**COMBINATORIAL**
**MATHEMATICS**
**YEAR**

**February 1969 - June 1970**

# ON THE FOUNDATIONS OF COMBINATORIAL THEORY III.

## Theory of Binomial Enumeration

by

Ronald Mullin
University of Waterloo
Waterloo,
Ontario

Gian-Carlo Rota
Massachusetts Institute
of Technology
Cambridge, Massachusetts

Department of Statistics
University of North Carolina at Chapel Hill
Institute of Statistice Mimeo Series No. 600.24
April 1970

*This is the final version of lecture notes for a series of lectures given by the second author at the Summer Study Institute on Combinatorial Mathematics and Its Applications in the Natural Sciences, held on June 2-27, 1969, in Chapel Hill, and sponsored by the U.S. Air Force Office of Scientific Research under Grant No. AFOSR-68-1406 and the U.S. Army Research Office, Durham, under Grant No. DA-ARO-D-31-124-G910.*

# ON THE FOUNDATIONS OF COMBINATORIAL THEORY

## III. Theory of Binomial Enumeration

by

Ronald Mullin

and

Gian-Carlo Rota

1. Introduction
2. Reluctant Functions and Trees
3. Fundamentals
4. Expansions
5. Closed Forms
6. The Automorphism Theorem
7. Umbral Notation
8. The Exponential Polynomials
9. Laguerre Polynomials
10. A Glimpse of Combinatorics
11. Bibliography

## 1. Introduction.

The present work is born from the interplay of two seemingly disparate branches of combinatorial theory. The first is the classical calculus of finite differences, which has been in the past more often related to numerical analysis than to problems of enumeration. In the calculus of finite differences, there occur several sequences of polynomials which are used in interpolation, numerical quadrature, and several other connections. Typical of such sequences of polynomials are the lower factorials

$$(1) \qquad p_n(x) = (x)_n = x(x-1)\ldots(x-n+1), \ m = 0,1,2,$$

and the upper factorials

$$(2) \qquad p_n(x) = x^{(n)} = x(x+1)\ldots(x+n-1), \ n = 0,1,2, \quad .$$

Less well known, but equally significant polynomial sequences are the Abel polynomials, studied by Abel, Hurwitz and others:

$$(3) \qquad p_n(x) = x(x-an)^{n-1}, \ n = 0,1,2,\ldots$$

and the exponential polynomials, studied by Touchard and others,

$$(4) \qquad \varphi_n(x) = \sum_{k\geq 0} S(n,k)x^k,$$

where $S(n,k)$ denote the familiar Stirling numbers of the second kind. Another significant sequence is the Laguerre polynomials

$$(5) \qquad L_n(x) = \sum_{k\geq 0} \frac{n!}{k!} \binom{n-1}{k-1}(-x)^k,$$

which have an extensive literature. These sequences of polynomials, as well as a large number of other sequences that have arisen in classical analysis and combinatorics, share a common property: that of being of *binomial type*. We say that a sequence of polynomials $p_n(x)$, where $p_n(x)$ is of exactly of degree of $n$, is of binomial type when it satisfies the sequences of identities

$$(6) \qquad p_n(x+y) = \sum_{k\geq 0} \binom{n}{k} p_k(x)p_{n-k}(y), \; n = 0,1,2,\ldots \quad .$$

It will be shown in the course of this study (and it is verified without difficulty using the results below) that each one of the sequences of polynomials mentioned above is of binomial type.

This work is a study of certain analytic or (more suggestively) algebraic-combinatorial properties of sequences of polynomials of binomial type. The main problem we aim at is the following: given two sequences $p_n(x)$ and $q_n(x)$, both of binomial type, there clearly exist coefficients $c_{nk}$, the so-called connection constants,

$$p_n(x) = \sum_{k>0} c_{nk} q_k(x) \tag{7}$$

which express one sequence of polynomials in terms of the other. Our problem is to determine as efficiently as possible the coefficients $c_{nk}$ in terms of minimal data on the polynomials $p_n(x)$ and $q_n(x)$. A few classical instances of this problem are given below.

In trying to solve this problem we were led to develop a systematic theory of polynomial sequences of binominal type. The main novelties we introduce in this theory are, first, a systematic use of operator methods as against less efficient generating function methods, which were used almost exclusively in the past, and secondly a solution of the connection problem stated above, which eluded past workers in the field, and which we believe to be remarkably simple.

Patches and bits of the theory developed in this work can be found in the literature of the last 50 years, starting with the work of Pincherle and Amaldi in 1900, following through the Danish and Italian schools of calculus of finite differences, culminating with the work of the great Danish actuarialist Steffensen. The statement (though not, alas, the proof) of Theorem 4 below is due to him. A few other results, such as the Expansion Theorem, where at least intuited by Pincherle and his school. But, we believe that our notion of umbral operator (a term introduced by Sylvester and extensively used by the invariant theorists and by E.T. Bell, though never correctly defined), together with our solution of the connection constants problem that it yields, gives a new direction to the calculus of finite differences, even for workers interested in purely analytic matters.

It turns out that there is a second and entirely different point of view from which the theory of polynomials of binomial types can be looked at. Each of the polynomial sequences listed above can be interpreted as counting the number of ways of placing "balls" into "boxes", subject to various restrictions. This ties in with the classical theory of distributions and occupancy, which can be alternatively considered as making words out of an alphabet, subject to various restrictions on the successions of letters. More precisely, we are given a set $S$ with $n$ elements and a set $X$ with $x$ elements, and we consider functions from the set $S$ to the set $X$ subject to various restrictions. The restrictions are such that they do not limit the range of the functions but only the domain. Thus, for example, the lower factorial powers (1) count the number of one-to-one functions from a set of $n$ elements to a set of $x$ elements. Similarly, the upper factorials

$$(8) \qquad x^{(n)} = x(x+1)(x+2)\ldots(x+n-1)$$

count the different ways of placing the balls $S$ into the boxes $x$ when a linear ordering is to be chosen of the balls within each box.

In the same vein, the Abel polynomials

$$(9) \qquad x(x-an)^{n-1}, \; n = 0,1,\ldots, \; an<x,$$

can be considered in combinatorial terms. Indeed, consider a circle of circumference $x$, and a set of $n$ arcs each of length $a$ and each having the same radius of curvature as the circle. If we drop the arcs randomly on the circumference of the circle then the probability that no two arcs overlap is easily seen to be

$$(10) \qquad \frac{x(x-na)^{n-1}}{x^n} .$$

Thus the Abel polynomials "count" the ways (i.e., the measure, since this case is continuous) in which the arcs may be placed without overlapping.

Whenever we count a set of functions from a set of $S$ to a set $X$, subject to restrictions on the domain, then, letting $p_n(x)$ be the number of such functions, we see immediately that $p_n(x)$ is a polynomial and that the sequence $p_n(x)$ must be of binomial type. Thus, sequences of polynomials of binomial type arise naturally as the unifying concept in the theory of distribution and occupancy.

Accordingly, the present study will be divided into two parts. In the first (the present) part we concentrate on the analytic properties of polynomial sequences of binomial type; the relationship to problems of distributions and occupancy is discussed only in Sections 2 and 10, and is meant only as an introduction to the second part. It turns out that every sequence of binomial type with positive integral coefficients can be associated to a counting problem of a certain class of "reluctant" functions, as defined in the next Section. In the second part of this work we shall intrepret the analytic results derived here in purely combinatorial, that is, set-theoretic terms.

Perhaps the most satisfying results of this investigation are, first, the unexpected relations of sequences of binomial type with problems of enumeration of rooted labeled trees, (Section 2), and secondly, the solution of the problem of the connection constants, which has deep combinatorial implications.

In several special cases, classical analysis has already answered the problem of the connection constants. For example, we have

$$x^n = \sum_{k \geq 0} S(k,n)(x)_k \tag{11}$$

$$(x)_n = \sum_{k \geq 0} s(k,n)x^k \tag{12}$$

$$x^{(n)} = \sum_{k \geq 0} |s(k,n)|x^k \tag{13}$$

where $s(k,n)$ and $S(k,n)$ are the Stirling number of the first and second kind. Another example is

$$x^{(n)} = \sum_{k \geq 0} \frac{n!}{k!} \binom{n-1}{k-1}(x)_k \tag{14}$$

(15) $$L_n(x) = \sum_{k\geq 0} (-1)^k \frac{n!}{k!} \binom{n-1}{k-1} x^k$$

where $L_n(x)$ are the Laguerre polynomials.

We hope that this introuction has given an idea of the scope of the present investigation. In the next Section we briefly outline some combinatorial connections, thereafter to dismiss them in favor of the analytic theory, until Section 10.

We thank M. Shelor and S. Smoliar for help in the preparation of this manuscript, and the Statistics Department of the University of North Carolina, as well as the U.S. Army Research Center at the University of Wisconsin, and Florida Atlantic University for giving us the opportunity to present these results to an actively contributing audience.

## 2. Reluctant Functions.

Given a function $f: S \to X$, where from now on $S$ will be a finite set with $n$ elements and $X$ will be a finite set of $x$ elements, we can associate with it "functorially" two objects: the range of $f$, namely, the sub-set of elements of $X$ which are images of some elements of $S$ under the function $f$; and the coimage of $f$, which consists of the partition of the set $X$ defined by the following equivalence relation: an element $a$ of $X$ is equivalent to an element $b$ of $S$ if and only if $f(a) = f(b)$. Thus, the coimage of $f$ is a partition of the set $S$.

We are now going to rather drastically generalize these concepts.

We define a reluctant function from $S$ to $X$ as follows. It is a function $f$ from $S$ to the disjoint union $S \cup X$, subject to the following restriction. For every element $s \in S$, the element $f(f(s))$ is defined if and only if $f(s) \in S$; similarly, $f(f(f(s)))$ is defined if and only if $f(f(s)) \in S$, etc. Our requirements is that only a finite number of terms of the sequence $s$, $f(s)$, $f(f(s))$, $f(f(f(s)))$,... be well-defined. A more suggestive, if less precise, way of stating the same condition is the following: for every element $s \in S$, the "orbit" $s, f(s), f(f(s)), f(f(f(s))), \ldots$ of $s$ under iteration of the function $f$ "eventually" ends up in $X$, where it stops. Thus, one might say that $f$ "reluctantly" maps $S$ into $X$.

The range of a reluctant function $f$ will consist of those elements of $X$ which are images of some element of $S$, just like in the case of an ordinary function. On the other hand, we need to generalize the notion of coimage of an ordinary function, as defined above, to the newly introduced concept of a reluctant function. Whereas the coimage of an ordinary function is simply a partition of the set $S$, the coimage a reluctant function is going to be more than a partition of $S$. In fact, for every element $x$ of $X$ which is in the range of the reluctant function $f$, the inverse image of the element $x$ is defined as the set of all elements $s$ of $S$ such that the sequence of its successors $f(s), f(f(s)), \ldots$ eventually ends up in $X$. The inverse images of distinct elements of $X$ are disjoint subsets of $S$. Thus, to every reluctant function there is associated a partition of the set $S$, just like in the case of an ordinary function. However, within each block of such a partition there is a natural structure of a *forest of rooted trees* describing the "history" of the elements of that block before they end up in X. Thus, we are led to define the coimage of a reluctant function to be a partition of ths set S, together with a structure of a rooted forest (i.e., set of rooted trees) defined on each block of the partition. Each rooted forest covering one block of the coimage is the "inverse image" — in the generalized sense just described — of an element x of X.

Note that each block of the coimage can be further partitioned into the connected components, that is, the trees, of the rooted forest. The resulting partition is a refinement of the coimage and has the additional property that each block has the structure of a rooted tree. We call this finer partition $\pi$ of S, together with the structure of *rooted tree* (See Harary or Moon for definitions) on each block of $\pi$, the *pre-image* of the reluctant function f (recall that a rooted tree is a partially ordered set). Thus, the coimage of f is obtanied by "piecing togehter" all those blocks of the *pre-image* of f which are "eventually mapped" to the same element x in X.

Clearly, the pre-image of any reluctant function is a *rooted labeled forest* on the set S, following classical terminology. Given any rooted labeled forest L on S, with $k$ blocks, that is, consisting of $k$ rooted trees, there are evidently $x^k$ reluctant functions whose pre-image is the forest $L$.

By way of example, let us consider the set of all reluctant functions from $S$ to $X$ (notice that our use of the word "from" and "to" is not strictly correct, but is nevertheless suggestive so we shall keep using it). Let $c_{nk}$ be the number of rooted labeled forests with $k$ blocks on the set $S$. Then the number of reluctant functions from $S$ to $X$ is evidently given by the polynomial

$$(1) \qquad \sum_{k\geq 0} c_{nk}x^k = A_n(x).$$

It is easy to see, by a simple combinatorial argument which imitates the standard set-theoretic proof of the binomial theorem, that the sequence of polynomials $A_n(x)$ is of binomial type. It is less obvious, and it will trivially follow from the present theory (see Section 10) that the polynomials $A_n(x)$ are given by the expression

$$(2) \qquad A_n(x) = x(x+n)^{n-1},$$

that is, that they are a special case of the Abel polynomials, corresponding to $a = -1$. This gives immediately the classical result of Cayley counting the number of rooted trees, since rooted trees correspond to reluctant functions having as pre-image a partition with one block, and so are the coefficients $c_{n1}$ in (2), which equal $n^{n-2}$.

We define a binomial class $B$ of reluctant functions as follows. To every set $S$ and set $X$ we assign a set $F(S,X)$ of reluctant functions from $S$ to $X$. The assignment is "functorial" – or, in combinatorial language, "unlabeled". This means that isomorphisms of the sets $S$ with $S_0$ and $X$ with $X_0$ induce a natural isomorphism of the sets $F(S,X)$ with $F(S_0,X_0)$. Thus, if the polynomials $p_n(x)$ denotes the size of the sets $F(S,X)$, the function $p_n(x)$ depends only on the size $n$ the set $S$ and the size $x$ of the set $X$.

We come now to the crucial condition. In set-theoretic terms, the condition states that there is a natural isomorphism

$$(4) \qquad F(S, X\oplus Y) = \sum_{A\subseteq S} F(A,X)\otimes F(S-A,Y).$$

Here, $\oplus$ and $\Sigma$ denote disjoint sum of sets, $\otimes$ denotes product of sets, and $=$ stands for natural isomorphism.

The variable $A$ ranges over all subsets of the set $S$. We set (for good reasons) $F(\emptyset, X) = 1$ for all non-empty sets $X$.

Taking the sizes of both sides of (*) we obtain the equation

$$p_n(x+y) = \sum_{k\geq 0} \binom{n}{k} p_k(x) p_{n-k}(y),$$

which expresses the fact that the polynomials $p_n(x)$ are a sequence of binomial type.

Roughly speaking, condition (*) states that by "piecing togehter" two reluctant functions in the family $B$, we again obtain a reluctant function in the family. It is a generalized set-theoretic version of the binomial theorem.

Two important ways of defining binomial classes $B$ of reluctant functions are the following. Let $T$ be a family of rooted trees (it is immaterial whether they are labeled or unlabeled). The family $B(T)$ will consist of all reluctant functions whose pre-images are labeled forests on $S$ each of whose components is isomorphic to a tree in the family $T$. Clearly $B(T)$ is a binomial class of reluctant functions. In the example considered above, the family $T$ consisted of all rooted trees.

Thus, we see that the enumeration of labeled forests is closely connected with the theory of polynomials of binomial type. The family T can be specifed in innumerable ways, which will be considered in the second part of the present work. For the moment, we shall give some illustrations that show that the classical polynomials listed in Section 1 can be interpreted as enumerating binomial classes of reluctant functions. We have already seen above that the Abel polynomials can be interpreted as enumerating the binomial class of all reluctant functions, as least for $a = -1$. A somewhat more elaborate argument would show that all the other Abel polynomials, for $a$ a negative integer, enumerate other binomial classes of reluctant functions

Perhaps the simplest example is given by the sequence $x^n$. This enumerates the binomial class $B(T)$, where $T$ consists of a single tree, with one root.

Another interesting example is the sequence of Laguerre polynomials $L_n(-x)$. These enumerate the binomial class $B(T)$, where $T$ is the set of all linearly ordered rooted trees. We leave the easy verification of this fact to the reader.

A fourth example comes for the inverses of the Abel polynomials, considered in Section 10, namely, functional digraphs, enumerated by the polynomials

$$p_n(x) = \sum_{k\geq 0} \binom{n}{k} k^{n-k} x^k,$$

which do not appear at first sight to be of binomial type. We prove that they are, by showing that they enumerate a binomial class B(T). Simply take T to be the family of all rooted trees, all of whose branches have length at most two!

Given a binomial class B(T) of reluctant functions, we can consider the subclass of those functions having the property that their coimage coincides with their pre-image. We denote this subclass by $B_m(T)$, and call it the monomorphic class associated with B(T); it generalizes the notion of a one-to-one function.

The monomorphic class associated with $x^n$ consists precisely of all one-to-one functions, enumerated by the lower factorials $(x)_n$. The monomorphic class associated with the Laguerre polynomials turns out to be enumerated by the upper factorials $x^{(n)}$ (as follows from the combinatorial interpretation of $x^{(n)}$ given above).

We state without proof (but the proof is easy) an important result about monomorphic classes. If the sequence

$$p_n(x) = \sum_{k=0}^{n} a_{nk} x^k$$

enumerates the binomial class $B(T)$, then the sequence of polynomials

$$q_n(x) = \sum_{k=0}^{n} a_{nk} (x)_k$$

enumerates the monomorphic class $B_m(T)$. This fact makes formula (14) of the preceeding Section immediately obvious, and a similar interpretation can be given to (11).

The substitution of $(x)_k$ for $x^k$ is an instance of *umbral substitution*, studied generally in Section 7. It will be seen in the second part of this work that the general umbral substituions of one basic sequence into another have combinatorial interpretations in terms of "piecing together" trees and other set-theoretic operations.

These examples such suffice to orient the reader to the combinatorial aspect of the theory we are about to develop. The notion of reluctant function does not exhaust the interpretation of sequences of polynomials of binomial type. For example it does not interpret combinatorially those sequences of polynomials of binomial type which have negative or non-integral coefficients. Nevertheless, we shall see in the second part of this work that all sequences of polynomial type with non-negative coefficients can be set-theoretically (or probabilistically) interpreted by a generalization of the notion of reluctant function, whereas those with negative coefficients can be interpreted by sieving methods (Möbius inversions, etc.). There is also an obvious connection with the theory of compound Poisson processes.

Apologizing for this sketchy introduction, we proceed to begin the analytic theory.

## 3. Fundamentals.

Throughout this paper, we shall be concerned with the algebra (over a field of characteristic zero) of all polynomials in one variable, to be denoted $\underline{P}$.

By a polynomial sequence we shall denote a sequence of polynomials $p_i(x)$, $i = 0,1,2,\ldots$ where $p_i(x)$ is exactly of degree $i$, for all $i$.

A polynomial sequence is said to be of binomial type if it satisfies the infinite sequence of identities

$$p_n(x+y) = \sum_{k\geq 0} \binom{n}{k} p_k(x) p_{n-k}(y), \quad n = 0,1,2,\ldots$$

All the polynomial sequences mentioned above are of binomial type. For some sequences, such as $x^n$, this is a trivial observation, but for others, such as the Abel and Touchard polynomials, the verification that they are of binomial type will be a consequence (a rather simple one, to be sure) of our theory.

Our study will revolve primarily around the study of linear operators on $\underline{P}$ considered as a vector space. Henceforth, all operators we consider will be tacitly assumed to be linear. We denote the action of an operator $T$ or the polynomial $p(x)$ by $Tp(x)$; this notation is not strictly correct; a correct version is $(Tp)(x)$. However, this notational license results in greater readability. By way of orientation, we list some of the operators of frequent occurrence in the theory of binomial enumeration. The most important are the shift operators. A shift operator, written $E^a$, is an operator which translates the argument of a polynomial by $a$, where $a$ is an element of the field, that is, $E^a p(x) = p(x+a)$.

An operator $T$ which commutes with all shift operators is called a shift-invariant operator, i.e.,

$$TE^a = E^a T .$$

The following are important examples of shift-invariant operators:

(i) Identity operator $I: x^n \to x^n$.

(ii) Differentiation operator $D: x^n \to nx^{n-1}$.

(iii) Difference operator $\Delta = E-I: (x)_n \to n(x)_{n-1}$, where we write $E$ in place of $E^1$, where $1$ is the identity of the field.

(iv) The Abel operator $DE^a = E^a D: x(x-na)^{n-1} \to nx(x-(n-1)a)^{n-1}$.

(v) Bernoulli operator $J: p(x) \to \int_x^{x+1} p(t)dt$.

(vi) Backward difference operator $\nabla = I-E^{-1}$: $x^{(n)} \to nx^{(n-1}$

(vii) Laguerre operator $L$: $p(x) \to -\int_0^\infty e^{-t}p'(x+t)dt$.

(viii) Hermite operator $H$: $p(x) \to \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} e^{-t^2/2}p(x+t)dt$.

(ix) Central difference operator
$\delta = E^{1/2}-E^{-1/2}$: $p(x) \to p(x+ 1/2)-p(x- 1/2)$.

(x) Euler (mean) operator $M = (1/2)(I+E)$: $p(x) \to (1/2)(p(x)+ p(x+1))$.

We define a delta operator, usually denoted by the letter $Q$, as a shift-invariant operator for which $Qx$ is a non-zero constant.

The derivative, difference, backward difference, central difference, Laguerre, and Abel operators are delta operators.

Delta operators possess many of the properties of the derivative operator, as we proceed to show.

Lemma 1: If $Q$ is a delta operator, then $Qa = 0$ for every constant $a$.

Proof: Since $Q$ is shift invariant, then

$$QE^a x = E^a Qx.$$

By the linearity of $Q$,

$$QE^a x = Q(x+a) = Qx+Qa = c+Qa,$$

since $Qx$ is equal to some non-zero constant $c$ by definition. But also

$$E^a Qx = E^a c = c$$

and so $c+Qa = c$. Hence $Qa = 0$, Q.E.D.

Lemma 2: If $p(x)$ is a polynomial of degree $n$ and $Q$ is a delta operator, then $Qp(x)$ is a polynomial of degree $n-1$.

Proof: It is sufficient to prove the conclusion for the special case $p(x) = x^n$, that is, to show that the polynomial $r(x) = Qx^n$ is of degree $n-1$ (exactly). From the binomial theorem and the linearity of $Q$ we have

$$Q(x+a)^n = \sum_{k\geq 0} \binom{n}{k} a^k Qx^{n-k}.$$

Also by the shift invariance of $Q$

$$Q(x+a)^n = QE^a x^n = E^a Qx^n = r(x+a)$$

so that

$$r(x+a) = \sum_{k\geq 0} \binom{n}{k} a^k Q x^{n-k}.$$

Putting $x = 0$, we have $r$ expressed as a polynomial in $a$:

$$r(a) = \sum_{k\geq 0} \binom{n}{k} a^k [Qx^{n-k}]_{x=0}.$$

The coefficient of $a^n$ is

$$[Qx^{n-n}]_{x=0} = [Q1]_{x=0} = 0$$

by Lemma 1. Further, the coefficient of $a^{n-1}$ is

$$\binom{n}{n-1} [Qx^{n-n+1}]_{x=0} = n[Qx]_{x=0} = nc\neq 0.$$

Hence $r$ is of degree $n-1$, Q.E.D.

Let $Q$ be a delta operator. A polynomial sequence $p_n(x)$ is called the sequence of *basic polynomials* for $Q$ if:

(1) $p_o(x) = 1$

(2) $p_n(0) = 0$ whenever $n>0$

(3) $Qp_n(x) = np_{n-1}(x)$

Using Lemma 2, it is easily shown by induction that *every delta operator has a unique sequence of basic polynomials associated with it*. For example, the basic polynomials for the deriative operator are $x^n$.

We shall now see that severval properties of the polynomial sequence $x^n$ can be generalized to an arbitrary sequence of basic polynomials. The first property we noticed about $x^n$ was that it was of binomial type. This turns out to be true for every sequence of basic polynomials, and is one of our basic results.

<u>Theorem 1</u>.

(a) If $p_n(x)$ is a basic sequence for some delta operator $Q$, then it is a sequence of polynomials of binomial type.

(b) If $p_n(x)$ is a sequence of polynomials of binomial type, then it is a basic sequence for some delta operator.

<u>Proof</u>:

(a) Iterating property (3) of basic polynomials, we see that

$$Q^k p_n(x) = (n)_k p_{n-k}(x)$$

and hence that for $k = n$,

$$[Q^n p_n(x)]_{x=0} = n!$$

while

$$[Q^k p_n(x)]_{x=0} = 0, \ k<n.$$

Thus, we may express $p_n(x)$ in the following form:

$$p_n(x) = \sum_{k\geq 0} \frac{p_k(x)}{k!} [Q^k p_n(x)]_{x=0}.$$

Since any polynomial $p(x)$ is a linear combination of the basic polynomials $p_n(x)$, this expression also holds for all polynomials $p(x)$, i.e.,

$$p(x) = \sum_{k\geq 0} \frac{p_k(x)}{k!} [Q^k p(x)]_{x=0}.$$

Now suppose $p(x)$ is the polynomial $p_n(x+y)$. Then

$$p_n(x+y) = \sum_{k\geq 0} \frac{p_k(x)}{k!} [Q^k p_n(x+y)]_{x=0}.$$

But

$$\begin{aligned}[Q^k p_n(x+y)]_{x=0} &= [Q^k E^y p_n(x)]_{x=0} \\ &= [E^y Q^k p_n(x)]_{x=0} \\ &= [E^y (n)_k p_{n-k}(x)]_{x=0} \\ &= (n)_k p_{n-k}(y)\end{aligned}$$

and so

$$p_n(x+y) = \sum_{k\geq 0} \binom{n}{k} p_k(x) p_{n-k}(y)$$

which means that $p_n(x)$ is of binomial type.

(b) Conversely, suppose $p_n(x)$ is a sequence of binomial type. Putting $y = 0$ in the binomial identity, we have

$$\begin{aligned} p_n(x) &= \sum_{k\geq 0} \binom{n}{k} p_k(x) p_{n-k}(0) \\ &= p_n(x)p_o(0) + np_{n-1}(x)p_1(0) + \ldots \quad . \end{aligned}$$

Since each $p_i(x)$ is exactly of degree $i$, it follows that $p_o(0) = 1$ (and hence $p_o(x) = 1$) and $p_i(0) = 0$ for all other $i$. Thus properties (1) and (2) of basic sequences are satisfied.

We now find a delta operator for which such a sequence $p_n(x)$ is the sequence of basic polynomials. Let $Q$ be the operator defined by the property that $Qp_o(x) = 0$ and $Qp_n(x) = np_{n-1}(x)$ for $n\geq 1$. Clearly $Qx$ must be a non-zero constant. Hence all that remains to be shown is that $Q$ is shift-invariant.

As before we may trivially rewrite the generalized binomial theorem in terms of Q:

$$p_n(x+y) = \sum_{k\geq 0} \frac{p_k(x)}{k!} Q^k p_n(y)$$

and, by linearity, this may be extended to all polynomials:

$$p(x+y) = \sum_{k\geq 0} \frac{p_k(x)}{k!} Q^k p(y).$$

Now replace p by Qp and interchange x and y on the right to get

$$(Qp)(x+y) = \sum_{k\geq 0} \frac{p_k(y)}{k!} Q^{k+1} p(x).$$

But

$$(Qp)(x+y) = E^y(Qp)(x) = E^y Qp(x)$$

and

$$\begin{aligned} \sum_{k\geq 0} \frac{p_k(y)}{k!} Q^{k+1} p(x) &= Q[\sum_{k\geq 0} \frac{p_k(y)}{k!} Q^k p(x)] \\ &= Q(p(x+y)) \\ &= QE^y p(x). \end{aligned}$$

Thus we have

$$E^yQp(x) = QE^yp(x),$$

for all polynomials $p(x)$, i.e., $Q$ is shift-invariant, Q.E.D.

## 4. Expansions.

We shall study next the various ways of expressing a shift-invariant operator in terms of a delta operator and its powers. The difficulties caused by convergence questions are minimal, and we shall get around them in the easiest possible way.

Consider a sequence of shift-invariant operators $T_n$ on $\underline{P}$. We say that the sequence converges to $T$, written $T_n \to T$, if for every polynomial $p(x)$ the sequence of polynomials $T_n p(x)$ converges pointwise to the polynomial $Tp(x)$. The convergence of an infinite series of operators is to be understood accordingly.

The following theorem generalizes the Taylor expansion theorem to arbitrary delta operators and basic polynomials.

Theorem 2. (First Expansion Theorems). Let $T$ be a shift-invariant operator, and let $Q$ be a delta operator with basic set $p_n(x)$. Then

$$T = \sum_{k \geq 0} \frac{a_k}{k!} Q^k .$$

where

$$a_k = [Tp_k(x)]_{x=0} .$$

Proof: Since the polynomials $p_n(x)$ are of binomial type then, as usual, we rewrite the binomial formula as

$$p_n(x+y) = \sum_{k\geq 0} \frac{p_k(y)}{k!} Q^k p_n(x).$$

Now we may regard this as a polynomial in the variable $y$ and apply $T$ to both sides to get:

$$Tp_n(x+y) = \sum_{k\geq 0} \frac{Tp_k(y)}{k!} Q^k p_n(x).$$

Again, by linearity, this expression can be extended to all polynomials $p$. After doing this and setting $y$ equal to zero we get

$$Tp(x) = \sum_{k\geq 0} \frac{[Tp_k(y)]_{y=0}}{k!} Q^k p(x) \qquad \text{Q.E.D.}$$

Obviously, the best-known example of this Theorem is when $T = I$ and $Q = D$; then $p_n(x) = x^n$, and we have Taylor's expansion. A second example is Newton's expansion, which has three forms. If $Q = \Delta$, then $p_n(x) = (x)_k$ and the coefficients are $a_k = [T(x)_k]_{x=0}$. If $Q = \nabla$ then $p_n(x) = x^{(n)}$ and $a_k = [Tx^{(k)}]_{x=0}$. The basic polynomials for $Q = \delta = E^{1/3}-E^{-1/2}$ will be determined later.

The following remark will be used occasionally:

__Lemma:__ If $Q$ is a delta operator, and $p(x)$, $q(x)$ any polynomials, then

$$[p(Q)q(x)]_{x=0} = [q(Q)p(x)]_{x=0}.$$

__Proof:__ By linearity, we need only consider the cases when $q(x) = p_k(x)$ and $p(Q) = Q^n$, where $p_k(x)$ are the basic polynomials of $Q$. But it is easy to see that the relation holds in this case. Q.E.D.

As a further example of the use of the expansion theorem, we derive the classical Newton-Cotes formulas of numerical integration. We wish to find an expansion, in terms of $\Delta$, of the Bernoulli operator $J_r$ defined by:

$$J_r p(x) = \int_x^{x+r} p(t)dt.$$

Noting that $J_r$ is a shift-invariant, we have the identities

$$\begin{aligned} J_r &= \frac{(I+\Delta)^r-1}{\Delta} \cdot \frac{\Delta}{D} \\ &= \frac{(I+\Delta)^r-I}{\Delta} \cdot J_1 \end{aligned}$$

which reduces the problem to finding an expansion of $J_1$ in terms of $\Delta$. Using the First Expansion Theorem, this is fairly simple:

$$J = \sum_{k\geq 0} \frac{a_k}{k!} \Delta^k$$

where

$$a_k = [J(x)_k]_{x=0} = \int_0^1 (x)_k dx$$

where we note that the $a_k$ are the Bernoulli numbers of the second kind. $J_2$ evaluated in this way gives Simpson's rule:

$$\int_x^{x+2h} p(t)dt = 2h(1+\Delta+ \frac{1}{6} \Delta^2 + \frac{1}{180} \Delta^4 + \frac{1}{180} \Delta^5 + \ldots)p(x).$$

A final example is the classical Euler's transformation

$$\sum_{k\geq 0} (-1)^k f(k) = 1/2 \sum_{n\geq 0} \frac{(-1)^n}{2^n} \Delta^n f(0)$$

Which follows from the identities:

$$\begin{aligned} \sum_{k>0} (-1)^k E^k &= \frac{I}{I+E} \\ &= \frac{I}{2I+\Delta} \\ &= 1/2 \frac{I}{I+ 1/2 \Delta} \\ &= 1/2 \sum_{n\geq 0} \frac{(-1)^n}{2^n} \Delta^n. \end{aligned}$$

Of course, in this case we are disregarding convergence questions.

We now turn our attention to the Abel polynomials. The delta operator in this case is $E^aD$. Thus, the Abel polynomials are basic polynomials and hence are of binomial type. Therefore, by Thm. 1 we have proved Abel's identity:

$$(x+y)(x+y-na)^{n-1} = \sum_{k\geq 0} \binom{n}{k} x(x-ka)^{k-1} y(y-(n-r)a)^{n-k-1},$$

not easily proved by direct methods. We can use the Expansion Theorem to get an Abel expansion of $e^x$. Indeed, we do get the following beautiful expansion

$$e^x = \sum_{k\geq 0} \frac{x(x-ka)^{k-1}}{k!} e^{ka},$$

convergent for $a<0$.

Theorem 3. Let $Q$ be a delta operator, and let $F$ be the ring of formal power series in the variable $t$, over the same field. Then there exists an isomorphism from $F$ onto the ring $\Sigma$ of shift-invariant operators, which carries

$$f(t) = \sum_{k\geq 0} \frac{a_k}{k!} \text{ into } \sum_{k\geq 0} \frac{a_k}{k!} Q^k.$$

**Proof:** The mapping is already linear and by the Expansion Theorem, it is onto. Therefore, all we have to verify is that the map preserves products. Let $T$ be the shift-invariant operator corresponding to the formal power series $f(t)$ and let $S$ be the shift-invariant operator corresponding to

$$g(t) = \sum_{k\geq 0} \frac{b_k}{k!} t^k.$$

We must verify that

$$[TSp_n(x)]_{x=0} = \sum_{k\geq 0} \binom{n}{k} a_k b_{n-k}$$

where $p_n(x)$ are the basic polynomials of $Q$. Now

$$[TSp_r(x)]_{x=0} = [(\sum_{k\geq 0} \frac{a_k}{k!} Q^k \sum_{n\geq 0} \frac{b_n}{n!} Q^n) p_r(x)]_{x=0}$$

$$= [\sum_{k\geq 0} \sum_{n\geq 0} \frac{a_k b_n}{k!n!} Q^{k+n} p_r(x)]_{x=0}.$$

But $p_n(0) = 0$ for $n>0$ and $p_0(x) = 1$. Hence, it follows that the only non-zero terms of the double sum occur when $n = r-k$. Thus

$$[TSp_r(x)]_{x=0} = [\sum_{k\geq 0} \frac{a_k b_{r-k}}{k!(r-k)!} Q^r p_r(x)]_{x=0}$$

$$= [\sum_{k\geq 0} \frac{a_k b_{r-k}}{k!(r-k)!} r! p_0(x)]_{x=0}$$

$$= \sum_{k\geq 0} \binom{r}{k} a_k b_{r-k}, \qquad \text{Q.E.D.}$$

Corollary 1. A shift-invariant operator $T$ is invertible if and only if $T1\neq 0$.

In the following, we shall write $P = p(Q)$, where $P$ is a shift-invariant operator and $p(t)$ is a formal power series, to indicate that the operator $P$ corresponds to the formal power series $p(t)$ under the isomorphism of Theorem 3. Note that $p(0) = 0$ and $p'(0)\neq 0$ whenever $P$ is a shift-invariant delta operator. For such formal power series, a unique inverse formal power series $p^{-1}(t)$ exists.

Corollary 2. Let $Q$ be a delta operator with basic polynomials $p_n(x)$, and let $q(D) = Q$. Let $q^{-1}(t)$ be the inverse formal power series. Then

$$\sum_{n\geq 0} \frac{p_n(x)}{n!} u^n = e^{xq^{-1}(u)}.$$

Proof: Expand $E^a$ in terms of $Q$. The coefficient $a_n$ are $p_n(a)$. Hence

$$\sum_{n\geq 0} \frac{p_n(a)}{n!} Q^n = E^a,$$

a formula which can be considered as a generalization of Taylor's formula, and which specializes (for example for $Q = \Delta$ it gives Newton's expansion) to several classical expansions. Now use the Isomorphism Theorem, with $D$ as the delta operators. We get

$$\sum_{n\geq 0} \frac{p_n(a)}{n!} q(t)^n = e^{at},$$

whence the conclusion, upon setting $u = q(t)$ and $a = x$, Q.E.D.

As an aside, we remark at this point a possibly useful connection between basic polynomials and orthogonal polynomials:

Proposition. Let $p_n(x)$, $n = 0,1,2,\ldots$ be a sequence of polynomials of binomial type. Then there exists a unique inner product $(p(x), q(x))$, on the vector space $\underline{P}$ of all polynomials $p(x)$, under which the sequence $p_n(x)$ is an orthogonal sequence and $(p_n(x), p_n(x)) = n!$. Under this inner product we have

$$[Q^n p(x)]_{x=0} = (p(x), p_n(x)/\sqrt{n!}$$

so that

$$p(x) = \sum_{n\geq 0}^{\infty} \frac{p_n(x)}{n!} [Q^n p(x)]_{x=0} = \sum_{n\geq 0} \frac{p_n(x)}{n!} (p(x),p_n(x)).$$

Proof: Let $T$ be the (uniquely defined)operator mapping $p_n(x)$ to $x^n$, for all $n$. Define the inner product as follows:

$$(p(x),q(x)) = [(Tp)(Q)q(x)]_{x=0}.$$

An argument similar to the proof of the Lemma preceding Theorem 2 shows that this bilinear form is symmetric (set $p(x) = p_n(x)$ and $q(x) = p_k(x)$), and that $p_n(x)$ is orthogonal to $p_k(x)$ for $k\neq n$. Finally,

$$(p_n(x),p_n(x)) = [(Tp_n)(Q)p_n(x)]_{x=0} =$$

$$[Q^n p_n(x)]_{x=0} = n!,$$

which shows that the bilinear form is positive definite.

It is trivially verified that $[Q^n p(x)]_{x=0} = (p(x),p_n(x))/\sqrt{n!}$. Thus the Expansion Theorem, in the form

$$q(a) = [E^a q(x)]_{x=0} = \sum_{n\geq 0} \frac{q_n(a)}{n!} [Q^n q(x)]_{x=0}$$

is the same as the orthogonal expansion of $q(x)$ relative to the above inner product, Q.E.D.

We note that for the Laguerre polynomials, discussed below, the inner product just introduced reduces to the classical inner product making the Laguerre polynomials an orthogonal set.

Note that for the operators (i), (ii), (iii), (iv), (vi), (vii), (ix) described at the beginning of this Section the polynomials defined there are the basic sets, as will be shown in the course of this study.

## 5. Closed Forms.

We now introduce a class of linear operators of an altogether different kind. Let $p(x)$ be a polynomial in the parameter $x$. Multiplying each term of $p(x)$ by a factor $x$, i.e., replacing each occurrence of $x^n$ by $x^{n+1}$, $n \geq 0$, -- we obtain a new polynomial in $x$ which we may denote $xp(x)$. The first $x$ in this expression may be regarded as a linear operator since it represents a linear transformation of polynomial into polynomials. We call this the multiplication operator and we denote it by the parameter $\underline{x}$ underlined. Thus, $\underline{x}: p(x) \mapsto xp(x)$. Note that the operator $\underline{x}$ is not shift-invariant.

Before proceeding further, it should be noted that $E^a p(x) = p(x+a)$ is a polynomial in the formal parameter $x+a$. Since the multiplication operator is not shift-invariant, we have the operator identity:

$$E^a \underline{x} = (\underline{x+a}) E^a,$$

where $\underline{x+a}: p(x) \mapsto (x+a)p(x)$.

Proposition 1. If $T$ is a shift-invariant operator, then

$$T' = T\underline{x} - \underline{x}T$$

is also a shift-invariant operator.

The proof is a straightforward verification. We call $T'$ the Pincherle derivative of the operator $T$.

We saw in the previous Section, as a special case of the Expansion Theorem, that any shift-invariant operator $T$ can be expressed as an expansion in the delta operator $D$, i.e., $T = \sum_{k\geq 0} \frac{a_k}{k!} D$ where $a_k = [Tx^k]_{x=0}$. Further, by the isomorphism theorem, (Theorem 3) the formal power series corresponding to $T$ is $\sum_{k\geq 0} \frac{a_k}{k!} t^k = f(t)$. We call $f(t)$ the indicator of $T$.

Proposition 2. If $T$ has indicator $f(t)$, then $T'$ has $f'(t)$ as its indicator.

Proof: Straightforward verification of coefficients by Theorem 3.

We note in passing Pincherle's Formula:

$$Tx^n p(x) = \sum_{k\geq 0} \binom{n}{k} x^{n-k} T^{(k)} p(x).$$

Note that by the isomorphism theorem of the preceding Section, we also have

$$(TS)' = T'S+TS'.$$

Proposition 3. $Q$ is a delta operator if and only if $Q = DP$ for some shift-invariant operator $P$, where $P^{-1}$ exists.

Proof: If $Q$ is a delta operator, then it can be written

$$Q = \sum_{k\geq 0} \frac{a_k}{k!} D^k$$

where

$$a_k = [Qx^k]_{x=0}.$$

But

$$a_0 = [Q1]_{x=0} = 0$$

$$a_1 = [Qx]_{x=0} \neq 0$$

by definition of a delta operator. Thus if we set

$$P = \sum_{k\geq 0} \frac{a_{k+1}}{(k+1)!} D^k$$

then the conclusion follows at once.

Conversely, suppose $Q = DP$ where $P$ is shift-invariant and $P^{-1}$ exists. Since $D$ and $P$ are shift-invariant, then $Q$ must be also. Further, shift-invariant operators commute (by Theorem 3), so that

$$Qx = DPx = PDx = P1 \neq 0,$$

since $P1 \neq 0$ for an invertible shift-invariant operator. Hence $Q$ is a delta operator. Q.E.D.

Theorem 4 (Closed forms for basic polynomials). If $p_n(x)$ is a sequence of basic polynomials for the delta operator $Q = DP$, then

(1) $p_n(x) = Q'P^{-n-1}x^n$

(2) $p_n(x) = P^{-n}x^n - (P^{-n})x^{n-1}$

(3) $p_n(x) = xP^{-n}x^{n-1}$

(4) (Rodrigues-type formula) $p_n(x) = x(Q')^{-1}p_{n-1}(x)$.

Proof: We shall first show that (1) and (2) define the same polynomial sequence:

$$Q'P^{-n-1} = (DP)'P^{-n-1}$$

$$= (D'P+DP')P^{-n-1}.$$

Thus, if we can show that $q_n(0) = 0$ for $n>0$, we will complete the proof that $q_n(x)$ is the sequence of basic polynomials for $Q$, and it will follow that they will satisfy formulas (1), (2), and (3). Now, from the equivalence of equations (1), (2), (3) we see that

$$q_n(x) = xP^{-n}x^{n-1}$$

and hence $q_n(0) = 0$ for $n \geq 0$. Thus (1), (2), and (3) have been proven, and $q_n(x) = p_n(x)$.

To prove (4), we first invert formula (1), getting:

$$x^n = (Q')^{-1}P^{n+1}p_n(x).$$

Notice that $Q'$ is invertible, as is easily verified. Inserting this into the right side of formula (3) we get:

$$\begin{aligned} p_n(x) &= xP^{-n}(Q')^{-1}P^n p_{n-1}(x) \\ &= x(Q')^{-1}p_{n-1}(x) \end{aligned}$$

which is the Rodrigues-type formula, Q.E.D.

The following formulas, numbered (5) and (6), relate the basic polynomials of two different delta operators in an analogous way. Their proof is immediate.

Corollary. Let $R = DS$ and $Q = DP$ be delta operators with basic polynomials $r_n(x)$ and $p_n(x)$, respectively, where $S^{-1}$ and $P^{-1}$ exists. Then:

$$(5) \qquad p_n(x) = Q'(R')^{-1}P^{-n-1}S^{n+1}r_n(x)$$

$$(6) \qquad p_n(x) = x(RQ^{-1})^n x^{-1} r_n(x).$$

Example 1. The Abel polynomials are the basic polynomials of the Abel operator $E^aD$. Indeed from formula (3):

$$\begin{aligned} p_n(x) &= xE^{-an}x^{n-1} \\ &= x(x-an)^{n-1}. \end{aligned}$$

Example 2. The lower factorials $(x)_n$ are the sequence of basic polynomials for the lower difference operator $\Delta = E-I$. Since $\Delta' = E$, the Rodrigues formula (4) gives immediately

$$p_n(x) = xE^{-1}p_{n-1}(x)$$

which by iteration gives the lower factorial power $(x)_n$.

Note that the basic polynomials for the central difference operator $\delta$ can be obtained from (6) and $\Delta$ much as the Abel polynomial were obtained from (3).

## 6. The Automorphism Theorem.

Let $\mathcal{A}(\underline{P})$ be the algebra of all linear operators on the algebra of all polynomials $\underline{P}$. Let $\Sigma$ be the subalgebra of shift-invariant operators on $\underline{P}$. We now prove our main result.

Theorem 5. Let $T$ be an operator in $\mathcal{A}(\underline{P})$, not necessarily shift-invariant. Let $P$ and $Q$ be delta operators with basic polynomials $p_n(x)$ and $q_n(x)$, respectively. Assume that

$$Tp_n(x) = q_n(x), \quad \text{for all } n \geq 0,$$

then $T^{-1}$ exists and

(a) the map $S \to TST^{-1}$ is an automorphism of the algebra $\Sigma$.

(b) $T$ maps every sequence of basic polynomials into a sequence of basic polynomials.

(c) Let $P = p(D)$ and $Q = q(D)$, where $p(t)$ and $q(t)$ are formal power series. Let the delta operator $R$ have formal power series expansion $r(D)$ and basic polynomials $r_n(x)$. Then

$$Tr_n(x) = s_n(x)$$

is a sequence of basic polynomials for the delta operator

$$S = r(p^{-1}(q(D))),$$

where $p^{-1}$ is the inverse formal power series of $p(t)$, that is $p(p^{-1}(t)) = p^{-1}(p(t)) = t$.

<u>Proof</u>:

(a) We have the string of identities:

$$\begin{aligned} TPp_n(x) &= T(np_{n-1}(x)) \\ &= nTp_{n-1}(x) \\ &= nq_{n-1}(x) \\ &= Qq_n(x) \\ &= QTp_n(x) \end{aligned}$$

and since every polynomial is a linear combination of the basic polynomials, by linearity, we infer that $TPp(x) = QTp(x)$ for all polynomials $p(x)$, that is, $TP = QT$. It is clear that T is invertible, since it maps polynomials of degree $n$ into polynomials of degree $n$, for all $n$. Hence

$$TPT^{-1} = Q$$

whence

$$TP^nT^{-1} = Q^n$$

for all $n \geq 0$. Let $S$ be any shift-invariant operator and let the expansion of $S$ in terms of $P$ be

$$S = \sum_{n \geq 0} \frac{a_n}{n!} P^n.$$

Then

$$TST^{-1} = T\left(\sum_{n \geq 0} \frac{a_n}{n!} P^n\right)T^{-1} = \sum_{n \geq 0} \frac{a_n}{n!} Q^n \qquad \text{(I)}$$

and thus $TST^{-1}$ is a shift-invariant operator. Furthermore the map $S \mapsto TST^{-1}$ is onto since any shift-invariant operator can be expanded in terms of $Q$. Thus, the map is an automorphism, as claimed.

Remark: We have also shown that $T$ maps delta operators into delta operators, since for delta operators the constant co-efficient $a_0$ vanishes.

(b) Let $s_n(x) = Tr_n(x)$ and let $S = TRT^{-1}$. By the remark above, $S$ is a delta operator since $R$ is.

Also

$$\begin{aligned} Ss_n(x) &= TRT^{-1}s_n(x) \\ &= TRr_n(x) \\ &= nTr_{n-1}(x) \\ &= ns_{n-1}(x). \end{aligned}$$

To complete the proof that $s_n(x)$ are the basic polynomials of $S$ we need only show that $s_n(0) = 0$ for $n>0$. Now we can write

$$r_n(x) = \sum_{k\geq 1} a_k p_k(x)$$

since $a_0 = 0$ because $R$ is a delta operator, and hence $r_n(0) = 0$. Hence

$$Tr_n(x) = \sum_{k\geq 1} a_k q_k(x) = s_n(x)$$

so that

$$s_n(0) = 0, \quad n>0,$$

as desired.

(c) Now $Q$ and $R$ can be written as power series in $P$, say $Q = f(P)$ and $R = g(P)$. In equation (I) above let

$$R = g(P) = \sum_{k\geq 0} \frac{a_n}{n!} P^n;$$

then

$$S = TRT^{-1} = g(Q)$$

and therefore

$$R = g(P) = g(p(D))$$

and

$$S = g(Q) = g(q(D)).$$

Finally we see that

$$r(D) = g(p(D))$$

$$g(D) = r(p^{-1}(D))$$

and

$$S = g(Q) = r(p^{-1}(q(D))), \qquad \text{Q.E.D.}$$

## 7. Umbral Notation.

In order to simplify the complex notation which has been appearing in many of the above formulas, we will make use and for the first time make rigorous the "umbral calculus" or "symbolic notation" first devised by Sylvester and later used informally by many authors. If $\{a_n(x)\}$ is a polynomial sequence then we simply note that there is a unique linear operator $L$ on $\underline{P}$ such that $L(x^n) = a_n(x)$. We say that $L$ is the umbral representation of the sequence $\{a_n(x)\}$. In particular, an operator $T$ with the properties specified in the preceding Theorem will be called an umbral operator.

If $f(x)$ is a polynomial then we use the notation $f(\underline{a}(x))$ to denote the image of $f(x)$ under the operator $L$. For example, $\underline{a}(x)$ denotes $a_1(x)$, while $[\underline{a}(x)]^2$ denotes $a_2(x)$. Similarly, $[\underline{a}(x)+b][\underline{a}(x)+c]$ denotes $a_2(x)+(b+c)a_1(x)+bc$. This is in essence the umbral notation, which we signify by boldface lettering.

Loosely speaking, umbral notation is a simple technique using exponents to denote subscripts. For example, the defining property for a polynomial sequence to be of binomial type

$$p_n(x+y) = \sum_{k\geq 0} \binom{n}{k} p_k(x) p_{n-k}(y)$$

can be restated umbrally as

$$\underline{p}^n(x+y) = [\underline{p}(x)+\underline{p}(y)]^n.$$

Note that, in view of our definition in terms of the operator L, this identity has a well-defined meaning.

Theorem 6. If P and Q are delta operators with basic sequences $p_n(x)$ and $q_n(x)$, and expansions $P = p(D)$ and $Q = q(D)$, then the umbral composition

$$r_n(x) = p_n(\underline{q}(x))$$

is the sequence of basic polynomials for the delta operator

$$R = p(q(D)).$$

Proof: Let T be the umbral operator defined by

$$Tx^n = q_n(x).$$

By the Automorphism Theorem of the preceeding Section, it follows that T takes any basic sequence into another basic sequence. Now if

$$p_n(x) = \sum_{i=0}^{n} a_i x^i$$

then

$$\begin{aligned} Tp_n(x) &= T(\sum_{i=0}^{n} a_i x^i) \\ &= \sum_{k=0}^{n} a_i Tx^i \\ &= \sum_{k=0}^{n} a_i q_i(x) \\ &= p_n(\underline{q}(x)). \end{aligned}$$

Thus $r_n(x)$ is a sequence of basic polynomials and by the Automorphism Theorem, it is the basic sequence for

$$R = TPT^{-1} = p(q(D)), \qquad \text{Q.E.D.}$$

Corollary: If $p_n(x)$ is a sequence of basic polynomials then there exists a basic sequence $q_n(x)$ such that

$$p_n(\underline{q}(x)) = x^n.$$

We say that $q_n(x)$ is the inverse sequence of $p_n(x)$.

Theorem 7. (Summation Formula). Suppose $p_n(x)$ and $q_n(x)$ are the basic sequences for the delta operators P and Q respectively. If $q_n(x)$ is inverse to $p_n(x)$, then

$$p_n(x) = \sum_{k\geq 0} \frac{x^k}{k!} [Q^k x^n]_{x=0}.$$

The proof is similar to the preceeding and is left to the reader.

We are now in a position to solve the problem stated in the Introduction: given basic sequences $p_n(x)$ and $q_n(x)$, with delta operators $P = p(D)$ and $Q = q(D)$, how are the coefficients $c_{nk}$

$$q_n(x) = \sum_{k\geq 0} c_{nk} p_k(x)$$

linking the $p_n(x)$ to the $q_n(x)$, the so-called *connection constants*, to be determined? The answer is dismayingly simple. Consider the polynomials

$$r_n(x) = \sum_{k\geq 0} c_{nk} x^k,$$

and consider the umbral operator $T$ defined by

$$Tx^n = p_n(x).$$

Then clearly

$$q_n(x) = Tr_n(x) = r_n(\underline{p}(x)),$$

so that $r_n(x)$ are of binomial type and $R = r(D)$ being their delta operator, we find $q(t) = r(p(t))$, or $r(t) = q(p^{-1}(t))$. Theorem 4 then provides explicit expressions for the $r_n(x)$. One couldn't expect a simpler answer.

As an example, consider the connection constants between $q_n(x) = x^n$ and $p_n(x) = (x)_n$. Here $q(t) = t$ and $p(t) = e^t-1$. Thus, $r(t) = \log(1+t)$ and, as we shall see below, the polynomials $r_n(x)$ turn out to be the exponential polynomials $\varphi_n(x)$, discussed below.

As a second example, let $p_n(x) = (x)_n$ and $q_n(x) = x^{(n)}$. An easy computation shows that $r(t) = t/(t-1)$, whose basic polynomials are the Laguerre polynomials, also discussed below.

An instructive example the reader may work out for himself — thereby obtaining a number of classical and new identities, is to take $p_n(x) = x(x-na)^{n-1}$ and $q_n(x) = x(x-nb)^{n-1}$ for $a \neq b$. These examples could be multiplied ad infinitum, and a great number of combinatorial identities in the literature can be seen to fall into the simple pattern we have just outlined.

Remark. It can be shown that every automorphism of the algebra $\Sigma$ is of the form $S \to TST^{-1}$ for some umbral operator $T$, but this fact will not be needed, so we omit the proof.

## 8. The Exponential Polynomials.

The exponential polynomials, studied by Touchard and other authors, are a good testing ground for the theory developed so far. We shall see that their basic properties and the identities they satisfy are almost trivial consequences of the theory.

Consider the sequence of lower factorials $(x)_n$, which as we have seen is the basic sequence for the delta operator $\Delta = e^D - I$. In this case the inverse sequence is the sequence of basic polynomials for the operator $Q = \log(I+D)$. We denote these polynomials by $\varphi_n(x)$; these are the exponential polynomials.

From the Corollary above we have umbrally

$$\underline{\varphi}(\underline{\varphi}-1)(\underline{\varphi}-2)\dots(\underline{\varphi}-n+1) = x^n.$$

Further by the summation formula

$$\varphi_n(x) = \sum_{k\geq 0} \frac{x^k}{k!}[\Delta^k x^k]_{x=0}$$

$$= \sum_{k\geq 0} S(k,n)x^k,$$

where $S(n,k)$ denote the Stirling numbers of the second kind.

Now let us apply the Rodrigues formula to see what we get. Since $Q = \log(I+D)$, we have $Q' = (I+D)^{-1}$ and hence

$$\varphi_n(x) = x(Q')^{-1}\varphi_{n-1}(x)$$

$$= x(I+D)\varphi_{n-1}(x)$$

$$= x\varphi_{n-1}(x)+x\varphi'_{n-1}(x),$$

which is the recursion formula for the exponential polynomials.

The next property of these exponential polynomials which we shall prove is expressed umbrally as

$$\varphi_{n+1}(x) = x(\underline{\varphi}+1)^n.$$

Let $T$ be the umbral operator that takes $(x)_n$ into $x^n$, so that $Tx^n = \varphi_n(x)$. Hence

$$Tx(x-1)_{n-1} = xx^{n-1}$$

or changing $n$ to $n+1$

$$Tx(x-1)_n = xx^n = xT(x)_n$$

can be rewritten as

$$TxE^{-1}(x)_n = xT(x)_n.$$

We can extend this by linearity to all polynomials $p(x)$ so that

$$TxE^{-1}p(x) = xTp(x).$$

Replacing $p(x)$ by $p(x+1)$ we have

$$TxE^{-1}p(x+1) = xTp(x+1).$$

Hence $$Txp(x) = xTp(x+1).$$

Since $Tx^n = \varphi_n(x)$ then $Tp(x) = p(\underline{\varphi}(x))$ and it follows that

$$\begin{aligned} \underline{\varphi}(x)p(\underline{\varphi}(x)) &= Txp(x) \\ &= xTp(x+1) \\ &= xp(\underline{\varphi}(x)+1). \end{aligned}$$

If we let $p(x) = x^n$ then

$$\varphi_{n+1}(x) = [\underline{\varphi}(x)]^{n+1} = \underline{\varphi}(x)[\underline{\varphi}(x)]^{n}$$

$$= x[\varphi(x)+1]^{n}$$

which is what we wanted to prove.

In a similar vein one can prove the remarkable Dobinský-type formula:

$$\varphi_n(x) = e^{-x} \sum_{k\geq 0} \frac{x^k k^n}{k!} ,$$

which we shall leave as an exercise to the reader.

## 9. Laguerre Polynomials.

As a further example of the above theory, we shall develop some properties of the Laguerre polynomials. The Laguerre operator $L$ is defined by

$$Lp(x) = -\int_0^\infty e^{-x}p'(x+t)dt.$$

It is a delta operator and as such has a sequence of basic polynomials which we shall call $L_n(x)$. By straightforward calculation , we find that the expansion of $L$ in terms of $D$ has coefficients

$$[Lx^n]_{x=0} = n!, \ n\geq 1$$

$$= 0, \ n=0$$

and hence we find that

$$L = \frac{D}{D-I} .$$

Hence from formula (3) of Theorem 4 we have

$$(*) \qquad L_n(x) = x(D-I)^n x^{n-1}.$$

Since for all polynomial $p(x)$ we also have

$$e^{x}De^{-x}p(x) = e^{x}(e^{-x}p'(x)-e^{-x}p(x))$$

$$= (D-I)p(x)$$

then $e^{x}De^{-x} = D-I$ and hence

$$e^{x}D^{n}e^{-x} = (D-I)^{n}.$$

Therefore we obtain the classical Rodrigues formula,

$$L_n(x) = xe^{x}D^{n}e^{-x}x^{n-1}.$$

From formula (*) we find by binomial expansion that

$$L_n(x) = \sum_{k=1}^{n} \frac{n!}{k!} \binom{n-1}{k-1} (-x)^k$$

where the coefficients

$$\frac{n!}{k!} \binom{n-1}{k-1}$$

are known as the Lah numbers. Our notation for the polynomials $L_n$ corresponds to the notation in Bateman for the polynomials $L_n^{(-1)}$, that is,

$$L_n^{(-1)}(x) = \frac{1}{n!} L_n(x).$$

We now come to the most important fact about the Laguerre polynomials. The indicator of L is

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

and hence

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

Thus, by the Automorphism Theorem we infer that the Laguerre polynomials are a self-inverse set. Thus, we have as an immediate consequence the beautiful identity

$$x^n = \sum_{k=1}^{n} \frac{n!}{k!} \binom{n-1}{k-1} (-1)^k L_k(x) = L_n(\underline{L}(x)).$$

Other identities concerning Laguerre polynomials stem from the fact that

$$L = \frac{D}{D-I} .$$

Since $L_n(x)$ are the basic polynomials of L we have

$$\frac{D}{D-I} L_n(x) = nL_{n-1}(x)$$

and hence the classical recursion formula

$$L_n'(x) = n(D-I)L_{n-1}(x).$$

In fact, if we expand $\frac{D}{D-I}$ into series form

$$\frac{D}{D-I} = -D \cdot \frac{I}{I-D} = -D(I+D+D^2+\dots)$$

we can use this to get other known recursion formulas.

## 10. A Glimpse of Combinatorics.

Although we intend to leave most of the combinatorial applications of the preceeding theory to the second part of this work, we shall outline two typical results which we hope will orient the reader to applications to problems of enumeration, typical of the second part of this work.

Theorem 8. Let $P$ be an invertible shift-invariant operator. Let $p_n(x)$ be a sequence of basic polynomials satisfying

$$[x^{-1}p_n(x)]_{x=0} = n[P^{-1}p_{n-1}(x)]_{x=0},$$

for all $n>0$. Then $p_n(x)$ is the sequence of basic polynomials for the delta operator $Q = DP$.

Proof: Define the operator $Q$ by $Q1 = 0$ and

$$Qp_n(x) = np_{n-1}(x)$$

and extending by linearity. Note that $Q$ is shift-invariant. In terms of $Q$, the preceding identity can be rewritten in the form

$$[x^{-1}p_n(x)]_{x=0} = [P^{-1}Qp_n(x)]_{x=0}.$$

By linearity, this extends to an identity for all polynomials $p(x)$ — an argument we have often used in this work. Thus, recalling that $[x^{-1}p(x)]_{x=0} = [Dp(x)]_{x=0}$ whenever $p(0) = 0$, we have

$$[Dp(x)]_{x=0} = [P^{-1}Qp(x)]_{x=0}$$

for all polynomials $p(x)$, including those for which $p(0) \neq 0$. Setting $p(x) = q(x+a)$ we obtain, using the shift-invariance of $P$ and $Q$,

$$\begin{aligned} Dq(a) &= [P^{-1}QE^{a}q(x)]_{x=0} \\ &= [E^{a}P^{-1}Qq(x)]_{x=0} \\ &= P^{-1}Qq(a) \end{aligned}$$

for all constants $a$. But this means that $D = P^{-1}Q$, or $Q = DP$, Q.E.D.

<u>Corollary 1</u>. Given any sequence of constants $c_{n1}$, $n = 1,2,\ldots$, there exists a unique sequence of basic polynomials $p_n(x)$ such that $[x^{-1}p_n(x)]_{x=0} = c_{n1}$, that is,

$$p_n(x) = \sum_{k\geq 1} c_{nk}x^{k}, \; n = 1,2,\ldots .$$

Corollary 2. Let $g(x)$ be the indicator of $Q$ in the above. Then $g = f^{-1}$, where $f(t) = \sum_{k\geq 0} c_{k,1} \frac{t^k}{k!}$.

Proof: From above

$$D = QP^{-1} = \sum_{k\geq 0} c_{k,1} \frac{Q^k}{k!} = f(Q)$$

and the result follows.

We now give some applications of the above theory.

Application 1. Let $t_{n,k}$ be the number of forests of rooted labeled trees (i.e., trees with a distinguished vertex) with $n$ vertices and $k$ components, then

$$A_n(x) = \sum_{k\geq 0} t_{n,k} x^k = x(x+n)^{n-1}.$$

Proof: Since $t_{n,1}$ is the number of rooted trees on $n$ vertices, then $t_{n,1} = nA_{n-1}(1)$ since each such tree on $n$ vertices may be obtained by mapping a forest on $n-1$ vertices onto a single new root vertex. The resulting root may be labeled in $n$ ways, i.e., either by using a new symbol or by using one of the $n-1$ old symbols and replacing it by the new symbol. But this relation may be written

$$[x^{-1} A_n(x)]_{x=0} = n[EA_{n-1}(x)]_{x=0}$$

and hence the delta operator for $A_n$ is $DE^{-1}$ by Theorem 8. Thus the associated polynomials are the Abel polynomials $x(x+n)^{n-1}$.

Corollary (Cayley). The number of labeled trees on $n$ vertices is $n^{n-2}$.

Proof: Since the number of rooted labeled trees is $n^{n-1}$ the number of unrooted trees is $n^{n-2}$ since each free tree can be labeled in $n$ ways.

Application 2. Let $S_n$ be a symmetric group on $n$ symbols and let $c_{n,k}$ be the number of group elements which consist of precisely $k$ cycles. If $C_n(x) = \sum_{k\geq 0} c_{n,k}x^k$ then $C_n(x) = x^{(n)}$.

Proof: We note that in this case $C_{n,1} = (n-1)!$ which is clearly the number of group elements consisting of just one cycle, and thus by Corollary 2 this is the required sequence.

Functional Digraphs. A digraph, D, (with loops permitted) on $n$ symbols is a functional digraph if and only if it satisfies the following two postulates,

1) each component of D contains precisely one consistently directed circuit; and

2) each non-circuit edge is directed towards the circuit contained in its component.

An idempotent is a functional digraph all of whose components contain a distinguished vertex which meets every edge of that component.

<u>Application 3</u>. The polynomial $p_n(x) = \sum_{k\geq 0} \binom{n}{k} k^{n-k} x^k$ is of binomial type. Let $h_{n,k}$ be the number of idempotent on $n$ symbols with precisely $k$ components. Then $h_{n,k} = \binom{n}{k} k^{n-k}$ since the $k$ distinguished vertices, $V$, may be chosen in $\binom{n}{k}$ ways and the remaining $n-k$ points may be directed into $V$ in $k^{n-k}$ ways. However, we may also view each idempotent as a structure generated by its components. It is interesting to note that the coefficients $h_{n,1} = n$ and the associated delta operator has indicator $f^{[-1]}(t)$ where $f(t) = te^t$. Thus these polynomials are the inverse sequence of the Abel polynomials. Several identities for them may be derived in much the same way as we related the exponential polynomials to the lower factorials in Section 6.

Anticipating some developments in the second part of this paper, we may state the following principle. In order to enumerate by a sequence $c_{n1}$ a class of rooted trees, graded by the number of vertices, one forms the associated basic set, which will enumerate a class of reluctant functions, and then proceedsto apply Theorem 8 or a variant of it, which will reflect the "composition rule" of such class of trees. The connection constants between two polynomial sequences enumerating sequences of reluctant functions have a combinatorial interpretation in terms of "piecing together" one set of trees in terms of another. Thus our starting point in the second part of this work will be: given two families $F_1$ and $F_2$ of rooted labeled forests, in how many ways can be a member of $F_2$ be "pieced together" from members of $F_1$? The simplest case of this is Cayley's theorem above, where $F_1$ consists of a single edge and $F_2$ consists of all labeled rooted forests.

## 11. Bibliography

Rota, G.-C., On the Foundations of Combinatorial Theory, I, Theory of Möbius functions, Zeit. für. Wahr., 2 (1964), 340-368.

Crapo, H.H. and Rota, G.-C., On the Foundation of Combinatorial Theory, II, Combinatorial Geometries, M.I.T. Press (to appear).

Boas, R.P. and Buck, R.C., Poynomial Expansions of Entire Functions, Springer, New York, 1964.

Martin, W.T., On Expansions in Terms of a Certain General Class of Functions, American J. Math., 58 (1936), 407-420.

Pringsheim, A., Zur Geschichte des Taylorschen Lehrsatzes, Bibliotheca Math., (3) 1 (1900), 433-479.

Steffensen, J.F., The Poweroid, an Extension of the Mathematical Notion of Power, Acta Mathematica, 73 (1941), 333-366.

Touchard, J., Nombres Exponentiels et Nombres de Bernoulli, Canad, J. Math., 8 (1956), 305-320.

Riordan, J., Combinatorial Identities, Wiley, New York, 1968.

Al-Salaam, W.A., Operational Expressions for the Laguerre and other Polynomials, Duke Math J., 31 (1964), 127-142.

Carlitz, L. and Riordan, J., The Divided Central Differences of Zero, Canadian J. Math., 15 (1963), 94-100.

Hurwitz, A., Ueber Abel's Verallgemeinerung der binomischen Formel, Acta Math., 26 (1902), 199-203.

Lah, I., Eine neue Art von Zahlen, ihre Eigneschaften ... Mittelungsbl. Math. Stat., 7 (1955), 203-216.

Riordan, J., Inverse Relations and Combinatorial Identities, American Math. Monthly, 71 (1964), 485-498.

Rota, G.-C., The Number of Partitions of a Set, American Math. Monthly, 71 (1964), 498-504.

Moon, J.W., Counting Labelled Trees, A Survey of Methods and Results, A publication of the Department of Mathematics of the University of Alberta, 1969.

Pincherle, S. and Amaldi, Le Operazioni Distributive, Zanichelli, Bologna, 1900.

Pincherle, S., Opere Scalte, 2 vols., Cremonese, Rome, 1954.

Sylvester, J.J., Collected Mathematical Papers, Cambridge, 1904-1912, 4 vols.

Bell, E.T., Postualational Bases for the Umbral Calculus, American J. Math., 62 (1940), 717-724.

Harary, F., A Seminar in Graph Theory, Holt, New York, 1967.

Hall, Marshall, Combinatorial Theory, Blaisdell, Waltham 1967.

Harary, Frank, Graph Theory, Addison-Wesley, Reading, 1969.

Los Alamos Scientific Laboratory
Brookhaven National Laboratory
The Rand Corporation
University of Waterloo
University of North Carolina
University of Colorado
Florida Atlantic University
The Rockefeller University
Massachusetts Institute of Technology