^2}}\right)^{n+2j}$$

or

$$(1+\sqrt{1-t^2})^n = \sum_{j=0}^{\infty} (-2)^n \binom{-n}{j} \left(\frac{-t}{1+\sqrt{1-t^2}}\right)^{2j}.$$

From Rainville, [4, p. 70] one can deduce

$$(1+\sqrt{1-z})^{-\alpha} = \sum_{k=0}^{\infty} \frac{\alpha(\alpha+2k-1)_{k-1}}{2^{\alpha+2k}k!} z^k \tag{9.1}$$

for all α. Hence

$$f(t)^k = \sum_{j=0}^{\infty} \frac{(-1)^k k(k+2j-1)_{j-1}}{2^{k+2j}j!} t^{2j+k}. \tag{9.2}$$

Thus we have

$$x^n = \sum_{k=0}^{n} \frac{\langle f(t)^k \mid x^n\rangle}{c_k} p_k(x),$$

where by (9.2)

$$\langle f(t)^k \mid x^n\rangle = 0 \qquad \text{if} \quad n-k \text{ is odd}$$

and

$$\langle f(t)^{n-2j} \mid x^n\rangle = \frac{(-1)^n(n-2j)(n-1)_{j-1}}{2^n j!} c_n$$

and so

$$x^n = \sum_{j=0}^{[n/2]} \frac{(-1)^n(n-2j)(n-1)_{j-1}}{2^n j!} \frac{c_n}{c_{n-2j}} p_{n-2j}(x).$$

Now we turn to some recurrence formulas. From the fact that $f(t)\, p_n(x) = (c_n/c_{n-1})\, p_{n-1}(x)$ we have

$$\sqrt{1-t^2}\, p_n(x) = (tf(t)+1)\, p_n(x)$$
$$= \frac{c_n}{c_{n-1}}\, tp_{n-1}(x) + p_n(x) \tag{9.3}$$

and

$$\sqrt{1-t^2}\, p_n(x) = (1+\sqrt{1-t^2})\, p_n(x) - p_n(x)$$
$$= \frac{c_n}{c_{n+1}}\,(1+\sqrt{1-t^2})\, f(t)\, p_{n+1}(x) - p_n(x)$$
$$= -\frac{c_n}{c_{n+1}}\, tp_{n+1}(x) - p_n(x). \tag{9.4}$$

Equating the two expressions gives the recurrence

$$\frac{c_n}{c_{n+1}}\, tp_{n+1}(x) + \frac{c_n}{c_{n-1}}\, tp_{n-1}(x) + 2p_n(x) = 0 \tag{9.5}$$

Notice that (9.5) holds for any Sheffer sequence using $f(t)$ as its delta series.

We may obtain another recurrence from (6.4). We observe first that

$$\frac{f(t)}{tf'(t)} = \sqrt{1-t^2}$$

and so

$$np_n(x) = xD\,\sqrt{1-t^2}\, p_n(x)$$

and using (9.3) gives

$$(xD-n)\, p_n(x) + \frac{c_n}{c_{n-1}}\, x\,Dt\, p_{n-1}(x) = 0. \tag{9.6}$$

Also, Eq. (6.6) becomes

$$(n-xD)\, p_n(x) = 2xD \sum_{k=1}^{[n/2]} \frac{c_n}{c_{2k}c_{n-2k}}\,(-1)^k\, p_{n-2k}(x).$$

Of course, Eqs. (9.5) and (9.6) may be used to derive other recurrences.

Now let us turn our attention to a specific form for c_n, namely,

$$c_n = \frac{1}{\binom{-\lambda}{n}},$$

where $-\lambda$ is not a non-negative integer. Let us collect some preliminary results. We have

$$\frac{c_n}{c_{n-1}} = \frac{n}{-\lambda - n + 1}$$

and

$$tx^n = \frac{n}{-\lambda - n + 1} x^{n-1}$$

One can easily verify that

$$t = -(\lambda + xD)^{-1} D. \tag{9.7}$$

Also,

$$\theta_t x^n = -(\lambda + n) x^{n+1}$$

and so

$$\theta_t = -x(\lambda + xD).$$

Next we have

$$\begin{aligned} \varepsilon_y(t) &= \sum_{k=0}^{\infty} \binom{-\lambda}{k} y^k x^k \\ &= (1 + yt)^{-\lambda}. \end{aligned}$$

The generating function now becomes

$$\left(1 - \frac{2yt}{1+t^2}\right)^{-\lambda} = \sum_{k=0}^{\infty} \binom{-\lambda}{n} p_k(y)\, t^k$$

or

$$(1 + t^2)^{\lambda}(1 - 2yt + t^2)^{-\lambda} = \sum_{k=0}^{\infty} \binom{-\lambda}{n} p_k(y)\, t^k.$$

[Note the similarity with the Gegenbauer generating function.] The conjugate representation yields

$$p_n(x) = \sum_{j=0}^{[n/2]} \frac{\binom{-\lambda}{n-2j}\binom{2j-n}{j}}{\binom{-\lambda}{n}} (-2x)^{n-2j}.$$

Referring to (9.7) we see that (9.5) becomes

$$\binom{-\lambda}{n+1} Dp_{n+1}(x) + \binom{-\lambda}{n-1} Dp_{n-1}(x) - 2\binom{-\lambda}{n} (\lambda + xD)\, p_n(x) = 0 \tag{9.8}$$

which holds for any Sheffer sequence using $f(t)$ as its delta series. Equation (9.6) becomes (using $(xD)^2 = xD + x^2D^2$)

$$(x^2D^2 + (\lambda - n + 1)\, xD - \lambda n)\, p_n(x) + \frac{n}{\lambda + n - 1}\, xD^2 p_{n-1}(x) = 0$$

Sheffer Sequences: Gegenbauer Case

Let $s_n(x)$ be Sheffer for the pair $(g(t), f(t))$ where

$$g(t) = \left(\frac{2}{1 + \sqrt{1 - t^2}}\right)^{\lambda_0}.$$

It is easy to see that

$$g(\bar{f}(t)) = (1 + t^2)^{\lambda_0}$$

and so the generating function is

$$(1 + t^2)^{\lambda - \lambda_0}(1 - 2yt + t^2)^{-\lambda} = \sum_{k=0}^{\infty} \binom{-\lambda}{k} s_k(y)\, t^k.$$

Thus when $\lambda_0 = \lambda$, the polynomials $\binom{-\lambda}{k} s_k(x)$ are the Gegenbauer polynomials.

To obtain the conjugate representation we notice that

$$\frac{1}{g(\bar{f}(t))}\, \bar{f}(t)^k = (-2)^k\, t^k (1 + t^2)^{-k - \lambda_0}$$

and a simple computation yields

$$s_n(x) = \sum_{j=0}^{[n/2]} \frac{\binom{-\lambda}{n-2j}\binom{-\lambda_0 + 2j - n}{j}}{\binom{-\lambda}{n}} (-2x)^{n-2j}.$$

To apply the Expansion Theorem we observe that

$$g(t)\, f(t)^k = 2^{\lambda_0}(-t)^k (1 + \sqrt{1 - t^2})^{-\lambda_0 - k}$$

and from (9.1) we obtain

$$g(t)\,f(t)^k = \sum_{j=0}^{\infty} \frac{(-1)^k(\lambda_0 + k)(\lambda_0 + k + 2j - 1)_{j-1}}{2^{k+2j}j!}\,t^{2j+k}.$$

Hence

$$x^n = \sum_{k=0}^{[n/2]} \frac{\binom{-\lambda}{n-2k}}{\binom{-\lambda}{n}} \frac{(-1)^n(\lambda_0 + n - 2k)(\lambda_0 + n - 1)_{k-1}}{2^n k!}\,s_{n-2k}(x).$$

We may use the techniques of Section 8 to relate Sheffer sequences for different choices of λ_0. Suppose $s_n(x)$ is Sheffer for $(g_0(t), f(t))$ and suppose $r_n(x)$ is Sheffer for $(g_1(t), f(t))$ where

$$g_0(t) = \left(\frac{2}{1 + \sqrt{1 - t^2}}\right)^{\lambda_0}$$

and

$$g_1(t) = \left(\frac{2}{1 + \sqrt{1 - t^2}}\right)^{\lambda_1}.$$

If

$$r_n(x) = \sum_{k=0}^{n} a_{n,k} s_k(x),$$

then by Theorem 8.5, $t_n(x) = \sum_{k=0}^{n} a_{n,k} x^k$ is Sheffer for the pair

$$\left(\frac{g_1(\bar{f}(t))}{g_0(\bar{f}(t))}, f(\bar{f}(t))\right)$$

which is the pair

$$((1 + t^2)^{\lambda_1 - \lambda_0}, t).$$

Thus

$$\begin{aligned} t_n(x) &= (1 + t^2)^{\lambda_0 - \lambda_1}\,x^n \\ &= \sum_{k=0}^{\infty} \binom{\lambda_0 - \lambda_1}{k} t^{2k} x^n \\ &= \sum_{k=0}^{[n/2]} \frac{\binom{-\lambda}{n-2k}}{\binom{-\lambda}{n}} \binom{\lambda_0 - \lambda_1}{k} x^{n-2k} \end{aligned}$$

and finally

$$r_n(x) = \sum_{k=0}^{[n/2]} \frac{\binom{-\lambda}{n-2k}}{\binom{-\lambda}{n}} \binom{\lambda_0 - \lambda_1}{k} s_{n-2k}(x).$$

We now come to the recurrence formulas. First we need

$$\frac{g'(t)}{g(t)} = \frac{\lambda_0 t}{\sqrt{1-t^2}(1+\sqrt{1-t^2})}$$
$$= -\lambda_0 t f'(t)$$

and

$$\frac{f(t)}{f'(t)} = t\sqrt{1-t^2}.$$

Then (6.4) gives

$$n s_n(x) = xD\sqrt{1-t^2}\, s_n(x) + \frac{c_n}{c_{n-1}} \lambda_0 t s_{n-1}(x).$$

Since (9.3) holds for $s_n(x)$ we obtain

$$(n - xD)(\lambda + xD)\, s_n(x) - \frac{n}{\lambda + n - 1} (\lambda_0 + xD)\, D s_{n-1}(x) = 0.$$

In case $\lambda = \lambda_0$, for the Gegenbauer polynomials this becomes

$$xDs_n(x) + \frac{n}{\lambda + n - 1} Ds_{n-1}(x) - ns_n(x) = 0. \tag{9.9}$$

We remark that Eqs. (9.8) and (9.9) are independent and can be used to derive all of the recurrences for Gegenbauer polynomials appearing in Rainville [4] including the second-order differential equation.

Finally, Eq. (6.6) gives

$$\binom{-\lambda}{n} (n - xD)\, s_n(x) = 2(xD + \lambda_0) \sum_{k=0}^{[n/2]} \binom{-\lambda}{2k} \binom{-\lambda}{n-2k} (-1)^k s_{n-2k}(x).$$

Sheffer Sequences: Chebyshev Case

If $\lambda = 1$, then

$$c_n = (-1)^n$$

and

$$tx^n = -x^{n-1}$$

and so

$$tp(x) = (-x)^{-1}\, p(x).$$

The recurrence (9.5) becomes

$$2xp_n(x) + p_{n+1}(x) + p_{n-1}(x) = 0. \tag{9.10}$$

The Sheffer sequence $T_n(x)$ for the pair $(g(t), f(t))$ where

$$g(t) = \frac{1}{\sqrt{1-t^2}}$$

is related to the Chebyshev polynomials of the first kind. In fact we have

$$g(\bar{f}(t)) = \frac{1+t^2}{1-t^2}$$

and so the generating function for $s_n(x)$ is

$$(1-t^2)(1-2yt+t^2)^{-1} = \sum_{k=0}^{\infty} (-1)^k\, T_k(x)\, t^k.$$

From $T_n(x) = g^{-1}(t)\, p_n(x)$ we get $T_n(x) = \sqrt{1-t^2}\, p_n(x)$ and (9.3) gives

$$T_n(x) = x^{-1}p_{n-1}(x) + p_n(x) \tag{9.11}$$

and (9.4) gives

$$T_n(x) = -x^{-1}p_{n+1}(x) - p_n(x). \tag{9.12}$$

The conjugate representation for $p_n(x)$

$$p_n(x) = \sum_{j=0}^{[n/2]} \binom{2j-n}{j} (-2x)^{n-2j}$$

then gives direct formulas for $T_n(x)$.

The Sheffer sequence $U_n(x)$ for the pair $(g(t), f(t))$ where

$$g(t) = \frac{2-2\sqrt{1-t^2}}{t^2} = -2\frac{f(t)}{t}$$

is related to the Chebyshev polynomials of the second kind. We have

$$g(\bar{f}(t)) = 1 + t^2$$

and so

$$(1 - 2yt + t^2)^{-1} = \sum_{k=0}^{\infty} (-1)^k U_k(x) t^k.$$

Now

$$\begin{aligned} U_n(x) &= g^{-1}(t) \, p_n(x) \\ &= -\frac{1}{2} \left(\frac{f(t)}{t}\right)^{-1} p_n(x) \\ &= \frac{1}{2} \left(\frac{f(t)}{t}\right)^{-1} f(t) \, p_{n+1}(x) \\ &= \tfrac{1}{2} t p_{n+1}(x) \\ &= -\tfrac{1}{2} x^{-1} p_{n+1}(x) \end{aligned}$$

and so we see that the Chebyshev polynomials of the second kind are intimately related to the associated sequence $p_n(x)$.

Combining (9.11) and (9.12) with (9.13) gives equations connecting $T_n(x)$ and $U_n(x)$,

$$T_n(x) = -2U_{n-2}(x) - 2xU_{n-1}(x),$$
$$T_n(x) = 2U_n(x) + 2xU_{n-1}(x),$$

and hence

$$T_n(x) = U_n(x) - U_{n-2}(x).$$

We conclude with the connection-constants problem

$$U_n(x) = \sum_{k=0}^{n} a_{n,k} T_k(x).$$

From Theorem 8.5 we see that $t_n(x) = \sum_{k=0}^{n} a_{n,k} x^k$ is Sheffer for $(1 - t^2, t)$ and so

$$\begin{aligned} t_n(x) &= (1 - t^2)^{-1} x^n \\ &= \sum_{j=0}^{\infty} (-1)^j t^{2j} x^n \\ &= \sum_{j=0}^{[n/2]} (-1)^j x^{n-2j}. \end{aligned}$$

Thus $a_{n,n-2j}=(-1)^j$ and $a_{n,n-2j+1}=0$ so

$$U_n(x)=\sum_{j=0}^{[n/2]}(-1)^j\,T_{n-2j}(x).$$

10. Examples: Jacobi and Others

In this section we study the delta series

$$f(t)=\frac{1+t-\sqrt{1+2t}}{t}$$

$$=\frac{t}{1+t+\sqrt{1+2t}}.$$

Associated Sequence

First we have

$$\bar{f}(t)=\frac{2t}{(1-t)^2},$$

$$f'(t)=\frac{1}{t\sqrt{1+2t}}f(t).$$

We shall denote the associated sequence for $f(t)$ by $p_n(x)$. The generating function is

$$\varepsilon_y\left(\frac{2t}{(1-t)^2}\right)=\sum_{k=0}^{\infty}\frac{p_k(y)}{c_k}t^k.$$

The conjugate representation for $p_n(x)$ is

$$p_n(x)=\sum_{k=0}^{n}\frac{\langle 2^k t^k(1-t)^{-2k}\mid x^n\rangle}{c_k}x^k.$$

Now

$$\langle 2^k t^k(1-t)^{-2k}\mid x^n\rangle=\sum_{j=0}^{\infty}\binom{-2k}{j}2^k(-1)^j\langle t^{j+k}\mid x^n\rangle$$

$$=\binom{-2k}{n-k}2^k(-1)^{n-k}c_n$$

and so

$$p_n(x)=\sum_{k=0}^{n}\binom{-2k}{n-k}\frac{c_n}{c_k}(-1)^{n-k}(2x)^k.$$

Our first recurrence comes from the equations

$$\begin{aligned}\sqrt{1+2t}\,p_n(x) &= (1+t+\sqrt{1+2t})\,p_n(x)-(1+t)\,p_n(x)\\ &= \frac{c_n}{c_{n+1}}(1+t+\sqrt{1+2t})f(t)\,p_{n+1}(x)-(1+t)\,p_n(x)\\ &= \frac{c_n}{c_{n+1}}\,tp_{n+1}(x)-(1+t)\,p_n(x) \qquad (10.1)\end{aligned}$$

and

$$\begin{aligned}\sqrt{1+2t}\,p_n(x) &= (1+t-tf(t))\,p_n(x)\\ &= (1+t)\,p_n(x)-\frac{c_n}{c_{n-1}}\,tp_{n-1}(x) \qquad (10.2)\end{aligned}$$

From these equations we obtain

$$\frac{c_n}{c_{n+1}}\,tp_{n+1}(x)+\frac{c_n}{c_{n-1}}\,tp_{n-1}(x)-2(1+t)\,p_n(x)=0 \qquad (10.3)$$

which holds for all Sheffer sequences using the delta series $f(t)$.

We obtain another recurrence from (6.4) by noticing that

$$\frac{f(t)}{tf'(t)}=\sqrt{1+2t}$$

and so

$$np_n(x)=xD\sqrt{1+2t}\;p_n(x).$$

Using (10.2) we obtain

$$\frac{c_n}{c_{n+1}}\,xDt\,p_{n+1}(x)-xD(1+t)\,p_n(x)-np_n(x)=0. \qquad (10.4)$$

Sheffer Sequence: Jacobi Case

We shall take

$$\begin{aligned}c_n &= \frac{(1+\alpha)^{(n)}}{((1+\alpha+\beta)/2)^{(n)}((2+\alpha+\beta)/2)^{(n)}}\\ &= \frac{2^{2n}(1+\alpha)^{(n)}}{(1+\alpha+\beta)^{(2n)}}.\end{aligned}$$

Since

$$\frac{c_n}{c_{n-1}} = \frac{4(\alpha+n)}{(\alpha+\beta+2n-1)(\alpha+\beta+2n)},$$

one can readily check that

$$t = 4(1+\alpha+\beta+2xD)^{-1}(2+\alpha+\beta+2xD)^{-1}\,x^{-1}(\alpha+xD). \quad (10.5)$$

Also,

$$\begin{aligned}\varepsilon_y(t) &= \sum_{k=0}^{\infty} \frac{((1+\alpha+\beta)/2)^{(k)}((2+\alpha+\beta)/2)^{(k)}}{(1+\alpha)^{(k)}}\, y^k t^k \\ &= {}_2F_1\left[\begin{matrix}\frac{1+\alpha+\beta}{2}, \frac{2+\alpha+\beta}{2}; & \\ & ;\ yt \\ 1+\alpha; & \end{matrix}\right].\end{aligned}$$

We denote by $J_n(x)$ the Sheffer sequence for the pair $(g(t), f(t))$ where

$$g(t) = \left(\frac{2}{1+\sqrt{1+2t}}\right)^{1+\alpha+\beta}.$$

Then since

$$g(\bar{f}(t)) = (1-t)^{1+\alpha+\beta},$$

the generating function for $J_n(x)$ is

$$\begin{aligned}&(1-t)^{-1-\alpha-\beta}\,{}_2F_1\left[\begin{matrix}\frac{1+\alpha+\beta}{2}, \frac{2+\alpha+\beta}{2}; & \\ & \frac{2yt}{(1-t)^2} \\ 1+\alpha; & \end{matrix}\right] \\ &= \sum_{k=0}^{\infty} \frac{((1+\alpha+\beta)/2)^{(k)}((2+\alpha+\beta)/2)^{(k)}}{(1+\alpha)^{(k)}}\, J_k(y)\, t^k.\end{aligned}$$

Referring to Rainville [4, p. 256], where the classical Jacobi polynomials are denoted by $P_n^{(\alpha,\beta)}(x)$, we have

$$J_n(x) = \frac{(1+\alpha+\beta)^{(n)}}{(1+\alpha)^{(n)}}\, c_n P_n^{(\alpha,\beta)}(x+1).$$

The conjugate representation for $J_n(x)$ is obtained from

$$\begin{aligned}\langle g(\bar{f}(t))^{-1}\bar{f}(t)^k \mid x^n\rangle &= \langle 2^k t^k(1-t)^{-1-\alpha-\beta-2k} \mid x^n\rangle \\ &= \sum_{j=0}^{\infty} \binom{-1-\alpha-\beta-2k}{j} (-1)^j\, 2^k \langle t^{j+k} \mid x^n\rangle \\ &= \binom{-1-\alpha-\beta-2k}{n-k} (-1)^{n-k}\, 2^k c_n \\ &= \binom{\alpha+\beta+n+k}{n-k} 2^k c_n\end{aligned}$$

and gives

$$\begin{aligned}J_n(x) &= \sum_{k=0}^{n} \binom{\alpha+\beta+n+k}{n-k} \frac{c_n}{c_k} (2x)^k \\ &= \sum_{k=0}^{n} \binom{\alpha+\beta+n+k}{n-k} \frac{(\alpha+n)_{n-k}}{(\alpha+\beta+2n)_{2n-2k}}\, 2^{2n-k} x^k.\end{aligned}$$

We wish to express x^n as a linear combination of $J_k(x)$. From Corollary 2 to the Expansion Theorem (Theorem 5.2) we have

$$x^n = \sum_{k=0}^{n} \frac{\langle g(t) f(t)^k \mid x^n\rangle}{c_k} J_k(x).$$

Now by (9.1) we have

$$\begin{aligned}(1+t+\sqrt{1+2t})^{-\mu} &= 2^{\mu}(1+\sqrt{1+2t})^{-2\mu} \\ &= \sum_{j=0}^{\infty} \frac{\mu(2\mu+2j-1)_{j-1}}{2^{\mu+j-1} j!} (-t)^j\end{aligned}$$

and so

$$\begin{aligned}g(t) f(t)^k &= 2^{1+\alpha+\beta} t^k (1+t+\sqrt{1+2t})^{-1-\alpha-\beta-k} \\ &= \sum_{j=0}^{\infty} (-1)^j \frac{(1+\alpha+\beta+k)(1+2\alpha+2\beta+2k+2j)_{j-1}}{2^{k+j-1} j!} t^{k+j}\end{aligned}$$

and

$$\langle g(t) f(t)^k \mid x^n\rangle = \frac{(-1)^{n-k}(1+\alpha+\beta+k)(1+2\alpha+2\beta+2n)_{n-k-1}}{2^{n-1}(n-k)!} c_n.$$

Thus

$$x^n = \sum_{k=0}^{n} \frac{(-1)^{n-k}(1+\alpha+\beta+k)(1+2\alpha+2\beta+2n)_{n-k-1}}{2^{n-1}(n-k)!} \frac{c_n}{c_k} J_k(x)$$
$$= \sum_{k=0}^{n} \frac{(1+\alpha+\beta+k)(1+2\alpha+2\beta+2n)_{n-k-1}(\alpha+n)_{n-k}}{(n-k)!\,(\alpha+\beta+2n)_{2n-2k}}$$
$$\times (-1)^{n-k}\, 2^{n+2k+1}\, J_k(x).$$

We conclude our discussion of the polynomials $J_n(x)$ with some recurrence formulas. Equation (10.3) holds for $J_n(x)$ where t is given by (10.5). Let us derive Eq. (6.6) of Theorem 6.7. From $\bar{f}(t) = 2t(1-t)^{-2}$ we obtain

$$\left(\frac{\bar{f}(t)}{t}\right)^{-1} \bar{f}(t) = \frac{1+t}{1-t}$$
$$= 1 + 2 \sum_{k=1}^{\infty} t^k$$

and so

$$b_0 = c_0,$$
$$b_k = 2c_k.$$

Also, from $g(\bar{f}(t)) = (1-t)^{1+\alpha+\beta}$ we obtain

$$\frac{t[g(\bar{f}(t))]'}{g(\bar{f}(t))} = -(1+\alpha+\beta)\frac{t}{1-t}$$
$$= -(1+\alpha+\beta) \sum_{k=1}^{\infty} t^k$$

and so

$$d_0 = 0,$$
$$d_k = -(1+\alpha+\beta)\, c_k.$$

Therefore Eq. (6.6) becomes

$$(n - xD)\, J_n(x) = \sum_{k=1}^{n} \frac{c_n}{c_k c_{n-k}} (2c_k xD + (1+\alpha+\beta)c_k)\, J_{n-k}(x)$$
$$= (1+\alpha+\beta+2xD) \sum_{k=1}^{n} \frac{c_n}{c_{n-k}} J_{n-k}(x).$$

11. Examples: The q-Case

In this section we shall briefly discuss the q-theory. We take

$$c_n = \frac{(1-q)(1-q^2)\cdots(1-q^n)}{(1-q)^n}.$$

Then

$$\frac{c_n}{c_{n-1}} = \frac{1-q^n}{1-q}$$

and so

$$tx^n = \frac{1-q^n}{1-q} x^{n-1}$$
$$= \frac{x^n-(qx)^n}{x-qx}$$

and so

$$tp(x) = \frac{p(x)-p(qx)}{x-qx}. \tag{11.1}$$

The q-binomial coefficient is

$$\binom{n}{k}_q = \frac{c_n}{c_k c_{n-k}}$$
$$= \frac{(1-q)\cdots(1-q^n)}{(1-q)\cdots(1-q^k)(1-q)\cdots(1-q^{n-k})}.$$

Thus we have

$$\varepsilon_y(t)\, x^n = \sum_{k=0}^{n} \binom{n}{k}_q y^k x^{n-k}.$$

From (3.5) and the equations

$$\begin{aligned} \partial_t t^k &= \frac{1-q^k}{(1-q)^k} t^{k-1}, \\ \partial_t f(t) &= \frac{f(t)-f(qt)}{t-qt} \end{aligned} \tag{11.2}$$

we obtain

$$y\varepsilon_y(t) = \frac{\varepsilon_y(t)-\varepsilon_y(qt)}{t-qt}$$

or

$$\varepsilon_y(qt) = (1 - (1 - q)\, yt)\, \varepsilon_y(t). \tag{11.3}$$

Sheffer sequences for the delta series $f(t) = t$ satisfy

$$ts_n(x) = \frac{1 - q^n}{1 - q} s_{n-1}(x)$$

and in view of (11.1) we get

$$s_n(x) - s_n(qx) = (1 - q^n)\, xs_{n-1}(x). \tag{11.4}$$

We define the sequence $[x]_{a,n}$ by

$$\begin{aligned} [x]_{a,0} &= 1, \\ [x]_{a,n} &= (x - a)(x - qa) \cdots (x - q^{n-1}a) \end{aligned}$$

and write $[x]_{1,n}$ as $[x]_n$.

Then using (11.1) it is straightforward to verify that

$$\begin{aligned} t[x]_{a,n} &= \frac{1 - q^n}{1 - q} [x]_{a,n-1} \\ &= \frac{c_n}{c_{n-1}} [x]_{a,n-1} \end{aligned} \tag{11.5}$$

and so $[x]_{a,n}$ is Sheffer for the delta series $f(t) = t$. Therefore (11.4) gives

$$[x]_{a,n} - [qx]_{a,n} = (1 - q^n)\, x[x]_{a,n-1}.$$

Since

$$\begin{aligned} \langle \varepsilon_a(t) \mid [x]_{a,n} \rangle &= [a]_{a,n} \\ &= \delta_{n,0}, \end{aligned}$$

the sequence $[x]_{a,n}$ is Sheffer for the pair $(\varepsilon_a(t), t)$. From Theorem 5.5 and (11.5) we obtain

$$\begin{aligned} x^n &= \varepsilon_a(t)[x]_{a,n} \\ &= \sum_{k=0}^{\infty} \frac{a_k}{c_k} t^k [x]_{a,n} \\ &= \sum_{k=0}^{\infty} \binom{n}{k}_q a^k [x]_{a,n-k}. \end{aligned}$$

The generating function for $[x]_{a,n}$ is

$$\frac{1}{\varepsilon_a(t)}\varepsilon_y(t)=\sum_{k=0}^{\infty}\frac{(1-q)^k}{(1-q)\cdots(1-q^k)}[y]_{a,k}t^k.$$

Letting $y=0$ and noticing that $[0]_{a,k}=(-a)^k q^{\binom{k}{2}}$ gives

$$\frac{1}{\varepsilon_a(t)}=\sum_{k=0}^{\infty}\frac{(1-q)^k}{(1-q)\cdots(1-q^k)}q^{\binom{k}{2}}(-at)^k.$$

Since

$$\begin{aligned}\langle t^0 \mid [x]_{a,n}\rangle &= [0]_{a,n}\\ &= (-a)^n q^{\binom{n}{2}}\end{aligned}$$

and using Theorem 4.1 we get

$$\begin{aligned}\langle t^k \mid [x]_{a,n}\rangle &= \frac{c_n}{c_{n-k}}\langle t^0 \mid [x]_{a,n-k}\rangle\\ &= \frac{c_n}{c_{n-k}}(-a)^{n-k} q^{\binom{n-k}{2}}.\end{aligned}$$

Thus by Corollary 2 of the Expansion Theorem

$$\begin{aligned}[x]_{a,n} &= \sum_{k=0}^{n}\frac{\langle t^k \mid [x]_{a,n}\rangle}{c_k}x^k\\ &= \sum_{k=0}^{n}\binom{n}{k}_q (-a)^{n-k} q^{\binom{n-k}{2}}x^k.\end{aligned}$$

Replacing x by $\varepsilon_1(t)$ gives the formula

$$(\varepsilon_1(t)-1)\cdots(\varepsilon_1(t)-q^{n-1})=\sum_{k=0}^{n}\binom{n}{k}_q(-1)^{n-k}q^{\binom{n-k}{2}}\varepsilon_1(t)^k.$$

Applying this to a polynomial $p(x)$ gives a formula which appears frequently in the literature in the somewhat confusing form

$$\Delta^n p(x)=\sum_{k=0}^{n}\binom{n}{k}_q(-1)^{n-k}q^{\binom{n-k}{2}}p(x+k).$$

This same Corollary also gives

$$\begin{aligned}[x]_{b,n} &= \sum_{k=0}^{n} \frac{\langle \varepsilon_a(t)t^k \mid [x]_{b,n}\rangle}{c_k} [x]_{a,k} \\ &= \sum_{k=0}^{n} \binom{n}{k}_q [a]_{b,n-k}[x]_{a,k}.\end{aligned}$$

This is actually the q-Vandermonde convolution formula in disguise. To see this notice that

$$[ax]_{a,k} = a^k [x]_k$$

and so (11.6) gives, with x replaced by ax and $b=1$,

$$[ax]_n = \sum_{k=0}^{n} \binom{n}{k}_q a^k [a]_{n-k}[x]_k. \tag{11.7}$$

We wish to make the substitutions

$$\begin{aligned} x &= q^l, \\ a &= q^m, \\ ax &= q^{l+m}. \end{aligned}$$

Now if $i \leqslant j$ we have

$$\begin{aligned}[q^j]_i &= (q^j - 1)(q^j - q) \cdots (q^j - q^{i-1}) \\ &= q^{1+2+\cdots+i-1}(q^j - 1)(q^{j-1} - 1) \cdots (q^{j-i+1} - 1) \\ &= q^{\binom{i}{2}} \frac{(q^j - 1)\cdots(q-1)}{(q^{j-i} - 1)\cdots(q-1)} \\ &= q^{\binom{i}{2}} \frac{(q-1)^j}{(q-1)^{j-i}} \frac{c_j}{c_{j-i}} \\ &= q^{\binom{i}{2}} (q-1)^i \frac{c_j}{c_{j-i}}\end{aligned}$$

and if $i > j$, then $[q^j]_i = 0$. Using this in (11.7) gives

$$\begin{aligned} q^{\binom{n}{2}}(q-1)^n \frac{c_{l+m}}{c_{l+m-n}} = \sum_{k=0}^{n} & \frac{c_n}{c_k c_{n-k}} q^{km} q^{\binom{n-k}{2}} (q-1)^{n-k} \\ & \times \frac{c_n}{c_{m-n+k}} q^{\binom{k}{2}} (q-1)^k \frac{c_l}{c_{l-k}} \end{aligned}$$

which becomes, in view of $\binom{n-k}{2}+\binom{k}{2}-\binom{n}{2}=-k(n-k)$,

$$\binom{l+m}{n}_q=\sum_{k=0}^{n} q^{k(m-n+k)}\binom{l}{k}_q\binom{m}{n-k}_q. \tag{11.8}$$

Next we touch briefly on the q-Bernoulli polynomials. We define the q-integral by

$$\int_q x^n=\frac{c_n}{c_{n+1}}x^{n+1}$$
$$=\frac{1-q}{1-q^{n+1}}x^{n+1}$$

and

$$\int_{q\,0}^{y} x^n=\frac{1-q}{1-q^{n+1}}\,y^{n+1}.$$

The *q-Bernoulli polynomials* have the generating function

$$\frac{t}{\varepsilon_1(t)-1}\,\varepsilon_y(t)=\sum_{k=0}^{\infty}\frac{(1-q)^k}{(1-q)\cdots(1-q^k)}\,B_k(y)\,t^k.$$

That is, $B_n(x)$ is Sheffer for the pair

$$\left(\frac{\varepsilon_1(t)-1}{t},t\right).$$

By Theorem 5.5 we have

$$\frac{\varepsilon_1(t)-1}{t}B_n(x)=x^n.$$

But

$$\left\langle\frac{\varepsilon_1(t)-1}{t}\,\middle|\,x^n\right\rangle=\frac{c_n}{c_{n+1}}\left\langle\frac{\varepsilon_1(t)-1}{t}\,\middle|\,tx^{n+1}\right\rangle$$
$$=\frac{c_n}{c_{n+1}}\langle\varepsilon_1(t)-1\mid x^{n+1}\rangle$$
$$=\frac{c_n}{c_{n+1}}\,1^{n+1}$$
$$=\int_{q\,0}^{1}x\cdot x^n$$

and so

$$\left\langle \frac{\varepsilon_1(t)-1}{t} \,\middle|\, p(x) \right\rangle = {}_q\!\int_0^1 xp(x).$$

In particular,

$$\delta_{n,0} = {}_q\!\int_0^1 xB_n(x).$$

Corollary 2 of the Expansion Theorem is the *q-Euler–MacLaurin Expansion*

$$\begin{aligned} p(x) &= \sum_{k=0}^{\infty} \frac{(1-q)^k}{(1-q)\cdots(1-q^k)} \left\langle \frac{\varepsilon_1(t)-1}{t} t^k \,\middle|\, p(x) \right\rangle B_k(x) \\ &= \sum_{k=0}^{\infty} \frac{(1-q)^k}{(1-q)\cdots(1-q^k)} B_k(x) \; {}_q\!\int_0^1 t^k p(x). \end{aligned}$$

We can also connect the two sequences $B_n(x)$ and $[x]_n$,

$$\begin{aligned} B_n(x) &= \sum_{k=0}^{n} \frac{\langle \varepsilon_1(t)t^k \mid B_n(x)\rangle}{c_k} [x]_k \\ &= \sum_{k=0}^{n} \binom{n}{k}_q B_k(1)[x]_k \end{aligned}$$

and

$$\begin{aligned} [x]_n &= \sum_{k=0}^{n} \left\langle \frac{\varepsilon_1(t)-1}{t} t^k \,\middle|\, [x]_n \right\rangle \frac{B_k(x)}{c_k} \\ &= \sum_{k=0}^{n} \binom{n}{k}_q \left\langle \frac{\varepsilon_1(t)-1}{t} \,\middle|\, [x]_{n-k} \right\rangle B_k(x) \\ &= \sum_{k=0}^{n} \binom{n}{k}_q B_k(x) \; {}_q\!\int_0^1 x[x]_{n-k}. \end{aligned}$$

We now turn to the q-Leibniz formula. First we need a lemma.

LEMMA. *For any series $f(t)$ and polynomial $p(x)$,*

$$\langle \partial_t^n f(t) \mid p(q^j x)\rangle = q^{-jn} \langle \partial_t^n f(q^j t) \mid p(x)\rangle$$

Proof. By linearity we need only check this for $f(t) = t^k$ and $p(x) = x^m$. But then

$$\begin{aligned}\langle \partial_t^n t^k \mid (q^j x)^m \rangle &= \frac{c_k}{c_{k-n}} q^{jm} \langle t^{k-n} \mid x^m \rangle \\ &= \frac{c_k}{c_{k-n}} q^{j(k-n)} \langle t^{k-n} \mid x^m \rangle \\ &= q^{j(k-n)} \langle \partial_t^n t^k \mid x^m \rangle \\ &= q^{-jn} \langle \partial_t^n (q^j t)^k \mid x^m \rangle.\end{aligned}$$

The q-Leibniz formula is

$$\partial_t^n (f(t)\, g(t)) = \sum_{k=0}^{n} \binom{n}{k}_q q^{-k(n-k)}\, \partial_t^k f(t)\, \partial_t^{n-k} g(qt).$$

The proof consists of the following calculations, in which we use the q-Vandermonde convolution (11.8) and Theorem 5.9,

$$\begin{aligned}
&\langle \partial_t^n (f(t)\, g(t)) \mid x^m \rangle \\
&\quad = \langle f(t)\, g(t) \mid x^{n+m} \rangle \\
&\quad = \sum_{j=0}^{n} \binom{n+m}{j}_q \langle f(t) \mid x^j \rangle \langle g(t) \mid x^{n+m-j} \rangle \\
&\quad = \sum_{j=0}^{n+m} \sum_{k=0}^{n+m} \binom{n}{k}_q \binom{m}{j-k}_q q^{k(m-j+k)} \langle f(t) \mid x^j \rangle \langle g(t) \mid x^{n+m-j} \rangle \\
&\quad = \sum_{k=0}^{n+m} \binom{n}{k}_q \sum_{j=0}^{n+m} \binom{m}{j-k}_q q^{k(m-j+k)} \langle f(t) \mid x^j \rangle \langle g(t) \mid x^{n+m-j} \rangle \\
&\quad = \sum_{k=0}^{n+m} \binom{n}{k}_q \sum_{j=0}^{m} \binom{m}{j}_q q^{k(m-j)} \langle f(t) \mid x^{j+k} \rangle \langle g(t) \mid x^{n+m-j-k} \rangle \\
&\quad = \sum_{k=0}^{n+m} \binom{n}{k}_q \sum_{j=0}^{m} \binom{m}{j}_q q^{k(m-j)} \langle \partial_t^k f(t) \mid x^j \rangle \langle \partial_t^{n-k} g(t) \mid x^{m-j} \rangle \\
&\quad = \sum_{k=0}^{n+m} \binom{n}{k}_q \sum_{j=0}^{m} \binom{m}{j}_q \langle \partial_t^k f(t) \mid x^j \rangle \langle \partial_t^{n-k} g(t) \mid (q^k x)^{m-j} \rangle \\
&\quad = \sum_{k=0}^{n+m} \binom{n}{k}_q \sum_{j=0}^{m} \binom{m}{j}_q \langle \partial_t^k f(t) \mid x^j \rangle\, q^{-k(n-k)} \langle \partial_t^{n-k} g(q^k t) \mid x^{m-j} \rangle \\
&\quad = \sum_{k=0}^{n+m} \binom{n}{k}_q q^{-k(n-k)} \langle \partial_t^k f(t)\, \partial_t^{n-k} g(q^k t) \mid x^m \rangle \\
&\quad = \left\langle \sum_{k=0}^{n+m} \binom{n}{k}_q q^{-k(n-k)}\, \partial_t^k f(t)\, \partial_t^{n-k} g(q^k t) \,\middle|\, x^m \right\rangle.
\end{aligned}$$

12. Future Directions

We intend this paper to be the first in a series of papers. Let us give a brief description of the subject of two future papers.

First is the idea of replacing the algebra of polynomials P by the field of formal Laurent series of the form

$$p(x)=\sum_{k=-\infty}^{n} a_k x^k.$$

The entire theory goes through in this new setting. Each Sheffer sequence

$$s_n(x)=\sum_{k=0}^{n} a_{n,k} x^k$$

is replaced by a sequence of formal Laurent series

$$\bar{s}_n(x)=\sum_{k=-\infty}^{n} \bar{a}_{n,k} x^k,$$

where $\bar{a}_{n,k}=a_{n,k}$ for $n, k \geqslant 0$. For $n<0$, the sequence $\bar{s}_n(x)$ has been termed a factor sequence and several examples have appeared in the classical literature. However, for $n>0$ the sequence $\bar{s}_n(x)$ has never been studied.

The second direction for future work comes from the observation that the present theory is somehow "centralized" at 0. This is evident from the fact that t^0 is essentially evaluation at 0 and $\langle f(t) \mid p(x)\rangle=\langle t^0 \mid f(t)p(x)\rangle$. We may "decentralize" the umbral calculus as follows. Let α_0, α_1,... be a sequence of independent transcendentals. Then the role of the sequence t^k is taken by the sequence $\varepsilon_{\alpha_k}(t)\, t^k$, where $\varepsilon_{\alpha_k}(t)$ is evaluation at α_k. A large part of the present theory still goes through. Some interesting new polynomial sequences now come to light, for example the sequence

$$s_n(x)=(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)$$

and the Gončarov polynomials $G_n(x)$. The latter are defined as the unique polynomials for which $G_n^{(k)}(\alpha_k)=\delta_{n,k}$. Both these sequences are important in the theory of interpolation.

References

1. S. M. Roman and G.-C. Rota, The umbral calculus, *Adv. in Math.* **27** (1978), 95–188.
2. G.-C. Rota, D. Kahaner, and A. Odlyzko, Finite operator calculus, *J. Math. Anal. Appl.* **42** (1973), 684–760.

3. R. P. BOAS AND R. C. BUCK, "Polynomial Expansions of Analytic Functions," Academic Press, New York, 1964.
4. E. RAINVILLE, "Special Functions," Chelsea, New York, 1971.
5. S. M. ROMAN, The algebra of formal series, *Adv. in Math.* **31** (1979), 309–329 [Erratum **35** (1980), 274].
6. S. M. ROMAN, The algebra of formal series. II. Sheffer sequences, *J. Math. Anal. Appl.* **74** (1980), 120–143.
7. S. M. ROMAN, The algebra of formal series. III. Several variables, *J. Approx. Theory* **26** (1979), 340–381.
8. S. M. ROMAN, Polynomials, power series and interpolation, *J. Math. Anal. Appl.* **80** (1981), 333–371.
9. S. M. ROMAN, The formula of Faa di Bruno, *Amer. Math. Monthly* **87** (1980), 805–809.