

FUNCTIONAL COMPOSITION PATTERNS AND POWER SERIES REVERSION

BY

GEORGE N. RANEY

1. **Introduction.** In 1859, Cayley [7] solved the problem of finding the number of different products of given terms in a given order, under a non-associative multiplication. More recent references to the same problem are found in the writings of Jacobson [9], Becker [2], Motzkin [11], and Bourbaki [3; 4]. This paper will be concerned with a natural generalization of Cayley's problem, and will show that the solution to the generalized problem contains all of the combinatorial information needed to establish the well known formula of Lagrange for the reversion of power series.

To describe the problem, we consider expressions which are built from operator symbols and argument symbols, using a prefix notation for operators. Weights are assigned to the symbols in an expression, an argument symbol having the weight 0 and an n -ary operator symbol having the weight n . Expressions are of various types, the type of an expression depending only on the weights of the symbols in it and on the order in which they appear. The expression $(a+b)+c$, for example, is written $++abc$ and is of the type 22000, while $a+(b+c)$ is written $+a+bc$ and is of the type 20200. The expression $F(G(x, H(y, z), t), K(u))$ is of the type 230200010. Following P. C. Rosenbloom [13], we call those finite sequences of natural numbers which designate the types of expressions "words." Definitions and some special properties of these sequences are stated in §2.

The general problem mentioned above can now be stated: Let a_1, a_2, \dots be a sequence of natural numbers in which only a finite number of terms are different from zero. How many words are there in which there are exactly a_1 occurrences of 1, a_2 occurrences of 2, etc.? Since for words the number of occurrences of 0 is determined by the formula $a_0 = 1 + \sum_{i=1}^{\infty} (i-1)a_i$, it is not necessary to mention a_0 in stating the problem.

It is advantageous to enumerate not just single words, but *lists of words* having a specified composition and this is done in Theorem 2.2 below. Theorems 2.3 and 2.4 express some properties of the solution to this enumeration problem, and form the basis for the rest of the paper.

In §3 and §4 the combinatorial information developed in §2 is expressed in the form of equations which hold in the large algebra of a certain semi-group. In §5 the Lagrange inversion formula is obtained from these equations by using a suitable homomorphism from this large algebra to the algebra of

Received by the editors April 6, 1959.

formal power series. The results of §4 are applied in §6 to the solution of equations by series. In §7 the results of §2 are applied to some convolution formulas.

2. The number of lists of words having a specified composition. A finite sequence of the natural numbers $0, 1, 2, \dots$ will be called a *string*. The number of terms in the sequence will be called the *length* of the string. If the string σ has length m , the set of natural numbers which are less than m will be called the *domain* of σ (written $D(\sigma)$); we shall consider σ to be a function from this domain into the set of natural numbers. For each $k \in D(\sigma)$, let g_k be the permutation of $D(\sigma)$ which is defined as follows: if $x \in D(\sigma)$ and $x+k < m$, then $g_k(x) = x+k$, and if $x \in D(\sigma)$ and $x+k \geq m$, then $g_k(x) = x+k-m$. The permutations g_k form a cyclic group whose order is m , the length of σ . For each $k \in D(\sigma)$, the composite function σg_k is a string whose length is m ; the string σg_k will be called a *cyclic variant* of the string σ .

If σ is a string and i is a natural number, define $p_i(\sigma)$ to be the number of $x \in D(\sigma)$ for which $\sigma(x) = i$; then $\sum_{i=0}^{\infty} p_i(\sigma)$ is the length of σ . Define $r(\sigma)$, the *rank* of σ , to be $\sum_{i=0}^{\infty} (i-1)p_i(\sigma)$. If σ' is a cyclic variant of σ , then $r(\sigma') = r(\sigma)$.

If σ and τ are strings, define $\sigma\tau$ to be the string which results from the concatenation of σ and τ . For every natural number i , $p_i(\sigma\tau) = p_i(\sigma) + p_i(\tau)$; hence $r(\sigma\tau) = r(\sigma) + r(\tau)$. If $\gamma = \sigma\tau$ then σ is called a *head* of γ ; if τ is not empty, then σ is called a *proper head* of γ .

Certain strings are called *words*. The string of length 1 whose only term is 0 is a word. If n is a natural number other than 0 and if $\alpha_1, \alpha_2, \dots, \alpha_n$ are words, then the string $n\alpha_1\alpha_2 \dots \alpha_n$ is a word. P. C. Rosenbloom has shown [13] that for a string γ to be a word it is necessary and sufficient that $r(\gamma) = -1$ while for every proper head σ of γ , $r(\sigma) \geq 0$.

Certain strings are called *lists of words*. The empty string is a list of 0 words. For $n \geq 1$, if α is a word and β is a list of $n-1$ words, then $\alpha\beta$ is a list of n words. Using Rosenbloom's result one can prove inductively that for a string γ to be a list of n words it is necessary and sufficient that $r(\gamma) = -n$, while for every proper head σ of γ , $r(\sigma) > -n$.

THEOREM 2.1. *If n is a natural number other than zero and if σ is a string whose rank is $-n$, then there are exactly n natural numbers k in the domain of σ such that the string σg_k is a list of n words.*

Proof. Of all the heads of all the cyclic variants of σ let α be one whose rank is maximum and which has maximum length within that rank. Let σ_1 be a cyclic variant of σ such that α is a head of σ_1 .

Let τ be a proper head of σ_1 and suppose that $r(\tau) \leq -n$. If τ is a head of α , then for some γ , $\alpha = \tau\gamma$, and $r(\alpha) = r(\tau) + r(\gamma) \leq -n + r(\gamma)$. Since n is greater than zero, $r(\alpha) < r(\gamma)$. However, γ is a head of a cyclic variant of σ , so that this result would contradict the fact that α has maximum rank. Therefore τ is

not a head of α and $\sigma_1 = \alpha\delta\epsilon$ where $\alpha\delta = \tau$ and ϵ is not empty. Then $-n = r(\sigma) = r(\sigma_1) = r(\tau) + r(\epsilon) \leq -n + r(\epsilon)$. It follows that $0 \leq r(\epsilon)$, and $r(\alpha) \leq r(\epsilon) + r(\alpha) = r(\epsilon\alpha)$. Since $\epsilon\alpha$ is a head of a cyclic variant of σ , namely $\epsilon\alpha\delta$, and since $\epsilon\alpha$ is longer than α , either α does not have maximum rank or α does not have maximum length within that rank. The supposition that $r(\tau) \leq -n$ has led to a contradiction; thus it is proved that if τ is a proper head of σ_1 , then $r(\tau) > -n$.

There is a shortest head of σ_1 whose rank is -1 . Call it α_1 and let $\sigma_1 = \alpha_1\beta_1$. If β_i has been defined and has rank $i - n < 0$, then let α_{i+1} be the shortest head of β_i whose rank is -1 and let $\beta_i = \alpha_{i+1}\beta_{i+1}$. Continue in this manner until β_n has been defined, obtaining $\sigma_1 = \alpha_1\alpha_2 \cdots \alpha_n\beta_n$. Since $r(\alpha_1 \cdots \alpha_n) = r(\alpha_1) + \cdots + r(\alpha_n) = -n$, $\alpha_1 \cdots \alpha_n$ cannot be a proper head of σ_1 ; hence β_n is empty, and $\sigma_1 = \alpha_1 \cdots \alpha_n$. Let $\sigma_1 = \alpha_1\alpha_2 \cdots \alpha_{n-1}\alpha_n$, $\sigma_2 = \alpha_2 \cdots \alpha_{n-1}\alpha_n\alpha_1$, \cdots , $\sigma_n = \alpha_n\alpha_1\alpha_2 \cdots \alpha_{n-1}$. For $i = 1, \cdots, n$, each α_i is a word and each σ_i is a list of n words. Even though the σ_i are not necessarily all distinct, they are the strings σg_k for n distinct natural numbers k in the domain of σ .

If there is still another natural number k' in the domain of σ such that $\sigma g_{k'}$ is a list of n words, then $\sigma g_{k'}$ has the form $\beta\gamma\delta$ where β and δ are both non-empty and $\delta\beta = \alpha_i$ for some i . Then $r(\delta) \geq 0$, since α_i is a word. It follows that $r(\beta\gamma) \leq -n$, and this contradicts the assumption that $\sigma g_{k'}$ is a list of n words. This completes the proof of Theorem 2.1.

Let a_1, a_2, \cdots be an infinite sequence of natural numbers of which at most a finite number of terms are different from zero. We will say that the string σ has *composition* $(n; a_1, a_2, \cdots)$ if $r(\sigma) = -n$ and $p_i(\sigma) = a_i$ for $i = 1, 2, \cdots$.

THEOREM 2.2. *Let $L(n; a_1, a_2, \cdots)$ be the number of lists of n words having composition $(n; a_1, a_2, \cdots)$. Then*

$$(2.1) \quad mL(n; a_1, a_2, \cdots) = nM(a_0, a_1, a_2, \cdots),$$

where $m = n + \sum_{i=1}^{\infty} ia_i$, and $a_0 = n + \sum_{i=1}^{\infty} (i-1)a_i$, and $M(a_0, a_1, a_2, \cdots)$ is the multinomial coefficient $(\sum_{i=0}^{\infty} a_i)! / (\prod_{i=0}^{\infty} (a_i!))^{-1}$. In the case $m = n = 0$, $L(n; a_1, a_2, \cdots) = L(0; 0, 0, \cdots) = 1$.

Proof. Let S be the set of strings with composition $(n; a_1, a_2, \cdots)$ and let W be the set of lists of n words with composition $(n; a_1, a_2, \cdots)$. For every $\sigma \in S$, $p_0(\sigma) = a_0$, and, since $m = \sum_{i=0}^{\infty} a_i$, all strings in S have the length m . The number of strings in S is $M(a_0, a_1, a_2, \cdots)$. Let G be the cyclic group of permutations g_k for $0 \leq k < m$, and let P be the set of ordered pairs (g, σ) such that $g \in G$, $\sigma \in S$, and $\sigma g \in W$. By Theorem 2.1, each member of S is the second element of n different ordered pairs in P , if $n > 0$. Therefore, P has $nM(a_0, a_1, a_2, \cdots)$ members, if $n > 0$.

Since $L(n; a_1, a_2, \cdots)$ denotes the number of members of W , the Cartesian product $G \times W$ has $mL(n; a_1, a_2, \cdots)$ members. For $g \in G$, $\lambda \in W$,

define $\theta(g, \lambda) = (g, g^{-1}\lambda)$. Then θ is a one-one mapping from $G \times W$ onto P . Therefore $mL(n; a_1, a_2, \dots) = nM(a_0, a_1, a_2, \dots)$, if $n > 0$. If $n = 0$ and $a_i = 0$ for $i = 1, 2, \dots$, then $L(n; a_1, a_2, \dots) = 1$. If $n = 0$ and $a_i \neq 0$ for some $i > 0$, then $L(n; a_1, a_2, \dots) = 0$. This completes the proof of Theorem 2.2.

The problem of Cayley mentioned above is equivalent to that of finding $L(n; a_1, a_2, \dots)$ in the special case $n = 1$, $a_i = 0$ for $i \neq 2$. Application of Theorem 2.2 in this case leads to the result that $L(1; 0, k, 0, 0, \dots) = (k+1)^{-1}C_{2k, k}$.

THEOREM 2.3. *For all natural numbers m and n ,*

$$(2.2) \quad L(m+n; a_1, a_2, \dots) = \sum L(m; b_1, b_2, \dots) L(n; c_1, c_2, \dots),$$

where the summation extends over all ordered pairs of sequences b_1, b_2, \dots and c_1, c_2, \dots of natural numbers such that $b_i + c_i = a_i$ for $i = 1, 2, \dots$.

Proof. Let α be a list of $m+n$ words having composition $(m+n; a_1, a_2, \dots)$. Then there is exactly one head of α which is a list of m words. For let $\alpha = \beta\gamma\delta$ where β is a list of m words and γ is not empty. Then β is a proper head of $\beta\gamma$ having rank $-m$, so that $\beta\gamma$ cannot be a list of m words. If $\alpha = \tau\sigma$ where τ is a list of m words, then σ is a list of n words. If τ has composition $(m; b_1, b_2, \dots)$ and σ has composition $(n; c_1, c_2, \dots)$, then $b_i + c_i = a_i$ for $i = 1, 2, \dots$.

THEOREM 2.4. *For all natural numbers k ,*

$$(2.3) \quad L(k+1; a_1, a_2, \dots) = L(k; a_1, a_2, \dots) + \sum L(k+i; a_1, a_2, \dots, a_i-1, \dots),$$

where the summation extends over all natural numbers $i > 0$ for which $a_i > 0$.

Proof. If α is a list of $k+1$ words having composition $(k+1; a_1, a_2, \dots)$ then either $\alpha = 0\beta$ where β is a list of k words having composition $(k; a_1, a_2, \dots)$ or for some natural number $i > 0$, $\alpha = i\beta\gamma$ where β is a list of i words and γ is a list of k words, so that $\beta\gamma$ is a list of $k+i$ words having composition $(k+i; a_1, a_2, \dots, a_i-1, \dots)$.

3. Large semigroup algebras and induced homomorphisms.

DEFINITION 3.1. Let S be a semigroup which satisfies the following condition (D): For every $s \in S$, there exist only finitely many couples (t, u) of elements of S such that $t+u=s$. Let A be a ring which is commutative and has a unit element, e . Let A^S be the set of functions $f: S \rightarrow A$, provided with the structure of an algebra by the following rules: if $s \in S$, $f, g \in A^S$, and $a \in A$, then

$$(1) \quad (f+g)(s) = f(s) + g(s),$$

$$(2) \quad (af)(s) = af(s),$$

$$(3) \quad (fg)(s) = \sum f(t)g(u),$$

where the summation extends over the set of ordered pairs (t, u) such that

$t+u=s$. Following Bourbaki [5], we call A^S the large algebra of the semigroup S relative to the ring A .

DEFINITION 3.2. Let I be an indexing set, and let $F: I \rightarrow A^S$. If for every $s \in S$, the set of $i \in I$ such that $(F(i))(s) \neq 0$ is finite, then F is called a *summable family of elements of A^S* . If F is a summable family, then the function $(\sum F): S \rightarrow A$ is defined as follows: $(\sum F)(s) = \sum (F(i))(s)$, where the summation extends over the set of $i \in I$ such that $(F(i))(s) \neq 0$, and the sum is zero in case this set is empty. We may sometimes write $\sum_{i \in I} F(i)$ for $\sum F$.

DEFINITION 3.3. Let S and T be semigroups satisfying condition (D). Let A and B be commutative rings with units, and let B have no divisors of zero. Let the function $\theta: S \rightarrow T$ be such that if $x_1, x_2 \in S$, then $\theta(x_1 + x_2) = \theta(x_1) + \theta(x_2)$. Let the function $\beta: S \rightarrow B$ be such that if $x_1, x_2 \in S$, then $\beta(x_1 + x_2) = \beta(x_1)\beta(x_2)$. Let θ and β be such that for every $y \in T$, the set of $x \in S$ such that $\theta(x) = y$ and $\beta(x) \neq 0$ is finite. Let the function $h: A \rightarrow B$ be such that if $a_1, a_2 \in A$, then $h(a_1 + a_2) = h(a_1) + h(a_2)$ and $h(a_1 a_2) = h(a_1)h(a_2)$. Then define $H = H(\theta, \beta, h)$ as follows: $H: A^S \rightarrow B^T$ and if $f \in A^S$, $y \in T$, then $(Hf)(y) = \sum \beta(x)h(f(x))$, where the summation extends over the set of $x \in S$ such that $\theta(x) = y$ and $\beta(x) \neq 0$, and the sum is zero in case this set is empty.

With these definitions it can be established in a straightforward manner that H is a summability preserving homomorphism from A^S into B^T .

LEMMA 3.1. *In the notation of the preceding three definitions,*

- (1) $H(f+g) = (Hf) + (Hg)$, for $f, g \in A^S$,
- (2) $H(af) = h(a)(Hf)$, for $a \in A, f \in A^S$,
- (3) $H(fg) = (Hf)(Hg)$, for $f, g \in A^S$,
- (4) *If I is an indexing set and $F: I \rightarrow A^S$ is a summable family of elements of A^S , then the composite function $(HF): I \rightarrow B^T$ is a summable family of elements of B^T , and $H(\sum F) = \sum (HF)$.*

4. The special element L^* . Let P be the set of sequences of natural numbers of which at most a finite number of terms are different from zero. Let $\bar{p}, \bar{q} \in P$, $\bar{p} = (p_1, p_2, \dots)$, $\bar{q} = (q_1, q_2, \dots)$, and define $\bar{p} + \bar{q} = (p_1 + q_1, p_2 + q_2, \dots)$. With this addition operation, P is a semigroup satisfying condition (D), and if A is a commutative ring with unit, e , one can define the large algebra A^P .

Some special elements of P are the identity element $\bar{0} = (0, 0, \dots)$ and the "Kronecker delta" elements \bar{d}_i for $i \geq 1$, all of whose terms are zero except the i th term, which is 1.

Some special elements of A^P are the zero element 0^* , defined by $0^*(\bar{p}) = 0$ for every $\bar{p} \in P$, the unit element E_0^* , defined by $E_0^*(\bar{p}) = 0$ for $\bar{p} \neq \bar{0}$ and $E_0^*(\bar{0}) = e$. Other special elements of A^P are the elements E_i^* for $i > 0$, defined by $E_i^*(\bar{p}) = 0$ if $\bar{p} \neq \bar{d}_i$, $E_i^*(\bar{d}_i) = e$.

Powers of an element f^* of A^P are defined recursively in the usual way: $(f^*)^0 = E_0^*$, $(f^*)^{m+1} = (f^*)^m f^*$.

In terms of the operations on A^P the results of Theorems 2.3 and 2.4 can be restated in the following form:

THEOREM 4.1. *Let L^* be the element of A^P defined by*

$$L^*(\bar{p}) = L(1; p_1, p_2, \dots)e.$$

Then

$$(L^*)^n(\bar{p}) = L(n; p_1, p_2, \dots)e.$$

For any $\bar{p} \in P$, there exist only finitely many natural numbers i such that $(E_i^(L^*)^i)(\bar{p}) \neq 0$, and $L^*(\bar{p}) = \sum (E_i^*(L^*)^i)(\bar{p})$. In short,*

$$(4.1) \quad L^* = \sum_{i=0}^{\infty} E_i^*(L^*)^i.$$

Proof. The first assertion is proved by induction on n , using Theorems 2.2 and 2.3. The second assertion then follows from Theorem 2.4 and the observation that $(E_i^*(L^*)^i)(\bar{p}) = (L^*)^i(\bar{p} - \bar{d}_i)$ if $p_i > 0$ and $(E_i^*(L^*)^i)(\bar{p}) = 0$ if $p_i = 0$.

5. The Lagrange inversion formula. Under addition the set N of natural numbers $0, 1, 2, \dots$ is a semigroup satisfying the condition (D) of Definition 3.1. If A is a commutative ring with unit, e , then the large algebra A^N is the algebra of formal power series with coefficients in A . Special elements of A^N are the zero element $\bar{0}$, the unit element $\bar{1}$, and the element \bar{x} , defined respectively by $\bar{0}(n) = 0$ for all n , $\bar{1}(n) = 0$ for $n \neq 0$, $\bar{1}(0) = e$, $\bar{x}(n) = 0$ for $n \neq 1$ and $\bar{x}(1) = e$. Powers of the element \bar{x} satisfy the equations $\bar{x}^j(j) = e$, $\bar{x}^j(n) = 0$ for $n \neq j$.

Every element $\bar{z} \in A^N$ can be expressed in the form $\bar{z} = \sum_{n=0}^{\infty} a_n \bar{x}^n$, where $a_n = \bar{z}(n)$ and the summation is as in Definition 3.2. If \bar{z} is replaced by \bar{z}^m , this becomes $\bar{z}^m = \sum_{n=0}^{\infty} \bar{z}^m(n) \bar{x}^n$, where, by the Multinomial Theorem,

$$(5.1) \quad \bar{z}^m(n) = \sum M(p_0, p_1, p_2, \dots) a_0^{p_0} a_1^{p_1} a_2^{p_2} \dots,$$

the summation extending over all sequences (p_0, p_1, p_2, \dots) such that $\sum_{i=0}^{\infty} p_i = m$ and $\sum_{i=0}^{\infty} i p_i = n$.

If $\bar{z} \in A^N$ then $D\bar{z}$, the derivative of \bar{z} , is the element of A^N defined by $(D\bar{z})(n) = (n+1)\bar{z}(n+1)$. Higher derivatives satisfy the equation

$$(5.2) \quad (D^k(\bar{z}))(n) = (n+k)!(n!)^{-1}\bar{z}(n+k).$$

From a theorem on formal power series which corresponds to the implicit function theorem (see Bourbaki [6]) it follows that for each sequence a_0, a_1, \dots of elements of A there is exactly one $\bar{w} \in A^N$ such that

$$(5.3) \quad \bar{w} = \bar{x} \sum_{n=0}^{\infty} a_n \bar{w}^n.$$

The Lagrange inversion formula (see Pólya-Szegő [12]) describes this \bar{w} in the following way: Let

$$(5.4) \quad \bar{z} = \sum_{n=0}^{\infty} a_n \bar{x}^n.$$

Then

$$(5.5) \quad \bar{w} = \sum_{n=1}^{\infty} \frac{(D^{n-1}\bar{z}^n)(0)}{n!} \bar{x}^n.$$

Here it will be shown that this results directly from (4.1) by use of a suitably induced summability preserving homomorphism from A^P into A^N .

To see the connection between the problem of enumerating lists of words and the Lagrange inversion formula, note that by (5.2) and (5.1),

$$\frac{(D^{n-1}\bar{z}^n)(0)}{n!} = \frac{1}{n} \bar{z}^n(n-1) = \frac{1}{n} \sum M(p_0, p_1, p_2, \dots) a_0^{p_0} a_1^{p_1} \dots,$$

where the summation extends over all sequences (p_0, p_1, \dots) such that $\sum_{i=0}^{\infty} p_i = n$ and $\sum_{i=0}^{\infty} i p_i = n-1$. Then by Theorem 2.2,

$$(5.6) \quad \frac{(D^{n-1}\bar{z}^n)(0)}{n!} = \sum L(1; p_1, p_2, \dots) a_0^{p_0} a_1^{p_1} \dots,$$

where $p_0 = 1 + \sum_{i=1}^{\infty} (i-1)p_i$ and the summation extends over all $\bar{p} = (p_1, p_2, \dots) \in P$ such that $\sum_{i=1}^{\infty} i p_i = n-1$.

Let a_0, a_1, a_2, \dots be the sequence of elements of A which is given in (5.3). Let $\theta: P \rightarrow N$ be defined by $\theta(\bar{p}) = \sum_{i=1}^{\infty} i p_i$ where $\bar{p} = (p_1, p_2, \dots)$. Let $\beta: P \rightarrow A$ be defined by

$$\beta(\bar{p}) = a_0^{\sum_{i=1}^{\infty} (i-1)p_i} a_1^{p_1} a_2^{p_2} \dots.$$

Referring to Definition 3.3, let $B = A$ and let h be the identity mapping from A to A . By Lemma 3.1, the function H of Definition 3.3 is a summability preserving homomorphism from A^P into A^N .

Looking at the effect of H on some special elements, $HE_j^* = a_0^{j-1} a_j \bar{x}^j$, if $j \geq 1$. For

$$\begin{aligned} (HE_j^*)(n) &= \sum_{\sum i p_i = n} \beta(\bar{p}) E_j^*(\bar{p}) = \begin{cases} \beta(\bar{d}_j) e, & \text{if } n = j, \\ 0, & \text{if } n \neq j \end{cases} \\ &= \begin{cases} a_0^{j-1} a_j e, & \text{if } n = j, \\ 0, & \text{if } n \neq j \end{cases} = (a_0^{j-1} a_j \bar{x}^j)(n). \end{aligned}$$

From the fact that H is a homomorphism, $HE_0^* = \bar{1}$.

By (4.1), $L^* = \sum_{j=0}^{\infty} E_j^*(L^*)^j$. By Lemma 3.1, $HL^* = \sum_{j=0}^{\infty} HE_j^*(HL^*)^j = \bar{1} + \sum_{j=1}^{\infty} a_0^{j-1} a_j \bar{x}^j (HL^*)^j$. Multiplying both members by $a_0 \bar{x}$, $a_0 \bar{x} (HL^*) = \bar{x} \sum_{j=0}^{\infty} a_j (a_0 \bar{x} (HL^*))^j$. Defining $\bar{w} = a_0 \bar{x} (HL^*)$,

$$(5.7) \quad \bar{w} = \bar{x} \sum_{j=0}^{\infty} a_j \bar{w}^j.$$

Let $E_{\bar{p}}^*$ denote the element of A^P such that $(E_{\bar{p}}^*)(\bar{p}) = e$, $(E_{\bar{p}}^*)(\bar{q}) = 0$ if $\bar{q} \neq \bar{p}$. Then $L^* = \sum_{\bar{p} \in P} L(1; \bar{p}) E_{\bar{p}}^*$, and by Lemma 3.1, $HL^* = \sum_{\bar{p} \in P} L(1; \bar{p}) H E_{\bar{p}}^*$. The element $E_{\bar{p}}^*$ can be expressed as a finite product of powers of the E_i^* , so that

$$\begin{aligned} HL^* &= \sum_{\bar{p} \in P} L(1; \bar{p}) H \left(\prod_{p_i \neq 0} (E_i^*)^{p_i} \right) \\ &= \sum_{\bar{p} \in P} L(1; \bar{p}) \prod (H E_i^*)^{p_i} \\ &= \sum_{\bar{p} \in P} L(1; \bar{p}) \prod (a_0^{i-1} a_i \bar{x})^{p_i} \\ &= \sum_{\bar{p} \in P} L(1; \bar{p}) a_0^{\sum (i-1)p_i} a_1^{p_1} a_2^{p_2} \cdots \bar{x}^{\sum i p_i} \\ &= \sum_{r=0}^{\infty} \bar{x}^r \left(\sum_{\sum i p_i = r} L(1; \bar{p}) a_0^{\sum (i-1)p_i} a_1^{p_1} a_2^{p_2} \cdots \right). \end{aligned}$$

Letting $p_0 = 1 + \sum_{i=1}^{\infty} (i-1)p_i$ and multiplying both members by $a_0 \bar{x}$, $\bar{w} = a_0 \bar{x} (HL^*) = \sum_{r=0}^{\infty} \bar{x}^{r+1} \left(\sum_{1 + \sum i p_i = r+1} L(1; \bar{p}) a_0^{p_0} a_1^{p_1} a_2^{p_2} \cdots \right)$. Replacing $r+1$ by n and using (5.6),

$$(5.8) \quad \bar{w} = \sum_{n=1}^{\infty} \frac{(D^{n-1} \bar{z}^n)(0)}{n!} \bar{x}^n.$$

Equations (5.4), (5.7), and (5.8) together constitute the Lagrange inversion formula.

6. The solution of equations by series. The memoir of Lagrange [10] in which the inversion formula was originally developed dealt with the problem of expressing the roots of equations by means of power series in the coefficients. The link between this problem and the problem of enumerating lists of words is found in the equation of Theorem 4.1:

$$(6.1) \quad L^* = \sum_{i=0}^{\infty} E_i^*(L^*)^i.$$

On one hand, (6.1) springs directly from the enumeration problem, since it is a corollary of Theorems 2.2, 2.3, and 2.4. On the other hand, (6.1) is an equation of fixpoint type whose solution is

$$(6.2) \quad L^* = \sum_{\bar{p} \in P} L(1; \bar{p})(E_1^*)^{p_1}(E_2^*)^{p_2} \cdots$$

Formal solutions of various equations of fixpoint type may be obtained from (6.1) and (6.2) by using homomorphisms from A^P to other large algebras of semigroups satisfying the condition (D).

For the trinomial equation

$$(6.3) \quad \bar{w} = 1 + \bar{x}\bar{w}^n,$$

for example, a homomorphism $H: A^P \rightarrow A^N$ different from that of the preceding section leads to the formal solution

$$(6.4) \quad \bar{w} = \sum_{k=0}^{\infty} L(1; k\bar{d}_n)\bar{x}^k = \sum_{k=0}^{\infty} \frac{1}{1 + (n-1)k} \binom{nk}{k} \bar{x}^k.$$

To induce this homomorphism, let $A = B$, let h be the identity mapping on A , and define $\theta: P \rightarrow N$ and $\beta: P \rightarrow A$ by specifying their values for the "Kroncker delta" elements \bar{d}_i , $i \neq 0$. Let $\theta(\bar{d}_i) = 0$ for $i \neq n$, $\theta(\bar{d}_n) = 1$. Let $\beta(\bar{d}_i) = 0$ for $i \neq n$, $\beta(\bar{d}_n) = e$. Then $H = H(\theta, \beta, h)$ has the following properties: $HE_0^* = \bar{1}$, $HE_j^* = 0$ for $j \neq 0$, n , $HE_n^* = \bar{x}$. Setting $\bar{w} = HL^*$, (6.1) goes over into (6.3) and (6.2) goes into (6.4). Theorem 2.2 provides the evaluation of the $L(1; k\bar{d}_n)$.

In like manner, using homomorphisms from A^P into algebras of formal double series, triple series, etc., one can obtain formal solutions of polynomial equations of fixpoint type having two, three, or more variable coefficients. This is closely related to work of Appell and Kampé de Fériet [1] on the solution of algebraic equations by general hypergeometric functions.

One can also apply the information contained in (6.1) and (6.2) to the algebra of formal Dirichlet series, or to the algebra of formal power series with coefficients which are integers modulo a prime.

7. Convolution formulas.

THEOREM 7.1. *Let A be a commutative ring with unit, and let N be the set of natural numbers $0, 1, 2, \dots$. Let the function $a: N \times N \rightarrow A$ have the property that*

$$(7.1) \quad a(m_1 + m_2, n) = \sum_{n_1 + n_2 = n} a(m_1, n_1)a(m_2, n_2), \quad \text{for } m_1, m_2, n \in N.$$

Let $b: N \times N \rightarrow A$ be defined by

$$(7.2) \quad b(m, n) = m(m+n)^{-1}a(m+n, n)$$

when m and n are not both zero, $b(0, 0) = 1$. Then

$$(7.3) \quad b(m_1 + m_2, n) = \sum_{n_1 + n_2 = n} b(m_1, n_1)b(m_2, n_2), \quad \text{for } m_1, m_2, n \in N.$$

Proof. The $a(m, n)$ are completely determined by the $a(1, n)$, using (7.1).

Let $a_n = a(1, n)$ for $n=0, 1, 2, \dots$, and as in (5.4) let $\bar{z} = \sum_{n=0}^{\infty} a_n \bar{x}^n$. By induction $\bar{z}^m = \sum_{n=0}^{\infty} a(m, n) \bar{x}^n$, and by the Multinomial Theorem, $a(m, n) = \sum M(p_0, p_1, p_2, \dots) a_0^{p_0} a_1^{p_1} a_2^{p_2} \dots$, the summation extending over all sequences (p_0, p_1, p_2, \dots) such that $\sum_{i=0}^{\infty} p_i = m$ and $\sum_{i=0}^{\infty} i p_i = n$. Then by (7.2), $b(m, n) = \sum m(m+n)^{-1} M(p_0, p_1, p_2, \dots) a_0^{p_0} a_1^{p_1} a_2^{p_2} \dots$, the summation extending over all sequences (p_0, p_1, p_2, \dots) such that $\sum_{i=0}^{\infty} p_i = m+n$ and $\sum_{i=0}^{\infty} i p_i = n$. By Theorem 2.2,

$$b(m, n) = \sum L(m; p_1, p_2, \dots) a_0^{p_0} a_1^{p_1} a_2^{p_2} \dots$$

A computation using Theorem 2.3 completes the proof of (7.3).

Expressing this in terms of generating functions (see [12]), if $\phi(\bar{x}) = \sum_{n=0}^{\infty} a_n \bar{x}^n$, $a_0 \neq 0$, then $(\phi(\bar{x}))^m = \sum_{n=0}^{\infty} a(m, n) \bar{x}^n$, and if $\psi(\bar{y}) = \phi(\bar{x})$ where $\bar{y} = \bar{x}(\phi(\bar{x}))^{-1}$, then $(\psi(\bar{y}))^m = \sum_{n=0}^{\infty} b(m, n) \bar{y}^n$.

If the function $c: A \times N \rightarrow A$ is now defined by the equation

$$(7.4) \quad \exp(\alpha \bar{x}) = \exp(\alpha \bar{y} \psi(\bar{y})) = \sum_{k=0}^{\infty} c(\alpha, k) \bar{y}^k,$$

it is found that

$$(7.5) \quad c(\alpha, k) = \sum_{m=0}^k \frac{\alpha^m b(m, k-m)}{m!} = \sum_{m=0}^k \frac{m}{k} \frac{\alpha^m a(k, k-m)}{m!}$$

and that

$$(7.6) \quad c(\alpha_1 + \alpha_2, n) = \sum_{n_1+n_2=n} c(\alpha_1, n_1) c(\alpha_2, n_2)$$

for $\alpha_1, \alpha_2 \in A, n \in N$.

Recently, H. W. Gould [8] has given generalized versions of the Vandermonde and Abel convolution formulas, and has suggested a single convolution formula which includes both of these as special cases. The formula reads: Let

$$(7.7) \quad \begin{aligned} G(\alpha, 0; \beta, \gamma) &= 1, \\ G(\alpha, n; \beta, \gamma) &= \frac{\alpha}{n!} \prod_{m=1}^{n-1} (\alpha + \beta m - \gamma m) \end{aligned}$$

for $\alpha, \beta, \gamma \in A, n \in N, n \geq 1$. The empty product is taken to be 1, so that $G(\alpha, 1; \beta, \gamma) = \alpha$. Then

$$(7.8) \quad G(\alpha_1 + \alpha_2, n; \beta, \gamma) = \sum_{n_1+n_2=n} G(\alpha_1, n_1; \beta, \gamma) G(\alpha_2, n_2; \beta, \gamma)$$

for $\alpha_1, \alpha_2, \beta, \gamma \in A, n \in N$.

Setting $\gamma=0$, $G(\alpha, n; \beta, 0) = \alpha/(\alpha+\beta n) \cdot (\alpha+\beta n)^n/(n!)$ and (7.8) becomes Abel's convolution. Setting $\gamma=1$, $G(\alpha, n; \beta, 1) = \alpha/(\alpha+\beta n) \cdot C_{\alpha+\beta n, n}$, and (7.8) becomes a convolution of the Vandermonde type.

To prove (7.8), let

$$(7.9) \quad \phi(x) = \phi(x; \beta, \gamma) = \frac{\gamma x e^{\beta x}}{e^{\gamma x} - 1}.$$

Let $z^{1/\delta} = e^{\gamma x}$. Then $1 - z^{1/\delta} + w z^{(1/\delta)(\beta/\gamma)} = 0$, where $w = \gamma x / \phi(x) = \gamma y$. Then (see Pólya-Szegő [12, Problem 212]),

$$z = 1 + \sum_{n=1}^{\infty} \binom{\delta + (\beta/\gamma)n - 1}{n-1} \frac{\delta}{n} w^n,$$

and letting $\alpha = \delta\gamma$,

$$z = e^{\alpha x} = 1 + \sum_{n=1}^{\infty} \frac{\alpha \gamma^{n-1}}{n} \binom{(\alpha/\gamma) + (\beta/\gamma)n - 1}{n-1} y^n.$$

By (7.4) we have that

$$c(\alpha, n) = \frac{\alpha}{n} \gamma^{n-1} \binom{(\alpha/\gamma) + (\beta/\gamma)n - 1}{n-1} = G(\alpha, n; \beta, \gamma)$$

for all n , and from (7.6) we obtain (7.8).

REFERENCES

1. P. Appell and J. Kampé de Fériet, *Fonctions hypergéométriques et hypersphériques, polynômes d'Hermite*, Paris, 1926, p. 400.
2. H. W. Becker, *Discussion of Problem 4277*, Amer. Math. Monthly vol. 56 (1949) pp. 697-699.
3. N. Bourbaki, *Théorie des ensembles*, Actualités Sci. Ind., no. 1212, 1954, pp. 50-53.
4. ———, *Théorie des ensembles*, Actualités Sci. Ind., no. 1243, 1956, p. 88.
5. ———, *Algèbre*, Actualités Sci. Ind., no. 1032, 1947, pp. 114-115.
6. ———, *Algèbre*, Actualités Sci. Ind., no. 1102, 1950, p. 64.
7. A. Cayley, *On the analytical forms called trees, second part*, Collected Mathematical Papers, vol. 4, Cambridge, 1891, pp. 114-115.
8. H. W. Gould, *Final analysis of Vandermonde's convolution*, Amer. Math. Monthly vol. 64 (1957) pp. 409-415.
9. N. Jacobson, *Lectures on abstract algebra*, vol. 1, New York, 1951, pp. 18-19.
10. J. L. Lagrange, *Oeuvres*, vol. 3, pp. 1-73.
11. Th. Motzkin, *Relations between hypersurface cross ratios, and a combinatorial formula for partitions of a polygon, for permanent preponderance, and for non-associative products*, Bull. Amer. Math. Soc. vol. 54 (1948) pp. 352-365.
12. G. Pólya and G. Szegő, *Aufgaben und Lehrsätze aus der Analysis*, New York, 1945, pp. 125-126.
13. P. C. Rosenbloom, *The elements of mathematical logic*, New York, 1950, pp. 152-157 and p. 205.

THE PENNSYLVANIA STATE UNIVERSITY,
UNIVERSITY PARK, PENNSYLVANIA