JOURNAL OF COMBINATORIAL THEORY (B) 18, 97–137 (1975)

On Elementary Calculus and the Good Formula

W. T. TUTTE

Department of Combinatorics, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada

Communicated by D. H. Younger

Received March 20, 1974

In the theory of combinatorial generating functions one can use the differential calculus without limits, and almost without limit.

1. INTRODUCTION

In my enumerative work I commonly use formal power series as generating functions. On occasion I am asked "But how do you know these series are convergent?" I reply "That does not matter; I am not trying to sum these series, only to determine the coefficients in some of them." But this answer is not always accepted as satisfactory. There is a widespread belief that nothing can be done with power series unless they are Taylor-made to fit analytic functions.

That belief is wrong. A rigorous theory of formal power series can be constructed, but accounts of it have not been too easy to find in the literature. Recently however there has appeared an expository article by Niven [4]. There is also some discussion of the problem in a still more recent paper by Gould [3]. The latter paper includes the somewhat unfortunate remark, not in tune with the rest of the exposition, that "our viewpoint can be either rigorous or formal". In a sense the whole of the present paper is a protest against the false opposition that this remark implies. When I knowingly use a non-rigorous argument I always describe it as informal. No, our viewpoint can be either analytic or formal, but in either case it ought to be rigorous.

The subject under discussion is not really more "formal" than other branches of mathematics. I prefer to think of it as elementary calculus, as distinct from the more complicated kind that involves the discussion of limiting processes.

In the following sections I present my own idea of a rigorous theory of formal power series, or elementary calculus. It is not meant to rival or supersede other accounts, but only to emphasize their message. I constructed it with the limited object of justifying to myself my own work with generating functions. No doubt much valuable material is omitted merely because I have not yet had occasion to use it. I have however often had occasion to use the Lagrange Formula, and sometimes I have needed the extension of it presented by Good in [2]. I therefore offer purely combinatorial proofs of the combinatorial parts of these theorems.

The guiding principle in this work is that under no circumstances is there to be any attempt to sum an infinite sequence of numbers ("elements of J" in the text). An infinite sum of power series is permissible only if it reduces to a finite sum for each individual coefficient.

2. Definitions

We start with a finite set X of undefined elements called *variables*. We also suppose given a commutative ring J having a unit element. Commonly J is the ring of integers, but the fields of rational, real and complex numbers are also in use. We refer to J as the *coefficient-ring*.

A *term* is a mapping T of X into the set of nonnegative integers. We refer to the integer $T(x)$ as the *degree* in T of the variable x. The sum of the degrees in T of the members of X is the *degree* dT of T. In particular we note the *constant* term, which has zero degree in every variable. Thus the degree of the constant term is zero. Let us denote the set of all terms by U.

The *carrier* of a term T is the set of all variables having positive degree in T. We say that T is *dependent* on the variables of its carrier, and *independent* of the other variables. Two terms are *independent* of one another if they have disjoint carriers. A term T is said to be *on* a subset W of X if its carrier is contained in W.

A *power series* (with respect to X and J) is a mapping P of U into J. The number $P(T)$ is the *coefficient* of the term T in P. The coefficient of the constant term is the *initial* of P. If $P(T)$ is nonzero we say that T is *active* in P. Let us denote the set of all power series, with respect to X and J, by V.

The *zero* power series is the power series in which each coefficient is zero, i.e., no term is active. A power series in which at most the constant term is active is a *constant* power series. It is then the *unit* power series if the coefficient of the constant term is 1. A power series with only a finite number of active terms is a *polynomial*. Its *degree* is the greatest integer that is the degree of some active term, except that the degree of the zero polynomial is taken to be zero. A polynomial P is *n-homogeneous* if all terms active in P have the same degree n. The zero polynomial is taken to be n-homogeneous for every nonnegative integer n. By a *homogeneous* polynomial we mean one that is n-homogeneous for some nonnegative integer n.

The *threshold* of a nonzero power series P is the least integer that is the degree of an active term. The zero power series has no threshold.

In elementary calculus we often find a symbol used with two distinct, though related, meanings. Thus the symbol 0 may represent the integer zero, the zero element of J, or the zero power series. The symbol c of a nonzero element of J may be used to denote the constant power series with that initial. The symbol x of a variable may be used to denote the term in which the degree of x is 1 and the degree of every other variable is zero. It may even be used to denote the power series in which the term x has coefficient 1 and every other term is inactive. However the current meaning of a symbol is usually clear from the context.

3. Multiplication of Terms

Let $T_1, T_2, \ldots, T_n$, where $n \geqslant 1$, be terms. Then there is a uniquely determined term T such that

$$T(x) = \sum_{j=1}^{n} T_j(x) \tag{1}$$

for each variable x. We call T the *product* of the n terms T_j, and we write

$$T = \prod_{j=1}^{n} T_j = T_1 T_2 \cdots T_n . \tag{2}$$

An empty product of terms is by convention taken to be the constant term.

It is clear from the above definition that multiplication of terms is commutative and associative. We have the following obvious relation between degrees.

$$d\left\{\prod_{j=1}^{n} T_j\right\} = \sum_{j=1}^{n} dT_j . \tag{3}$$

The constant term is the unit of the above multiplication, and it is therefore often denoted by the symbol 1. If x and y are variables we naturally speak of the product of the term x and the term y as "the term xy", or if $y = x$ as "the term x^2". In fact a term T is often denoted by the symbol

$$x^{T(x)} y^{T(y)} \cdots t^{T(t)}, \tag{4}$$

where $\{x, y, \ldots, t\}$ is the carrier of T.

A *factor* of a term T is a term A such that $A(x) \leqslant T(x)$ for each variable x. Thus a term A is a factor of T if and only if there is a term A^* such that $AA^* = T$. If such a term A^* exists it is uniquely determined, and we call it the *cofactor* of A with respect to T. In particular T itself and the constant term are factors of T, each being the cofactor of the other.

Returning to Eq. (2) we observe that the T_j are factors of T. We refer to the sequence $(T_1, T_2, ..., T_n)$ as a *resolution* of T into n factors. Such a resolution can be found for each positive n. For example we can put $T_1 = T$ and $T_j = 1$ wheenver $j > 1$. It is important to note the fact, obvious from the above definitions, that a term T has only a finite number of distinct factors. Hence T has only a finite number of resolutions into n factors, for each positive integer n.

4. Addition of Power Series

A *collection* of power series is a pair (Y, f), where Y is a finite or countably infinite set, and f is a mapping of Y into V. We call Y the *index-set* of the collection. If Y is a finite set we say that (Y, f) is a finite collection. In what follows we often denote the power series $f(y)$ by P_y. We may then write the collection as $(Y, y \rightarrow P_y)$.

Let T be any term. We write $K(T; Y, f)$ for the set of all $y \in Y$ such that T is active in $f(y)$. We describe the collection (Y, f) as *summable* if $K(T; Y, f)$ is a finite set for each term T.

4.1. *Every finite collection is summable.* (Trivial).

4.2. *Let (Y, f) be a collection of power series. For each positive integer n let there be only a finite number of elements y of Y such that $f(y)$ is nonzero and has threshold less than n. Then (Y, f) is summable.*

Proof. Let T be any term. Write its degree as $n - 1$. By the definition of a threshold there are only a finite number of elements y of Y such that T is active in $f(y)$.

If (Y, f) is summable there is a uniquely determined power series P satisfying

$$P(T) = \sum_{y \in K} P_y(T) \tag{5}$$

for each term T, where $K = K(T; Y, f)$. We call P the *sum* of the collection (Y, f), and we write

$$P = \sum_{y \in Y} f(y) = \sum_{y \in Y} P_y . \tag{6}$$

We note that the summation on the extreme right of (6) specifies both the index-set Y and the mapping f, the latter by the rule $f(y) = P_y$. When we say that such a summation is *well defined* we mean that the corresponding collection is summable. We make the convention that the assertion of any formula includes by implication the assertion that all summations of power series occurring in it are well defined under the prevailing conditions.

Let us denote the threshold of a nonzero power series P by $\mathrm{Th}(P)$.

4.3. *Let (Y, f) be a summable collection, and let P be its sum. Then any term active in P is active in $f(y)$ for some $y \in Y$. Hence, if P is nonzero,*

$$\mathrm{Th}(P) \geqslant \mathrm{Min}\{\mathrm{Th}(P_y)\} \tag{7}$$

where the minimum is taken over all $y \in Y$ such that P_y is nonzero.

The first part of this proposition follows from the definition of a sum, and then the remainder follows from the definition of a threshold.

If Y is the set of all integers from j to k, or the set of all integers from j upward, we can use the customary notation and write the sum P of (6) as

$$\sum_{y=j}^{k} P_y \qquad \text{or} \qquad \sum_{y=j}^{\infty} P_y\,, \tag{8}$$

respectively. On occasion the first symbol is used with $k < j$. It is then interpreted as the zero power series. The first and second summations of (8) are also denoted by

$$P_j + P_{j+1} + \cdots + P_k$$

and

$$P_j + P_{j+1} + P_{j+2} + \cdots,$$

respectively.

Our addition of power series is commutative, that is it does not depend on any particular order that may be assigned to the elements of Y.

If P and Q are power series then the sum $P + Q$ is P if and only if Q is the zero power series. As is usual for additive systems we write $P + P = 2P$. Similarly $P + P + P = 3P$, and so on.

It is often convenient to change the index-set of a collection by making use of the following theorem.

4.4. *Let (T, f) be any summable collection of power series. Let W be a* *set, and let g be a* 1–1 *mapping of W onto Y. Then the collection* (W, fg) *is summable, and moreover*

$$\sum_{w \in W} fg(w) = \sum_{y \in Y} f(y). \tag{9}$$

Proof. Let T be any term. An element w of W belongs to $K(T; W, fg)$ if and only if T is active in $fg(w)$, that is if and only if T is active in $f(y)$ where $y = g(w)$. This is so if and only if $g(w)$ belongs to $K(T; Y, f)$. This observation establishes the summability of (W, fg). Moreover T has the same coefficient in $fg(w)$ as in $f(y)$, these being merely two names of the same power series. Hence T has equal coefficients in the two sides of (9).

In the next investigation W denotes any subset of Y. We write f_W for the restriction of f to W, the mapping of W into V such that $f_W(w) = f(w)$ for each $w \in W$. We then have a collection (W, f_W), and it is clear that

$$K(T; W, f_W) = W \cap K(T; Y, f) \tag{10}$$

for each term T. We deduce the following proposition.

4.5. *If* (Y, f) *is a summable collection and* $W \subseteq Y$, *then* (W, f_W) *is a summable collection.*

Let us define a *decomposition* of Y as a finite or countably infinite set Γ of subsets of Y, these subsets being disjoint and having Y as their union. We say that Γ is *admissible* with respect to f if the collection (γ, f_γ) is summable for each $\gamma \in \Gamma$. If Γ is admissible with respect to f we can define a mapping ϕ of Γ into V by the following rule.

$$\phi(\gamma) = \sum_{y \in \gamma} f(y). \tag{11}$$

Then we have a collection (Γ, ϕ) of power series. We call it the collection *induced* on Γ by f.

4.6. *Let* (Y, f) *be a summable collection, and let* Γ *be any decomposition of* Y. *Then* Γ *is admissible with respect to* f.

Let (Γ, ϕ) *be the collection induced on* Γ *by* f. *Then* (Γ, ϕ) *is summable. Moreover*

$$\sum_{\gamma \in \Gamma} \phi(\gamma) = \sum_{y \in Y} f(y). \tag{12}$$

Proof. The admissibility of Γ is a consequence of 4.5.

Let T be any term. Write $\phi(\gamma) = Q_\gamma$, $f(y) = P_y$ and $K(T; Y, f) = K$. Let L denote the set of all $\gamma \in \Gamma$ such that $K \cap \gamma$ is nonnull. Then $K(T; \Gamma, \phi) \subseteq L$, by 4.3.

Since K is finite it meets only a finite number of members of Γ. Accordingly T is active in Q_γ for only a finite number of members γ of Γ, by 4.3. Hence (Γ, ϕ) is summable. Moreover

$$\sum_{\gamma \in L} Q_\gamma(T) = \sum_{\gamma \in L} \left\{ \sum_{y \in K \cap \gamma} P_y(T) \right\} = \sum_{y \in K} P_y(T),$$

by (10) and (11). Since $K(T; \Gamma, \phi) \subseteq L$ it follows that T has equal coefficients in the two sides of (12). The theorem follows.

We say that we get from the expression on the left of (12) to that on the right by *removing brackets*, and from the expression on the right to that on the left by *inserting* brackets. The operation of removing brackets can be applied to a double sum such as

$$\sum_{r \in R} \left\{ \sum_{z \in Z(r)} P_{r,z} \right\}, \tag{13}$$

the summations involved being well defined. We define Y to be the set of all ordered pairs (r, z) such that $r \in R$ and $z \in Z(r)$. We write $\gamma(r)$ for the set of all pairs $(s, z) \in Y$ such that $s = r$. The sets $\gamma(r)$, $r \in R$, are then the member-sets of a decomposition Γ of Y. We can accordingly apply (12) provided that the collection $(Y, (r, z) \to P_{r,z})$ is summable.

We can initially reject from R all elements r for which $Z(r)$ is null. We then have a 1–1 correspondence between R and Γ, to which we apply 4.4. We obtain the following rule.

4.7. *If all the summations are well defined, then*

$$\sum_{r \in R} \left\{ \sum_{z \in Z(r)} P_{r,z} \right\} = \sum_{(r,z) \in Y} P_{r,z}$$

where each $P_{r,z}$ is a power series and Y is the set of all ordered pairs (r, z) such that $r \in R$ and $z \in Z(r)$.

If $Y = (1, 2, 3)$, then two applications of 4.6 give

$$(P_1 + P_2) + P_3 = \sum_{y \in Y} P_y = P_1 + (P_2 + P_3).$$

Thus 4.6 includes the associative law of addition.

5. Multiplication of Power Series

Let $P_1, P_2, \ldots, P_n$ be power series, $(n \geqslant 1)$. We define the *product*

$$\prod_{j=1}^{n} P_j = P_1 P_2 \cdots P_n$$

as the uniquely determined power series P in which the coefficient of an arbitrary term T is given by

$$P(T) = \sum_R P_1(T_1)\, P_2(T_2) \cdots P_n(T_n), \tag{14}$$

where the summation is over all resolutions $R = (T_1, T_2, \ldots, T_n)$ of T into n factors.

It is clear from this definition that multiplication of power series is commutative and associative.

If $P_1, P_2, \ldots, P_n$ are all equal to Q we write their product as Q^n. On occasion we use the symbol Q^0, interpreting it to mean the unit power series.

By the above definition a product PQ of power series is equal to Q if P is the unit power series, and to 0 if P is the zero power series.

5.1. *If P_j is an m_j-homogeneous polynomial for each j, then $P_1 P_2 \ldots P_n$ is an m-homogeneous polynomial, where*

$$m = \sum_{j=1}^{n} m_j .$$

This follows from (3) and (14).

5.2. *If $P = P_1 P_2 \ldots P_n$ and P is nonzero, then each of the power series P_j is nonzero. Moreover*

$$\mathrm{Th}(P) \geqslant \sum_{j=1}^{n} \mathrm{Th}(P_j). \tag{15}$$

Proof. If $P_j = 0$ then $P(T) = 0$ for each term T, by (14). Hence each P_j must be nonzero. The remainder of the theorem follows from (3) and the definition of a threshold.

5.3. *Suppose J to have no divisors of zero. Suppose $P = P_1 P_2 \ldots P_n$ where each of the factors P_j is nonzero. Then P is nonzero. Moreover*

$$\mathrm{Th}(P) = \sum_{j=1}^{n} \mathrm{Th}(P_j).$$

Proof. We enumerate the variables as $x_1, x_2, \ldots, x_m$. For each power series P_j we define the *leading term* S_j as that term of degree $\mathrm{Th}(P_j)$ that is active in P_j, has the highest degree in x_1 consistent with this condition, has the highest degree in x_2 consistent with these conditions, and so on. It is evidently uniquely determined. Write

$$S = \prod_{j=1}^{n} S_j .$$

Then S has only one resolution $(T_1, T_2, \ldots, T_n)$ into n factors such that T_j is active in P_j for each suffix j, namely the resolution $(S_1, S_2, \ldots, S_n)$. It follows from (14) that $P(S)$ is the product of the n numbers $P_j(S_j)$, which are all nonzero.

It follows from the condition imposed on J that S is active in P, and therefore that P is nonzero. But dS is the sum of the thresholds of the power series P_j, by (3). The theorem now follows from (15).

5.4. *Let (Y, f), (Y, g) and (Y, h) be collections of power series with the same index-set Y. Suppose further that $h(y) = f(y)\, g(y)$ for each $y \in Y$. Let (Y, f) be summable. Then (Y, h) is summable.*

Proof. If a term T is active in $h(y)$ then some factor of T is active in $f(y)$, by (14). But T has only a finite number of factors and each of these is active in $f(y)$ for only a finite number of members y of Y. Hence $K(T; Y, h)$ is finite.

It can be shown that multiplication is distributive over addition. More generally we have the following theorem.

5.5. *Let (Y, f), where $f(y) = P_y$, be any summable collection of power series. Let Q be any power series. Then*

$$\sum_{y \in Y} Q P_y = Q \left\{ \sum_{y \in Y} P_y \right\}. \tag{16}$$

Proof. The implied assertion that the expression on the left is well defined is a consequence of 5.4.

Let T be any term. Let Z be the finite set that is the union of the sets $K(A; Y, f)$, taken for all factors A of T. We shall use the symbol $A \mid T$ to denote that A is a factor of T.

The coefficient of T on the left of (16) is

$$\sum_{y \in Z} \left\{ \sum_{A \mid T} Q(A)\, P_y(A^*) \right\},$$

where A^* is the cofactor of A in T, by the definitions of sum and product.

Since we are now dealing not with power series but with elements of J the above expression can be written also as

$$\sum_{A|T} \left\{ Q(A) \sum_{y\in Z} P_y(A^*) \right\}.$$

But this is the coefficient of T on the right of (16).

5.6. *Let (Y, f) and (Z, g) be summable collections of power series. Let B be the set of all ordered pairs (y, z) such that $y \in Y$ and $z \in Z$. Let h be the mapping of B into V such that $h(y, z) = f(y)\, g(z)$ for each $(y, z) \in B$. Then the collection (B, h) is summable. Moreover*

$$\sum_{(y,z)\in B} f(y)\, g(z) = \left\{ \sum_{y\in Y} f(y) \right\} \left\{ \sum_{z\in Z} g(z) \right\}.$$

Proof. Let T be any term. If T is active in $f(y)\, g(z)$ then some factor of T is active in $f(y)$ and some factor of T is active in $g(z)$, by (14). Since T has only a finite number of factors this can happen for only a finite number of members of B. Hence (B, h) is summable. Moreover

$$\sum_{(y,z)\in B} f(y)\, g(z) = \sum_{y\in Y} \left\{ \sum_{z\in Z} f(y)\, g(z) \right\},$$

by 4.6 and 4.7,

$$= \sum_{y\in Y} \left\{ f(y) \sum_{z\in Z} g(z) \right\} = \left\{ \sum_{y\in Y} f(y) \right\} \left\{ \sum_{z\in Z} g(z) \right\},$$

by two applications of 5.5.

It is by applying 5.5 that we obtain such well-known results as

$$(P + Q)^2 = (P + Q)\, P + (P + Q)\, Q = P^2 + 2PQ + Q^2,$$
$$(P + Q)^3 = (P + Q)^2(P + Q) = P^3 + 3P^2Q + 3PQ^2 + Q^3,$$

for arbitrary power series P and Q. Continuing along these lines we find, by way of a familiar induction, that power series satisfy the Binomial Theorem for any positive integral index.

We often encounter symbols of the form kP, where $k \in J$ and $P \in V$. We then interpret kP as the product of P and the constant power series with initial k. By (14) we have

$$(kP)(T) = k(P(T)) \tag{17}$$

as a relation between coefficients, valid for each term T. By (17) we have

$$(jk)\, P = j(kP)), \tag{18}$$

and it does not matter whether the j and k of this identity are regarded as elements of J or as constant power series.

In Section 4 we encountered the product nP of a power series P by a positive integer n. It is clear from the definitions that this is identical with $n'P$, where n' is the element of J corresponding to n. By this we mean that n' is the sum of n elements of J each equal to the unit element 1.

If k is the negative of the unit element of J we call kP the *negative* of the power series P and write it also as $-P$. Evidently this negative of P is the unique solution for Q of the equation $P + Q = 0$. The *difference* $P - Q$ of two power series P and Q, in that order, is by definition $P + (-Q)$.

If T is any term we may speak of "the power series T". This is the power series in which T is the only active term, with coefficient 1. The product of the power series T by an element k of J is the "power series kT". Here T has coefficient k, and no other term is active. Using these conventions we can assert the following identity.

$$P = \sum_{T \in U} P(T) \cdot T. \tag{19}$$

Here $P(T) \cdot T$ means the product of the power series T by the coefficient $P(T)$ of the term T in the power series P. It is valid for any P, for the summation on the right is clearly well defined, and each term T has the same coefficient on each side of the equation.

If the term T is the product of the terms A and B it follows from (14) that the power series T is the product of the power series A and B. We deduce the identity

$$T = \prod_{x \in X} x^{T(x)}. \tag{20}$$

Here the symbol on the left denotes the power series T, and $T(x)$ is as usual the degree of the variable x in the term T. But $x^{T(x)}$ means the power series x raised to the power $T(x)$. Formula (20) remains valid when we restrict the product to variables in the carrier of the term T.

We can obtain some useful formulas by inserting brackets into (19). For example we can define $\Lambda(P, n)$ as the sum of the collection of the power series $P(T) \cdot T$, taken for all terms T such that $dT = n$. We note that $\Lambda(P, n)$ is an n-homogeneous polynomial. We call it the nth *homogeneous component* of P. We then have

$$P = \sum_{n=0}^{\infty} \Lambda(P, n) \tag{21}$$

by 4.6. We say that Formula (21) resolves the power series P into its homogeneous components.

Let H be any subset of X. Write $K = X - H$. Each term T has a unique expression as a product AB, with A on H and B on K. Using 4.6 we deduce from (19) that

$$P = \sum_{A \text{ on } H} \left\{ \sum_{B \text{ on } K} P(AB) \cdot AB \right\}$$

$$= \sum_{A \text{ on } H} \left\{ A \sum_{B \text{ on } K} P(AB) \cdot B \right\}.$$

We define the *multiplier* $M_H(P, A)$ of A in P, with respect to H, by

$$M_H(P, A) = \sum_{B \text{ on } K} P(AB) \cdot B. \tag{22}$$

We can now write

$$P = \sum_{A \text{ on } H} M_H(P, A) \cdot A. \tag{23}$$

Here the expression under the summation sign is a product of two power series, $M_H(P, A)$ and A.

5.7. *Let P and Q be power series, and let H be a subset of X. Then, for each term A on H,*

$$M_H(PQ, A) = \sum_{C|A} M_H(P, C)\, M_H(Q, C^*), \tag{24}$$

where C^* is the cofactor of C in A.

Proof. Applying 5.6 to (23) we find that

$$PQ = \sum M_H(P, C)\, M_H(Q, B) \cdot CB,$$

where the sum is over all ordered pairs (C, B) such that C and B are terms on H. We now insert brackets, collecting together all term-products CB such that $CB = A$. We find that

$$PQ = \sum_{A \text{ on } H} \left\{ \sum_{C|A} M_H(P, C)\, M_H(Q, C^*) \cdot A \right\}.$$

Since all the terms active in $M_H(P, C)\, M_H(Q, C^*)$ are on $X - H$ the theorem follows.

Variations on the notation of (19) are often used to express power series, especially when there is only one variable x. We may for example be confronted with the expression

$$P = 1 + 2x + 3x^2 + 4x^3 + \cdots. \tag{25}$$

This is the power series in which the term x^n, the nth power of the term x, has coefficient $n + 1$. We may be instructed that such an expression is what is meant by a "formal power series," and that the crosses do not denote true addition but are there merely to separate the terms. However we can interpret the symbol x in (25) as meaning the power series x, and then the crosses indicate genuine addition as we have defined it for power series.

6. Substitutions

For the purpose of this section we require a second set X' of variables. This gives rise to a new set U' of terms and a new set V' of power series. In special cases X' may be identical with X, or one of X and X' may be a proper subset of the other. The same coefficient-ring J is used for V and V'.

A *substitution* from X to X' is a mapping θ of X into V'. It is a *regular* substitution if $\theta(x)$ has zero initial for each $x \in X$.

If $T \in U$ we define $\theta(T)$ as the power series

$$\prod_{x \in X} (\theta(x))^{T(x)}$$

of V', an empty product being as usual interpreted as the unit power series.

Now suppose $P \in V$. We have a collection (U, g) of power series, where $g(T) = P(T)\,\theta(T)$ for each $T \in U$. We say that θ is an *admissible* substitution with respect to P if the collection (U, g) is summable. If this condition holds we denote the power series

$$Q = \sum_{T \in U} P(T)\,\theta(T)$$

of V' by $\theta(P)$.

Since all finite collections are summable we have the following.

6.1. *If θ is a substitution from X to X', and P is a polynomial in V, then θ is admissible with respect to P.*

6.2. *Let θ be any regular substitution from X to X'. Let T be a term of U such that $\theta(T)$ is nonzero. Then the threshold of $\theta(T)$ is at least equal to the degree of T.*

To prove this we observe that $\theta(T)$ is a product of dT power series, each with threshold at least 1. The theorem follows, by 5.2.

6.3. *Let θ be any regular substitution from X to X'. Let P be any power series of V. Then θ is admissible with respect to P.*

Proof. If T and T' are terms of U and U', respectively, and T' is active in $\theta(T)$, then $dT' \geqslant dT$ by 6.2. Hence a given term T' of U' can be active in $P(T)\,\theta(T)$ for only a finite number of terms T of U.

6.4. *Let P and Q be power series of V. Let θ be a substitution from X to X' that is admissible with respect to both P and Q. Then θ is admissible with respect to both $P + Q$ and PQ. Moreover $\theta(P+Q) = \theta(P) + \theta(Q)$ and $\theta(PQ) = \theta(P)\,\theta(Q)$.*

Proof. Under the conditions stated a given term T' of U' is active in $P(T)\,\theta(T)$ or $Q(T)\,\theta(T)$ for only finitely many terms T of U. Consequently it is active in $(P(T) + Q(T))\,\theta(T)$ for only finitely many terms T of U. Accordingly θ is admissible with respect to $P + Q$. Moreover

$$\begin{aligned}\theta(P+Q) &= \sum_{T\in U} \{P(T) + Q(T)\}\,\theta(T)\\ &= \sum_{T\in U} P(T)\,\theta(T) + \sum_{T\in U} Q(T)\,\theta(T), \qquad \text{by 4.7,}\\ &= \theta(P) + \theta(Q).\end{aligned}$$

Using 5.6 we find that

$$\theta(P)\,\theta(Q) = \sum P(A)\,\theta(A)\,Q(B)\,\theta(B),$$

where the sum is over all ordered pairs (A, B) of elements of U. We now use 4.6 to bring together all pairs (A, B) having the same product AB. Since it is clear from the definitions that $\theta(AB) = \theta(A)\,\theta(B)$ for arbitrary terms A and B of U we then have

$$\theta(P)\,\theta(Q) = \sum_{T\in U} \left\{\sum_{A|T} P(A)\,Q(A^*)\right\} \theta(T),$$

where A^* is the cofactor of A in T,

$$= \sum_{T\in U} \{\text{Coefficient of } T \text{ in } PQ\}\,\theta(T).$$

Since the last expression is well defined, by 4.6 and 5.6, we deduce that θ is admissible with respect to PQ, and that $\theta(P)\,\theta(Q) = \theta(PQ)$.

6.5. *Let (Y, f), where $f(y) = P_y$, be a summable collection of power series of V, and let θ be any regular substitution from X to X'. Then*

$$\theta\left\{\sum_{y\in Y} P_y\right\} = \sum_{y\in Y} \theta(P_y). \tag{26}$$

Proof. We use 6.3 to establish admissibility.

$$\theta\left\{\sum_{y\in Y} P_y\right\} = \sum_{T\in U}\left\{\sum_{y\in K(T)} P_y(T)\right\}\theta(T),$$

where $K(T) = K(T; Y, f)$,

$$= \sum_{T\in U}\left\{\sum_{y\in K(T)} P_y(T)\,\theta(T)\right\}, \qquad \text{by 5.5,}$$

$$= \sum_{T\in U}\left\{\sum_{y\in Y} P_y(T)\,\theta(T)\right\},$$

since $P_y(T) = 0$ when y is not in $K(T)$.

Consider any term T' of U'. It is active in $\theta(T)$ for only a finite number of terms T of U, by 6.2. For each such T there are only a finite number of elements y of Y such that $P_y(T)$ is nonzero. Hence T' is active in $P_y(T)\,\theta(T)$ for only a finite number of pairs (y, T). Accordingly it is permissible, by 4.6, to remove the brackets in the last double sum, and then to insert brackets in a different way. We find that

$$\theta\left\{\sum_{y\in Y} P_y\right\} = \sum_{y\in Y}\left\{\sum_{T\in U} P_y(T)\,\theta(T)\right\}$$

$$= \sum_{y\in Y} \theta(P_y).$$

We now note some very simple examples of regular substitutions. Let us describe a substitution θ from X to X' as *advariant* if for each variable x of X there is a variable x' of X' such that $\theta(x)$ is the power series x'.

Suppose first that $X = X'$. There is an advariant substitution θ such that $\theta(x)$ is the power series x for each $x \in X$. We call this the *identical* substitution from X to X. By (19), 6.4 and 6.5 we have $\theta(P) = P$ for each $P \in V$.

A less trivial advariant substitution θ from X to X is associated with a permutation π of the set X. We define $\theta(x)$, $x \in X$, as the power series $\pi(x)$. Two power series P and $\theta(P)$ can then be said to be related by a permutation of the variables, and a power series P such that $P = \theta(P)$ can be called *symmetrical* with respect to the permutation π.

Another case arises when X is a subset of X'. There is an advariant substitution θ from X to X' such that $\theta(x)$ is the power series x of V' for each $x \in X$. We call θ the *direct embedding* of V into V'. We seldom bring ourselves to distinguish between the power series P of V and the power series $\theta(P)$ of V'. They are represented by the same formula (19); the only question is whether we are to consider the variables appearing in that formula as members of X or as members of the wider set X'.

Let us say that a substitution θ from X to X is *conservative* with respect to a variable x if $\theta(x)$ is the power series x. Thus the identical substitution is conservative for every variable. Consider a substitution θ from X to X transforming one variable x into a constant power series k, but conservative for every other variable. Applying θ is often called "assigning the particular value k to x". It is a hazardous operation when k is nonzero, for then θ is not regular and we cannot use 6.3 to guarantee its admissibility. But θ is admissible with respect to every $P \in V$ when $k = 0$, by 6.2.

A power series is often called a *function* of the variables. It may be denoted by some such expression as $f(x_1, x_2, \ldots, x_n)$, where the x_j are the variables. Then if θ is some admissible substitution to the same set of variables we may find $\theta(x_j)$ written as $g_j(x_1, x_2, \ldots, x_n)$. We may then replace each symbol x_j in the expression $f(x_1, x_2, \ldots, x_n)$ by the corresponding symbol $g_j(x_1, x_2, \ldots, x_n)$, and regard the resulting expression as a representation of $\theta(P)$.

P, a power series of V, is said to be *dependent* on the variable $x \in X$ if some term active in P has a positive degree in x. Some simplification of the above notation can be made by ignoring variables of which the function under consideration is independent. For example if our variables are x, y, and z, and $g_2(x, y, z)$ is independent of y and z, we may agree to write that function simply as $g_2(x)$. As a concession to rigour we may make some mention of the direct embedding from $\{x\}$ to $\{x, y, z\}$.

7. Substitutional Equations

Let X, X', U, U', V, V' be as in Section 6. Let the variables of X be enumerated as $x_1, x_2, \ldots, x_n$. For each x_j let there be given $z_j \in X'$, $P_j \in V'$ and $Q_j \in V$. We may ask: is there a substitution θ from X to X' such that

$$\theta(x_j) = P_j + z_j\theta(Q_j) \tag{27}$$

for each suffix j? ($1 \leqslant j \leqslant n$). We refer to (27) as a set of *substitutional equations* for the unknown substitution θ. When $n = 1$ the substitutional equation is usually solved by the Lagrange Formula. When $n > 1$ we can use the more general Good Formula. We defer the discussion of these formulas to a later section. Here we strive only to show that if (27) has a solution for θ then that solution is unique. To qualify as a solution a substitution θ must of course be admissible with respect to each of the power series Q_j.

Let us say that two power series P and Q of V are *n-equivalent*, where n is a given positive integer, if $P(T) = Q(T)$ for each term T of U such that $dT < n$. For example we can define the *n-truncation* P_n of P as the power series of V such that $P_n(T) = P(T)$ when $dT < n$ and $P_n(T) = 0$ when $dT \geqslant n$. Evidently P_n is a polynomial whose degree is less than n, and it is n-equivalent to P.

The two following theorems are immediate consequences of the definitions of sum and product.

7.1. *Let (Y, f) and (Y, g) be two summable collections of power series of V, with the same index-set Y. Let n be a positive integer such that $f(y)$ and $g(y)$ are n-equivalent for each $y \in Y$. Then the sums*

$$\sum_{y \in Y} f(y) \qquad \textit{and} \qquad \sum_{y \in Y} g(y)$$

are n-equivalent.

7.2. *Let P_1, P_2, Q_1 and Q_2 be power series of V, and let n be a positive integer such that P_1 is n-equivalent to P_2 and Q_1 is n-equivalent to Q_2. Then P_1Q_1 is n-equivalent to P_2Q_2.*

The latter theorem has the following powerful variation, another immediate consequence of the definition of a product.

7.3. *Let P_1, P_2, Q_1 and Q_2 be power series of V, and let n be a positive integer such that P_1 is $(n + 1)$-equivalent to P_2 and Q_1 is n-equivalent to Q_2. Suppose further that P_1 and P_2 have zero initials. Then P_1Q_1 is $(n + 1)$-equivalent to P_2Q_2.*

7.4. *Let θ_1 and θ_2 be two solutions for θ of the substitutional equations (27). Then θ_1 is identical with θ_2.*

Proof. Let us say that θ_1 and θ_2 are *n-equivalent* for a given positive integer n if $\theta_1(x_j)$ and $\theta_2(x_j)$ are n-equivalent power series of V' for each $x_j \in X$. We note that θ_1 and θ_2 are necessarily 1-equivalent, for the constant terms of $\theta_1(x_j)$ and $\theta_2(x_j)$ are each equal to the constant term of P_j, by (27).

Suppose θ_1 and θ_2 are distinct. Then there must be a positive integer N

such that θ_1 and θ_2 are N-equivalent but not $(N + 1)$-equivalent. We apply the foregoing theorems with V' replacing V. Since $\theta(Q_j)$ is by definition a sum of products, each of an element of J and power series of the form $\theta(x_k)$ it follows from 7.1 and 7.2 that $\theta_1(Q_j)$ and $\theta_2(Q_j)$ are N-equivalent for each j. Hence $z_j\theta_1(Q_j)$ and $z_j\theta_2(Q_j)$ are $(N + 1)$-equivalent for each j, by 7.3. It follows from (27) and 7.1 that $\theta_1(x_j)$ and $\theta_2(x_j)$ are $(N + 1)$-equivalent for each j. But this is contrary to the choice of N. We conclude that θ_1 and θ_2 are the same substitution.

Presumably the above argument could be developed into an existence theorem for θ. But the results of Section 11 will make this unnecessary.

8. Differentiation

Let X, U and V be as in Section 2.

Corresponding to each variable x of X we define a *differential operator* D_x. This is a mapping of V into V. For each $P \in V$ we define $D_x(P)$ as follows. The coefficient of any term T in $D_x(P)$ is the coefficient of the term xT in P, multiplied by the integer $T(x) + 1$. We call $D_x(P)$ the *derivative* of P with respect to x. We observe that the coefficients in P of the terms in which the degree of x is zero play no part in the definition of the derivative. We have indeed the following rule.

8.1. *If P is independent of x, then $D_x(P) = 0$.*

We note two other immediate consequences of the definition.

8.2. *If $D_x(P) = 0$ for each $x \in X$, then P is a constant power series.*

8.3. *If P is an n-homogeneous polynomial with $n \geqslant 1$, then $D_x(P)$ is an $(n - 1)$-homogeneous polynomial.*

8.4. The Sum Rule. *Let (Y, f) be a summable collection of power series. Then*

$$D_x\left\{\sum_{y \in Y} f(y)\right\} = \sum_{y \in Y} D_x(f(y)).$$

Proof. Let f_1 be the mapping of Y into V such that $f_1(y) = D_x(f(y))$ for each y. A term T can be active in $D_x(f(y))$ only if the term xT is active in $f(y)$. Thus

$$K(T;\, Y, f_1) \subseteq K(xT;\, Y, f). \tag{28}$$

We deduce that the summation on the right of 8.4 is well defined. Moreover, writing $K = K(xT; Y, f)$, we find that the coefficient of T on the left of 8.4 is

$$(T(x) + 1) \left\{ \sum_{y \in K} P_y(xT) \right\} = \sum_{y \in K} (T(x) + 1)\, P_y(xT),$$

where $P_y = f(y)$. But this can be written as

$$\sum_{y \in K} \{\text{Coefficient of } T \text{ in } D_x(f(y))\},$$

and this is equal to the coefficient of T on the right of 8.4.

For a power series P multiplied by a positive integer n the Sum Rule gives

$$D_x(nP) = nD_x(P). \tag{29}$$

8.5. THE PRODUCT RULE. *Let $P = P_1 P_2 \ldots P_n$ be a product of n power series $P_j \cdot (n \geqslant 1)$. For each integer i satisfying $1 \leqslant i \leqslant n$ let us define $P_j^{(i)}$ to be P_j if $i \neq j$, and to be $D_x(P_j)$ if $i = j$. Then*

$$D_x(P) = \sum_{i=1}^{n} \left\{ \prod_{j=1}^{n} P_j^{(i)} \right\}. \tag{30}$$

Proof. Let T be any term. Its coefficient in $D_x(P)$ is the coefficient of xT in P, multiplied by the integer $T(x) + 1$. Let R denote any resolution $(T_1, T_2, \ldots, T_n)$ of xT into n factors. Then we can write the coefficient of T in $D_x(P)$ as

$$\sum_{i=1}^{n} \left\{ \sum_{R} T_i(x)\, P_1(T_1)\, P_2(T_2) \cdots P_n(T_n) \right\}.$$

The resolutions R such that $T_i(x) \neq 0$ (for a fixed i) are in 1–1 correspondence with the resolutions $S = (S_1, S_2, \ldots, S_n)$ of T into n factors, the rule being that $T_j = S_j$ if $j \neq i$, and $T_i = xS_i$. Using 8.1 we deduce that the coefficient of T in $D_x(P)$ is

$$\sum_{i=1}^{n} \left\{ \text{Coefficient of } T \text{ in } \prod_{j=1}^{n} P_j^{(i)} \right\}.$$

But this is the coefficient of T on the right of (30). The theorem follows.

As a special case of the Product Rule we have

$$D_x(P^n) = nP^{n-1} D_x(P), \qquad n \geqslant 1. \tag{31}$$

By repeated application of D_x to a product PQ of power series we find that

$$D_x(PQ) = PD_x(Q) + D_x(P) \cdot Q, \tag{32}$$

$$D_x^2(PQ) = PD_x^2(Q) + 2D_x(P) \cdot D_x(Q) + D_x^2(P) \cdot Q, \tag{33}$$

and so on. Continuing in this way and using a familiar induction we arrive at the following rule.

8.6. THE LEIBNITZ FORMULA. *If P and Q are power series of V, and n is a positive integer, then*

$$D_x^n(PQ) = \sum_{j=0}^{n} \binom{n}{j} D_x^j(P) \cdot D_x^{n-j}(Q).$$

Applying (32) to the case in which P is a constant power series with initial k we find

$$D_x(kQ) = kD_x(Q) \tag{34}$$

by 8.1. This is consistent with (29).

Often differential operators with respect to two or more variables occur in the same formula. It may then be helpful to bear in mind the following rule of commutation.

8.7. *If x and z are variables of X, then*

$$D_xD_z(P) = D_zD_x(P)$$

for each $P \in V$.

Proof. We may suppose x and z distinct since otherwise there is nothing to prove. But then the coefficient of an arbitrary term T on either side of the equation is the coefficient of xzT in P, multiplied by $(T(x) + 1)(T(z) + 1)$.

8.8. THE CHAIN RULE. *Let P be any power series of V, and let θ be any regular substitution from X to a set X′ of variables. Let t be any member of X′. Then*

$$D_t(\theta(P)) = \sum_{x \in X} \{\theta(D_x(P)) \cdot D_t(\theta(x))\}. \tag{35}$$

Proof. Let T be any term of U. We apply the Product Rule to the product $\theta(T)$ (see Section 6). The result can be written as

$$D_t(\theta(T)) = \sum_{x \in X} \theta(D_x(T)) \cdot D_t(\theta(x)), \tag{36}$$

where T on the right is the power series T. Hence

$$D_t(\theta(P)) = D_t \left\{ \sum_{T \in U} P(T)\, \theta(T) \right\}$$

$$= \sum_{T \in U} \{P(T)\, D_t(\theta(T))\}, \qquad \text{by 8.4,}$$

$$= \sum_{T \in U} \left\{ \sum_{x \in X} P(T)\, \theta(D_x(T))\, D_t(\theta(x)) \right\}.$$

The new set X' of variables gives rise to a corresponding set U' of terms and a set V' of power series. If a term A of U' is active in $P(T)\,\theta(D_x(T))\,D_t(\theta(x))$ then some factor of A is active in $\theta(D_x(T))$. This happens for only a finite number of terms T of U, by 6.2. We may therefore apply 4.6 to obtain

$$D_t(\theta(P)) = \sum_{x \in X} \left\{ \sum_{T \in U} P(T)\, \theta(D_x(T))\, D_t(\theta(x)) \right\}$$

$$= \sum_{x \in X} \left\{ \sum_{T \in U} P(T)\, \theta(D_x(T)) \right\} D_t(\theta(x)),$$

by 5.5 and 6.3,

$$= \sum_{x \in X} \theta \left\{ \sum_{T \in U} P(T)\, D_x(T) \right\} D_t(\theta(x)),$$

by 6.5. The application of 8.4 now completes the proof.

A substitution θ from X to a set X' of variables will be called *nonsingular* if either it is regular or $\theta(x)$ is nonzero for each $x \in X$. In all other cases θ is *singular*.

8.9. *Let θ be a nonsingular substitution from X to X'. Let θ be admissible with respect to a power series P of V. Then θ is admissible with respect to $D_x(P)$, for each $x \in X$.*

Proof. By 6.3 we may suppose θ to be nonregular.

Write $F = D_x(P)$. Then for any term T of U we have

$$F(T)\,\theta(T)\,\theta(x) = (T(x) + 1)\, P(xT)\,\theta(xT) \qquad \text{by 6.4.} \tag{37}$$

By hypothesis it follows that a term T' of U' is active in $F(T)\,\theta(T)\,\theta(x)$ for only finitely many $T \in U$.

Enumerate the members of X' as $x_1', x_2', \ldots, x_m'$.

Let W' be the set of all terms T' of U' such that T' is active in $F(T)\,\theta(T)$ for infinitely many terms T of U. Assume W' nonnull.

Given any subset U_1' of U' we can define a unique *lowest term* of U_1' as the term of U_1' with the least degree in x_1', subject to this condition the least degree in x_2', and so on. Let A' be the lowest term of W', and B' the lowest term of the set of terms of U' active in $\theta(x)$. The nonsingularity of θ ensures the existence of B'. It is clear that the term $A'B'$ is active in $F(T)\,\theta(T)\,\theta(x)$ for infinitely many T, contrary to our earlier result. Accordingly W' is null, and θ is admissible with respect to F.

9. Reciprocals

Let X, U and V be as in Section 2. The reciprocal P^{-1} of a power series P is defined by the property

$$P^{-1}P = 1. \tag{38}$$

Such a power series P^{-1} may not exist for a given P. It is easy to show for example that the power series x, where x is a variable, has no reciprocal. But if a reciprocal P^{-1} of P exists it is unique; if Q were another we would have $Q = Q(PP^{-1}) = (QP)\,P^{-1} = P^{-1}$. If P^{-1} exists, and k is any positive integer, then P^k has the reciprocal $(P^{-1})^k$, which we write also as P^{-k}.

9.1. *If $x \in X$ then the power series $1 - x$ has a reciprocal, given by the following formula.*

$$(1 - x)^{-1} = \sum_{j=0}^{\infty} x^j. \tag{39}$$

We verify this by multiplying the power series on the right by $(1 - x)$. The familiar formula

$$D_x\{(1 - x)^{-n}\} = n(1 - x)^{-(n+1)} \tag{40}$$

remains valid. We establish it in the case $n = 1$ by applying the product rule to $(1 - x)(1 - x)^{-1}$, and then we extend to an arbitrary n by applying the product rule to $((1 - x)^{-1})^n$. By differentiating both sides of (39) we can show that the power series of (25) is $(1 - x)^{-2}$.

Theorem 9.1 can be generalized as follows.

9.2. *Let P be a power series of V, with zero initial. Then the power series $1 - P$ has a reciprocal, given by the following formula.*

$$(1 - P)^{-1} = \sum_{j=0}^{\infty} P^j. \tag{41}$$

We obtain this result by applying the appropriate regular substitution to both sides of the identity

$$(1 - x) \sum_{j=0}^{\infty} x^j = 1.$$

The operation is justified by 6.4 and 6.5.

If three power series P, Q, and R of V are related by the formula $P = QR$ we may refer to R as the *quotient* of P by Q and write $R = P/Q$. If Q is nonzero and J has no divisors of zero there can be no other such quotient, for if also $P = QR_1$ we have $Q(R_1 - R) = 0$ and therefore $R_1 = R$ by 5.3.

If a power series Q has a reciprocal then the quotient P/Q exists for each power series P, being equal to PQ^{-1}. But we cannot assert the existence of a quotient for arbitrary power series P and Q of V, even when Q is restricted to be nonzero.

If J is an integral domain then the power series of V constitute an integral domain. We can construct the corresponding quotient field and contemplate the extension of our elementary calculus to the quotients of that field. We then indeed discuss quotients P/Q corresponding to any pairs (P, Q) such that Q is nonzero, but in general these quotients are not themselves power series, just as a rational number is not in general an integer.

10. A Combinatorial Identity

Let X, U, and V be as in Section 2. If $W \subseteq X$ we denote the number of members of W by $\beta(W)$.

A *cyclic operator* is a pair $L = (W, \pi)$ such that $W \subseteq X$ and π is a permutation of W. We write $\alpha(L)$ for the number of cycles of π. If x is a variable we write $e(L, x) = 1$ or 0 according as x is or is not in W.

Let (Y, f) be any finite collection of power series of V. Let g be any mapping of Y into X. If $x \in X$ we write $g^{-1}(x)$ for the set of all $y \in Y$ such that $g(y) = x$. Thus the sets $g^{-1}(x)$, $x \in X$, are disjoint and have Y as their union. Some or all of them may be null. We write s for the number of members of Y, and $s(x)$ for the number of members of $g^{-1}(x)$. Thus

$$\sum_{x \in X} s(x) = s. \tag{42}$$

For each $x \in X$ we now write

$$S(x) = \prod_{y \in g^{-1}(x)} f(y). \tag{43}$$

If $g^{-1}(x)$ is null we interpret $S(x)$ as the unit power series.

We define the application of the cyclic operator $L = (W, \pi)$ to $S(x)$ as follows. If x is not in W then

$$LS(x) = S(x),$$

but if x is in W then

$$LS(x) = D_{\pi(x)}S(x).$$

We say that L *conforms* to g (L cf g) if $g^{-1}(x)$ is nonnull whenever $x \in W$.

If L conforms to g we use the symbol $D_{g,L}$ to denote the product

$$\prod_{x \in X} D_x^{s(x)-e(L,x)}$$

of differential operators. We note that the index $s(x) - e(L, x)$ is never negative. Thus if P is a power series of V then $D_{g,L}(P)$ is the power series obtained from it by applying $s - \beta(W)$ differential operators, of which exactly $s(x) - e(L, x)$ are identical with D_x for each $x \in X$. The order in which these operators are applied does not matter, by 8.7. If $s - \beta(W) = 0$, that is if $D_{g,L}$ is a null product of differential operators, we take $D_{g,L}(P)$ to be P.

Let λ be any power series of V. Then we make the following definition.

$$Q(\lambda; f, g) = \sum_{L \text{ cf } g} (-1)^{\alpha(L)} D_{g,L} \left\{\lambda \prod_{x \in X} LS(x)\right\}. \tag{44}$$

We proceed to study the power series $Q(\lambda; f, g)$. Sometimes we shall have to consider certain minor changes in f and g. Suppose for example that y' is a specified element of Y, and that P' is a power series of V. We can then define a mapping f' of Y into V such that $f'(y') = P'$ but $f'(y) = f(y)$ whenever $y \in Y - \{y'\}$. Keeping Y and g fixed we may then construct $Q(\lambda; f', g)$ as above. In order to convey more information by the symbolism we may write f' as $(f, y' \rightarrow P')$, and $Q(\lambda; f', g)$ as $Q(\lambda; (f, y' \rightarrow P'), g)$.

We begin by noting two trivial propositions.

10.1. *If* $s = 0$ *then* $Q(\lambda; f, g) = \lambda$.

10.2. *If* $\lambda = 0$, *or if* $f(y) = 0$ *for some* $y \in Y$, *then*

$$Q(\lambda; f, g) = 0.$$

In the case of 10.1 it should be noted that L can conform to g only if W is the null subset of X.

We continue with two linearity theorems, these being immediate consequences of 5.5 and 8.4.

10.3 *Let (R, ρ) be a summable collection of power series of V. Then*

$$Q\left(\sum_{r\in R} \rho(r); f, g\right) = \sum_{r\in R} Q(\rho(r); f, g).$$

10.4. *Let (R, ρ) be a summable collection of power series of V, and let F be its sum. Let y' be a specified member of Y. Then*

$$Q(\lambda; (f, y' \to F), g) = \sum_{r\in R} Q(\lambda; (f, y' \to \rho(r)), g).$$

10.5. *Let $f(y')$ be a constant power series for some specified $y' \in Y$. Write $Y' = Y - \{y'\}$. Let f' and g' be the restrictions of f and g, respectively, to Y'. Then*

$$Q(\lambda; f, g) = f(y')\, D_{g(y')} Q(\lambda; f', g').$$

Proof. We set beside $S(x)$ the corresponding product $S'(x)$, defined in terms of f' and g'. Thus $S(x) = S'(x)$ if y' is not in $g^{-1}(x)$, and $S(x) = f(y')\, S'(x)$ otherwise. We observe that if a cyclic operator L conforms to g' it conforms also to g. On the other hand if L conforms to g it conforms also to g', unless $g(y')$ is in W and $g^{-1}(g(y'))$ has no member other than y'. Accordingly we can rewrite (44) as

$$Q(\lambda; f, g) = f(y') \sum_{L\,\mathrm{cf}\,g'} (-1)^{\alpha(L)}\, D_{g,L} \left\{\lambda \cdot \prod_{x\in X} LS'(x)\right\},$$

by (34). But if L conforms to g' we have

$$D_{g,L} = D_{g(y')} D_{g',L}\,.$$

The theorem follows.

10.6. *If $z \in X$, then*

$$D_z Q(\lambda; f, g) = Q(D_z\lambda; f, g) + \sum_{y'\in Y} Q(\lambda; (f, y' \to D_z(f(y'))), g).$$

We obtain this result by applying D_z to each side of (44), using 8.7 and the Product Rule.

Let x be any member of X. We say that a power series P of V is *basic* on x if it can be written as kx, where $k \in J$ and the symbol x denotes the power series x. We then call k the *weight* of P.

10.7. *Let s be nonzero. Let λ be constant and let $f(y)$ be basic for each $y \in Y$. Let $g^{-1}(x)$ have at most one member for each $x \in X$. Then*

$$Q(\lambda; f, g) = 0.$$

Proof. Let X_1 be the set of all $x \in X$ such that $s(x) = 1$. Then if a cyclic operator $L = (W, \pi)$ conforms to g we have $W \subseteq X_1$. By 10.2 we can assume that no $f(y)$ is zero.

Suppose that either there exists $y' \in Y$ such that $f(y')$ is basic on some $x' \in X - X_1$, or there are two distinct elements y_1 and y_2 of Y such that $f(y_1)$ and $f(y_2)$ are basic on the same member x_1 of X_1. Then by the conditions of the theorem there exists $z \in X_1$ such that no $f(y)$ is basic on z. But on the right of (44) the operator D_z is applied, either directly to some $S(x)$ or to the product $\lambda \Pi L S(x)$. Hence the theorem holds, by 8.1.

In the remaining case there is a uniquely determined permutation σ of X_1 such that for each $x \in X_1$ the set $g^{-1}(x)$ has a unique element $y(x)$, and $f(y(x))$ is basic on $\sigma(x)$. Then if $L = (W, \pi)$ conforms to g and makes a nonzero contribution to the right of (44) the cycles of π must be cycles of σ. Apart from the factor $(-1)^{\alpha(L)}$ each such operator L makes the same contribution H, the product of the weights of the basic power series $f(y)$. So by (44) we have

$$Q(\lambda; f, g) = \sum_{j=0}^{h} \binom{h}{j} (-1)^j H = 0,$$

where h is the number of cycles of σ.

We go on to derive some successive extensions of 10.7, culminating in the general combinatorial identity 10.10.

10.8. *Let s be nonzero. Let λ be constant and let $f(y)$ be basic for each $y \in Y$. Then*

$$Q(\lambda; f, g) = 0.$$

Proof. If $s = 1$ the theorem holds, by 10.7. Assume as an inductive hypothesis that the theorem is true whenever s is less than some integer $q > 1$, and consider the case $s = q$.

By 10.7 we may assume that there is a variable x such that $s(x) \geqslant 2$. Let y_1 and y_2 be two distinct elements of $g^{-1}(x)$. We write $Y_1 = Y - \{y_1\}$. We define g_1 to be the restriction of g to Y_1, and we define a mapping f_1 of Y_1 into V in the following way: $f_1(y) = f(y)$ unless $y = y_2$, but $f_1(y_2) = f(y_1) f(y_2)$. Then, by (44),

$$Q(\lambda; f, g) = D_x Q(\lambda; f_1, g_1).$$

By alternate applications of 10.6 and 10.5 we deduce that $Q(\lambda; f, g)$is a sum of power series of the form $kQ(\mu; f', g')$, where $k \in J$, μ is constant and $f'(y)$ is basic for each relevant y. Moreover each of these expressions corresponds to a value of s less than q.

We deduce that the theorem holds when $s = q$. It follows in general by induction.

10.9. *Let s be nonzero. Let λ be constant and let $f(y)$ be a homogeneous polynomial for each $y \in Y$. Then*

$$Q(\lambda; f, g) = 0.$$

Proof. If possible choose λ, Y, f and g so that the theorem fails, so that s has the least value consistent with this condition, and so that the sum n of the degrees of the polynomials $f(y)$, $y \in Y$, has the least value consistent with this.

Let n_y denote the degree of the polynomial $f(y)$. We may assume $n_y \geqslant 1$ for each y, for otherwise the theorem holds by 10.5 and the choice of s.

If $n_y = 1$ for each y it follows from 10.4 that $Q(\lambda; f, g)$ is a sum of power series of the form $Q(\lambda; f', g)$, where $f'(y)$ is basic for each y. Then the theorem holds, by 10.8.

In the remaining case $n > s$. Hence $Q(\lambda; f, g)$ is a homogeneous polynomial of degree at least 1, or is zero, by 5.1 and 8.3. However by 10.2, 10.6 and the choice of n we have

$$D_z Q(\lambda; f, g) = 0,$$

for each $z \in X$. Hence $Q(\lambda; f, g)$ is a constant power series, by 8.2. We deduce that $Q(\lambda; f, g) = 0$.

We conclude that the proposed choice of λ, Y, f, and g is impossible.

10.10. *Let s be nonzero and let λ be constant. Then*

$$Q(\lambda; f, g) = 0.$$

Proof. A power series can be represented as the sum of its homogenous components (21). Hence, starting with 10.9, we can use 10.4 to establish Theorem 10.10 for cases in which successively fewer of the power series $f(y)$ are restricted to be homogeneous polynomials. Since Y is finite the theorem follows.

Now let M be any subset of Y. Let the restrictions of f and g to M be denoted by f_M and g_M, respectively, and let the restrictions of f and g

to $Y - M$ be denoted by F_M and G_M respectively. Let λ and μ be power series of V. We make the following definition.

$$R(\lambda, \mu; f, g) = \sum_{M \in Y} Q(\lambda; f_M, g_M)\, Q(\mu; F_M, G_M). \tag{45}$$

For each $z \in X$ we have

$$\begin{aligned} D_z R(\lambda, \mu; f, g) = &\sum_{M \in Y} D_z Q(\lambda; f_M, g_M) \cdot Q(\mu; F_M, G_M) \\ &+ \sum_{M \in Y} Q(\lambda; f_M, g_M) \cdot D_z Q(\mu; F_M, G_M), \end{aligned} \tag{46}$$

by the Sum and Product Rules. We apply 10.6 to evaluate $D_z Q(\lambda; f_M, g_M)$ and $D_z Q(\mu; F_M, G_M)$. We then collect together all expressions involving $D_z(f(y'))$ for the same $y' \in Y$. The result can be written as follows.

10.11. *For each* $z \in X$,

$$\begin{aligned} D_z R(\lambda, \mu; f, g) = &R(D_z\lambda, \mu; f, g) + R(\lambda, D_z\mu; f, g) \\ &+ \sum_{y' \in Y} R(\lambda, \mu; (f, y' \to D_z(f(y'))), g). \end{aligned}$$

We proceed to state our second major combinatorial identity in this section.

10.12. $R(\lambda, \mu; f, g) = Q(\lambda\mu; f, g)$.

Proof. We begin with a Lemma that establishes the theorem in one special case.

LEMMA. *If all the power series* λ, μ, $f(y)$ *are homogeneous polynomials, then the theorem holds.*

Proof of the Lemma. We denote the degrees of the polynomials λ, $\mu, f(y)$ by n_λ, n_μ, and n_y, respectively, and we write

$$n = \sum_{y \in Y} n_y + n_\lambda + n_\mu. \tag{47}$$

If possible choose, λ, μ, Y, f and g so that the theorem fails, s has the least value consistent with this condition, and n has the least value consistent with these requirements.

If $s = 0$ the theorem is true by 10.1. We conclude that in fact $s > 0$. We can deduce from 10.2 that λ, μ and the $f(y)$ are nonzero power series.

If λ is constant then

$$R(\lambda, \mu; f, g) = \lambda Q(\mu; f, g) = Q(\lambda\mu; f, g),$$

by 10.1, 10.10, (45) and (34). We must therefore suppose $n_\lambda > 0$, and similarly $n_\mu > 0$.

Suppose $n_{y'} = 0$ for some $y' \in Y$. By applying 10.5 to the right of (45), and then using (46), we find that

$$R(\lambda, \mu; f, g) = f(y')\, D_{g(y')} R(\lambda, \mu; f', g'),$$

where f' and g' are as for 10.5,

$$= f(y')\, D_{g(y')} Q(\lambda\mu; f', g'),$$

by the choice of s,

$$= Q(\lambda\mu; f, g),$$

by 10.5. From this contradiction we deduce that in fact $n_y \geqslant 1$ for each $y \in Y$.

We now have $n \geqslant s + 2$. Then $R(\lambda, \mu; f, g)$ and $Q(\lambda\mu; f, g)$ are each k-homogeneous polynomials, each for some $k \geqslant 2$, by 5.1 and 8.3. Moreover these two theorems enable us to deduce from 10.11 and the choice of n that

$$\begin{aligned} D_z R(\lambda, \mu; f, g) &= Q(\mu D_z\lambda; f, g) + Q(\lambda D_z\mu; f, g) \\ &\quad + \sum_{y' \in Y} Q(\lambda\mu; (f, y' \to D_z(f(y'))), g) \\ &= Q(D_z(\lambda\mu); f, g) \\ &\quad + (D_z Q(\lambda\mu; f, g) - Q(D_z(\lambda\mu); f, g)), \end{aligned}$$

by 8.7, 10.3 and the Product Rule,

$$= D_z Q(\lambda\mu; f, g).$$

Since this result holds for each $z \in X$ we deduce that $R(\lambda, \mu; f, g) - Q(\lambda\mu; f, g)$ is a constant C, by 8.2. But each of $R(\lambda, \mu; f, g)$ and $Q(\lambda\mu; f, g)$ is a k-homogeneous polynomial for some $k > 0$. Hence $C = 0$.

We conclude that the proposed choice of λ, μ, Y, f and g is impossible; the Lemma is true.

The rest of the proof of 10.12 resembles that of 10.10. Starting with the Lemma we use 10.3 and 10.4, to establish Theorem 10.12 for cases in which successively fewer of the power series λ, μ, $f(y)$ are restricted to be homogeneous polynomials. Since Y is finite the theorem follows.

11. The Good Formula

In this section we take J to be the field of rational numbers, or one of its extensions. X, U, and V are as in Section 2. We enumerate the variables in X as $x_1, x_2, \ldots, x_n$, and for each suffix j, from 1 to n, we distinguish a power series F_j of V.

We use the symbols C, C' and C'' to denote n-vectors of nonnegative integers. These are to be written in full as $(c_1, c_2, \ldots, c_n)$, $(c_1', c_2', \ldots, c_n')$, and $(c_1'', c_2'', \ldots, c_n'')$, respectively.

Let λ be any power series of V. We define a power series $Q(\lambda; C)$ as follows. It is $Q(\lambda; f, g)$, where Y and g are such that $g^{-1}(x_j)$ has exactly c_j members, and where $f(y) = F_j$ for each $y \in g^{-1}(x_j)$, for each suffix j. As in Section 10 we have to consider cyclic operators of the form $L = (W, \pi)$. One of these conforming to g, that is satisfying $c_j > 0$ for each $x_j \in W$, may equally well be said to *conform* to C. Analogously we replace the symbol $D_{g,L}$ by $D_{C,L}$.

11.1. *Let λ be constant. Then $Q(\lambda; C) = \lambda$ if C is a zero vector, and $Q(\lambda; C) = 0$ in all other cases* (by 10.1 and 10.10).

11.2.

$$\sum_{(C',C'')} \left[\prod_{j=1}^{n} \left\{ \frac{c_j!}{c_j'!\, c_j''!} \right\} Q(\lambda; C')\, Q(\mu; C'') \right] = Q(\lambda\mu; C),$$

where the summation is over all ordered pairs (C', C'') such that $C' + C'' = C$.

The latter result is a special case of 10.12.

With each $x_j \in X$ let us associate a second variable z_j of X. The variables $z_1, z_2, \ldots, z_n$ need not be all distinct.

Let H be any power series of V. Consider the formal sum

$$E(H) = \sum_{C} \chi(C)\, Q(H; C), \tag{48}$$

where

$$\chi(C) = \prod_{j=1}^{n} \left\{ \frac{z_j^{c_j}}{c_j!} \right\}. \tag{49}$$

We note that the summation on the right of (48) is well defined; a term T of U can be active in $\chi(C)\, Q(H; C)$ only if

$$dT \geqslant \sum_{j=1}^{n} c_j, \tag{50}$$

and this can happen for only a finite number of vectors C. Hence (48) does indeed define a power series $E(H)$ of V.

11.3. *If H is constant, then $E(H) = H$* (by 11.1).

11.4. *Let (R, ρ) be a summable collection of power series of V. Let its sum be H. Then*

$$E(H) = \sum_{r \in R} E(\rho(r)).$$

Proof. We can write

$$E(H) = \sum_{C} \chi(C) \left\{ \sum_{r \in R} Q(\rho(r); C) \right\},$$

by 10.3. But a term T of U can be active in $\chi(C)\, Q(\rho(r); C)$ for only a finite number of vectors C, and with a fixed C for only a finite number of elements r of R. So, by 4.7,

$$E(H) = \sum_{r \in R} \left\{ \sum_{C} \chi(C)\, Q(\rho(r); C) \right\}$$

$$= \sum_{r \in R} E(\rho(r)).$$

11.5. *Let H and K be power series of V. Then*

$$E(H)\, E(K) = E(HK).$$

Proof. $E(H)\, E(K) = \sum \chi(C')\, \chi(C'')\, Q(H; C')\, Q(K; C'')$, by 5.6, where the summation is over all ordered pairs (C', C'') of n-vectors of non-negative integers. By 4.6 we can insert brackets to collect together all such pairs with the same sum C.

$$E(H)\, E(K) = \sum_{C} \chi(C) \left\{ \sum_{C'+C''=C} \prod_{j=1}^{n} \left\{ \frac{c_j!}{c_j'!\, c_j''!} \right\} Q(H; C')\, Q(K; C'') \right\}$$

$$= \sum_{C} \chi(C)\, Q(HK; C), \qquad \text{by 11.2,}$$

$$= E(HK).$$

When H is the power series x_j let us write $E(H) = \xi_j$. We define θ as the substitution from X to X such that $\theta(x_j) = \xi_j$ for each suffix j.

11.6. *θ is a regular substitution.*

This is because the initial of the power series ξ_j is that of

$$\chi(0)\, Q(x_j\,;0) = x_j\,,$$

by (48) and 11.1.

11.7. *Let H be any power series of V. Then*

$$E(H) = \theta(H).$$

Proof. The theorem is true when H is the power series x_j, $(1 \leqslant j \leqslant n)$, by the definition of θ. Hence by 11.3 and 11.5 it is true for the power series T, for each term T of U. Using (19) we find from 11.4 and 11.5 that it is true for a general power series. Theorems 6.4 and 6.5 are also needed to justify the argument.

11.8. $$\theta(x_j) = x_j + z_j\theta(F_j),$$

for each suffix j.

Proof.

$$\begin{aligned}\theta(x_1) &= \sum_C \chi(C)\, Q(x_1\,;C)\\ &= \sum_C \chi(C) \sum_L (-1)^{\alpha(L)}\, D_{C,L} \left\{x_1 \prod_{j=1}^{n} L(F_j^{c_j})\right\}.\end{aligned}$$

Here L must conform to C. The symbol $F_j^{c_j}$ corresponds to $S(x_j)$ in the sense of (43).

W evaluate separately the contribution to the sum on the right of those vectors C for which $c_1 = 0$. For these the order of the symbols x_1 and $D_{C,L}$ can be reversed, by 8.1 and (32). Equivalently we may say that, for each such C,

$$Q(x_1\,;C) = x_1Q(1;C).$$

Hence, by 11.1, the contribution of all such vectors C is x_1.

We can now write

$$\begin{aligned}&\theta(x_1) - x_1\\ &= \sum_C \left(\frac{z_1}{c_1+1}\right) \chi(C) \sum_L (-1)^{\alpha(L)}\, D_{x_1} D_{C,L} \left\{x_1 L(F_1^{c_1+1}) \prod_{j=2}^{n} L(F_j^{c_j})\right\},\end{aligned}$$

thus arranging that the sum on the right is still taken over all n-vectors C of nonnegative integers.

We classify the cyclic operators $L = (W, \pi)$ in two sets N_1 and N_2, according as x_1 is or is not in W. In the first case the product $D_{x_1}D_{C,L}$ of differential operators involves D_{x_1} to the power c_1, in the second case to the power $c_1 + 1$. Using the Leibnitz Formula 8.6 and the Commutative Law 8.7 we find that (if $t = \pi(x_1)$),

$$\begin{aligned}
\theta(x_1) - x_1 \\
&= \sum_C \left(\frac{z_1}{c_1+1}\right) \chi(C)\, x_1 Q(1; C') \\
&+ \sum_C \left(\frac{z_1}{c_1+1}\right) \chi(C) \sum_{L\in N_1} (-1)^{\alpha(L)}\, c_1 D_{C,L} \left\{D_t(F_1^{c_1+1}) \prod_{j=2}^{n} L(F_j^{c_j})\right\} \\
&+ \sum_C \left(\frac{z_1}{c_1+1}\right) \chi(C) \sum_{L\in N_2} (-1)^{\alpha(L)}\, (c_1+1)\, D_{C,L} \left\{F_1^{c_1+1} \prod_{j=2}^{n} L(F_j^{c_j})\right\},
\end{aligned}$$

where C' is obtained from C by increasing the first component c_1 by 1.

We note that

$$D_t(F_1^{c_1+1}) = ((c_1 + 1)/c_1)\, F_1 D_t(F_1^{c_1})$$

if $c_1 \neq 0$, by (31). Moreover $D_t(F_1^{c_1}) = 0$ if $c_1 = 0$, by 8.1. Hence, using 11.1, we find that

$$\begin{aligned}
\theta(x_1) - x_1 &= z_1 \sum_C \chi(C) \sum_L (-1)^{\alpha(L)}\, D_{C,L} \left\{F_1 \prod_{j=1}^{n} L(F_j^{c_j})\right\} \\
&= z_1 \sum_C \chi(C)\, Q(F_1\,; C) \\
&= z_1\theta(F_1).
\end{aligned}$$

A similar argument applies with x_1 replaced by any other member x_k of X. We thus have the required result.

We now have an explicit formula for a solution θ of the substitutional equations

$$\theta(x_j) = x_j + z_j\theta(F_j). \tag{51}$$

We summarize our results so far as follows.

11.9. *There is a unique solution θ for the substitutional equations* (51). *It is a regular substitution, and its effect on any power series H of V is given by*

$$\theta(H) = \sum_C \chi(C)\, Q(H; C), \tag{52}$$

where

$$Q(H;\,C) = \sum_{L} (-1)^{\alpha(L)}\, D_{C,L} \left\{ H \cdot \prod_{j=1}^{n} L(F_j^{c_j}) \right\}. \tag{53}$$

As usual L must conform to C. The uniqueness of the solution is a consequence of 7.4.

The next two theorems state some properties of the substitution θ given by (52) and (53).

11.10. *Let ϕ be a nonsingular substitution from X to X that is admissible with respect to each F_j, and that is conservative with respect to each z_j. Then if ϕ is admissible with respect to some $H \in V$ it is admissible also with respect to $\theta(H)$.*

Proof. Since ϕ is conservative with respect to each z_j we have $\phi(z_j P) = z_j \phi(P)$ by 6.4 whenever ϕ is admissible with respect to the power series P. By repetition of this result we have

$$\phi(\chi(C) \cdot P) = \chi(C)\, \phi(P) \tag{54}$$

for each n-vector C.

Write

$$R_C = \chi(C)\, Q(H;\, C).$$

By (53), 6.4, and 8.9 the substitution ϕ is admissible with respect to R_C. We can write also

$$R_C = \sum_{T \in U} K_C(T) \cdot T$$

where $K_C(T) \in J$, by (19). We observe that $\chi(C)$ is a factor of the term T whenever $K_C(T)$ is nonzero.

Consider the power series $K_C(T) \cdot \phi(T)$. A term T' of U can be active in it for only a finite number of vectors C. For by the preceding observations $\chi(C)$ must be a factor of each term active in $K_C(T) \cdot \phi(T)$. But for any such C the term T' is active in $K_C(T) \cdot \phi(T)$ for only a finite number of terms T, by the admissibility of ϕ with respect to R_C. So T' is active in $K_C(T) \cdot \phi(T)$ for only a finite number of pairs (C, T).

We can now define a power series P by the rule

$$P = \sum K_C(T) \cdot \phi(T),$$

where the well-defined summation is over all ordered pairs (C, T).

Let N_T denote the set of all n-vectors C such that $K_C(T)$ is nonzero. It is a finite set, since $\chi(C)$ must divide T whenever C is in N_T. Then

$$P = \sum_{T\in U} \left\{ \sum_{C\in N_T} K_C(T) \cdot \phi(T) \right\}, \qquad \text{by 4.6,}$$

$$= \sum_{T\in U} \left\{ \sum_{C\in N_T} K_C(T) \right\} \phi(T), \qquad \text{by 5.5,}$$

$$= \sum_{T\in U} \{\text{Coefficient of } T \text{ in } \theta(H)\}\ \phi(T).$$

This summation being well defined we deduce that ϕ is admissible with respect to $\theta(H)$. Moreover $P = \phi(\theta(H))$.

11.11. *Let ϕ be as in* 11.10. *Then there is a substitution ψ from X to X with the following property: if ϕ is admissible with respect to some $H \in V$, then ψ is admissible with respect to H, and moreover $\psi(H) = \phi(\theta(H))$.*

Proof. We define ψ as the substitution from X to X such that

$$\psi(x_j) = \phi(\theta(x_j))$$

for each $x_j \in X$. To justify this definition we note that ϕ is admissible with respect to $\theta(x_j)$, by 6.1 and 11.10. It is also admissible with respect to $\theta(S)$, where S is any term of U, by 6.1 and 11.10.

If ϕ is admissible with respect to $H \in V$, then it is admissible with respect to $\theta(H)$, by 11.10, and we have

$$\phi(\theta(H)) = \sum_{T\in U} \{\text{Coefficient of } T \text{ in } \theta(H)\} \cdot \phi(T)$$

$$= \sum_{T\in U} \left\{\text{Coefficient of } T \text{ in } \sum_{S\in U} H(S)\ \theta(S)\right\} \cdot \phi(T).$$

There are only finitely many S such that T is active in $H(S)\ \theta(S)$. Let them constitute a subset U_T of U. Then

$$\phi(\theta(H)) = \sum_{T\in U} \left\{ \sum_{S\in U_T} \{\text{Coefficient of } T \text{ in } H(S)\ \theta(S)\}\ \phi(T) \right\}.$$

Now a term T' of U is active in

$$\{\text{Coefficient of } T \text{ in } H(S)\ \theta(S)\}\ \phi(T)$$

for only a finite number of terms T, and therefore for only a finite number of pairs (S, T), by the admissibility of ϕ with respect to $\theta(S)$. Using 4.6 we find that

$$\begin{aligned}\phi(\theta(H)) &= \sum_{S\in U}\left\{\sum_{T\in U}\{\text{Coefficient of } T \text{ in } H(S)\,\theta(S)\}\,\phi(T)\right\}\\ &= \sum_{S\in U}\phi(H(S)\,\theta(S))\\ &= \sum_{S\in U} H(S)\,\phi(\theta(S)) = \sum_{S\in U} H(S)\,\psi(S),\end{aligned}$$

by 6.4, applied to θ, ϕ and ψ. Since our summations are well defined we conclude that ψ is admissible with respect to H, and that $\psi(H) = \phi(\theta(H))$.

We can now assert a generalization of 11.9.

11.12. *For each $x_j \in X$ let there be defined $z_j \in X$ and $F_j \in V$. Let ϕ be a nonsingular substitution from X to X which is admissible with respect to each F_j, and which is conservative with respect to each z_j. Then the substitutional equations*

$$\psi(x_j) = \phi(x_j) + z_j\psi(F_j), \qquad 1 \leqslant j \leqslant n, \tag{55}$$

have a unique solution for ψ. If ϕ is admissible with respect to some $H \in V$, then ψ is admissible with respect to H, and moreover

$$\psi(H) = \sum_{C} \chi(C)\,\phi\{Q(H; C)\}. \tag{56}$$

Proof. We consider the regular substitution θ defined by (52). It satisfies (51). By 6.4, 11.10, and 11.11 we can apply ϕ to each side of (51) and obtain (55), with ψ as in 11.11. Thus the ψ of 11.11 is a solution of the substitutional equations (55). It is a unique solution by 7.4. To obtain (56) we apply ϕ to each side of (52), using (54) and 11.11.

When we are given a set of substitutional equations such as (55) it is convenient to classify the variables as *conservative* or *transitional*; x_j is conservative if $F_j = 0$ and transitional otherwise. We assume that there is at least one transitional variable, and we enumerate the transitional variables as $x_1, x_2, \ldots, x_m$. If x_j is a conservative variable we have $\psi(x_j) = \phi(x_j)$. The nature of z_j is then of no importance, but for convenience of exposition we may take it to be identical with one of the variables $z_1, z_2, \ldots, z_m$.

Referring to (52) and (53) we see that if x_j is conservative then the contribution of an n-vector C to the right of (52) is zero whenever c_j is nonzero. We can therefore restrict the sum in (52) to *effective* n-vectors C, in which $c_j = 0$ whenever $j > m$. We may as well replace each such C by the corresponding m-vector $S = (c_1, c_2, ..., c_m)$. The operator $D_{C,L}$ can equally well be written as $D_{S,L}$, for D_{x_j} will always have zero index in it when $j > m$. $\chi(S)$ can be defined by the product on the right of (49) with m replacing n. Similarly $Q(H; S)$ can be defined by the sum on the right of (53), with $D_{C,L}$ written as $D_{S,L}$ and with m replacing n. The statement that $L = (W, \pi)$ conforms to the effective n-vector C means that W consists of transitional variables only. We can express this fact by saying that L *conforms* to the corresponding S.

Using the new notation we rewrite 11.12 as follows.

11.13. *For each $x_j \in X$ with $j \leqslant m$ let there be defined $z_j \in X$ and a nonzero $F_j \in V$. Let ϕ be a nonsingular substitution from X to X which is admissible with respect to each F_j and conservative with respect to each z_j. ($j \leqslant m$). Then the equations*

$$\psi(x_j) = \phi(x_j) + z_j\psi(F_j), \qquad (1 \leqslant j \leqslant m), \tag{57}$$

have a unique solution for the substitution ψ, provided that ϕ and ψ are required to be conservative with respect to each x_j such that $j > m$. If ϕ is admissible with respect to some $H \in V$, then ψ is admissible with respect to H, and moreover

$$\psi(H) = \sum_S \chi(S)\, \phi\{Q(H; S)\}, \tag{58}$$

where

$$Q(H; S) = \sum_L (-1)^{\alpha(L)}\, D_{S,L} \left\{H \cdot \prod_{j=1}^{m} L(F_j^{c_j})\right\}. \tag{59}$$

Here L must conform to S.

To prove this we extend the definitions trivially by putting $F_j = 0$ and $z_j = z_1$ when $m < j \leqslant n$. We have already required $\psi(x_j) = x_j = \phi(x_j)$ when $j > m$, and this can be regarded as extending the list of equations (57) to all values of j. It is now only necessary to apply 11.12 to the extended equations, and convert to the new notation.

To put the above result into a form resembling Good's Theorem 12 we must make the further assumption that each of the variables z_j, $j \leqslant m$, is conservative. Thus each is some x_k with $k > m$. It then follows from (8.1) and the Product Rule that

$$D_{x_j}(z_k P) = z_k D_{x_j}(P)$$

whenever $j \leqslant m$, $k \leqslant m$ and $P \in V$. We are thus free to transfer variables z_j from one side of a differential operator to the other.

Now a term T of U can be active in

$$\chi(S)\ \psi\left\{(-1)^{\alpha(L)}\ D_{S,L}\left\{H\cdot\prod_{j=1}^{m}L(F_j^{c_j})\right\}\right\}$$

only if dT is at least equal to the sum of the components of S. This happens for only finitely many vectors S, and for each of these we have to consider only finitely many cyclic operators L. We can therefore remove and reinsert brackets in our formula for $\psi(H)$, given by (58) and (59), in accordance with 4.6 and 4.7. In this way we can collect together all expressions in which $D_{S,L}$ is the same differential operator. It is convenient to write

$$\Delta_S = D_{x_1}^{c_1}D_{x_2}^{c_2}\cdots D_{x_m}^{c_m}. \tag{60}$$

We now write $Z = \{x_1, x_2, ..., x_m\}$. Applying the proposed transformation we find that

$$\psi(H) = \sum_S \chi(S)\ \phi\left\{\sum_L (-1)^{\alpha(L)}\ \Delta_S\left\{H\right.\right.$$

$$\left.\left.\times\prod_{x_j\in W} z_jD_{\pi(x_j)}\{(F_j^{c_j+1})/(c_j+1)\}\cdot\prod_{x_j\in Z-W}\{F_j^{c_j}\}\right\}\right\}$$

$$= \sum_S \chi(S)\ \phi\left\{\Delta_S\left\{H\cdot\prod_{j=1}^{m}F_j^{c_j}\cdot B\right\}\right\},$$

where

$$B = \sum_L (-1)^{\alpha(L)}\prod_{x_j\in W}\{z_jD_{\pi(x_j)}F_j\},$$

by 5.5 and 8.4. But B is the determinant of the $n\times n$ matrix M for which the entry in the ith row and jth column is

$$\delta(i,j) - z_iD_{x_j}(F_i).$$

Here $\delta(i,j)$ is 1 or 0 according as i and j are equal or unequal. We can now conclude the section with the following result, the Good Formula.

11.14. *Under the conditions of* 11.13, *and with the further restriction that each* z_j, $(j\leqslant m)$, *is conservative, we can replace* (58) *and* (59) *by the following equation.*

$$\psi(H) = \sum_S \chi(S)\ \phi\left\{\Delta_S\left\{H\cdot\prod_{j=1}^{m}F_j^{c_j}\cdot\det(M)\right\}\right\}. \tag{61}$$

12. Applications and Special Cases

We continue with the notation of Section 11. In particular J is still to be the field of rational numbers or one of its extensions.

We ought to notice the special case of 11.13 in which F_j is the unit power series for each $j \leqslant m$. Then the only cyclic operator $L = (W, \pi)$ that can make a nonzero contribution on the right of (59) is the one for which W is null. Equations (58) and (59) then reduce to

$$\psi(H) = \sum_{S} \chi(S)\, \phi\{\Delta_S(H)\}. \tag{62}$$

This is the Taylor Formula. The requirement that ϕ shall be admissible with respect to each F_j is trivially satisfied, but the requirement that ϕ shall be admissible with respect to H is still effective.

Let us next consider Theorem 11.13 in the special case $m = 1$. We drop the suffix, writing c, x, z and F for c_1, x_1, z_1 and F_1, respectively. Then if $c > 0$ we have, by (59) and the Product Rule,

$$\begin{aligned} Q(H; S) &= D_x{}^c\{HF^c\} - D_x^{c-1}\{HD_xF^c\} \\ &= D_x^{c-1}\{F^c \cdot D_x(H)\}. \end{aligned}$$

We thus arrive at the Lagrange Formula.

12.1. *If* $m = 1$ *in* 11.13, *then* (58) *and* (59) *can be replaced by the following equation.*

$$\psi(H) = \phi(H) + \sum_{c=1}^{\infty} (z^c/c!)\, \phi\{D_x^{c-1}\{F^c \cdot D_x(H)\}\}. \tag{63}$$

It seems appropriate to give one or two examples here, if only to enable the reader to relate the present terminology to his own. However I am writing this paper in the belief that the reader is familiar with the use of formal power series to obtain combinatorial identities in the form of relations between coefficients. Accordingly I have not hitherto given examples to illustrate the procedures. I have confined myself to showing that the familiar operations can be described, and the familiar theorems proved, in purely combinatorial terms.

Consider however the problem of determining γ_n, the number of nth powers of a general element x with respect to a nonassociative multiplication. It is discussed for example in [4]. We readily obtain the following rule.

$$\gamma_0 = 1, \qquad \sum_{j=0}^{n-1} \gamma_j \gamma_{n-j} = \gamma_n \qquad \text{when} \quad n > 0. \tag{64}$$

We introduce the power series

$$\Gamma = \sum_{j=0}^{\infty} \gamma_j z^j \tag{65}$$

in a single variable z. We then observe that (64) is equivalent to the assertion that Γ satisfies the equation

$$\Gamma = 1 + z\Gamma^2. \tag{66}$$

This can be treated as a substitutional equation, with the help of 11.13 and its specialization 12.1. We can keep to a single variable, making z serve for both z and x in (63). We have $F = x^2$. The substitution ϕ is defined by $\phi(x) = 1$. It is admissible with respect to F by 6.1. The substitution ψ is defined by $\psi(x) = \Gamma$. By (63) we have

$$\begin{aligned}\Gamma &= 1 + \sum_{c=1}^{\infty} (z^c/c!)\ \phi\{D_x^{c-1} x^{2c}\} \\ &= 1 + \sum_{c=1}^{\infty} (z^c/c!)\ \phi\{(2c)!\ x^{c+1}/(c+1)!\} \\ &= \sum_{c=0}^{\infty} \left\{\frac{z^c \cdot (2c)!}{c!\ (c+1)!}\right\}.\end{aligned}$$

Thus the coefficients γ_n are the Catalan numbers.

In another example, encountered in [1], it is found that there are two power series u and v, satisfying the equations

$$u = x/(1-v)^2, \qquad v = y/(1-u)^2, \tag{67}$$

where x and y are in X (see Section 9). The expression

$$uv(1 - u - v)$$

is then a power series, depending only on x and y, and it is required to determine the coefficients in it.

If we wish to quote 11.14 it is best to introduce two entirely new variables x_1 and x_2, to be the transitional ones. It is always permissible to regard the initial set of variables as a subset of a larger one. (See the discussion of direct embeddings in Section 6). We now define ϕ as the substitution that satisfies $\phi(x_1) = \phi(x_2) = 0$, and that is conservative for all the other variables, including of course x and y. We define ψ as the substitution that satisfies $\psi(x_1) = u$ and $\psi(x_2) = v$, and that is conservative with respect to all the other variables. We then put

$$F_1 = 1/(1-x_2)^2, \qquad F_2 = 1/(1-x_1)^2.$$

We put also

$$H = x_1 x_2 (1 - x_1 - x_2).$$

Now ϕ is a regular substitution. It is therefore admissible with respect to F_1, F_2, and H, by 6.3. Our problem becomes that of determining the power series $\psi(H)$, and we can solve it by applying (61). This is done, with minor changes of notation, in [1].

Other examples of the application of the Good Formula can be found in [2]. They are combinatorial, even though the formula itself is based on analytical arguments in that paper.

We conclude with an observation about coefficients. We have restricted J mainly to justify our divisions by $c_j!$. But in practice we seem to be operating only in the region of integral coefficients. This is clarified by the following theorem.

12.2. *Suppose that in* 11.13 *the coefficients in the power series* F_j *and* $\phi(x_j)$, $j \leqslant m$, *are all integers. Then the coefficients in the power series* $\psi(x_j)$ *are all integers.*

Proof. The theorem holds for terms of degree 0, for the initial of $\psi(x_j)$ is the same as that of $\phi(x_j)$ by (55). Suppose it true for all coefficients of degree $\leqslant n$. Then the coefficients in $\psi(F_j)$ are integers for all terms of degree up to and including n, by the definitions of sum, product and substitution. It follows from (55) that the coefficients in the power series $\psi(x_j)$ are integers for all terms of degree $\leqslant n + 1$. The theorem follows by induction.

This result suggests that it might be possible to rewrite Sections 11 and 12 so as to avoid the restriction on J. However I have not yet had occasion to apply the Lagrange–Good theory to power series with general coefficient-rings.

References

1. W. G. Brown and W. T. Tutte, On the enumeration of rooted nonseparable planar maps, *Can. J. Math.* **16** (1964), 572–577.
2. I. J. Good, Generalizations to several variables of Lagrange's expansion, with applications to stochastic processes, *Proc. Cambridge Phil. Soc.* **56** (1960), 367–380.
3. H. W. Gould, Coefficient identities for powers of Taylor and Dirichlet series, *Amer. Math. Monthly* **81** (1974), 3–14.
4. I. Niven, Formal power series, *Amer. Math. Monthly*, **76** (1969), 871–889.