Wiley Classics Library

HALL, JR.

Combinatorial Theory

Second Edition

This page intentionally left blank

Combinatorial Theory

This page intentionally left blank

MARSHALL HALL, JR.
Emory University

Combinatorial Theory

Second Edition

Wiley Classics Library Edition Published 1998

A Wiley-Interscience Publication
JOHN WILEY & SONS, INC.
New York • Chichester • Weinheim • Brisbane • Singapore • Toronto

This text is printed on acid-free paper. ♾

Wiley Classics Library edition published 1998.

Library of Congress Cataloging in Publication Data:

Hall, Marshall, 1910–
Combinatorial theory.
(Wiley-Interscience series in discrete mathematics)
"A Wiley-Interscience publication."
Bibliography: p.
Includes index.
1. Combinatorial analysis. I. Title. II. Series.
QA164.H3 1986 511'.6 85-26799
ISBN 0-471-09138-3
ISBN 0-471-31518-4 (Paperback)

10 9 8 7 6 5 4 3 2 1

Dedicated to the memory of
Herbert John Ryser

This page intentionally left blank

Preface

Combinatorics has been very active since the appearance of the first edition of this book in 1967. Important new results and new methods have both broadened and unified the subject.

In the period 1972–1975 Richard Wilson proved the major asymptotic result on block designs. For a design D with parameters (v, b, r, k, λ) the elementary relations $bk = vr$ and $r(k-1) = \lambda(v-1)$ must hold. He shows that, subject to these conditions, given k and λ, a design always exists for sufficiently large v. He treated the even broader case of pairwise balanced designs in which block sizes k_i, $i = 1, \ldots, m$ are permitted. These results are given in Section 15.6.

In 1926 B. L. van der Waerden conjectured that, for a doubly stochastic matrix (nonnegative entries summing to 1 in every row and column) of size, the minimum value of the permanent would be $n!/n^n$, this being achieved with every entry being $1/n$. The truth of this was proved by Egorychev in 1980. The proof of this is given in Section 5.4.

Perhaps challenged by the statement in the first edition that "no infinite class of Williamson matrices has been found," Richard Turyn in 1972 found an infinite class. This is given in Section 14.3.

The theory of error-correcting codes has been growing by leaps and bounds. Its relationship to design theory was brought out in 1973 by a major paper by F. J. MacWilliams, N. J. A. Sloane, and J. W. Thompson. This connection and some of its consequences are the subject of the new Chapter 17.

This book is not an encyclopaedia. In particular, the recent proof of the famous four-color conjecture is not included. But a concious effort has been made to bring it up to date in the areas it covers.

Many people have been helpful. A partial list is: H. J. Ryser, W. H. Mills, R. M. Wilson, W. G. Bridges, R. Mena, E. F. Assmus, J. L. Hayden, R. Calderbank, N. J. A. Sloane, R. J. McEliece, D. Knuth, S. S. Shrikhande, Navin Singhi, Clement Lam, J. H. van Lint, J. I. Hall, J. J. Seidel, and Robert Roth.

MARSHALL HALL, JR.

Atlanta, Georgia
January 1986

This page intentionally left blank

Preface to the First Edition

Combinatorial theory is the name now given to the subject formerly called "combinatorial analysis" or "combinatorics," though these terms are still used by many people. Like many branches of Mathematics, its boundaries are not clearly defined, but the central problem may be considered that of arranging objects according to specified rules and finding out in how many ways this may be done. If the specified rules are very simple, then the chief emphasis is on the enumeration of the number of ways in which the arrangement may be made. If the rules are subtle or complicated, the chief problem is whether or not such arrangements exist, and to find methods for constructing the arrangements. An intermediate area is the relationship between related choices, and a typical theorem will assert that the maximum for one kind of choice is equal to the minimum for another kind.

The text is divided into three major segments. The first four chapters deal with problems of enumeration. Chapters 5 through 9 deal with the intermediate area of theorems on choice. Chapters 10 through 16 are concerned with the existence and construction of designs.

The theory of enumeration is covered extensively in the classical work of Major P. A. MacMahon, *Combinatorial Analysis*, London, Vol. I, 1915, Vol. II, 1916, and in the recent book by John Riordan, *An Introduction to Combinatorial Analysis*, John Wiley & Sons, Inc., New York, 1958. The treatment of this subject in the first four chapters of this book is relatively brief, and does not attempt to match the scope of these books. H. J. Ryser in the Carus monograph, *Combinatorial Mathematics*, 1963, gives a brief but elegant account of the theorems on choice and the construction and existence of block designs.

Many people have been helpful to me in preparing this book. These include Dr. Leonard Baumert, Professor Robert Dilworth, Dr. Karl Goldberg, Professor Donald Knuth, Dr. Morris Newman, and Professor A. W. Tucker. Special thanks are due to Professors Garrett Birkhoff, Robert Greenwood, and Herbert Ryser, who read the entire manuscript and gave me many helpful suggestions. In preparation of the manuscript and correction of clerical errors, the assistance of Mrs. Kay Hardt, Dr. Allen Pfeffer, and Mr. Robert McEliece was invaluable.

I am indebted to the Literary Executor of the late Sir Ronald A. Fisher, F.R.S., Cambridge, to Dr. Frank Yates, F.R.S., Rothamsted, and to Messrs. Oliver & Boyd Ltd., Edinburgh, for permission to quote a portion of text from their book *Statistical Tables for Biological, Agricultural, and Medical Research.* I also wish to thank Dr. C. R. Rao of the Indian Statistical Institute and the editors of *Sankhyā* for permission to quote from the paper "A Study of BIB Designs with Replications 11 to 15."

MARSHALL HALL, JR.

Pasadena, California

Contents

This page intentionally left blank

Combinatorial Theory

This page intentionally left blank

1

Permutations and Combinations

1.1. DEFINITIONS

A permutation is an ordered selection of objects from a set S.
A combination is an unordered selection of objects from a set S.

We may or may not permit repetition in our permutations and combinations. Thus, selecting two letters from the three letters a, b, c, we have nine permutations with repetitions permitted:

$$aa, ab, ac, ba, bb, bc, ca, cb, cc.$$

We have six permutations without repetitions:

$$ab, ac, ba, bc, ca, cb.$$

We have six combinations with repetitions permitted:

$$aa, bb, cc, ab, ac, bc,$$

and three combinations without repetitions:

$$ab, ac, bc.$$

The number of permutations of n things taken r at a time, without repetition, written ${}_nP_r$, is easily evaluated. For in a permutation $a_1a_2\cdots a_r$ we may choose a_1 as any of the n objects, a_2 as any one of the remaining $(n-1)$ objects, and having chosen $a_1a_2\cdots a_i$, we may take a_{i+1} as any one of the $(n-i)$ remaining objects. Hence,

$$ {}_nP_r = n(n-1)\cdots(n-r+1) = \frac{n!}{(n-r)!} = (n)_r. \tag{1.1.1}$$

A combination of n things taken r at a time without repetition, say, $a_1 a_2 \cdots a_r$, will lead to $r!$ different permutations, namely, all $r!$ permutations of $a_1, \ldots, a_r$. Hence the number of combinations of n things taken r at a time, written ${}_nC_r$, is given by

$$ {}_nC_r = \frac{{}_nP_r}{r!} = \frac{n!}{(n-r)!r!} = \binom{n}{r}, \tag{1.1.2} $$

which is the familiar binomial coefficient. Indeed, in the product $(x + y)^n = (x + y) \cdots (x + y)$, the coefficient of the term $x^r y^{n-r}$ is the number of ways of choosing r of the factors $x + y$, from which we take an x, and then y from the remaining $n - r$ factors $x + y$. We note that

$$ {}_nC_r = {}_nC_{n-r}. \tag{1.1.3} $$

The number of permutations of n things taken r at a time, repeats permitted, is n^r, since in $a_1 a_2 \cdots a_r$ there are n choices for each of $a_1, a_2, \ldots, a_r$ in turn.

To find the number of combinations of n things taken r at a time with repeats permitted, we cannot simply divide n^r by an appropriate factor, since different combinations may yield a different number of permutations. Thus, taking combinations of a, b, c, d, e three at a time, the combination abc gives six permutations, the combination aab gives three permutations, and the combination aaa gives only one permutation. Here we use the device of counting a different set, which is in one-to-one correspondence with our given set. To a given combination (say, bbd), let us adjoin the entire set $abcde$ and write the whole set in order $abbbcdde$ and then insert marks separating the different letters thus: $a \mid bbb \mid c \mid dd \mid e$. In general, to a combination of r letters, repeats permitted, from a set of n letters adjoin all n letters and write in order the set of $(n + r)$ letters and then insert $(n - 1)$ marks between the different letters. Thus, with $(n + r)$ positions to be filled and $(n + r - 1)$ spaces between these positions, we are to insert $(n - 1)$ marks. The number of ways of doing this is

$$ \binom{n + r - 1}{n - 1} = \binom{n + r - 1}{r}. \tag{1.1.4} $$

There is a one-to-one correspondence between the ways of inserting the $(n - 1)$ marks in the $(n + r - 1)$ spaces and the combinations with repeats of n things taken r at a time. Hence, this number is $\binom{n+r-1}{r}$. The expression for the number of combinations, n things taken r at a time, without repeats and with repeats, are similar in form. Thus, for five things taken three at a time, the numbers are, respectively,

$$ \frac{5 \cdot 4 \cdot 3}{1 \cdot 2 \cdot 3} \quad \text{and} \quad \frac{5 \cdot 6 \cdot 7}{1 \cdot 2 \cdot 3}, $$

where the factors in the numerators decrease in one case and increase in the other.

The number of combinations with repeats permitted of n things taken r at a time is the number of solutions $(x_1, x_2, \ldots, x_n)$ in nonnegative integers x_i of

$$r = x_1 + x_2 + \cdots + x_n, \tag{1.1.5}$$

where x_i is the number of times the ith object is included in the combination. This suggests another, but similar, evaluation of the number of combinations. Put $y_i = x_i + 1$, $i = 1, \ldots, n$. Then (1.1.5) becomes

$$n + r = y_1 + y_2 + \cdots + y_n, \tag{1.1.6}$$

and, the number of solutions $(y_1, \ldots, y_n)$ of (1.1.6) in positive integers y_i is clearly the same as the number of solutions of (1.1.5) in nonnegative integers. If we take $(n + r)$ dots and place $(n - 1)$ marks in the $(n + r - 1)$ spaces between the dots, we may take y_1 as the number in the first set of dots, y_2 as the number in the second set, and so on. Thus, again we see that the number of solutions of (1.1.6) is $\binom{n+r-1}{n-1}$, and this is in turn the number of nonnegative solutions of (1.1.5), and so it is the number of combinations of n things r at a time with repeats permitted.

As a further application of this method we may find the number of combinations of $1, 2, \ldots, n$ taken r at a time without including repeats or consecutive numbers. Let us list $1, 2, \ldots, n$ in order and put a mark after each number selected. If there are x_1 numbers before the first mark, x_2 between the first and second mark, and finally x_{r+1} after the last mark, then these determine the choice and

$$n = x_1 + x_2 + \cdots + x_{r+1}, \tag{1.1.7}$$

where $x_1 \geqslant 1, x_2 \geqslant 2, \ldots, x_r \geqslant 2$, and $x_{r+1} \geqslant 0$. We now write

$$n - r + 2 = x_1 + (x_2 - 1) + \cdots + (x_r - 1) + (x_{r+1} + 1), \tag{1.1.8}$$

giving a representation of $n - r + 2$ as a sum of $(r + 1)$ positive integers, and this number is $\binom{n-r+1}{r}$, the number of ways of putting r marks in $(n - r + 1)$ spaces.

There are an enormous number of identities involving binomial coefficients, a few of which follow:

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n; \tag{1.1.9a}$$

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = \begin{cases} 0, & n > 0 \\ 1, & n = 0; \end{cases} \tag{1.1.9b}$$

$$\sum_{k=1}^{n} k(-1)^k \binom{n}{k} = \begin{cases} 0, & n > 1 \\ k, & n = 1; \end{cases} \tag{1.1.9c}$$

$$\sum_{k=r}^{n} (-1)^k \binom{k}{r}\binom{n}{k} = \begin{cases} 0, & n > r \\ (-1), & n = r. \end{cases} \tag{1.1.9d}$$

These may be derived from the relation

$$(1 + x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k.$$

To obtain (1.1.9a) and (1.1.9b), put $x = 1$ and $x = -1$, respectively. For (1.1.9c), differentiate with respect to x and then put $x = -1$. For (1.1.9d), differentiate r times with respect to x, divide by $r!$, and put $x = -1$.

If we have permutations $a_1 a_2 \cdots a_n$ of n objects, of which b_1 are of one kind, b_2 of a second kind, and b_i of an ith kind for i running to r, where naturally $b_1 + b_2 + \cdots + b_r = n$, we may first replace the b_i objects of the ith kind by distinct objects in every case, and then we have $n!$ permutations. But by identifying the like objects, we have counted each permutation $b_1!b_2!\cdots b_r!$ times. Hence, the number of permutations is

$$\frac{n!}{b_1!b_2!\cdots b_r!}, \qquad b_1 + b_2 + \cdots + b_r = n, \tag{1.1.10}$$

the familiar multinomial coefficient.

1.2. APPLICATIONS TO PROBABILITY

In a given situation let us suppose that there are n possible outcomes, which we label $x_1, x_2, \ldots, x_n$, and which are mutually exclusive. We assign to the outcome x_i a number $p_i = p(x_i)$, where p_i is a real number, $p_i \geqslant 0$, and $p_1 + p_2 + \cdots + p_n = 1$. If an event E occurs along with the possibilities $x_{i_1}, \ldots, x_{i_m}$ and not otherwise, we define the probability of E as $p(E) = p_{i_1} + \cdots + p_{i_m}$. The assignment of the initial probabilities $p_1, p_2, \ldots, p_n$ is not a mathematical problem, but is an estimate of the relative likelihoods of the different outcomes, and in any actual case the validity of the mathematical calculation of $p(E)$ depends on the correctness of this assignment.

There are many practical situations in which it seems reasonable to consider the n outcomes as equally likely, and so we take $p_1 = p_2 \cdots = p_n = 1/n$. In this case the probability of the event E occurring along with m possible outcomes, but no others, is $p(E) = m/n$. In such a situation the calculation of $p(E)$ becomes the purely combinatorial problem of calculating m, the number of possible outcomes yielding the event E. In throwing at random a die whose

six faces are numbered from 1 to 6, it seems reasonable to assume that any face is as likely to come on top as any other, if the die is of uniform density. In this case, we take $p_1 = p_2 = \cdots = p_6 = 1/6$, where p_i is the probability that the face numbered i will come up. If we are merely interested in whether or not a 6 comes up, we consider only two possible outcomes, putting $p = 1/6$ as the probability for a 6 and a $p' = 5/6$ as the probability of not getting a 6.

Let us suppose we have N urns numbered from 1 to N. We are to place at random n balls in the urns, where $n < N$. We ask the probability that each of the urns numbered 1 to n will contain exactly one ball. This probability depends on two things: (1) whether the balls are distinguishable or indistinguishable, and (2) whether there is an exclusion principle that does not allow a second ball to be placed in an urn that already contains one ball. If the n balls are distinguishable and there is no exclusion principle, there will be N^n ways of placing the n balls in the N urns. There will be $n!$ ways of placing them in the urns numbered $1, \ldots, n$, placing one in each of these urns. With these conventions, the probability is

$$p(E) = \frac{n!}{N^n}. \tag{1.2.1}$$

If the balls are distinguishable and there is an exclusion principle, the first ball may be placed in any one of N urns, the next in any one of $(N - 1)$ urns, and the ith in any one of $(N - i + 1)$ urns, whence the number of ways of placing the n balls in the N urns is ${}_NP_n = N(N - 1)\cdots(N - n + 1)$. They may be placed in urns $1, \ldots, n$ in $n!$ ways, and under these conventions, the probability is

$$p(E) = \frac{n!}{{}_NP_n} = \frac{1}{\binom{N}{n}}. \tag{1.2.2}$$

If the balls are indistinguishable and there is no exclusion principle, we are asking for the solutions of $x_1 + x_2 + \cdots + x_N = n$ in nonnegative integers x_i, where x_i is the number of balls placed in the ith urn. This, as we noted in the preceding section, is the number of combinations of N things taken n at a time with repeats permitted, and is $\binom{N+n-1}{n}$. Exactly one of these is the solution $x_1 = x_2 = \cdots = x_n = 1$, $x_{n+1} = x_{n+2} = \cdots = x_N = 0$; in this case our probability is

$$p(E) = \frac{1}{\binom{N+n-1}{n}}. \tag{1.2.3}$$

From the physical standpoint, "indistinguishable" means that one combination is as likely as another.

If the balls are indistinguishable and there is an exclusion principle, the number of ways of placing the balls is merely the number of combinations of N things taken n at a time without repeats, and this is ${}_NC_n = \binom{N}{n}$. The choice of the first n urns is a single one of these combinations, and here the probability is

$$p(E) = \frac{1}{\binom{N}{n}}. \tag{1.2.4}$$

Note that this is the same as (1.2.2), so that with an exclusion principle, the probability is the same whether the balls are distinguishable or not.

In statistical physics we consider a collection of n particles, which may be protons, electrons, mesons, neutrons, neutrinos, or photons, each of which may be in any of N "states," which may be energy levels. The macroscopic state of the system of the n particles is a vector $x = (x_1, x_2, \ldots, x_N)$, where x_i is the number of particles in the ith state. The probability of any single macroscopic state depends on whether or not the particles are distinguishable and whether or not the particles obey the Pauli exclusion principle, which says that no two (indistinguishable) particles may be in the same state. If the particles are considered distinguishable and do not obey the exclusion principle, the probability of any single macroscopic state is given by (1.2.1) and the particles are said to obey the Maxwell–Boltzmann statistics. If the particles are indistinguishable and do not obey the exclusion principle, the probability is given by (1.2.3), and they are said to obey the Bose–Einstein statistics. If they are indistinguishable and do obey the exclusion principle, the probability is given by (1.2.4), and the particles are said to obey the Fermi–Dirac statistics. Electrons, protons, and neutrons obey Fermi–Dirac statistics. Photons and pi-mesons obey Bose–Einstein statistics. The case (1.2.2) of distinguishable particles with an exclusion principle does not arise in physics.

At high temperatures, when the number N is large and the different microscopic states are approximately equally likely, the Fermi–Dirac and Bose–Einstein statistics are essentially the same as the classical Maxwell–Boltzmann statistics. At low temperatures, the low-energy levels are more likely than the high-energy ones, and then the preceding models must be modified accordingly.

PROBLEMS

1. Prove

$$\sum_{i=0}^{m} \binom{r}{i}\binom{s}{m-i} = \binom{r+s}{m}.$$

Hint: $(1 + x)^r(1 + x)^s = (1 + x)^{r+s}$. Give an alternate proof of this identity

by considering the number of ways of choosing a committee of m people out of a group of r men and s women.

2. A flag is to be designed with 13 horizontal stripes colored red, white, or blue, subject to the condition that no stripe be of the same color as the one above it. In how many ways may this be done?

3. How many positive integers less than 10^n (in the decimal scale) have their digits in nondecreasing order?

4. In how many ways may n identical gifts be given to r children (a) under no restriction and (b) if each child must receive at least one gift?

5. A hand of five cards is selected from a deck of $4n$ cards that contains four different suits, each with n cards, $n \geqslant 5$, numbered $1, \ldots, n$. Rank in order of increasing frequency, depending on the value of n, the following hands: a straight flush (five consecutively numbered cards of the same suit), four of a kind (four cards having the same number), full house (three cards of one number, the other two of another number), flush (five cards of the same suit), straight (five consecutively numbered cards), three of a kind (three cards of the same number), two pair (two cards of one number, two others of a second number), and a pair (two cards of a number).

2

Inversion Formulae

2.1. THE PRINCIPLE OF INCLUSION AND EXCLUSION. MÖBIUS INVERSION

Suppose we have N objects and a number of properties $P(1), \ldots, P(n)$. Let N_i be the number of objects with property $P(i)$ and, more generally, $N_{i_1 i_2 \cdots i_r}$ the number of objects with properties $P(i_1)$, $P(i_2), \ldots,$ and $P(i_r)$. Then we assert that the number of objects $N(0)$ with none of the properties is given by the inversion formula

$$N(0) = N - \sum N_i + \sum_{i_1 < i_2} N_{i_1 i_2} + \cdots$$
$$+ (-1)^s \sum_{i_1 < i_2 \cdots < i_s} N_{i_1 i_2 \cdots i_s} + \cdots + (-1)^n N_{12 \cdots n}. \tag{2.1.1}$$

We now prove this. An object with none of the properties is counted once in the term N and does not contribute to the remaining terms. An object A with the property $P(j)$ is counted once in N and once in N_j, and so contributes 1 to the term N, -1 to the term $-\sum_i N_i$, and thus contributes $1 - 1 = 0$ to the right-hand side of (2.1.1). An object A with exactly r properties, say, $j_1, \ldots, j_r$, contributes 1 to the sum

$$\sum_{i_1 < \cdots < i_s} N_{i_1 i_2 \cdots i_s}, \qquad \text{when } s \leqslant r$$

for every choice of $i_1, \ldots, i_s$ from $j_1, \ldots, j_r$—that is, for $\binom{r}{s}$ choices. Hence, A contributes to the right-hand side of (2.1.1) exactly

$$1 - \binom{r}{1} + \binom{r}{2} + \cdots + (-1)^s \binom{r}{s} + \cdots + (-1)^r \binom{r}{r} = (1 - 1)^r = 0. \tag{2.1.2}$$

Thus, the right-hand side of (2.1.1) counts each element with no properties exactly once, and every other element zero times; hence its value is $N(0)$, as was to be proved. Use of the formula (2.1.1) is sometimes called the method of inclusion and exclusion.

In the same way we may find the number $N(r)$ of objects with exactly r properties. This is given by

$$N(r) = \sum_{i_1 < \cdots < i_r} N_{i_1 \cdots i_r} + \cdots + (-1)^{s-r} \binom{s}{r} \sum_{i_1 < \cdots < i_s} N_{i_1 \cdots i_s} + \cdots \quad (2.1.3)$$

On the right-hand side of (2.1.3) an object with exactly r properties is counted once in the first term and is not counted in the other terms. An object with exactly t properties, where $t > r$, contributes $(-1)^{s-r}\binom{s}{r}\binom{t}{s}$ to the term

$$(-1)^{s-r} \binom{s}{r} \sum_{i_1 < \cdots < i_s} N_{i_1 \cdots i_s}.$$

But

$$\sum_{s=r}^{t} (-1)^{s-r} \binom{s}{r}\binom{t}{s} = 0 \quad (2.1.4)$$

from the relation (1.1.9d), and so the relation (2.1.3) is proved.

As an application of the method of inclusion and exclusion we consider the problem of derangements. How many permutations $a_1, a_2, \ldots, a_n$ of $1, 2, \ldots, n$ are there

$$\begin{matrix} 1, 2, \ldots, & i, \ldots, n \\ a_1, a_2, \ldots, & a_i, \ldots, a_n \end{matrix} \quad (2.1.5)$$

such that we have $a_i \neq i$ for every $i = 1, 2, \ldots, n$? Here we take the N objects as the $n!$ permutations $a_1, a_2, \ldots, a_n$ and the property $P(i)$ as $a_i = i, i = 1, \ldots, n$. Then $N_{i_1 i_2 \cdots i_r} = (n - r)!$, this being the number of permutations fixing r specified numbers. Furthermore, for $\sum N_{i_1 i_2 \cdots i_r}$ there are $\binom{n}{r}$ summands, this being the number of ways of choosing $i_1, i_2, \ldots, i_r$ from $1, 2, \ldots, n$. Applying (2.1.1) we have

$$N(0) = n! - n \cdot (n-1)! + \binom{n}{2}(n-2)! + \cdots$$

$$+ (-1)^r \binom{n}{r} (n-r)! + \cdots + (-1)^n \cdot 1. \quad (2.1.6)$$

We may rewrite this in the form

$$N(0) = n! \left(1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^r \cdot \frac{1}{r!} + \cdots + \frac{(-1)^n}{n!}\right). \quad (2.1.7)$$

We recognize

$$1 - 1 + \frac{1}{2!} - \frac{1}{3!} \cdots$$

as the initial terms of an infinite series whose value is e^{-1}. The infinite series is alternating, and the first omitted term is $(-1)^{n+1}/(n+1)!$. From this we see that $N(0)$ differs from $n!/e$ by less than $1/(n+1)$, and so $n!/e$ is an extremely good approximation to the number of derangements of n letters.

If we ask not only the number of derangements of $1, 2, \ldots, n$ but also the number of permutations $a_1 a_2 \cdots a_n$ of $1, 2, \ldots, n$ for which $a_i = i$ in exactly r instances for each value of $r = 0, 1, \ldots, n$, the problem is known as the *problème des rencontres*. The solution is an easy extension of the problem of derangements. We may choose r numbers from $1, \ldots, n$ in $\binom{n}{r}$ ways, and having chosen these we multiply by the number of derangements of the remaining $(n - r)$ letters. This gives the number of permutations with exactly r agreements $a_i = i$ as

$$N(r) = \frac{n!}{r!}\left(1 - 1 + \frac{1}{2!} + \cdots + (-1)^{n-r} \cdot \frac{1}{(n-r)!}\right). \tag{2.1.8}$$

This could also have been found from the rule (2.1.3).

For an additional type of inversion formula we turn to an arithmetical function, the Möbius function $\mu(n)$. This is defined for positive integers n. If $n > 1$, then n has a unique factorization as a product of prime powers

$$n = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r}, \tag{2.1.9}$$

where the p's are different primes. We define $\mu(n)$ by the rules

$$\begin{aligned} &\mu(1) = 1, \\ &\mu(n) = 0, \qquad \text{if any } e_i > 1 \text{ in (2.1.9)}, \\ &\mu(n) = (-1)^r, \qquad \text{if } e_1 = e_2 = \cdots = e_r = 1 \text{ in (2.1.9)}. \end{aligned} \tag{2.1.10}$$

Lemma 2.1.1.

$$\sum_{d|n} \mu(d) = \begin{cases} 1, & \text{if } n = 1 \\ 0, & \text{if } n > 1, \end{cases}$$

the sum being over all positive divisors d of n.

Proof. If $n = 1$, then $d = 1$ is the only divisor, and $\mu(1) = 1$. If $n > 1$ and n is given by (2.1.9), write $n^* = p_1 p_2 \cdots p_r$. Then a divisor d of n that is not a divisor of n^* will have a multiple prime factor and we have $\mu(d) = 0$. Hence,

$$\sum_{d|n} \mu(d) = \sum_{d|n^*} \mu(d). \tag{2.1.11}$$

But $\sum_{d|n^*} \mu(d)$ is easily evaluated as

$$1 - r + \binom{r}{2} + \cdots + (-1)^k \binom{r}{k} + \cdots = (1-1)^r = 0, \tag{2.1.12}$$

since there are $\binom{r}{k}$ divisors that are the product of k distinct primes and for each of which $\mu(d) = (-1)^k$. Thus, our lemma is proved.

Theorem 2.1.1 (Möbius inversion formula). *Let $f(n)$ and $g(n)$ be functions defined for every positive integer n satisfying*

$$f(n) = \sum_{d|n} g(d). \tag{2.1.13a}$$

Then we may invert this relation to express g in terms of f by the rule

$$g(n) = \sum_{d|n} \mu(d) f\left(\frac{n}{d}\right). \tag{2.1.13b}$$

The second relation also implies the first.

Proof. We have

$$f\left(\frac{n}{d}\right) = \sum_{d'|n/d} g(d') \qquad \text{for every } d|n.$$

Hence,

$$\sum_{d|n} \mu(d) f\left(\frac{n}{d}\right) = \sum_{d|n} \mu(d) \cdot \sum_{d'|n/d} g(d'). \tag{2.1.14}$$

Let us write $n = dd'n_1$. Then, for a fixed d', d ranges over the divisors of n/d'. Hence,

$$\sum_{d|n} \mu(d) \cdot \sum_{d'|(n/d)} g(d') = \sum_{d'|n} g(d') \sum_{d|(n/d')} \mu(d) = g(n), \tag{2.1.15}$$

since the sum

$$\sum_{d|(n/d')} \mu(d) = 0$$

by the lemma, except for $d' = n$. Thus, the right-hand side of (2.1.14) simplifies

to $g(n)$ and our theorem is proved. Similarly, given (2.1.13b), we may substitute in the right-hand side of (2.1.13a) and find that it simplifies to $f(n)$, proving (2.1.13a).

Möbius inversion may be used to enumerate circular partitions. If letters $a_1, a_2, \ldots, a_n$ are arranged in a circle with a_1 following a_n, then any one of the linear sequences $a_2, a_3, \ldots, a_n, a_1$; $a_3, \ldots, a_n, a_1, a_2$; $\ldots$; $a_n, a_1, \ldots, a_{n-1}$ may be thought of as determining the same circular sequence. But not all n linear sequences corresponding to the same circular sequence need be different. If, for a divisor d of n, the sequence $a_1, a_2, \ldots, a_n$ consists of a sequence of d letters $a_1, a_2, \ldots, a_d$ repeated n/d times, the linear sequences repeat after the first d. With each circular sequence of length n we may associate a unique minimum period d such that the circular sequence consists of n/d repetitions of a sequence of d letters. Furthermore, each circular sequence of length d and period d where $d|n$ may be repeated n/d times to give a circular sequence of length n and period d. Each of these sequences corresponds to exactly d different linear words of length n. If there are r different letters, there are r^n linear permutations $a_1 a_2 \cdots a_n$. If $M(d)$ is the number of circular sequences of length and period d, then $dM(d)$ is the number of linear sequences of length n corresponding to them. This gives us the equation

$$\sum_{d|n} dM(d) = r^n. \tag{2.1.16}$$

If we take $f(x) = r^x$ and $g(x) = xM(x)$, we may apply Möbius inversion to (2.1.16) and obtain

$$nM(n) = \sum_{d|n} \mu(d) r^{n/d}, \tag{2.1.17}$$

whence

$$M(n) = \frac{1}{n} \sum_{d|n} \mu(d) r^{n/d}. \tag{2.1.18}$$

This gives the number of circular permutations of length and period n. If we wish the total number of circular permutations of length n, this number is $T(n)$:

$$T(n) = \sum_{d|n} M(d). \tag{2.1.19}$$

If we wish the total number of circular permutations of n objects in which the number of objects of each kind is specified (say, b_i of the ith kind, $i = 1, \ldots, r$, where $b_1 + b_2 + \cdots + b_r = n$), we recall that the number of linear permutations is the multinomial coefficient

$$\frac{n!}{b_1! \cdots b_r!}, \qquad b_1 + b_2 + \cdots + b_r = n. \tag{2.1.20}$$

Here a circular permutation of this kind of length n and period n/d will have d a divisor of all $b_1, b_2, \ldots, b_r$, or, what amounts to the same thing, d a divisor of $(b_1, \ldots, b_r)$, the greatest common divisor of $b_1, \ldots, b_r$. Thus, if $M(b_1, \ldots, b_r)$, $b_1 + b_2 + \cdots + b_r = n$, is the number of circular permutations of length and period n with b_i objects of type i, $i = 1, \ldots, r$, the same argument given above yields

$$M(b_1, \ldots, b_r) = \frac{1}{n} \sum_{d \mid (b_1, \ldots, b_r)} \mu(d) \frac{(n/d)!}{(b_1/d)! \cdots (b_r/d)!}. \tag{2.1.21}$$

The *problème des ménages* is as follows: A hostess wishes to place n couples at a circular table so that men and women are in alternate places, but so that no husband will sit on either side of his wife. In how many ways may this be done? It is easy to see that it cannot be done with fewer than three couples, but for three or more couples it may be done.

Let us first place the women at alternate places, designating them by numbers $1, 2, \ldots, n$ in circular order. Let the place to the left of the ith woman and on the right of the $(i + 1)$th be numbered i, giving number n to the place between the nth woman and the first. Then the first husband can sit anywhere except in the nth or first place, and the ith husband anywhere except the $(i - 1)$th or the ith. If husband number a_i sits in place i, then $a_1 a_2 \cdots a_n$ is a permutation of $1, 2, \ldots, n$ and in the array

$$\begin{array}{ccccc} 1 & 2 & \cdots & n-1 & n \\ n & 1 & \cdots & n-2 & n-1 \\ a_1 & a_2 & \cdots & a_{n-1} & a_n \end{array} \tag{2.1.22}$$

we see our condition is precisely that the permutation $a_1 a_2 \cdots a_n$ must be discordant with the first two rows. We are thus in a problem of inclusion and exclusion with properties $P(1)$: $a_1 = 1, \mu, \ldots$; $P(i)$: $a_i = i - 1, \mu, \ldots$; $P(n)$: $a_n = n - 1, \mu$. If $P(i)$ is true for r values $a_{i_1}, \ldots . a_{i_r}$, there will be $(n - r)!$ ways of completing the permutation. Thus, we must first calculate the number of ways of having $P(i)$ true for r values, or, as we shall say, the number of ways of having r hits. The number of ways of having one hit is $2n$, and by the circular symmetry of (2.1.22) we may suppose this to be either the n in the first column or the n in the nth column. We now list the remaining numbers, writing the columns one after another, giving either

$$1, 1, 2, 2, 3, \ldots, n-2, n-2, n-1,$$

or (2.1.23)

$$1, 2, 2, 3, 3, \ldots, n-2, n-1, n-1.$$

Our remaining choices of $(r - 1)$ numbers are restricted by saying that in the

arrays in (2.1.23) we may not choose two consecutive values because consecutive choices amount either to taking the same number twice or to taking both elements in a column of (2.1.22). This is the number of ways of choosing $(r-1)$ objects, no two consecutive, from a row of $2n-3$. This has been evaluated in Chapter 1 and is $\binom{2n-r-1}{r-1}$. This number is to be multiplied by $2n$ for the first choice, but the same set of r values could be obtained by regarding any one of the r values as the first, and so we must divide by r. Thus, our number is

$$\frac{2n}{r}\binom{2n-r-1}{r-1} = \frac{2n}{2n-r}\binom{2n-r}{r}. \tag{2.1.24}$$

We may now apply formula (2.1.1) and find as our answer that the number of solutions U_n of permutations discordant both with $1, 2, \ldots, n$ and $2, 3, \ldots, n, 1$ is given by

$$U_n = n! - 2n\cdot(n-1)! + \cdots + (-1)^r \frac{2n}{2n-r}\binom{2n-r}{r}(n-r)! + \cdots + (-1)^n \cdot 2. \tag{2.1.25}$$

From this relation we may derive the recursion

$$(n-2)U_n = n(n-2)U_{n-1} + nU_{n-2} + 4(-1)^{n+1}, \qquad n \geqslant 4. \tag{2.1.26}$$

This recursion can be proved without much difficulty. For $r = 0$ and 1, the terms in $(n-2)U_n$ and $n(n-2)U_{n-1}$ are equal. For $r = 2, \ldots, n-1$, we have the identity involving the rth term of U_n and U_{n-1} and the $(r-2)$th term of U_{n-2}:

$$\begin{aligned}(n-2)\frac{2n}{2n-r}\binom{2n-r}{r}(n-r)! &= n(n-2)\frac{2(n-1)}{2n-r-2}\binom{2n-r-2}{r}(n-r-1)! \\ &\quad + n\cdot\frac{2(n-2)}{2n-r-2}\binom{2n-r-2}{r-2}(n-r)!\end{aligned} \tag{2.1.27}$$

Finally, the term with $r = n$ in $(n-2)U_n$ and that with $r = n-2$ in nU_{n-2} combine, so that

$$(n-2)(-1)^n\cdot 2 = n(-1)^{n-2}\cdot 2 + 4(-1)^{n+1}. \tag{2.1.28}$$

This proves the validity of the recursion (2.1.26).

2.2. PARTIALLY ORDERED SETS AND THEIR MÖBIUS FUNCTIONS

A *partially ordered set* P is a system $P\{\ldots, x, y, \ldots\}$ of elements with an ordering relation $x \geqslant y$ (read "x includes y") that holds for certain pairs of elements and an equality $x = y$, such that the following axioms hold:

PO 1. $x \geqslant x$ *for every* x *of* P.
PO 2. *If* $x \geqslant y$ *and* $y \geqslant z$, *then* $x \geqslant z$.
PO 3. *If* $x \geqslant y$ *and* $y \geqslant x$, *then* $x = y$.

A *simply ordered set* or *chain* also satisfies:

PO 4. *If* x, y *are elements of* P, *then either* $x \geqslant y$ *or* $y \geqslant x$.

We write $y \leqslant x$ as an alternate form of $x \geqslant y$, and $x > y$ (or $y < x$) if $x \geqslant y$ (or $y \leqslant x$) and $x \neq y$.

Partial ordering is a very general concept. Two particular cases of interest are:

1. The elements of P are all the subsets of a finite set T, where we write 0 for the void subset and 1 for the set T itself, and $y \leqslant x$ means that y is a subset of x.
2. The elements of P are the positive integers and $y \leqslant x$ means that y divides x.

It is easy to check the validity of the axioms in both instances.

If T is a subset of a partially ordered set P, then an element x of P such that $x \leqslant t$ for every t of T is called a *lower bound* of T. If z is a lower bound of T such that $x \leqslant z$ for every lower bound x of T, then z is called a *greatest lower bound* of T. From PO 3 it follows that if T has a greatest lower bound, it is necessarily unique. Similarly, if $x \geqslant t$ for every t of T, x is called an *upper bound* of T, and if z is an upper bound of T such that $x \geqslant z$ for every upper bound x, then z is called a *least upper bound* of T and again is clearly unique if it exists. If P itself has a greatest lower bound, this is called its *zero element*, and if it has a least upper bound this is called the *all element* (or sometimes the unit element). An *interval* $[x, y]$ where $x \leqslant y$ is the set of elements w such that $x \leqslant w \leqslant y$. If x and y are the only elements in the interval $[x, y]$, we say that y *covers* x. A partially ordered set P is said to be *locally finite* if the number of elements in every interval $[x, y]$ is finite.

The Möbius function and Möbius inversion were defined for functions over locally finite partially ordered sets originally by L. Weisner [1] and P. Hall [2]. This idea was greatly expanded by G. C. Rota [1]. A brief treatment is given here, based on Rota's work.

We consider a class of real-valued functions $f(x, y)$ defined for $x, y \in P$, a locally finite partially ordered set. We require that $f(x, y) = 0$ if $x \not\leqslant y$. The sum of two such functions, as well as multiplication by scalars, is defined as

usual. The product $h = fg$ is defined as follows:

$$h(x, y) = \sum_{x \leqslant z \leqslant y} f(x, z)g(z, y), \qquad x, y \text{ fixed.} \tag{2.2.1}$$

This product is well defined, since the sum on the right is finite, P being locally finite. Under the operations of sum, scalar product, and the product rule $h = fg$ of (2.2.1), the functions $f(x, y)$ define the *incidence algebra* $A(P)$ of P. It is easy to verify that the multiplication defined for $A(P)$ is associative and distributive and that $A(P)$ has an identity, the Kronecker delta function $\delta(x, x) = 1, \delta(x, y) = 0$ if $x \neq y$.

Lemma 2.2.1. *A function $f(x, y)$ of $A(P)$ has both a left and a right inverse if and only if $f(x, x) \neq 0$ for every x of P.*

Proof. In (2.2.1) take $h(x, y) = \delta(x, y)$. Now, given f, we wish to solve for g. Since this requires $1 = \delta(x, x) = f(x, x)$ for every x, the condition $f(x, x) \neq 0$ for every x of P is clearly necessary. Thus, suppose $f(x, x) \neq 0$ for every x. Then $g(x, x) = f(x, x)^{-1}$ for every x. To evaluate $g(x, y)$ with $x < y$, we may assume inductively that we have already found $g(z, y)$ for every z satisfying $x < z \leqslant y$. Then

$$h(x, y) = \delta(x, y) = 0 = \sum_{x \leqslant z \leqslant y} f(x, z)g(z, y),$$

whence

$$-f(x, x)g(x, y) = \sum_{x < z \leqslant y} f(x, z)g(z, y), \tag{2.2.2}$$

and we may find $g(x, y)$, since $f(x, x) \neq 0$ and all terms of the finite sum on the right are known. Thus, f has a right inverse. Similarly applying our induction to terms $x \leqslant z < y$, we may use (2.2.1), interchanging the roles of f and g to show that f has a left inverse. But if $fg_1 = 1 = \delta(x, y)$ and $g_2 f = 1$, then by a familiar argument, $g_2 = g_2 1 = g_2(fg_1) = (g_2 f)g_1 = 1g_1 = g_1$ and the left and right inverses are the same.

Definition. *Let P be a locally finite partially ordered set and $A(P)$ its incidence algebra. The zeta function $\zeta(x, y)$ of $A(P)$ is that function for which $\zeta(x, y) = 1$ for $x \leqslant y$, $\zeta(x, y) = 0$ otherwise. The Möbius function $\mu(x, y)$ of $A(P)$ is the inverse of the zeta function.*

Since $\zeta(x, x) = 1 \neq 0$ for every x, by our lemma $\zeta(x, y)$ has an inverse function $\mu(x, y)$, which is both a right and left inverse of it. Hence, we have

$$\mu(x, x) = 1, \qquad \text{for every } x \text{ of } P. \tag{2.2.3}$$

For $x < y$, we have

$$\mu(x, y) = -\sum_{x \leqslant z < y} \mu(x, z), \qquad x < y \text{ fixed.} \tag{2.2.4}$$

$$\mu(x, y) = -\sum_{x < z \leqslant y} \mu(z, y), \qquad x < y \text{ fixed.} \tag{2.2.5}$$

Here (2.2.4) expressed the Möbius function as the left inverse of zeta, and (2.2.5) expresses it as the right inverse.

The Möbius inversion theorem follows.

Theorem 2.2.1. *Let P be a locally finite partially ordered set with a zero element 0. Let $f(x)$ be given for all x of P and let $g(x)$ be determined from $f(x)$ by the rule*

$$g(x) = \sum_{y \leqslant x} f(y), \qquad \text{all } x \text{ of } P. \tag{2.2.6}$$

Then, if $\mu(y, z)$ is the Möbius function of P, we have

$$f(x) = \sum_{y \leqslant x} g(y)\mu(y, x), \qquad \text{all } x \text{ of } P. \tag{2.2.7}$$

Proof. Since every interval $[0, x]$ is finite, the sums in (2.2.6) and (2.2.7) are well defined. For a fixed x, consider the sum

$$S = \sum_{y \leqslant x} g(y)\mu(y, x) = \sum_{y \leqslant x} \left(\sum_{z \leqslant y} f(z) \right) \mu(y, x), \tag{2.2.8}$$

where we have substituted from (2.2.6). Now interchange the order of summation to get

$$\begin{aligned} S &= \sum_{z \leqslant x} f(z) \sum_{y \leqslant x} \mu(y, z) = \sum_{z} f(z)\zeta(z, y) \sum_{y \leqslant x} \mu(y, x) \\ &= \sum_{z \leqslant x} f(z) \sum_{z \leqslant y \leqslant x} \zeta(z, y)\mu(y, x) = \sum_{z \leqslant x} f(z)\delta(z, x) = f(x). \end{aligned} \tag{2.2.9}$$

Since $S = f(x)$, we have proved (2.2.7), the conclusion of our theorem.

Let us now determine the Möbius function for the two special cases mentioned at the beginning of this section.

In case 1, P is the partially ordered set of all subsets of a finite set T, ordered by inclusion. Here we assert that for $x \leqslant y$,

$$\mu(x, y) = (-1)^{n(y) - n(x)}, \tag{2.2.10}$$

where $n(x)$, $n(y)$ are respectively the number of elements of T in x and in y.

The assertion is certainly true when $n(y) - n(x) = 0$ or 1. By induction assume (2.2.10) to be true for $n(y) - n(x) \leqslant r - 1$ and consider a case with $n(y) - n(x) = r$. Then (2.2.4) becomes

$$\mu(x, y) = -1 + \binom{r}{1} - \binom{r}{2} + \cdots - \binom{r}{j}(-1)^j + \cdots - \binom{r}{r-1}(-1)^{r-1}, \quad (2.2.11)$$

since there are $\binom{r}{j}$ z's with $x \leqslant z < y$ with $n(z) - n(x) = j$, namely, the subsets of T obtained by adjoining to xj of the r elements of y not in x. Comparison of (2.2.11) with the binomial expansion of $(1 - 1)^r = 0$ gives $\mu(x, y) = (-1)^r$, as was to be shown.

Let T be the integers $1, 2, \ldots, n$ and let properties $P(1), P(2), \ldots, P(n)$ be associated with these integers. Let K be a set of N elements, each of which has the properties $P(i)$, $i \in x$ for some subset x of T. Let $f(x)$ be the number of elements of K having exactly the properties $P(i)$, $i \notin x$, x a subset of T. Then, if we put

$$g(x) = \sum_{y \leqslant x} f(y), \quad (2.2.12)$$

the function $g(x)$ is the number of elements of K having all the properties $P(i)$ for $i \notin x$ and possibly others. Here, for $x = T$, the inversion (2.2.7) gives us

$$f(T) = g(T) - \sum_{n(y)=n-1} g(y) + \cdots + (-1)^j \sum_{n(y)=n-j} g(y) + \cdots + (-1)^n \sum_{n(y)=0} g(y). \quad (2.2.13)$$

But here $f(T) = N(0)$ is the number of elements having none of the properties $g(T) = N$, since this counts all elements having properties of the void set and possibly others. If $n(y) = n - j$, then $g(y)$ counts all elements having the j properties not in y and possibly others. But this shows that (2.2.13) is the principle of inclusion and exclusion of (2.1.1).

In case 2, P is the partially ordered set of the positive integers, where $x \leqslant y$ means that x divides y.

Here in a segment $[x, y]$, if $x \leqslant z \leqslant y$, then $z = xd$, where $d|(y/x)$, and so this segment corresponds to the divisors of y/x. Note that the integer 1 is the zero element of P. Comparison of (2.2.4) with the lemma preceding Theorem 2.1.1 shows that $u(x, y) = \mu(y/x)$ in this case. Thus, Theorem 2.1.1 on Möbius inversion is the special case of Theorem 2.2.1 for P, the positive integers partially ordered by division.

PROBLEMS

1. Let A be the $n \times n$ matrix with zeros down the main diagonal and 1's

elsewhere. The determinant of A is $(-1)^{n-1}(n-1)$. (This will be shown in Section 10.2.) Of the $n!$ terms in the expansion of the determinant of A, how many are $+1$, -1, 0, respectively?

2. Given the array

$$\begin{array}{cccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 2 & 3 & 4 & 0 & 6 & 7 & 8 & 9 & 5 \\ a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9, \end{array}$$

in how many ways can we choose $a_0, \ldots, a_9$ as a permutation of $0, \ldots, 9$ so that no column of the array will have a repeated number?

3. Show that U_n in (2.1.25) is approximately $n!/e^2$.

4. The function $\Lambda(n)$ is defined for positive integers n by the rule

$$\sum_{d|n} \Lambda(d) = \log n.$$

Prove that $\Lambda(n) = \log p$ if $n = p^e$, p a prime, and $\Lambda(n) = 0$ otherwise.

5. Find the Möbius functions of the two partially ordered sets P_1, P_2 with five elements 0, a, b, c, 1, where: (a) in P_1, $0 \leqslant a \leqslant 1$, $0 \leqslant b \leqslant 1$, $0 \leqslant c \leqslant 1$, and there are no further inclusions; (b) in P_2, $0 \leqslant a \leqslant 1$, $0 \leqslant b \leqslant c \leqslant 1$, and there are no further inclusions.

3

Generating Functions and Recursions

3.1. RULES AND PROPERTIES

If $u_0, u_1, u_2, \ldots, u_n, \ldots$ is a sequence of numbers, we may associate with this sequence a *generating function* $g(x)$ by the rule

$$g(x) = u_0 + u_1 x + u_2 x^2 + \cdots + u_n x^n + \cdots. \tag{3.1.1}$$

If this series has a circle of convergence with a radius $R > 0$, then it may happen that the properties of the function $g(x)$ enable us to evaluate the coefficients u_n (or at least give estimates of their order of magnitude) or perhaps find other information of value. If $h(x)$ is the generating function of the sequence $v_0, v_1, v_2, \ldots, v_n, \ldots$, then

$$h(x) = v_0 + v_1 x + v_2 x^2 + \cdots + v_n x^n + \cdots. \tag{3.1.2}$$

If we add (3.1.1) multiplied by c, and (3.1.2) multiplied by d, we have

$$cg(x) + dh(x) = (cu_0 + dv_0) + (cu_1 + dv_1)x + \cdots + (cu_n + dv_n)x^n + \cdots, \tag{3.1.3}$$

and if we multiply, we have

$$g(x)h(x) = w_0 + w_1 x + w_2 x^2 + \cdots + w_n x^n + \cdots, \tag{3.1.4}$$

where for every $n = 1, 2, 3, \ldots,$

$$w_n = u_0 v_n + u_1 v_{n-1} + \cdots + u_{n-1} v_1 + u_n v_0. \tag{3.1.5}$$

Even if the series for $g(x)$ and $h(x)$ are not convergent, we may regard (3.1.3), (3.1.4), and (3.1.5) as defining formal operations on formal series. In these terms we easily verify that the addition, multiplication by scalars, and series multiplication satisfy the associative, commutative, and distributive laws. Furthermore, if $u_0 \neq 0$ and if we take $v_0 = u_0^{-1}$, we may use (3.1.5) to determine