Discrete Mathematics 37 (1981) 193-202
North-Holland Publishing Company

FROM SETS TO FUNCTIONS: THREE ELEMENTARY EXAMPLES

S.A. JONI*

Carnegie-Mellon University, Pittsburgh, PA 15213, USA

G.-C. ROTA**

Massachusetts Institute of Technology, Cambridge, MA 02139, USA

B. SAGAN

University of Michigan, Ann Arbor, MI 48109, USA

Received 12 August 1980

A sequence of binomial type is a basis for $\mathbb{R}[x]$ satisfying a binomial-like identity, e.g. powers, rising and falling factorials. Given two sequences of binomial type, the authors describe a totally combinatorial way of finding the change of basis matrix: to each pair of sequences is associated a poset whose Whitney numbers of the 1st and 2nd kind give the entries of the matrix and its inverse.

1. Introduction

In the paper "On the foundations of combinatorial theory III. Theory of binomial enumeration" [7], Mullin and Rota developed the theory of polynomials of binomial type (definition below). They observed that in many cases these sequences of polynomials and the linear relations between them, given by the so-called connection constants, could be treated set-theoretically. But the problem of extending this approach to all sequences of binomial type remains open.

In this note—with a view towards understanding the tools required for such a set-theoretic approach—we will treat completely three sequences of binomial type and their connection constants: the power sequence, lower, and upper factorials. None of our formulae are new, but some of the underlying combinatorics is. Specifically, it will be seen that each pair of the above sequences is associated with a partially ordered set and certain 'functions' (actually special cases of Mullin and Rota's reluctant functions). Then the two inverse connection constants formulae for the given sequence pair will be obtained in one case by summing over the poset, and in the other by differentiating i.e. Möbius inversion. Although these polynomial identities are proved only for positive integral values of the variable x, it immediately follows that they hold for all $x \in \mathbb{C}$.

* Research partially supported by NSF Grant MCS77-01947.
** Research partially supported by NSF Grant MCS78-02743.

The most difficult part of this program is to guess the correct poset. We hope that our work will eventually throw some light on this mysterious step.

2. Preliminaries

We first consider three sequences of polynomials which are of common occurrence in combinatorics. The first sequence is the *power sequence*

$$x^n, \quad n=0,1,2,\ldots \tag{2.1}$$

which counts the number of functions from a set with n objects to a set with x objects. Informally we often say that x^n counts the number of ways of placing n distinguishable balls into x distinguishable boxes (the *occupancy* interpretation) or else the number of words of length n from an alphabet with x letters (the *distribution* interpretation).

The second sequence is the *lower factorial sequence*

$$(x)_n = x(x-1)\cdots(x-n+1), \quad n=0,1,2,\ldots \tag{2.2}$$

which counts the number of one-to-one functions from a set with n objects to a set with x objects. In the occupancy interpretation, $(x)_n$ counts the number of ways of placing n distinguishable balls into x distinguishable boxes such that each box contains at most one ball. In the distribution interpretation, we see that $(x)_n$ counts the number of words of length n made from an alphabet with x letters such that no word has a repeated letter.

Our third sequence is the *upper factorial sequence*

$$(x)^n = x(x+1)\cdots(x+n-1), \quad n=0,1,2,\ldots \tag{2.3}$$

Unlike the first two sequences, these polynomials do not count functions. Instead, they count the number of *dispositions* from a set with n objects to a set with x objects. Dispositions can be visualized (occupancy interpretation) as all ways of placing n distinguishable flags on x distinguishable flagpoles. It is easy to see that $(x)^n$ counts all such arrangements: first we have x choices of a flagpole for the first flag. If flag 2 is on the same pole as flag 1, then we can place it above or below flag 1. Otherwise, it is on one of the $x-1$ remaining poles. Thus there are $x+1$ choices for flag 2. Similarly, there are $x+2$ choices for flag 3 and, in general, $x+k-1$ choices for flag k.

More precisely, let $[n]$ denote the set $\{1,2,\ldots,n\}$. Then a disposition, d, is a function from $[n]$ to $[x]$ *together* with a linear order on the elements of each pre-image $d^{-1}(y)$, $y\in[x]$. For example, one possible disposition from [5] to [2] is

$$d^{-1}(1)=2,3; \qquad d^{-1}(2)=5,1,4$$

or

$$d(2,3)=1; \qquad d(5,1,4)=2.$$

Note that this is different from the disposition

$$d'(3,2)=1; \qquad d'(5,1,4)=2$$

since the linear order of 2 and 3 is reversed.

Let $(p_n(x))$ and $(q_n(x))$ be two polynomial sequences such that for all $n \geq 0$, $\deg(p_n(x)) = \deg(q_n(x)) = n$. By elementary linear algebra, there exists two sequences of *connection constants* $(c_{n,k})$ and $(d_{n,k})$ such that for all n

$$p_n(x) = \sum_k c_{n,k} q_k(x) \tag{2.4}$$

and

$$q_n(x) = \sum_k d_{n,k} p_k(x). \tag{2.5}$$

When the polynomial sequences in question are of *binomial type* [7], that is, when for all n

$$p_n(x+y) = \sum_k \binom{n}{k} p_k(x) p_{n-k}(y) \tag{2.6}$$

these connection constants have been extensively studied; operator theoretic methods yield formulae for the constants [2, 8]. It is well known that the sequences introduced in (2.1), (2.2), and (2.3) are all of binomial type. As stated in Section 1, the purpose of this paper is to give set-theoretic proofs of the corresponding connection coefficient formulae. In each case we shall give a direct counting argument for one of the connection formulae (2.4), (2.5) and realize the inverse formula via Möbius inversion over a suitably constructed poset.

So that this work be reasonably self-contained, we shall state the Möbius inversion theorem. The reader is referred to [1], [3], [4], and [9] for further study. A poset P is said to have a 0 if it has a unique minimal element. All our posets will have a 0.

Möbius inversion theorem. *Let P be a given poset having a 0, and let f and g be maps from P to a field F such that for all $\tau \in P$*

$$g(\tau) = \sum_{\sigma \geq \tau} f(\sigma). \tag{2.7}$$

Then there exists a unique function $\mu : P \to F$ such that

$$f(0) = \sum_{\sigma \in P} \mu(\sigma) g(\sigma). \tag{2.8}$$

The function μ is called the *Möbius function* of P.

Given $\sigma \in P$, the *interval* $[0, \sigma]$ is the subposet $\{\tau \in P \mid \tau \leq \sigma\}$. P is said to satisfy the *chain condition* if for each $\sigma \in P$, all maximally total ordered subsets of

$[0, \sigma]$ have the same cardinality. In this case the *rank* of σ, $r(\sigma)$, is defined as one less than the number of elements in a maximal chain (totally ordered subset) of $[0, \sigma]$. For these posets we also have *Whitney numbers of the first and second kind*, $w_k(P)$ and $W_k(P)$ respectively, defined by

$$w_k(P) = \sum_{\substack{\sigma \in P \\ r(\sigma)=k}} \mu(\sigma) \tag{2.9}$$

and

$$W_k(P) = \sum_{\substack{\sigma \in P \\ r(\sigma)=k}} 1. \tag{2.10}$$

3. Powers and lower factorials

The simplest connection constants are these associated with the polynomial sequences (x^n) and $((x)_n)$. One set of constants is given by the classical formula

$$x^n = \sum_k S(n, k)(x)_k. \tag{3.1}$$

Here $S(n, k)$ denotes a *Stirling number of the second kind*, and counts the number of partitions of $[n]$ into k parts or blocks. The $S(n, k)$ are the Whitney numbers of the second kind for the poset Π_n, where Π_n is the poset of all partitions of $[n]$ ordered by refinement. Thus in Π_n we say that $\sigma \leq \pi$ if each block of π is the union of blocks of σ. We repeat a well-known proof of formula (3.1). Corresponding to each function f mapping $[n]$ to $[x]$, there exists a unique partition π_f of $[n]$. π_f is obtained by placing i, j in the same block of π_f if and only if $f(i) = f(j)$. Let π be an element of Π_n and suppose π has k blocks. Since $(x)_k$ counts the number of one-to-one maps of $[k]$ to $[x]$, there are $(x)_k$ maps f from $[n]$ to $[x]$ such that $\pi_f = \pi$. Therefore we see that

$$x^n = \sum_{\pi \in \Pi_n} (x)_{\nu(\pi)} \tag{3.2}$$

where $\nu(\pi)$ is the number of blocks of π. Collecting all terms such that the value of $\nu(\pi) = k$ in formula (3.2) immediately yields (3.1).

The inverse connection formula corresponding to (3.1) is

$$(x)_n = \sum_k s(n, k)x^k \tag{3.3}$$

where the $s(n, k)$ denote the *Stirling numbers of the first kind*. We derive this formula by using Möbius inversion on Π_n as follows. For all $\pi \in \Pi_n$, set

$$f(\pi) = (x)_{\nu(\pi)} \quad \text{and} \quad g(\pi) = x^{\nu(\pi)}.$$

Then formula (3.2) can be rewritten as

$$g(0)=\sum_{\pi\geq 0} f(\pi) \tag{3.4}$$

where the finest partition 0 is the partition each of whose blocks contain only one element, i.e. $1/2/\cdots/n$. Since for all $\sigma\in\Pi_n$ the poset $P_\sigma=\{\pi\in\Pi_n \mid \pi\geq\sigma\}$ is isomorphic to $\Pi_{\nu(\sigma)}$, formula (3.4) can be generalized to

$$g(\sigma)=\sum_{\pi\geq\sigma} f(\pi). \tag{3.5}$$

Consequently, using the Mobius inversion theorem (formula (2.8)), we have

$$(x)_n=f(0)=\sum_{\pi\in\Pi_n}\mu(\pi)x^{\nu(\pi)}=\sum_k x^k \sum_{\substack{\pi\in\Pi_n\\ \nu(\pi)=k}}\mu(\pi) \tag{3.6}$$

However as it is well known and easily shown [9] that the Whitney numbers of the first kind for Π_n are the Stirling numbers of the first kind, and our proof is complete.

The very simple pattern of this example will be seen to hold for our other two cases although the corresponding posets will, of course, be different.

4. Lower factorials and upper factorials

The first connection constants formula for $((x)^n)$ and $((x)_n)$ is

$$(x)^n=\sum_k \frac{n!}{k!}\binom{n-1}{k-1}(x)_k. \tag{4.1}$$

A *linear partition*, λ, is a partition of $[n]$ together with a total order on the numbers in each block. The blocks themselves are unordered. Let $\mathcal{L}_n$ denote the collection of all linear partitions of $[n]$ and let $\nu(\lambda)$ denote the number of blocks of λ. In Section 3 we saw that each function f from $[n]$ to $[x]$ can be thought of as a pair (g,π) where g is a one-to-one map from $[k]$ to $[x]$ and π is a partition of $[n]$ into k blocks. Similarly, each disposition can be thought of as a pair consisting of a one-to-one function g mapping $[k]$ to $[x]$, and a linear partition λ of $[n]$ into k blocks. Thus, since $(x)^n$ counts the total number of dispositions of $[n]$, we have

$$(x)^n=\sum_{\lambda\in\mathcal{L}_n}(x)_{\nu(\lambda)}. \tag{4.2}$$

To obtain the number of linear partitions of $[n]$ into k blocks, we first note that there are $n!\binom{n-1}{k-1}$ linear partitions of $[n]$ with k *ordered* blocks. This can be visualized as all the ways of placing $k-1$ slashes into the $n-1$ interior spaces of a linear arrangement (or permutation) of $[n]$. Since we wish to count linear partitions with *unordered* blocks we divide by $k!$, obtaining the desired $(n!/k!)\times\binom{n-1}{k-1}$. These numbers are known as the *Lah numbers* [6]. Thus, collecting terms in formula (4.2) according to the values of $\nu(\lambda)$ we obtain formula (4.1).

As in the previous section, we turn to proving the inverse connection constants formula by Möbius inversion. Our poset is the poset of linear partitions $\mathscr{L}_n$, where we set $\eta\leqslant\lambda$ if each block of λ can be obtained by the juxtaposition of blocks of η. Note that $\mathscr{L}_n$ itself is not a lattice since it has $n!$ maximal elements, corresponding to the $n!$ linear arrangements of $[n]$. Our inverse connection constants formula is

$$(x)_n=\sum_k(-1)^{n-k}\frac{n!}{k!}\binom{n-1}{k-1}(x)^k. \tag{4.3}$$

To prove formula (4.3) let us set for all $\lambda\in\mathscr{L}_n$,

$$f(\lambda)=(x)_{\nu(\lambda)}\quad\text{and}\quad g(\lambda)=(x)^{\nu(\lambda)}.$$

The minimal linear partition 0 is the linear partition $1/2/3/\cdots/n$. Formula (4.2) can be rewritten as

$$g(0)=\sum_{\lambda\geqslant 0}f(\lambda). \tag{4.4}$$

In fact for any η

$$g(\eta)=\sum_{\lambda\geqslant\eta}f(\lambda)$$

since $P_\eta=\{\lambda\in\mathscr{L}_n\mid\lambda\geqslant\eta\}$ is isomorphic to $\mathscr{L}_{\nu(\eta)}$. Therefore Möbius inversion gives (cf. (2.8))

$$(x)_n=\sum_{\lambda\in\mathscr{L}_n}\mu(\lambda)(x)^{\nu(\lambda)}. \tag{4.5}$$

Moreover, we claim that for all $\lambda\in\mathscr{L}_n$,

$$\mu(\lambda)=(-1)^{n-\nu(\lambda)}. \tag{4.6}$$

This is easily seen once we observe that for all $\lambda\in\mathscr{L}_n$ the interval $B_\lambda=[0,\lambda]$ is a Boolean lattice, and recall the well-known fact [9] that for all b in a Boolean lattice B $\mu(b)=(-1)^{\operatorname{rank}(b)}$. Since the rank of λ in B_λ is $n-\nu(\lambda)$, (4.6) is established, and combining this result with (4.5) gives

$$\begin{aligned}(x)_n&=\sum_{\lambda\in\mathscr{L}_n}(-1)^{n-\nu(\lambda)}(x)^{\nu(\lambda)}\\&=\sum_k(-1)^{n-k}\frac{n!}{k!}\binom{n-1}{k-1}(x)^k.\end{aligned}$$

Whence our proof is complete.

5. Upper factorials and powers

The connection constants formulae we shall consider in our penultimate section are

$$(x)^n = \sum_k |s(n, k)|\, x^k, \tag{5.1}$$

and

$$x^n = \sum_k (-1)^{n-k} S(n, k)(x)^k. \tag{5.2}$$

Once again, the key to our proof is the discovery of the right poset. We denote the collection of all permutations (or linear arrangements) of $[n]$ by $\mathfrak{S}_n$. Each permutation σ in $\mathfrak{S}_n$ has a unique decomposition into disjoint cycles where $(i_1, i_2, \ldots, i_r)$ is a cycle of σ if σ maps i_1 to i_2, i_2 to $i_3, \ldots,$ and i_r to i_1. *We shall make the convention that each cycle is written so that its leftmost entry is minimal.* i.e.,

$$i_1 = \min\{i_1, i_2, \ldots, i_r\}.$$

Since it is well known that $|s(n, k)|$ counts the number of permutations of $[n]$ with k cycles, formula (5.1) is established by demonstrating a bijection between dispositions of $[n]$ and pairs (σ, f) where for some $1 \leq k \leq n$, $\sigma \in \mathfrak{S}_n$ has k cycles, and f is a map from $[k]$ to $[x]$. We construct the disposition d from the pair (σ, f) as follows: let $c_1, c_2, \ldots, c_k$ denote the cycles of σ labeled so that $\min c_1 = 1 < \min c_2 < \cdots < \min c_k$, and let $\{j_1 < j_2 < \cdots < j_s\} = f^{-1}(y)$ for some $1 \leq y \leq x$. For $f^{-1}(y) \neq \{\emptyset\}$, we shall define d to map all elements of $\{c_{j_1}, c_{j_2}, \ldots, c_{j_s}\}$ to y. The linear order of $d^{-1}(y)$ is obtained by the juxtaposition $c_{j_s} c_{j_{s-1}} \cdots c_{j_2} c_{j_1}$ (where each cycle is written with smallest element first). For example, if $\sigma = (1, 6, 2)(3, 4)(5, 7)(8)(9)$, $f(1) = f(3) = f(4) = y_1$, and $f(2) = f(5) = y_2$, then the corresponding disposition d is given by

$$d(8, 5, 7, 1, 6, 2) = y_1 \quad \text{and} \quad d(9, 3, 4) = y_2.$$

Conversely, given a disposition d, the pair (σ, f) is constructed as follows. Find the block of d containing 1, say $d(\ldots, 1, i_1, \ldots, i_r) = y_1$. Then c_1, the first cycle of σ is $(1, i_1, \ldots, i_r)$, and $f(1) = y_1$. Next locate the block of d containing the smallest integer m in $[n] - \{1, i_1, \ldots, i_r\}$. Either

$$d(\ldots, m, v_1, \ldots, v_t) = y_2$$

or

$$d(\ldots, m, v_1, \ldots, v_t, 1, i_1, \ldots, i_r) = y_1.$$

In either case the second cycle of σ, c_2, is $(m, v_1, \ldots, v_t)$. Moreover, we set $f(2) = y_2$ in the first case or $f(2) = y_1$ in the second. Continuing our above example, if

$$d(8, 5, 7, 1, 6, 2) = y_1 \quad \text{and} \quad d(9, 3, 4) = y_2,$$

then $c_1=(1,6,2)$, $c_2=(3,4)$, $c_3=(5,7)$, $c_4=(8)$, $c_5=(9)$, $f(1)=f(3)=f(4)=y_1$, and $f(2)=f(5)=y_2$.

Since it is easily seen that the above two constructions are inverses of each other, our proof is complete.

To obtain formula (5.2) we consider the following partial order on $\mathfrak{S}_n$. Given σ, τ in $\mathfrak{S}_n$ we shall say that $\sigma \leq \tau$ if each cycle of σ (written with smallest element first) is composed of a string of consecutive integers from some cycle of τ. For example, $(12)(3) \leq (123)$, $(1)(23) \leq (123)$, but $(13)(2) \not\leq (123)$. The 0 of $\mathfrak{S}_n$, considered as a poset, is the identity permutation $(1)(2)\cdots(n)$, and we shall set $c(\sigma)=$ the number of cycles of σ. This given, formula (5.1) can be rewritten

$$(x)^n = \sum_{\sigma \geq 0} x^{c(\sigma)}. \tag{5.3}$$

This formula can be generalized to $P_\tau = \{\sigma \in \mathfrak{S}_n \mid \sigma \geq \tau\}$ and so Möbius inversion over $\mathfrak{S}_n$ gives

$$x^n = \sum_{\sigma \in \mathfrak{S}_n} \mu(\sigma)(x)^{c(\sigma)}. \tag{5.4}$$

We shall say that $\sigma \in \mathfrak{S}_n$ is *increasing* if each of its cycles increases from left to right i.e. if $(i_1, i_2, \ldots, i_s)$ is a cycle in σ, then $i_1 < i_2 < \cdots < i_s$. The Mobius function for $\mathfrak{S}_n$ is given in the following:

Lemma 5.1. *For each* $\sigma \in \mathfrak{S}_n$,

$$\mu(\sigma) = \begin{cases} (-1)^{n-c(\sigma)} & \text{if } \sigma \text{ is increasing,} \\ 0 & \text{otherwise.} \end{cases} \tag{5.5}$$

Proof. Given $\sigma \in \mathfrak{S}_n$, consider the interval $I_\sigma = [0, \sigma]$. It is easily verified that I_σ is a lattice. The atoms of I_σ (or elements of rank 1) correspond to the transpositions (i_r, i_{r+1}), where (i_r, i_{r+1}) is a substring of a cycle of σ and $i_r < i_{r+1}$. Thus, if σ is increasing the atoms of I_0 correspond to all of the possible $n - c(\sigma)$ transpositions. Hence I_0 is a Boolean lattice, and so $\mu(\sigma) = (-1)^{n-c(\sigma)}$. Moreover, if σ is not increasing, then in some cycle of σ we have a consecutive pair $(\ldots, i_r, i_{r+1}, \ldots)$ such that $i_r > i_{r+1}$. Hence σ is not the joint of atoms of I_σ, and consequently by Hall's theorem [5] $\mu(\sigma)=0$.

We apply Lemma 5.1 to formula (5.4) and obtain

$$\begin{aligned} x^n &= \sum_{\substack{\sigma \in \mathfrak{S}_n \\ \sigma \text{ increasing}}} (-1)^{n-c(\sigma)}(x)^{c(\sigma)} \\ &= \sum_k (-1)^{n-k}(x)^k \sum_{\substack{\sigma \text{ increasing} \\ c(\sigma)=k}} 1. \end{aligned} \tag{5.6}$$

Table 1

Polynomial	Expressed in terms of	Method	Poset
x^n	$(x)_n$	summation	Π_n, the lattice of partitions of $[n]$
$(x)_n$	x^n	inversion	
$(x)^n$	$(x)_n$	summation	$\mathscr{L}_n$, the poset of linear partitions of $[n]$
$(x)_n$	$(x)^n$	inversion	
$(x)^n$	x^n	summation	$\mathfrak{S}_n$, the poset of permutations of $[n]$
x^n	$(x)^n$	inversion	

Since the number of increasing permutations with k cycles is clearly equal to $S(n, k)$ (i.e. the number of partitions of $[n]$ with k blocks) our proof of formula (5.2) is complete.

6. Summary

We have summarized the results of Sections 3 to 5 in Table 1. To further emphasize the similarities between the three cases, we note the following properties of the posets P_n ($n = 1, 2, 3, \ldots$) associated with any given pair of sequences.

(1) Each P_n is a ranked poset having a 0.

(2) The constant coefficient formulae are obtained by collecting all terms at given rank. Hence the constants are just the Whitney numbers of the second kind (when summing over P_n) or of the first kind (from Möbius inversion).

(3) The Möbius function of P_n alternates from rank to rank. Here 'alternates' is taken in the weak sense where zero may be considered positive or negative.

(4) For any $\pi \in P_n$, the poset $P_\pi = \{\sigma \in P_n \mid \sigma \geqslant \pi\}$ is isomorphic to $P_{n-r(\pi)}$. It is this fact that permits us to use the Inversion theorem.

References

[1] H.H. Crapo, The Möbius function of a lattice, J. Combin. Theory 1 (A) (1966) 126–131.

[2] A.M. Garsia and S.A. Joni, A new expression for umbral operators and power series inversion, Proc. Am. Math. Soc. 64 (1977) 179–185.

[3] C. Greene, On the Möbius algebra of a partially ordered set, Adv. Math. 10 (1973) 177–187.

[4] M. Hall Jr., Combinatorial Theory (Blaisdell, Waltham, 1967).

[5] P. Hall, A contribution to the theory of groups of prime power order, Proc. London Math. Soc., II. Ser. 36 (1932) 39–95.

[6] I. Lah, Eine neue Art von Zahlen, ihre Eigenschaften . . . , Mittelungsbl. Math. Statist. 7 (1955) 203–213.
[7] R. Mullin and G.-C. Rota, On the foundations of combinatorial theory III. Theory of binomial enumeration, in: Graph Theory and Applications (Academic Press, New York, 1970).
[8] G.-C. Rota, Ed., Finite Operator Calculus (Academic Press, New York, 1975).
[9] G.-C. Rota, On the foundations of combinatorial theory I. Theory of Möbius functions, Z. Wahrsch. Verw. Gebiete 2 (1964) 340–368.