Chapter 8

Special Counting Sequences

We have considered several special counting sequences in the previous chapters. The counting sequence for permutations of a set of n elements is

$$0!, 1!, 2!, \ldots, n!, \ldots.$$

The counting sequence for derangements of a set of n elements is

$$D_0, D_1, D_2, \ldots, D_n, \ldots,$$

where D_n has been evaluated in Theorem 6.3.1. In addition, we have investigated the Fibonacci sequence

$$f_0, f_1, f_2, \ldots, f_n, \ldots,$$

and a formula for f_n has been given in Theorem 7.1.1. In this chapter, we study primarily six famous and important counting sequences: the sequence of Catalan numbers, the sequences of the Stirling numbers of the first and second kind, the sequence of the number of partitions of a positive integer n, and the sequences of the small and large Schröder numbers.

8.1 Catalan Numbers

The *Catalan sequence*[1] is the sequence

$$C_0, C_1, C_2, \ldots, C_n, \ldots,$$

where

$$C_n = \frac{1}{n+1}\binom{2n}{n}, \qquad (n = 0, 1, 2, \ldots)$$

[1]After Eugène Catalan (1814–1894).

is the nth *Catalan number.* The first few Catalan numbers are evaluated to be

$$\begin{array}{ll} C_0 = 1 & C_5 = 42 \\ C_1 = 1 & C_6 = 132 \\ C_2 = 2 & C_7 = 429 \\ C_3 = 5 & C_8 = 1430 \\ C_4 = 14 & C_9 = 4862. \end{array}$$

The Catalan number

$$C_{n-1} = \frac{1}{n}\binom{2n-2}{n-1}$$

arose in Section 7.6 as the number of ways to divide a convex polygonal region with $n+1$ sides into triangles by inserting diagonals that do not intersect in the interior. The Catalan numbers occur in several seemingly unrelated counting problems, and we discuss some of them in this section.[2]

Theorem 8.1.1 *The number of sequences*

$$a_1, a_2, \ldots, a_{2n} \tag{8.1}$$

of $2n$ *terms that can be formed by using exactly* n $+1$*s and exactly* n -1*s whose partial sums are always positive:*

$$a_1 + a_2 + \cdots + a_k \geq 0, \qquad (k = 1, 2, \ldots, 2n) \tag{8.2}$$

equals the n*th Catalan number*

$$C_n = \frac{1}{n+1}\binom{2n}{n}, \qquad (n \geq 0).$$

Proof. We call a sequence (8.1) of n $+1$s and n -1s *acceptable* if it satisfies (8.2) and *unacceptable* otherwise. Let A_n denote the number of acceptable sequences of n $+1$s and n -1s, and let U_n denote the number of unacceptable sequences. The total number of sequences of n $+1$'s and n -1's is

$$\binom{2n}{n} = \frac{(2n)!}{n!n!},$$

since such sequences can be regarded as the permutations of objects of two different types with n objects of one type (the $+1$s) and n of the other (the -1s). Hence,

$$A_n + U_n = \binom{2n}{n},$$

[2]For a list of 66 combinatorially defined sets that are counted by the Catalan numbers, see R. P. Stanley, *Enumerative Combinatorics Volume 2*, Cambridge University Press, Cambridge, 1999 (Exercise 6.19, pp. 219–229 and Solution, pp. 256–265). There the term *Catalania* or *Catalan mania* is introduced.

and we evaluate A_n by first evaluating U_n and then subtracting from $\binom{2n}{n}$.

Consider an unacceptable sequence (8.1) of n +1s and n −1s. Because the sequence is unacceptable, there is a *first* k such that the partial sum

$$a_1 + a_2 + \cdots + a_k$$

is negative. Because k is first, there are equal numbers of +1s and −1s preceding a_k. Hence we have

$$a_1 + a_2 + \cdots + a_{k-1} = 0$$

and

$$a_k = -1.$$

In particular, k is an odd integer. We now reverse the signs of each of the first k terms; that is, we replace a_i by $-a_i$ for each $i = 1, 2, \ldots, k$ and leave unchanged the remaining terms. The resulting sequence

$$a'_1, a'_2, \ldots, a'_{2n}$$

is a sequence of $(n + 1)$ +1s and $(n - 1)$ −1s. This process is reversible: Given a sequence of $(n + 1)$ +1s and $(n - 1)$ −1s, there is a first instance when the number of +1s exceeds the number of −1s (since there are more +1's than −1s). Reversing the signs of the +1s and −1s up to that point results in an unacceptable sequence of n +1s and n −1s. Thus, there are as many unacceptable sequences as there are sequences of $(n+1)$ +1s and $(n-1)$ −1s. The number of sequences of $(n+1)$ +1s and $(n+1)$ −1s is the number

$$\frac{(2n)!}{(n+1)!(n-1)!}$$

of permutations of objects of two types, with $n + 1$ objects of one type and $n - 1$ of the other. Hence,

$$U_n = \frac{(2n)!}{(n+1)!(n-1)!},$$

and, therefore,

$$\begin{aligned} A_n &= \frac{(2n)!}{n!n!} - \frac{(2n)!}{(n+1)!(n-1)!} \\ &= \frac{(2n)!}{n!(n-1)!}\left(\frac{1}{n} - \frac{1}{n+1}\right) \\ &= \frac{(2n)!}{n!(n-1)!}\left(\frac{1}{n(n+1)}\right) \\ &= \frac{1}{n+1}\binom{2n}{n}. \end{aligned}$$

□

There are many different interpretations of Theorem 8.1.1. We discuss two of them in the next examples. The first is a classical problem.

Example. There are $2n$ people in line to get into a theater. Admission is 50 cents.[3] Of the $2n$ people, n have a 50-cent piece and n have a \$1 dollar bill.[4] The box office at the theater rather foolishly begins with an empty cash register. In how many ways can the people line up so that whenever a person with a \$1 dollar bill buys a ticket, the box office has a 50-cent piece in order to make change? (After everyone is admitted, there will be n \$1 dollar bills in the cash register.)

First, suppose that the people are regarded as "indistinguishable"; that is, we simply have a sequence of n 50-cent pieces and n dollar bills, and it doesn't matter who holds which and where they are in the line. If we identify a 50-cent piece with a $+1$ and a dollar bill with a -1, then the answer is the number

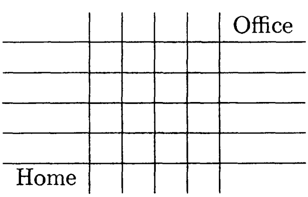
$$C_n = \frac{1}{n+1}\binom{2n}{n}$$

of acceptable sequences as defined in Theorem 8.1.1. Now suppose that the people are regarded as "distinguishable;" that is, we take into account who is who in the line. So we have n people holding 50-cent pieces and n holding dollar bills. The answer is now

$$(n)!(n!)\frac{1}{n+1}\binom{2n}{n} = \frac{(2n)!}{n+1}$$

since, with each sequence of n 50-cent pieces and n dollar bills, there are $n!$ orders for the people with 50-cent pieces and $n!$ orders for the people with dollar bills. □

Example. A big city lawyer works n blocks north and n blocks east of her place of residence. Every day she walks $2n$ blocks to work. (See the map below for $n = 4$.) How many routes are possible if she never crosses (but may touch) the diagonal line from home to office?



Each acceptable route either stays above the diagonal or stays below the diagonal. We find the number of acceptable routes above the diagonal and multiply by 2. Each

[3] This problem shows its age!

[4] A closer approximation to the current reality would be to have the theater charge \$5, and have n people with \$5 dollar bills and n with \$10 bills.

route is a sequence of n northerly blocks and n easterly blocks. We identify north with $+1$ and east with -1. Thus, each route corresponds to a sequence

$$a_1, a_2, \ldots, a_{2n}$$

of n $+1$s and n -1s, and in order to keep the route from dipping below the diagonal, we must have

$$\sum_{i=1}^{k} a_i \geq 0, \qquad (k = 1, \ldots, 2n).$$

Hence, by Theorem 8.1.1, the number of acceptable routes above the diagonal equals the nth Catalan number, and the total number of acceptable routes is

$$2C_n = \frac{2}{n+1}\binom{2n}{n}.$$

□

We next show that the Catalan numbers satisfy a particular homogeneous recurrence relation of order 1 (but with a nonconstant coefficient).[5] We have

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{1}{n+1}\frac{(2n)!}{n!n!}$$

and

$$C_{n-1} = \frac{1}{n}\binom{2n-2}{n-1} = \frac{1}{n}\frac{(2n-2)!}{(n-1)!(n-1)!}.$$

Dividing, we obtain

$$\frac{C_n}{C_{n-1}} = \frac{4n-2}{n+1}.$$

Therefore, the Catalan sequence is determined by the following recurrence relation and initial condition:

$$\begin{aligned} C_n &= \frac{4n-2}{n+1}C_{n-1}, \quad (n \geq 1) \\ C_0 &= 1. \end{aligned} \tag{8.3}$$

Previously we noted that $C_9 = 4862$. It follows from the recurrence relation (8.3) that

$$C_{10} = \frac{38}{11}\,C_9 = \frac{38}{11}(4862) = 16,796.$$

We now define a new sequence of numbers

$$C_1^*, C_2^*, \ldots, C_n^*, \ldots,$$

[5]This is in contrast to the usual way we have proceeded. Here we are starting with a formula and using it to obtain a recurrence relation.

which, in order to refer to them by name, we call the *pseudo-Catalan numbers.* The psuedo-Catalan numbers are defined in terms of the Catalan numbers as follows:

$$C_n^* = n!C_{n-1}, \qquad (n = 1, 2, 3, \ldots).$$

We have

$$C_1^* = 1!(1) = 1,$$

and, using (8.3) with n replaced by $n-1$, we obtain

$$\begin{aligned} C_n^* &= n!C_{n-1} \\ &= n!\frac{4n-6}{n}C_{n-2} \\ &= (4n-6)(n-1)!C_{n-2} \\ &= (4n-6)C_{n-1}^*. \end{aligned}$$

Thus, the pseudo-Catalan numbers are determined by the following recurrence relation and initial condition:

$$\begin{aligned} C_n^* &= (4n-6)C_{n-1}^*, \quad (n \geq 2) \\ C_1^* &= 1. \end{aligned} \tag{8.4}$$

Using this recurrence relation, we calculate the first few pseudo-Catalan numbers:

$$\begin{array}{ll} C_1^* = 1 & C_4^* = 120 \\ C_2^* = 2 & C_5^* = 1680 \\ C_3^* = 12 & C_6^* = 30240. \end{array}$$

The defining formula for the Catalan numbers and the definition of the pseudo-Catalan numbers imply the formula

$$C_n^* = (n-1)!\binom{2n-2}{n-1} = \frac{(2n-2)!}{(n-1)!}, \qquad (n \geq 1)$$

for the pseudo-Catalan numbers. This formula can also be derived from the recurrence relation (8.4).

Example. Let $a_1, a_2, \ldots, a_n$ be n numbers. By a *multiplication scheme* for these numbers we mean a scheme for carrying out the multiplication of $a_1, a_2, \ldots, a_n$. A multiplication scheme requires a total of $n-1$ multiplications between two numbers, each of which is either one of $a_1, a_2, \ldots, a_n$ or a partial product of them. Let h_n denote the number of multiplication schemes for n numbers. We have $h_1 = 1$ (this can be taken as the definition of h_1) and $h_2 = 2$, since

$$(a_1 \times a_2) \quad \text{and} \quad (a_2 \times a_1)$$

are two possible schemes. This example serves to show that the order of the numbers in the multiplication scheme is taken into consideration.[6] If $n = 3$, there are 12 schemes:

$$\begin{array}{lll}(a_1 \times (a_2 \times a_3)) & (a_2 \times (a_1 \times a_3)) & (a_3 \times (a_1 \times a_2)) \\ ((a_2 \times a_3) \times a_1) & ((a_1 \times a_3) \times a_2) & ((a_1 \times a_2) \times a_3) \\ (a_1 \times (a_3 \times a_2)) & (a_2 \times (a_3 \times a_1)) & (a_3 \times (a_2 \times a_1)) \\ ((a_3 \times a_2) \times a_1) & ((a_3 \times a_1) \times a_2) & ((a_2 \times a_1) \times a_3).\end{array}$$

Thus, $h_3 = 12$. Each multiplication scheme for three numbers requires two multiplications, and each multiplication corresponds to a set of parentheses. With the outside parentheses, each multiplication $\times$ can be identified with a set of parentheses. In general, each multiplication scheme can be obtained by listing $a_1, a_2, \ldots, a_n$ in some order and then inserting $n - 1$ pairs of parentheses so that each pair of parentheses designates a multiplication of two factors. But in order to derive a recurrence relation for h_n, we look at it in an inductive way. Each scheme for $a_1, a_2, \ldots, a_n$ can be gotten from a scheme for $a_1, a_2, \ldots, a_{n-1}$ in exactly one of the following ways:

(1) Take a multiplication scheme for $a_1, a_2, \ldots, a_{n-1}$ (which has $n-2$ multiplications and $n - 2$ sets of parentheses) and insert a_n on either side of either factor in one of the $n - 2$ multiplications. Thus, each scheme for $n - 1$ numbers gives $2 \times 2 \times (n - 2) = 4(n - 2)$ schemes for n numbers in this way.

(2) Take a multiplication scheme for $a_1, a_2, \ldots, a_{n-1}$ and multiply it on the left or right by a_n. Thus, each scheme for $n - 1$ numbers gives two schemes for n numbers in this way.

To illustrate, let $n = 6$ and consider the multiplication scheme

$$((a_1 \times a_2) \times ((a_3 \times a_4) \times a_5))$$

for a_1, a_2, a_3, a_4, a_5.[7] There are four multiplications in this scheme. We take any one of them, say, the multiplication of $(a_3 \times a_4)$ and a_5, and insert a_6 on either side of either of these two factors to get

$$\begin{array}{l}((a_1 \times a_2) \times (((a_6 \times (a_3 \times a_4)) \times a_5)) \\ ((a_1 \times a_2) \times (((a_3 \times a_4) \times a_6) \times a_5)) \\ ((a_1 \times a_2) \times ((a_3 \times a_4) \times (a_6 \times a_5))) \\ ((a_1 \times a_2) \times ((a_3 \times a_4) \times (a_5 \times a_6))).\end{array}$$

There are $4 \times 4 = 16$ schemes for $a_1, a_2, a_3, a_4, a_5, a_6$ obtained in this way. Besides these, we have two additional schemes in which a_6 enters into the final multiplication, namely,

[6] In more algebraic language, we are not allowed to use the commutative law ($a \times b$ is not to be replaced by $b \times a$), nor are we allowed to use the associative law ($a \times (b \times c)$ is not to be replaced by $(a \times b) \times c$).

[7] Which multiplication $\times$ corresponds to each set of parentheses in the preceding scheme?

$$(a_6 \times ((a_1 \times a_2) \times ((a_3 \times a_4) \times a_5))), \quad (((a_1 \times a_2) \times ((a_3 \times a_4) \times a_5)) \times a_6).$$

Thus, each multiplication scheme for five numbers gives 18 schemes for six numbers, and we have $h_6 = 18h_5$.

Let $n \geq 2$. Then, generalizing the foregoing analysis, we see that each of the h_{n-1} multiplication schemes for $n-1$ numbers gives

$$4(n-2) + 2 = 4n - 6$$

schemes for n numbers. We thus obtain the recurrence relation

$$h_n = (4n-6)h_{n-1}, \quad (n \geq 2),$$

which, together with the initial value $h_1 = 1$, determines the sequence $h_1, h_2, \ldots, h_n, \ldots$ This is the same type of recurrence relation with the same initial value satisfied by the pseudo-Catalan numbers (8.4). Hence,

$$h_n = C_n^* = (n-1)!\binom{2n-2}{n-1}, \quad (n \geq 1).$$

□

In the preceding example, suppose that we count only those multiplication schemes in which the n numbers are listed in the order $a_1, a_2, \ldots, a_n$. Thus, for instance, $((a_2 \times a_1) \times a_3)$ is no longer counted. Let g_n denote the number of multiplication schemes with this additional restriction. Then, since we consider only one of the $n!$ possible orderings, $h_n = n!g_n$, and hence

$$g_n = \frac{h_n}{n!} = \frac{C_n^*}{n!} = \frac{1}{n!}(n-1)!\binom{2n-2}{n-1} = \frac{1}{n}\binom{2n-2}{n-1} = C_{n-1}, \quad (n \geq 1), \tag{8.5}$$

showing that g_n is the $(n-1)$st Catalan number.

We can also derive a recurrence relation for g_n by using its definition as follows: In each scheme for $a_1, a_2, \ldots, a_n$ there is a final multiplication $\times$, and it corresponds to the outer parentheses. We thus have

$$((\text{scheme for } \ a_1, \ldots, a_k) \times (\text{scheme for } \ a_{k+1}, \ldots, a_n)),$$

where the $\times$ shown is the last multiplication. The multiplication scheme for $a_1, \ldots, a_k$ can be chosen in g_k ways, and the multiplication scheme for $a_{k+1}, \ldots, a_n$ can be chosen in g_{n-k} ways. Since k can be any of the numbers $1, 2, \ldots, n-1$, we have

$$g_n = g_1 g_{n-1} + g_2 g_{n-2} + \cdots + g_{n-1} g_1, \quad (n \geq 2). \tag{8.6}$$

This nonlinear recurrence relation, along with the initial condition $g_1 = 1$, uniquely determines the counting sequence

$$g_1, g_2, g_3, \ldots, g_n, \ldots .$$

The solution of the recurrence relation (8.6) that satisfies the initial condition $g_1 = 1$ is given by (8.5). Since $g_n = C_{n-1}$, we can also write

$$C_{n-1} = C_0C_{n-2} + C_1C_{n-3} + \cdots + C_{n-2}C_0, \ (n \geq 2),$$

and so

$$C_n = C_0C_{n-1} + C_1C_{n-2}C_1 + \cdots + C_{n-1}C_0 = \sum_{k=0}^{n-1} C_kC_{n-1-k} \ (n \geq 1). \qquad (8.7)$$

The recurrence relation (8.6) is the same recurrence relation that occurred in Section 7.6 in connection with the problem of dividing a convex polygonal region into triangles by means of its diagonals, where we showed by analytic means that its solution is C_{n-1}. Thus, we have a purely combinatorial derivation of the formula obtained in Section 7.6, and we conclude that *the number of ways to divide a convex polygonal region with* $n+1$ *sides into triangular regions by inserting diagonals that do not intersect in the interior is the same as the number of multiplication schemes for* n *numbers given in a specified order with the common value equal to the* $(n-1)$*st Catalan number.*

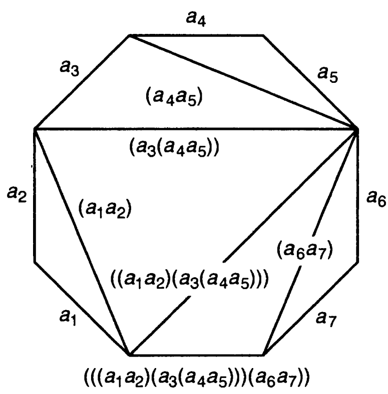


Figure 8.1

The correspondence between the multiplication schemes for the n numbers $a_1, a_2, ..$ and triangularizations of convex polygonal regions of $n+1$ sides is indicated in Figure 8.1 for $n = 7$, where we have suppressed the multiplication symbol. Each diagonal corresponds to one of the multiplications other than the last, with the base of the polygon corresponding to the last multiplication.

8.2 Difference Sequences and Stirling Numbers

Let

$$h_0, h_1, h_2, \ldots, h_n, \ldots \tag{8.8}$$

be a sequence of numbers. We define a new sequence

$$\Delta h_0, \Delta h_1, \Delta h_2, \ldots, \Delta h_n, \ldots, \tag{8.9}$$

called the *(first-order) difference sequence* of (8.8), by

$$\Delta h_n = h_{n+1} - h_n, \qquad (n \geq 0).$$

The terms of the difference sequence (8.9) are the differences of consecutive terms of the sequence (8.8). We can form the difference sequence of (8.9) and obtain the *second-order difference sequence*

$$\Delta^2 h_0, \Delta^2 h_1, \Delta^2 h_2, \ldots, \Delta^2 h_n, \ldots.$$

Here,

$$\begin{aligned}
\Delta^2 h_n &= \Delta(\Delta h_n) \\
&= \Delta h_{n+1} - \Delta h_n \\
&= (h_{n+2} - h_{n+1}) - (h_{n+1} - h_n) \\
&= h_{n+2} - 2h_{n+1} + h_n, \qquad (n \geq 0).
\end{aligned}$$

More generally, we can inductively define the pth-*order difference* sequence of (8.8) by

$$\Delta^p h_0, \Delta^p h_1, \Delta^p h_2, \ldots, \Delta^p h_n, \ldots \quad (p \geq 1),$$

where

$$\Delta^p h_n = \Delta(\Delta^{p-1} h_n).$$

Thus, the pth-order difference sequence is the first-order difference sequence of the $(p-1)$st-order difference sequence. We define the 0th-*order difference sequence* of a sequence to be itself; that is,

$$\Delta^0 h_n = h_n, \qquad (n \geq 0).$$

The *difference table* for the sequence (8.8) is obtained by listing the pth-order difference sequences in a row for each $p = 0, 1, 2, \ldots$:

$$\begin{array}{ccccccccc}
h_0 & & h_1 & & h_2 & & h_3 & & h_4 \quad \cdots \\
& \Delta h_0 & & \Delta h_1 & & \Delta h_2 & & \Delta h_3 \quad \cdots & \\
& & \Delta^2 h_0 & & \Delta^2 h_1 & & \Delta^2 h_2 \quad \cdots & & \\
& & & \Delta^3 h_0 & & \Delta^3 h_1 \quad \cdots & & &
\end{array}$$

The pth-order differences are in row p, with the sequence itself in row 0. (Thus, we start counting the rows with 0.)

Example. Let a sequence $h_0, h_1, h_2, \ldots, h_n, \ldots$ be defined by

$$h_n = 2n^2 + 3n + 1, \qquad (n \geq 0).$$

The difference table for this sequence is

$$\begin{array}{ccccccccccccc}
1 & & 6 & & 15 & & 28 & & 45 & & 66 & & 91 \; \cdots \\
& 5 & & 9 & & 13 & & 17 & & 21 & & 25 \; \cdots & \\
& & 4 & & 4 & & 4 & & 4 & & 4 \; \cdots & & \\
& & & 0 & & 0 & & 0 & & 0 \; \cdots & & & \\
& & & & & & \cdots & & & & & &
\end{array}$$

The third-order difference sequence in this case consists of all 0s and hence so do all higher-order differences sequences. □

We now show that if a sequence has the property that its general term is a polynomial of degree p in n, then the $(p+1)$th-order differences are all 0. When this happens, we may suppress all the rows of 0s after the first row of 0s.

Theorem 8.2.1 *Let the general term of a sequence be a polynomial of degree p in n:*

$$h_n = a_p n^p + a_{p-1} n^{p-1} + \cdots + a_1 n + a_0, \qquad (n \geq 0).$$

Then $\Delta^{p+1} h_n = 0$ for all $n \geq 0$.

Proof. We prove the theorem by induction on p. If $p = 0$, then we have

$$h_n = a_0, \text{a constant, for all } n \geq 0;$$

and hence,

$$\Delta h_n = h_{n+1} - h_n = a_0 - a_0 = 0, \qquad (n \geq 0).$$

We now suppose that $p \geq 1$ and assume that the theorem holds when the general term is a polynomial of degree at most $p - 1$ in n. We have

$$\begin{aligned}
\Delta h_n &= (a_p (n+1)^p + a_{p-1}(n+1)^{p-1} + \cdots + a_1 n + a_0) \\
&\quad -(a_p n^p + a_{p-1} n^{p-1} + \cdots + a_1 n + a_0).
\end{aligned}$$

By the binomial theorem,

$$\begin{aligned}
a_p (n+1)^p - a_p n^p &= a_p \left(n^p + \binom{p}{1} n^{p-1} + \cdots + 1 \right) - a_p n^p \\
&= a_p \binom{p}{1} n^{p-1} + \cdots + a_p.
\end{aligned}$$

From this calculation, we conclude that the pth powers of n cancel in Δh_n and that Δh_n is a polynomial in n of degree at most $p-1$. By the induction assumption,

$$\Delta^p(\Delta h_n) = 0, \qquad (n \geq 0).$$

Since $\Delta^{p+1} h_n = \Delta^p(\Delta h_n)$, it now follows that

$$\Delta^{p+1} h_n = 0, \qquad (n \geq 0).$$

Hence, the theorem holds by induction. □

Now suppose that g_n and f_n are the general terms of two sequences, and another sequence is defined by

$$h_n = g_n + f_n, \qquad (n \geq 0).$$

Then

$$\begin{aligned} \Delta h_n &= h_{n+1} - h_n \\ &= (g_{n+1} + f_{n+1}) - (g_n + f_n) \\ &= (g_{n+1} - g_n) + (f_{n+1} - f_n) \\ &= \Delta g_n + \Delta f_n. \end{aligned}$$

More generally, it follows inductively that

$$\Delta^p h_n = \Delta^p g_n + \Delta^p f_n, \qquad (p \geq 0)$$

and, indeed, if c and d are constants, it also follows that

$$\Delta^p(cg_n + df_n) = c\Delta^p g_n + d\Delta^p f_n, \qquad (n \geq 0) \tag{8.10}$$

for each integer $p \geq 0$. We refer to the property in (8.10) as the *linearity property* of differences.[8] From (8.10) we see that the difference table for the sequence of h_n's can be obtained by multiplying the entries of the difference table for the g_n's by c and multiplying the entries of the difference table for the f_n's by d, and then adding corresponding entries.

Example. Let $g_n = n^2 + n + 1$ and let $f_n = n^2 - n - 2$, $(n \geq 0)$. The difference table for the g_n's is

$$\begin{array}{ccccccccccc} 1 & & 3 & & 7 & & 13 & & 21 & \cdots \\ & 2 & & 4 & & 6 & & 8 & \cdots \\ & & 2 & & 2 & & 2 & \cdots \\ & & & 0 & & 0 & \cdots \end{array}$$

[8] In the language of linear algebra, the set of sequences forms a vector space, and Δ is a linear transformation on this vector space.

The difference table for the f_n's is

$$\begin{array}{ccccccccccc}
-2 & & -2 & & 0 & & 4 & & 10 & & \cdots \\
& 0 & & 2 & & 4 & & 6 & & \cdots & \\
& & 2 & & 2 & & 2 & & \cdots & & \\
& & & 0 & & 0 & & \cdots & & &
\end{array}$$

Let

$$\begin{aligned}
h_n = 2g_n + 3f_n &= 2(n^2+n+1) + 3(n^2-n-2) \\
&= 5n^2 - n - 4.
\end{aligned}$$

The difference table for the h_n's is obtained by multiplying the entries of the first difference table by 2 and the entries of the second difference table by 3 and then adding corresponding entries. The result is

$$\begin{array}{ccccccccccc}
-4 & & 0 & & 14 & & 38 & & 72 & & \cdots \\
& 4 & & 14 & & 24 & & 34 & & \cdots & \\
& & 10 & & 10 & & 10 & & \cdots & & \\
& & & 0 & & 0 & & \cdots & & &
\end{array}$$

$\square$

By its very definition, the difference table for a sequence $h_0, h_1, h_2, \ldots, h_n, \ldots$ is determined by the entries in row number 0. We next observe that the difference table is also determined by the entries along the left edge, the 0*th diagonal*—that is, by the numbers

$$h_0 = \Delta^0 h_0, \Delta^1 h_0, \Delta^2 h_0, \Delta^3 h_0, \ldots$$

along the leftmost diagonal of the difference table.[9] This property is a consequence of the fact that the entries on a diagonal (running from left to right) of the difference table are determined from those on the previous diagonal. For instance, the entries on the 1st diagonal are

$$\begin{aligned}
h_1 &= \Delta^0 h_1 = \Delta^1 h_0 + \Delta^0 h_0 = \Delta h_0 + h_0 \\
\Delta h_1 &= \Delta^2 h_0 + \Delta h_0 \\
\Delta^2 h_1 &= \Delta^3 h_0 + \Delta^2 h_0 \\
\cdots & \quad \cdots
\end{aligned}$$

If the 0th diagonal of a difference table contains only 0s, then the entire difference table contains only 0s. The next simplest 0th diagonal is one that contains only 0s except for one 1, say, in row p. (Thus there are p 0s preceding the 1.) From the fact

[9]This property is the discrete analogue of the fact that an analytic function $f(x)$ is determined (via its Taylor expansion) by the value of the function and all its derivatives at $x = 0 : f(0), f'(0), f''(0), \ldots$.

that the entries on the 0th diagonal in rows $p+1, p+2, \ldots$ are all 0, it is apparent that all the entries in rows $p+1, p+2, \ldots$ equal 0.

Suppose, for instance, $p = 4$. Thus, rows 5 and greater contain only 0s. Can we find the general term of a sequence such that the 0th diagonal of its difference table is

$$0, 0, 0, 0, 1, 0, 0, \ldots? \tag{8.11}$$

We use these entries on the left edge to determine a triangular portion of the difference table and obtain

$$\begin{array}{ccccccccc}
0 & & 0 & & 0 & & 0 & & 1 \\
& 0 & & 0 & & 0 & & 1 & \\
& & 0 & & 0 & & 1 & & \\
& & & 0 & & 1 & & & \\
& & & & 1. & & & &
\end{array}$$

Since row number 5 consists of all 0s, we look for a sequence whose nth term h_n is a polynomial in n of degree 4. From the portion of the difference table just computed, we see that

$$h_0 = 0, \ h_1 = 0, \ h_2 = 0, \ h_3 = 0, \text{ and } h_4 = 1.$$

Thus, if h_n is a polynomial of degree 4, it has roots $0, 1, 2, 3$, and hence

$$h_n = cn(n-1)(n-2)(n-3)$$

for some constant c. Since $h_4 = 1$, we must have

$$1 = c(4)(3)(2)(1) \text{ or, equivalently, } c = \frac{1}{4!}.$$

Accordingly, the sequence with general term

$$h_n = \frac{n(n-1)(n-2)(n-3)}{4!} = \binom{n}{4}, \qquad (n \geq 0)$$

has a difference table with 0th diagonal given by (8.11).

The same kind of argument shows that, more generally,

$$h_n = \frac{n(n-1)(n-2)\cdots(n-(p-1))}{p!} = \binom{n}{p}$$

is a polynomial in n of degree p whose difference table has its 0th diagonal equal to

$$\overbrace{0, 0, \ldots, 0}^{p}, 1, 0, 0, \ldots.$$

Using the linearity property of differences and the fact that the 0th diagonal of a difference table determines the entire difference table, and hence the sequence itself, we obtain the next theorem.

Theorem 8.2.2 *The general term of the sequence whose difference table has its* 0th *diagonal equal to*

$$c_0, c_1, c_2, \ldots, c_p, 0, 0, 0, \ldots, \qquad \textit{where } c_p \neq 0$$

is a polynomial in n *of degree* p *satisfying*

$$h_n = c_0\binom{n}{0} + c_1\binom{n}{1} + c_2\binom{n}{2} + \cdots + c_p\binom{n}{p}. \tag{8.12}$$

□

Combining Theorems 8.2.1 and 8.2.2, we see that every polynomial in n of degree p can be expressed in the form (8.12) for some choice of constants $c_0, c_1, \ldots, c_p$. These constants are uniquely determined. (See Exercise 10.)

Example. Consider the sequence with general term

$$h_n = n^3 + 3n^2 - 2n + 1, \qquad (n \geq 0).$$

Computing differences, we obtain

$$\begin{array}{ccccccc} 1 & & 3 & & 17 & & 49 \\ & 2 & & 14 & & 32 & \\ & & 12 & & 18 & & \\ & & & 6. & & & \end{array}$$

Since h_n is a polynomial in n of degree 3, the 0th diagonal of the difference table is

$$1, 2, 12, 6, 0, 0, \ldots.$$

Hence, by Theorem 8.2.2, another way to write h_n is

$$h_n = 1\binom{n}{0} + 2\binom{n}{1} + 12\binom{n}{2} + 6\binom{n}{3}. \tag{8.13}$$

Why would we want to write h_n in this way? Here's one reason. Suppose we want to find the partial sums

$$\sum_{k=0}^{n} h_k = h_0 + h_1 + \cdots + h_n.$$

Using (8.13), we see that

$$\sum_{k=0}^{n} h_k = 1\sum_{k=0}^{n}\binom{k}{0} + 2\sum_{k=0}^{n}\binom{k}{1} + 12\sum_{k=0}^{n}\binom{k}{2} + 6\sum_{k=0}^{n}\binom{k}{3}.$$

From (5.14) we know that

$$\sum_{k=0}^{n} \binom{k}{p} = \binom{n+1}{p+1}. \tag{8.14}$$

Hence,

$$\sum_{k=0}^{n} h_k = 1\binom{n+1}{1} + 2\binom{n+1}{2} + 12\binom{n+1}{3} + 6\binom{n+1}{4},$$

a very simple formula for the partial sums. □

The foregoing procedure can be used to find the partial sums of any sequence whose general term is a polynomial in n.

Theorem 8.2.3 *Assume that the sequence $h_0, h_1, h_2, \ldots, h_n, \ldots$ has a difference table whose 0th diagonal equals*

$$c_0, c_1, c_2, \ldots, c_p, 0, 0, \ldots.$$

Then

$$\sum_{k=0}^{n} h_k = c_0\binom{n+1}{1} + c_1\binom{n+1}{2} + \cdots + c_p\binom{n+1}{p+1}.$$

Proof. By Theorem 8.2.2, we have

$$h_n = c_0\binom{n}{0} + c_1\binom{n}{1} + \cdots + c_p\binom{n}{p}.$$

Using formula (8.14), we obtain

$$\begin{aligned}\sum_{k=0}^{n} h_k &= c_0\sum_{k=0}^{n}\binom{k}{0} + c_1\sum_{k=0}^{n}\binom{k}{1} + \cdots + c_k\sum_{k=0}^{n}\binom{k}{p}\\ &= c_0\binom{n+1}{1} + c_1\binom{n+1}{2} + \cdots + c_p\binom{n+1}{p+1}.\end{aligned}$$ □

Example. Find the sum of the fourth powers of the first n positive integers.

Let $h_n = n^4$. Computing differences, we obtain

$$\begin{array}{ccccccccc} 0 & & 1 & & 16 & & 81 & & 256 \\ & 1 & & 15 & & 65 & & 175 & \\ & & 14 & & 50 & & 110 & & \\ & & & 36 & & 60 & & & \\ & & & & 24. & & & & \end{array}$$

Because h_n is a polynomial of degree 4, the 0th diagonal of the difference table equals

$$0, 1, 14, 36, 24, 0, 0, \ldots.$$

Hence,

$$\begin{aligned} 1^4 + 2^4 + \cdots + n^4 &= \sum_{k=0}^{n} k^4 \\ &= 0\binom{n+1}{1} + 1\binom{n+1}{2} + 14\binom{n+1}{3} \\ &\quad +36\binom{n+1}{4} + 24\binom{n+1}{5}. \qquad \square \end{aligned}$$

In a similar way, we can evaluate the sum of the pth powers of the first n positive integers by considering the sequence whose general term is $h_n = n^p$. The preceding example treated the case $p = 4$.

The numbers that occur in the 0th diagonal of the difference tables are of combinatorial significance, and we now discuss them.

Let

$$h_n = n^p.$$

By Theorems 8.2.1 and 8.2.2, the 0th diagonal of the difference table for h_n has the form

$$c(p,0), c(p,1), c(p,2), \ldots, c(p,p), 0, 0, \ldots,$$

and it follows that

$$n^p = c(p,0)\binom{n}{0} + c(p,1)\binom{n}{1} + \cdots + c(p,p)\binom{n}{p}. \tag{8.15}$$

If $p = 0$, then $h_n = 1$, a constant, and (8.15) reduces to

$$n^0 = 1 = 1\binom{n}{0} = 1;$$

in particular,

$$c(0,0) = 1.$$

Since, if $p \geq 1$, n^p, as a polynomial in n, has a constant term equal to 0, we also have

$$c(p,0) = 0, \quad (p \geq 1).$$

We rewrite (8.15) by introducing a new expression. Let

$$[n]_k = \begin{cases} n(n-1)\cdots(n-k+1) & \text{if } k \geq 1 \\ 1 & \text{if } k = 0. \end{cases}$$

We note that $[n]_k$ is the same as $P(n,k)$, the number of k-permutations of n distinct objects (see Section 3.2), but we wish now to use the less cumbersome notation $[n]_k$. We also note that

$$[n]_{k+1} = (n-k)[n]_k.$$

Since

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{[n]_k}{k!},$$

we obtain

$$[n]_k = k!\binom{n}{k}.$$

Hence, (8.15) can be rewritten as

$$\begin{aligned} n^p &= c(p,0)\frac{[n]_0}{0!} + c(p,1)\frac{[n]_1}{1!} + \cdots + c(p,p)\frac{[n]_p}{p!} \\ &= \sum_{k=0}^{p} c(p,k)\frac{[n]_k}{k!} \\ &= \sum_{k=0}^{p} \frac{c(p,k)}{k!}[n]_k. \end{aligned}$$

Now we introduce the numbers

$$S(p,k) = \frac{c(p,k)}{k!}, \qquad (0 \le k \le p)$$

and in terms of them, (8.15) becomes

$$\begin{aligned} n^p &= S(p,0)[n]_0 + S(p,1)[n]_1 + \cdots + S(p,p)[n]_p \\ &= \sum_{k=0}^{p} S(p,k)[n]_k. \end{aligned}$$

The numbers $S(p,k)$ just introduced are called the *Stirling numbers*[10] *of the second kind.*[11] Since

$$S(p,0) = \frac{c(p,0)}{0!} = c(p,0),$$

we have

$$S(p,0) = \begin{cases} 1 & \text{if } p = 0 \\ 0 & \text{if } p \ge 1. \end{cases} \tag{8.16}$$

In (8.15), the coefficient of n^p on the left-hand side is 1, and on the right-hand side the coefficient is

$$\frac{c(p,p)}{p!}.$$

[10] After James Stirling (1692–1770).

[11] So there must be Stirling numbers of the first kind! We discuss them later in this section.

(Only the last term on the right side of (8.15) contributes to the coefficient of n^p, since the other terms are polynomials in n of degree less than p.) Thus, we have

$$S(p,p) = \frac{c(p,p)}{p!} = 1, \qquad (p \geq 0). \tag{8.17}$$

We now show that the Stirling numbers of the second kind satisfy a Pascal-like recurrence relation.

Theorem 8.2.4 *If $1 \leq k \leq p-1$, then*

$$S(p,k) = kS(p-1,k) + S(p-1,k-1).$$

Proof. We first observe that, were it not for the factor k in front of $S(p-1,k)$, we would have the Pascal recurrence. We have

$$n^p = \sum_{k=0}^{p} S(p,k)[n]_k \tag{8.18}$$

and

$$n^{p-1} = \sum_{k=0}^{p-1} S(p-1,k)[n]_k.$$

Thus,

$$\begin{aligned}
n^p = n \times n^{p-1} &= n\sum_{k=0}^{p-1} S(p-1,k)[n]_k \\
&= \sum_{k=0}^{p-1} S(p-1,k)n[n]_k \\
&= \sum_{k=0}^{p-1} S(p-1,k)(n-k+k)[n]_k \\
&= \sum_{k=0}^{p-1} S(p-1,k)(n-k)[n]_k + \sum_{k=0}^{p-1} kS(p-1,k)[n]_k \\
&= \sum_{k=0}^{p-1} S(p-1,k)[n]_{k+1} + \sum_{k=1}^{p-1} kS(p-1,k)[n]_k.
\end{aligned}$$

We replace k by $k-1$ in the left summation in the line directly above and obtain

$$\begin{aligned}
n^p &= \sum_{k=1}^{p} S(p-1,k-1)[n]_k + \sum_{k=1}^{p-1} kS(p-1,k)[n]_k \\
&= S(p-1,p-1)[n]_p + \sum_{k=1}^{p-1} \left(S(p-1,k-1) + kS(p-1,k)\right)[n]_k.
\end{aligned}$$

For each k with $1 \leq k \leq p-1$, comparing the coefficient of $[n]_k$ in this expression for n^p with the coefficient of $[n]_k$ in the expression (8.18), we obtain

$$S(p,k) = S(p-1,k-1) + kS(p-1,k).$$

□

The recurrence relation given in Theorem 8.2.4 and the initial values

$$S(p,0) = 0, \quad (p \geq 1) \text{ and } S(p,p) = 1, \quad (p \geq 0)$$

from (8.16) and (8.17) determine the sequence of Stirling numbers of the second kind $S(p,k)$. As for the binomial coefficients, we have a Pascal-like triangle for these Stirling numbers. (See Figure 8.2.)

$p\backslash k$	0	1	2	3	4	5	6	7	$\cdots$
0	1								
1	0	1							
2	0	1	1						
3	0	1	3	1					
4	0	1	7	6	1				
5	0	1	15	25	10	1			
6	0	1	31	90	65	15	1		
7	0	1	63	301	350	140	21	1	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Figure 8.2 The triangle of $S(p,k)$

Each entry $S(p,k)$ in the triangle, other than those on the vertical and diagonal sides of the triangle (these are the entries given by the initial values), is obtained by multiplying the entry in the row directly above it by k and adding the result to the entry immediately to its left in the row directly above it.

From the triangle of the Stirling numbers of the second kind, it appears that

$$\begin{aligned} S(p,1) &= 1, \ (p \geq 1) \\ S(p,2) &= 2^{p-1} - 1, \ (p \geq 2) \\ S(p,p-1) &= \binom{p}{2}, \ (p \geq 1). \end{aligned}$$

We leave the verification of these formulas as exercises. They are also readily verified using the combinatorial interpretation of the Stirling numbers of the second kind given in the next theorem.

Theorem 8.2.5 *The Stirling number of the second kind $S(p,k)$ counts the number of partitions of a set of p elements into k indistinguishable boxes in which no box is empty.*

Proof. First, we give an explanation of what indistinguishable means in this case. To say that the boxes are indistinguishable means that we can't tell one box from another. They all look the same. If, for instance, the contents of some box are the elements a, b, and c, then it doesn't matter which box it is. The only thing that matters is what the contents of the various boxes are, not *which* box holds what.

Let $S^*(p,k)$ denote the number of partitions of a set of p elements into k indistinguishable boxes in which no box is empty. We easily see that

$$S^*(p,p) = 1, \qquad (p \geq 0)$$

because, if there are the same number of boxes as elements, each box contains exactly one element (and remember, we can't tell one box from another), and

$$S^*(p,0) = 0, \qquad (p \geq 1)$$

because if there is at least one element and no boxes, there can be no partitions. If we can show that the numbers $S^*(p,k)$ satisfy the same recurrence relation as the Stirling numbers of the second kind; that is, if we can show that

$$S^*(p,k) = kS^*(p-1,k) + S^*(p-1,k-1), \qquad (1 \leq k \leq p-1)$$

then we will be able to conclude that $S^*(p,k) = S(p,k)$ for all k and p with $0 \leq k \leq p$.

We argue as follows: Consider the set of the first p positive integers $1, 2, \ldots, p$ as the set to be partitioned. The partitions of $\{1, 2, \ldots, p\}$ into k nonempty, indistinguishable boxes are of two types:

(1) those in which p is all alone in a box; and

(2) those in which p is not in a box by itself. Thus, the box containing p contains at least one more element.

In the case of type (1), if we remove p from the box that contains it, we are left with a partition of $\{1, 2, \ldots, p-1\}$ into $k-1$ nonempty, indistinguishable boxes. Hence, there are $S^*(p-1,k-1)$ partitions of $\{1, 2, \ldots, p\}$ of type (i).

Now consider a partition of type (2). Suppose we remove p from the box that contains it. Since p was not all alone in its box, we are left with a partition $A_1, A_2, \ldots, A_k$ of $\{1, 2, \ldots, p-1\}$ into k nonempty, indistinguishable boxes. We might now want to conclude that there are $S^*(p-1,k)$ partitions of type (2), but this is not so. The reason is that the partition $A_1, A_2, \ldots, A_k$ of $\{1, 2, \ldots, p-1\}$ which results upon the removal of p arises from k different partitions of $\{1, 2, \ldots, p\}$, namely, from

$$\begin{gathered} A_1 \cup \{p\}, A_2, \ldots, A_k, \\ A_1, A_2 \cup \{p\}, \ldots, A_k, \\ \vdots \\ A_1, A_2, \ldots, A_k \cup \{p\}. \end{gathered}$$

Put another way, after we delete p, we can't tell which box it came from; it could have been any one of the k boxes, since all boxes remain nonempty upon the removal of p. It follows that there are $kS^*(p-1,k)$ partitions of $\{1, 2, \ldots, k\}$ of type (2). Hence,

$$S^*(p,k) = kS^*(p-1,k) + S^*(p-1,k-1),$$

and the proof is complete. □

Now that we know that $S(p,k)$ counts the number of partitions of a set of p elements into k nonempty, indistinguishable boxes, we have no use for the notation $S^*(p,k)$ introduced in the proof of Theorem 8.2.5. It is now redundant.

We now use our combinatorial interpretation of the Stirling numbers of the second kind to obtain a formula for them. In doing so, we shall first determine the number $S^\#(p,k)$[12] of partitions of $\{1, 2, \ldots, k\}$ into k nonempty, *distinguishable* boxes.[13] Think of one box as colored red, one colored blue, one green, and so on. Now it not only matters which elements are together in a box, but which box it is. (Is it the red box, the blue box, the green one, . . . ?) Once the contents of the k boxes are known, we can color the k boxes in $k!$ ways. Thus,

$$S^\#(p,k) = k!S(p,k), \tag{8.19}$$

and it follows that

$$S(p,k) = \frac{1}{k!}S^\#(p,k).$$

(Note that (8.19) implies that the numbers $S^\#(p,k)$ are the same as the numbers $c(p,k)$ introduced earlier.) Thus, it suffices to find a formula for $S^\#(p,k)$, and this we do by applying the inclusion–exclusion principle of Chapter 6. Before doing so, we remark that the validity of (8.19) rests on the fact that each box is nonempty. If boxes were allowed to be empty, we could not multiply $S(p,k)$ by $k!$ to get $S^\#(p,k)$. If r of the boxes of a partition were empty, then it would give rise to only $\frac{k!}{r!}$ partitions into distinguishable boxes, because permuting empty boxes amongst themselves doesn't change anything.[14]

[12] We abandoned one notation and almost immediately introduce another. In mathematics, notation is important. It adds clarity when properly used; briefness is not its only virtue.

[13] Just when you're starting to feel comfortable with indistinguishable boxes, we change the rules and distinguish them.

[14] What we really have is a multiset with r objects of the same type (the empty set) and $k-r$ other different objects (the contents of the nonempty boxes).

Theorem 8.2.6 *For each integer k with $0 \le k \le p$, we have*

$$S^{\#}(p,k) = \sum_{t=0}^{k} (-1)^t \binom{k}{t} (k-t)^p;$$

hence,

$$S(p,k) = \frac{1}{k!} \sum_{t=0}^{k} (-1)^t \binom{k}{t} (k-t)^p.$$

Proof. Let U be the set of *all* partitions of $\{1, 2, \ldots, p\}$ into k distinguishable boxes $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k$. We define k properties $P_1, P_2, \ldots, P_k$, where P_i is the property that the ith box $\mathcal{B}_i$ is empty. Let A_i denote the subset of U consisting of those partitions for which box $\mathcal{B}_i$ is empty. Then

$$S^{\#}(p,k) = |\overline{A}_1 \cap \overline{A}_2 \cap \cdots \cap \overline{A}_k|.$$

We have

$$|U| = k^p$$

since each of the p elements can be put into any one of the k distinguishable boxes. Let t be an integer with $1 \le t \le k$. How many partitions of U belong to the intersection $A_1 \cap A_2 \cap \cdots \cap A_t$? For these partitions, boxes $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_t$ are empty and the remaining boxes $\mathcal{B}_{t+1}, \ldots, \mathcal{B}_k$ may or may not be empty. Thus, $|A_1 \cap A_2 \cap \cdots \cap A_t|$ counts the number of partitions of $\{1, 2, \ldots, p\}$ into $k-t$ distinguishable boxes and hence equals $(k-t)^p$. The same conclusion holds no matter which t boxes are assumed empty; that is,

$$|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_t}| = (k-t)^p$$

for each t-subset $\{i_1, i_2, \ldots, i_t\}$ of $\{1, 2, \ldots, k\}$. By the inclusion–exclusion principle (see formula (6.3)), we have

$$S^{\#}(p,k) = \sum_{t=0}^{k} (-1)^t \binom{k}{t} (k-t)^p.$$

□

The *Bell number*[15] B_p is the number of partitions of a set of p elements into nonempty, indistinguishable boxes. Now we do not specify the number of boxes, but since no box is to be empty, the number of boxes cannot exceed p. The Bell numbers are just the sum of the entries in a row of the triangle of Stirling numbers of the second kind (see Figure 8.2); that is,

$$B_p = S(p,0) + S(p,1) + \cdots + S(p,p).$$

[15] After E. T. Bell (1883–1960).

We therefore have

$$\begin{array}{ll} B_0 = 1 & B_4 = 15 \\ B_1 = 1 & B_5 = 52 \\ B_2 = 2 & B_6 = 203 \\ B_3 = 5 & B_7 = 877. \end{array}$$

The Bell numbers satisfy a recurrence relation, but not one of constant order.

Theorem 8.2.7 *If $p \geq 1$, then*

$$B_p = \binom{p-1}{0}B_0 + \binom{p-1}{1}B_1 + \cdots + \binom{p-1}{p-1}B_{p-1}.$$

Proof. We partition the set $\{1, 2, \ldots, p\}$ into nonempty, indistinguishable boxes. The box containing p also contains a subset X (possibly empty) of $\{1, 2, \ldots, p-1\}$. The set X has t elements, where t is some integer between 0 and $p-1$. We can choose a set X of size t in $\binom{p-1}{t}$ ways and partition the $p-1-t$ elements of $\{1, 2, \ldots, p-1\}$ that don't belong to X into nonempty, indistinguishable boxes in B_{p-1-t} ways. Hence,

$$B_p = \sum_{t=0}^{p-1} \binom{p-1}{t} B_{p-1-t}.$$

As t takes on the values $0, 1, \ldots, p-1$, so does $(p-1)-t$. Hence, we obtain

$$\begin{aligned} B_p &= \sum_{t=0}^{p-1} \binom{p-1}{(p-1)-t} B_t \\ &= \sum_{t=0}^{p-1} \binom{p-1}{t} B_t. \end{aligned}$$

□

The Stirling numbers of the second kind show us how to write n^p in terms of $[n]_0, [n]_1, \ldots, [n]_p$. The Stirling numbers of the first kind play the inverse role. They show us how to write $[n]_p$ in terms of $n^0, n^1, \ldots, n^p$.[16] By definition,

$$\begin{aligned} [n]_p &= n(n-1)(n-2)\cdots(n-p+1) \\ &= (n-0)(n-1)(n-2)\cdots(n-(p-1)). \end{aligned} \tag{8.20}$$

Thus,

[16]For those familiar with linear algebra, the polynomials of degree at most p with, say, real coefficients form a vector space of dimension $p+1$. Both $1, n, n^2, \ldots, n^p$ and $[n]_0 = 1, [n]_1, \ldots, [n]_p$ are a basis for this vector space. The Stirling numbers of the first and second kind show us how to express one basis in terms of the other.

(1) $[n]_0 = 1$,

(2) $[n]_1 = n$,

(3) $[n]_2 = n(n-1) = n^2 - n$,

(4) $[n]_3 = n(n-1)(n-2) = n^3 - 3n^2 + 2n$,

(5) $[n]_4 = n(n-1)(n-2)(n-3) = n^4 - 6n^3 + 11n^2 - 6n$.

In general, the product on the right in (8.20) has p factors. If we multiply it out, we obtain a polynomial involving the powers

$$n^p, n^{p-1}, \ldots, n^1, n^0 = 1$$

of n in which the coefficients alternate in sign; that is, we obtain an expression of the form

$$\begin{aligned} [n]_p &= s(p,p)n^p - s(p,p-1)n^{p-1} + \cdots + \\ &\qquad (-1)^{p-1}s(p,1)n^1 + (-1)^p s(p,0)n^0 \\ &= \sum_{k=0}^{p}(-1)^{p-k}s(p,k)n^k. \end{aligned} \tag{8.21}$$

The *Stirling numbers of the first kind* are the coefficients

$$s(p,k), \qquad (0 \leq k \leq p)$$

that occur in (8.21). It follows readily from (8.20) and (8.21) that

$$s(p,0) = 0, \qquad (p \geq 1)$$

and

$$s(p,p) = 1, \qquad (p \geq 0).$$

Thus, the Stirling numbers of the first kind satisfy the same initial conditions as the Stirling numbers of the second kind. But they satisfy a different recurrence relation, whose proof follows the same basic outline as that of Theorem 8.2.4.

Theorem 8.2.8 *If* $1 \leq k \leq p-1$, *then*

$$s(p,k) = (p-1)s(p-1,k) + s(p-1,k-1).$$

Proof. By (8.21), we have

$$[n]_p = \sum_{k=0}^{p}(-1)^{p-k}s(p,k)n^k. \tag{8.22}$$

Replacing p by $p-1$ in this equation, we also have

$$[n]_{p-1} = \sum_{k=0}^{p-1}(-1)^{p-1-k}s(p-1,k)n^k.$$

Next, we observe that

$$[n]_p = [n]_{p-1}(n-(p-1)).$$

Hence,

$$[n]_p = (n-(p-1))\sum_{k=0}^{p-1}(-1)^{p-1-k}s(p-1,k)n^k,$$

which, after rewriting, becomes

$$\sum_{k=0}^{p-1}(-1)^{p-1-k}s(p-1,k)n^{k+1} + \sum_{k=0}^{p-1}(-1)^{p-k}(p-1)s(p-1,k)n^k.$$

We replace k by $k-1$ in the first summation and obtain

$$[n]_p = \sum_{k=1}^{p}(-1)^{p-k}s(p-1,k-1)n^k + \sum_{k=0}^{p-1}(-1)^{p-k}(p-1)s(p-1,k)n^k.$$

Comparing the coefficient of n^k in this expression with the coefficient of n^k in the expression (8.22), we get

$$s(p,k) = s(p-1,k-1) + (p-1)s(p-1,k)$$

for each integer k with $1 \leq k \leq p-1$. $\square$

Like the Stirling numbers of the second kind, the Stirling numbers of the first kind also count something quite natural, and this is explained in the next theorem. Its proof is similar in structure to the proof of Theorem 8.2.5.

Theorem 8.2.9 *The Stirling number $s(p,k)$ of the first kind counts the number of arrangements of p objects into k nonempty circular permutations.*

Proof. We refer to the circular permutations in the statement of the theorem as circles. Let $s^{\#}(p,k)$ denote the number of ways to arrange p people in k nonempty circles. We have

$$s^{\#}(p,p) = 1, \qquad (p \geq 0)$$

because, if there are p people and p circles, then each circle contains one person.[17] We also have

$$s^{\#}(p,0) = 0, \qquad (p \geq 1)$$

[17]The right hand of each person holds the left hand of the same person.

because, if there is at least one person, any arrangement contains at least one circle. Thus, the numbers $s^{\#}(p,k)$ satisfy the same initial conditions as the Stirling numbers of the first kind. We now show that they satisfy the same recurrence relation; that is,

$$s^{\#}(p,k) = (p-1)s^{\#}(p-1,k) + s^{\#}(p-1,k-1).$$

Let the people be labeled $1, 2, \ldots, p$. The arrangements of $1, 2, \ldots, p$ into k circles are of two types. Those of the first type have person p in a circle by himself; there are $s^{\#}(p-1,k-1)$ of these. In the second type, p is in a circle with at least one other person. These can be obtained from the arrangements of $1, 2, \ldots, p-1$ into k circles by putting person p on the left of any one of $1, 2, \ldots, p-1$. Thus, each arrangement of $1, 2, \ldots, p-1$ gives $p-1$ arrangements of $1, 2, \ldots, p$ in this way, and hence there is a total of $(p-1)s^{\#}(p-1,k)$ arrangements of the second type. Hence, the number of arrangements of p people into k circles is

$$s^{\#}(p,k) = s^{\#}(p-1,k-1) + (p-1)s^{\#}(p-1,k).$$

It now follows that $s(p,k) = s^{\#}(p,k)$. $\square$

For emphasis, we note that what we have done in the proof of Theorem 8.2.9 is to partition the set $\{1, 2, \ldots, p\}$ into k nonempty, *indistinguishable* boxes and then arrange the elements in each of the boxes into a circular permutation.

8.3 Partition Numbers

A *partition of a positive integer* n is a representation of n as an unordered sum of one or more positive integers, called *parts*. Since the order of the parts is unimportant, we can always arrange the parts so that they are ordered from largest to smallest. The partitions of 1, 2, 3, 4, and 5 are, respectively,

$$\begin{aligned}
&1;\\
&2, 1+1;\\
&3, 2+1, 1+1+1;\\
&4, 3+1, 2+2, 2+1+1, 1+1+1+1;\\
&5, 4+1, 3+2, 3+1+1, 2+2+1, 2+1+1+1, 1+1+1+1+1.
\end{aligned}$$

A partition of n is sometimes written as

$$\lambda = n^{a_n} \ldots 2^{a_2} \ldots 1^{a_1}, \tag{8.23}$$

where a_i is a nonnegative integer equal to the number of parts equal to i. (This expression is purely symbolic; its terms are not exponentials nor is the expression a product.) When written in the form (8.23), the term i^{a_i} is usually omitted if $a_i = 0$. In this notation, the partitions of 5 are

$$5^1, 4^1 1^1, 3^1 2^1, 3^1 1^2, 2^2 1^1, 2^1 1^3, 1^5.$$

Let p_n denote the number of different partitions of the positive integer n, and for convenience, let $p_0 = 1$. The *partition sequence* is the sequence of numbers

$$p_0, p_1, \ldots, p_n, \ldots.$$

By the preceding discussion, we have $p_0 = 1, p_1 = 1, p_2 = 2, p_3 = 3, p_4 = 5$, and $p_5 = 7$. It is a simple observation (cf. (8.23)) that p_n equals the number of solutions in nonnegative integers $a_n, \ldots, a_2, a_1$ of the equation

$$na_n + \cdots + 2a_2 + 1a_1 = n.$$

Let λ be the partition $n = n_1 + n_2 + \cdots + n_k$ of n, where $n_1 \geq n_2 \geq \cdots \geq n_k > 0$. The *Ferrers diagram*, or simply *diagram*, of λ is a left-justified array of dots that has k rows with n_i dots in row i $(1 \leq i \leq k)$. For example, the diagram of the partition $10 = 4 + 2 + 2 + 1 + 1$ is

$$\begin{array}{llll} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & & \\ \bullet & \bullet & & \\ \bullet & & & \\ \bullet & & & \end{array}$$

The Ferrers diagram of a partition furnishes a geometric picture of a partition and can be helpful in visualizing identities involving the number of partitions of various types.

Theorem 8.3.1 *Let n and r be positive integers with $r \leq n$. Let $p_n(r)$ be the number of partitions of n in which the largest part is r, and let $q_n(r)$ be the number of partitions of $n - r$ in which no part is greater than r. Then*

$$p_n(r) = q_n(r).$$

Proof. We don't have a formula for the number of partitions of the two types in the theorem, but we can prove that the two numbers are equal by establishing a one-to-one correspondence between the two types of partitions. This is quite easy to do: Taking a partition of n with largest part equal to r and removing a part equal to r, we obtain a partition of $n - r$ with no part greater than r. The inverse operation is that of taking a partition of $n - r$ with no part greater than r and inserting a part equal to r, and this gives a partition of n in which the largest part equals r. (In terms of the Ferrers diagram, in the first instance, we remove the top row (containing r dots) of the diagram of the partition of n. and in the second instance, we put a new row of r

dots on the top of the diagram of $n - r$.) Hence we have a one-to-one correspondence proving that the two numbers are equal. $\square$

The *conjugate partition* of the partition λ of n is the partition λ^* whose diagram is obtained from the diagram of λ by interchanging rows with columns (flipping the diagram over the diagonal running from the upper left to the lower right). For example, the diagram of the conjugate of the partition $10 = 4 + 2 + 2 + 1 + 1$ is

$$\begin{array}{lllll} \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & & \\ \bullet & & & & \\ \bullet & & & & \end{array}$$

and thus the conjugate partition is $10 = 5 + 3 + 1 + 1$. The number of parts of the conjugate of a partition λ equals the largest part of λ. It should also be clear that the conjugate of the conjugate of a partition λ is λ itself; that is, $(\lambda^*)^* = \lambda$.

Let λ be the partition $n = n_1 + n_2 + \cdots + n_k$ of n. More formally, the conjugate partition λ^* of λ is the partition $n = n_1^* + n_2^* + \cdots + n_l^*$ of n $(l = n_1)$, where n_i^* is the number of parts of λ that are at least equal to i:

$$n_i^* = |\{j : n \geq i\}| \quad (i = 1, 2, \ldots, l).$$

Example. Let λ be the partition $12 = 4 + 4 + 2 + 2$ of 12, whose diagram is

$$\begin{array}{llll} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & & \\ \bullet & \bullet & & \end{array}.$$

The conjugate λ^* is also the partition $12 = 4 + 4 + 2 + 2$, implying that $\lambda^* = \lambda$. $\square$

A partition λ is a *self-conjugate partition* provided, as in the preceding example, $\lambda = \lambda^*$. Another self-conjugate partition is $10 = 5 + 2 + 1 + 1 + 1$. The Ferrers diagram of a self-conjugate partition is symmetric about the diagonal beginning at its upper right corner; if we reflect the diagram about this diagonal, there is no change in the diagram.

Theorem 8.3.2 *Let n be a positive integer. Let p_n^{s} equal the number of self-conjugate partitions of n, and let p_n^{t} be the number of partitions of n into distinct odd parts. Then*

$$p_n^{\mathrm{s}} = p_n^{\mathrm{t}}.$$

Proof. As in the proof of Theorem 8.3.1 we establish a one-to-one correspondence between the two types of partitions, thereby proving that their numbers are equal. The correspondence is most easily described in terms of the Ferrers diagram. Take a self-conjugate partition of n. The number of dots within the first row and column is an odd number; we remove these dots and then combine them into the first row of a new diagram. (Note that in removing the dots in the first row and column, we are left with the diagram of another self-conjugate partition.) The number of dots remaining in the second row and second column is a smaller odd number, and we remove them and then combine them into the second row of our new diagram. We continue like this until all the dots have been removed and put into the new diagram. The result is the Ferrers diagram of a partition of n into distinct odd parts. For example, consider the self-conjugate partition $15 = 5 + 4 + 3 + 2 + 1$, The above transformation is

$$\begin{array}{lcl} \bullet\ \bullet\ \bullet\ \bullet\ \bullet & & \\ \bullet\ \bullet\ \bullet\ \bullet & & \bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet \\ \bullet\ \bullet\ \bullet & \rightarrow & \bullet\ \bullet\ \bullet\ \bullet\ \bullet \\ \bullet\ \bullet & & \bullet \\ \bullet & & \end{array} .$$

Starting with any partition of n into distinct odd parts, we can reverse this transformation by bending at the middle the rows of its Ferrers diagram and fitting these bent rows inside one another in order to obtain a self-conjugate partition of n (in the preceding example, reverse the arrow). Thus we have a one-to-one correspondence, proving that $p_n^{\mathbf{s}} = p_n^{\mathbf{t}}$. $\square$

Another famous partition identity is the following identity of Euler.

Theorem 8.3.3 *Let n be a positive integer. Let $p_n^{\mathbf{o}}$ be the number of partitions of n into odd parts, and let $p_n^{\mathbf{d}}$ be the number of partitions of n into distinct parts. Then*

$$p_n^{\mathbf{o}} = p_n^{\mathbf{d}}.$$

Proof. We establish a one-to-one correspondence between the two types of partitions. Consider a partition of n into odd parts. If the parts are distinct (there aren't two copies of the same part), then we also have a partition of n into distinct parts. If there are two copies of the same part, say k and k, then we combine them into one part, $2k$. We continue to do this until all parts are distinct. Since each time we combine two parts we reduce the number of parts, this procedure certainly terminates and with a partition of n into distinct parts.[18]

We now have to show we can reverse our steps and get back to a partition of n into odd parts. So consider a partition of n into distinct parts. If all parts are odd, then we have a partition of n into odd parts. Otherwise there is at least one even part, and we spilt each even part into two equal parts. If now all parts are odd, we are

[18]Notice that (1) when we combine two equal parts, we create an even part, and (2) if there are several pairs of equal parts, it doesn't matter in what order we combine them; indeed we can do a "mass" combining, by combining each pair in one step. In general this leads to more equal pairs, at which time we do another mass combining, and so forth.

done. Otherwise, we take all of the new even parts and split them equally. At each stage we split at least one even number into two equal and smaller parts, and hence this procedure terminates and with a partition of n into odd parts. Thus we have a one-to-one correspondence between partitions of n into odd parts and partitions of n into distinct parts. $\square$

We illustrate the one-to-one correspondence in the proof of Theorem 8.3.3. Consider the partition of 32 given by

$$32 = 7+5+5+5+3+3+1+1+1+1.$$

The corresponding partition of 32 into distinct parts is obtained as follows:

$$\begin{aligned} 7+5+5+5+3+3+1+1+1+1 &\rightarrow 7+10+5+6+2+2 \\ &\rightarrow 7+10+5+6+4. \end{aligned}$$

The partition

$$32 = 11+9+6+4+2$$

into distinct parts corresponds to the partition of 32 into odd parts obtained as follows:

$$\begin{aligned} 11+9+6+4+2 &\rightarrow 11+9+3+3+2+2+1+1 \\ &\rightarrow 11+9+3+3+1+1+1+1+1+1. \end{aligned}$$

We now obtain an expression for the generating function of the sequence of partition numbers in the form of an infinite product.

Theorem 8.3.4

$$\sum_{n=0}^{\infty} p_n x^n = \prod_{k=1}^{\infty} (1-x^k)^{-1}$$

Proof. The expression on the right equals the product

$$(1+x+\cdots+x^{1a_1}+\cdots)(1+x^2+\cdots+x^{2a_2}+\cdots)(1+x^3+\cdots+x^{3a_3}+\cdots)\cdots.$$

A term x^n arises in this product by choosing a term x^{1a_1} from the first factor, x^{2a_2} from the second, x^{3a_3} from the third, and so on, with $1a_1+2a_2+3a_3+\cdots = n$. (Of course, all but a finite number of the a_i's equal 0; that is, the first term 1 is chosen from all but a finite number of the factors.) Thus, each partition of n contributes 1 to the coefficient of x^n, and the coefficient of x^n equals the number p_n of partitions of n. $\square$

Let $\mathcal{P}_n$ denote the set of all partitions of the positive integer n. There is a natural way to partially order the partitions in $\mathcal{P}_n$. (For this definition, it is notationally convenient to allow zero parts so that when we compare two partitions they have the same number of parts.) Let

$$\lambda : n = n_1 + n_2 + \cdots + n_k \quad (n_1 \geq n_2 \geq \cdots \geq n_k \geq 0)$$

and

$$\mu : n = m_1 + m_2 + \cdots + m_k \quad (m_1 \geq m_2 \geq \cdots \geq m_k \geq 0)$$

be two partitions of n. We say that λ *is majorized by* μ (or that μ *majorizes* λ) and write

$$\lambda \leq \mu,$$

provided that the partial sums for λ are at most equal to the corresponding partial sums for μ:

$$n_1 + \cdots + n_i \leq m_1 + \cdots + m_i \quad (i = 1, 2, \ldots, k).$$

It is straightforward to check that the relation of *majorization* is reflexive, antisymmetric, and transitive and hence is a partial order on $\mathcal{P}_n$.

Example. Consider the three partitions of 9:

$$\lambda : 9 = 5 + 1 + 1 + 1 + 1; \mu : 9 = 4 + 2 + 2 + 1; \nu : 9 = 4 + 4 + 1.$$

For the purpose of comparing all three of these partitions, we add trailing 0s to μ and ν, and think of μ as $9 = 4 + 2 + 2 + 1 + 0$ and ν as $9 = 4 + 4 + 1 + 0 + 0$. We have $\mu \leq \nu$ as

$$\begin{aligned}
&4 \leq 4,\\
&4 + 2 \leq 4 + 4,\\
&4 + 2 + 2 \leq 4 + 4 + 1,\\
&4 + 2 + 2 + 1 \leq 4 + 4 + 1 + 0.
\end{aligned}$$

On the other hand, λ and μ are incomparable as $4 < 5$ but $4 + 2 + 2 > 5 + 1 + 1$. Similarly, λ and ν are incomparable. □

In Section 4.3 we discussed the lexicographic order for n-tuples of 0s and 1s. The lexicographic order can also be used on partitions to produce a total order on $\mathcal{P}_n$ that turns out to be a linear extension of the partial order of majorization. Let $\lambda : n = n_1 + n_2 + \cdots + n_k$, $(n_1 \geq n_2 \geq \cdots \geq n_k)$, and $\mu : n = m_1 + m_2 + \cdots + m_k$, $(m_1 \geq m_2 \geq \cdots \geq m_k)$ be two different partitions of n. Then we say that λ precedes μ in the *lexicographic order*,[19] provided that there is an integer i such that $n_j = m_j$ for $j < i$ and $n_i < m_i$. For instance, the partition $12 = 4 + 3 + 2 + 2 + 1$ precedes the partition $12 = 4 + 3 + 3 + 1 + 1$ since, reading from left to right, $4 = 4$, $3 = 3$, but $2 < 3$. It is simple to verify that lexicographic order is a partial order on $\mathcal{P}_n$.

[19]The alphabet is the integers, with smaller integers preceding larger integers in the alphabet. Also, just as in the lexicographic order of n-tuples of 0s and 1s, we read "words" from left to right.

Theorem 8.3.5 *Lexicographic order is a linear extension of the partial order of majorization on the set $\mathcal{P}_n$ of partitions of a positive integer n.*

Proof. The fact that lexicographic order is a total order (each two partitions of n are comparable) follows almost immediately from its definition. We continue with the notation preceding the statement of the theorem. Let λ and μ be different partitions of n, with λ majorized by μ. Choose the first integer i such that $n_j = m_j$ for $j < i$ but $n_i \neq m_i$. Since

$$n_1 + \cdots + n_{i-1} + n_i \leq m_1 + \cdots m_{i-1} + m_i,$$

we conclude that $n_i < m_i$, and hence λ precedes μ in the lexicographic order. □

We conclude this section by stating without proof another famous partition identity of Euler, called *Euler's pentagonal number theorem.*[20]

Theorem 8.3.6 *Let n be a positive integer. Let p'_n be the number of partitions of n into an even number of distinct parts, and let p''_n be the number of partitions of n into an odd number of distinct parts. Then*

$$p'_n = p''_n + e_n,$$

where e_n is an error term given by $e_n = (-1)^j$ if n is of the form $j(3j \pm 1)/2$, and 0 otherwise.

Example. Let $n = 8$. Then the partitions of 8 into an even number of distinct parts are

$$7 + 1, 6 + 2, 5 + 3.$$

The partitions of 8 into an odd number of distinct parts are

$$8, 5 + 2 + 1, 4 + 3 + 1.$$

Thus $p'_8 = p''_8 = 3$. Now let $n = 7$. Then the partitions of 7 into an even number of distinct parts are:

$$6 + 1, 5 + 2, 4 + 3.$$

The partitions of 7 into an odd number of distinct parts are

$$7, 4 + 2 + 1.$$

Thus $p'_7 = 3 = 2 + 1 = p''_7 + 1$. Note that $7 = 2(3 \cdot 2 + 1)/2$ and thus $e_7 = (-1)^2 = 1$.□

[20] For a proof, see G. E. Andrews and K. Eriksson, *Integer Partitions*, Cambridge University Press, Cambridge, 2004.

8.4 A Geometric Problem

In this section we shall obtain a combinatorial geometric interpretation of the sum

$$h_n^{(k)} = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k} \quad (0 \leq k \leq n) \tag{8.24}$$

of the first $k+1$ binomial coefficients with upper argument equal to n—that is, the sum of the first $k+1$ numbers in row n of Pascal's triangle. For each fixed k, we obtain a sequence

$$h_0^{(k)}, h_1^{(k)}, h_2^{(k)}, \ldots, h_n^{(k)}, \ldots . \tag{8.25}$$

If $k = 0$, we have

$$h_n^{(0)} = \binom{n}{0} = 1,$$

and (8.25) is the sequence of all 1s. If $k = 1$, we obtain

$$h_n^{(1)} = \binom{n}{0} + \binom{n}{1} = n + 1.$$

If $k = 2$, we have

$$\begin{aligned} h_n^{(2)} &= \binom{n}{0} + \binom{n}{1} + \binom{n}{2} \\ &= 1 + n + \frac{n(n-1)}{2} \\ &= \frac{n^2+n+2}{2}. \end{aligned}$$

We also note that $h_0^{(k)} = 1$ for all k. We use Pascal's formula to determine the differences of (8.25):

$$\begin{aligned} \Delta h_n^{(k)} &= h_{n+1}^{(k)} - h_n^{(k)} \\ &= \tbinom{n+1}{0} + \tbinom{n+1}{1} + \cdots + \tbinom{n+1}{k} - \tbinom{n}{0} - \tbinom{n}{1} - \cdots - \tbinom{n}{k} \\ &= \left[\tbinom{n+1}{1} - \tbinom{n}{1}\right] + \cdots + \left[\tbinom{n+1}{k} - \tbinom{n}{k}\right] \\ &= \tbinom{n}{0} + \cdots + \tbinom{n}{k-1}. \end{aligned}$$

Hence,

$$\Delta h_n^{(k)} = h_n^{(k-1)}. \tag{8.26}$$

It is a consequence of (8.26) that the difference table for the sequence

$$h_0^{(k)}, h_1^{(k)}, h_2^{(k)}, h_2^{(k)}, \ldots, h_n^{(k)}, \ldots \tag{8.27}$$

can be obtained from the difference table for

$$h_0^{(k-1)}, h_1^{(k-1)}, h_2^{(k-1)}, \ldots, h_n^{(k-1)}, \ldots$$

by inserting (8.27) on top as a new row.

The number $h_n^{(k)}$ counts the number of subsets with at most k elements of a set with n elements. We now show that $h_n^{(k)}$ also has an interpretation as a counting function for a geometrical problem:

> $h_n^{(k)}$ *counts the number of regions into which* k*-dimensional space is divided by* n $(k-1)$*-dimensional hyperplanes in general position.*

We need to explain some of the terms in this assertion.

We start with $k = 1$. Consider a one-dimensional space, that is, a line. A zero-dimensional space is a point and n points in general position means simply that the points are distinct. If we insert n distinct points on the line, then the line gets divided into $n + 1$ parts, called regions. (See Figure 8.3, in which four points divide the line into 5 regions.)

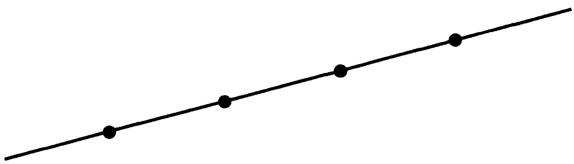


Figure 8.3

This result agrees with the definition of $h_n^{(1)}$ given in (8.24).

Now let $k = 2$, and consider n lines in a plane in general position. In this case, "general position" means that the lines are distinct and not parallel (so that each pair of lines intersects in exactly one point) and the points of intersection are all different—that is, no three of the lines meet in the same point. For n lines in general position in a plane, the number of points of intersection is $\binom{n}{2}$, since each pair of lines gives a different point. The number of regions into which a plane is divided by n lines in general position is given in the following table for $n = 0$ to 5.

Lines	Regions
0	1
1	2
2	4
3	7
4	11
5	16

This table is readily verified.

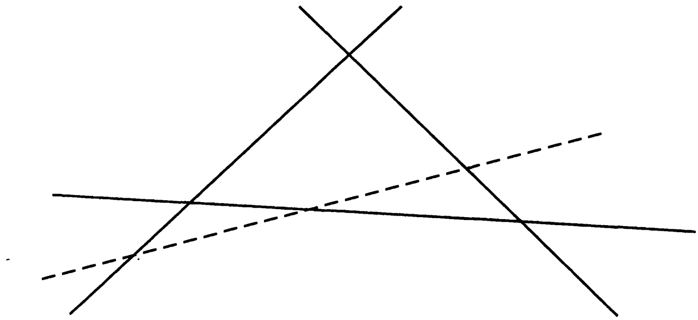


Figure 8.4

We now reason inductively. Suppose we have n lines in general position and we then insert a new line so that the resulting set of $n+1$ lines is in general position. The first n lines intersect the new line in n different points. The n points, as we have already verified, divide the new line into

$$h_n^{(1)} = n+1$$

parts. Each of these $h_n^{(1)} = n+1$ parts divides a region formed by the first n lines into two regions. (See Figure 8.4 for the case $n = 3$ in which the new line is the dashed line.) Hence, the number of regions is increased by $h_n^{(1)} = n+1$ in going from n lines to $n+1$ lines. But this is exactly the relation expressed by (8.26) for the case $k = 2$:

$$\Delta h_n^{(2)} = h_{n+1}^{(2)} - h_n^{(2)} = h_n^{(1)} = n+1.$$

Since $h_0^{(2)} = 1$, we conclude that

$$h_n^{(2)} = \binom{n}{0} + \binom{n}{1} + \binom{n}{2}$$

is the number of regions formed by n lines in general position in a plane.

The case $k = 3$ is similar. Consider n planes in 3-space in general position. *General position* now means that each pair of planes, but no three planes, meet in a line, and every three planes, but no four planes, meet in a point. We now insert a new plane so that the resulting set of $n+1$ planes is also in general position. The first n planes intersect the new plane in n lines in general position (because the planes are in general position). These n lines divide the new plane into $h_n^{(2)}$ planar regions, as determined previously for $k = 2$. Each of these $h_n^{(2)}$ planar regions divides a space region formed by the first n planes into two. Hence, the number of space regions is increased by $h_n^{(2)}$ in going from n planes to $n+1$ planes. This is exactly the relation expressed by (8.26) for the case $k=3$:

$$\Delta h_n^{(3)} = h_{n+1}^{(3)} - h_n^{(3)} = h_n^{(2)}.$$

Since $h_0^{(3)} = 1$ (zero planes divide space into one region, namely, all of space), we conclude that

$$h_n^{(3)} = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \binom{n}{3}$$

is the number of regions into which space is divided by n planes in general position in 3-space.

The same type of reasoning applies to higher dimensional space. The number of regions into which k-dimensional space is divided by n $(k-1)$-dimensional hyperplanes in general position equals

$$h_n^{(k)} = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k}.$$

We conclude by considering the case $k=n$. From our definition (8.24), we obtain

$$h_n^{(n)} = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = 2^n.$$

Our geometrical assertion in this case is that n hyperplanes in general position in n-dimensional space divide n-dimensional space into 2^n regions. Since there are only n $(n-1)$-dimensional hyperplanes, *general position* now means that the n hyperplanes have exactly one point in common. This fact is familiar to all, at least for the cases $k=1,2$, and 3. Consider the case $k=3$ of three-dimensional space. We can coordinatize the space by associating with each point a triple of numbers (x_1, x_2, x_3). The three coordinate planes $x_1 = 0, x_2 = 0$, and $x_3 = 0$ divide the space into $2^3 = 8$ quadrants. (Each quadrant is determined by prescribing signs to each of x_1, x_2, x_3.) More generally, n-dimensional space is coordinatized by associating an n-tuple of numbers $(x_1, x_2, \ldots, x_n)$ with each point. There are n coordinate planes, namely, those determined by $x_1 = 0$, $x_2 = 0, \ldots$, and $x_n = 0$. These planes divide n-dimensional space into the 2^n "quadrants" determined by prescribing signs to each of $x_1, x_2, \ldots, x_n$. One such quadrant is the so-called nonnegative quadrant $x_1 \geq 0, x_2 \geq 0, \ldots, x_n \geq 0$.

8.5 Lattice Paths and Schröder Numbers

In this section, we formalize the notion of a lattice path, which we have experienced in the Exercises in Chapter 2 and in an example in Section 8.1.

We consider the *integral lattice* of points in the coordinate plane with integer coordinates. Given two such points (p, q) and (r, s), with $p \geq r$ and $q \geq s$, a *rectangular* *lattice path* from (r, s) to (p, q) is a path from (r, s) to (p, q) that is made up of *horizontal steps* $H = (1, 0)$ and *vertical steps* $V = (0, 1)$. Thus, a rectangular lattice path from (r, s) to (p, q) starts at (r, s) and gets to (p, q) using unit horizontal and vertical segments.

Example. Figure 8.5 shows a rectangular lattice path from $(0,0)$ to $(7,5)$, consisting of seven horizontal steps (H) and five vertical steps (V). Given that the path starts at $(0,0)$, it is uniquely determined by the sequence

$$H, V, V, H, H, H, V, V, H, V, H, H$$

of seven H's and five V's. □

Theorem 8.5.1 *The number of rectangular lattice paths from* (r, s) *to* (p, q) *equals the binomial coefficient*

$$\binom{p-r+q-s}{p-r} = \binom{p-r+q-s}{q-s}.$$

Proof. The two binomial coefficients in the statement of the theorem are equal. A rectangular lattice path from (r, s) to (p, q) is uniquely determined by its sequence of $p-r$ horizontal steps H and $q-s$ vertical steps V, and every such sequence determines a rectangular lattice path from (r, s) to (p, q). Hence, the number of paths equals the number of permutations of $p-r+q-s$ objects of which $p-r$ are H's and $q-s$ are V's. From Section 3.4 we know this number to be the binomial coefficient

$$\binom{p-r+q-s}{p-r}.$$

□

Consider a rectangular lattice path from (r, s) to (p, q), where $p \geq r$ and $q \geq s$. Such a path uses exactly $(p-r) + (q-s)$ steps, and there is no loss in generality in assuming that $(r, s) = (0, 0)$. This is because we may simply translate (r, s) back to $(0,0)$ and (p, q) back to $(p-r, q-s)$ and obtain a one-to-one correspondence between rectangular lattice paths from (r, s) to (p, q) and those from $(0,0)$ to $(p-r, q-s)$. By Theorem 8.5.1, if $p \geq 0$ and $q \geq 0$, the number of rectangular lattice paths from $(0,0)$ to (p, q) equals

$$\binom{p+q}{p} = \binom{p+q}{q}.$$

We now consider rectangular lattice paths from $(0,0)$ to (p, q) that are restricted to lie on or below the line $y = x$ in the coordinate plane. We call such paths *subdiagonal rectangular lattice paths.* A subdiagonal rectangular lattice path from $(0,0)$ to $(9,9)$ is shown in Figure 8.6.

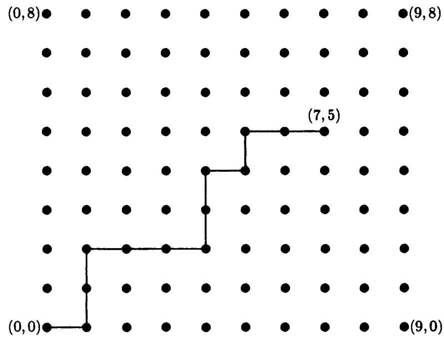


Figure 8.5

In Section 8.1 we proved the next theorem.

Theorem 8.5.2 *Let n be a nonnegative integer. Then the number of subdiagonal rectangular lattice paths from $(0,0)$ to (n,n) equals the nth Catalan number*

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

□

More generally, we can count the number of subdiagonal rectangular lattice paths from $(0,0)$ to (p,q) whenever $p \geq q$. Of course, if $q > p$, there can be no subdiagonal rectangular lattice paths from $(0,0)$ to (p,q), since such a lattice path would have to cross the diagonal.

Theorem 8.5.3 *Let p and q be positive integers with $p \geq q$. Then the number of subdiagonal rectangular lattice paths from $(0,0)$ to (p,q) equals*

$$\frac{p-q+1}{p+1}\binom{p+q}{q}.$$

Proof. For the proof, we generalize the proof given in Section 8.1, which showed that the Catalan number C_n counts the number of subdiagonal rectangular lattice paths from $(0,0)$ to (n,n), and, in particular, the proof of Theorem 8.1.1. To obtain our answer, we determine the number $l(p,q)$ of rectangular lattice paths γ from $(0,0)$ to (p,q) that cross the diagonal, and then subtract $l(p,q)$ from the total number $\binom{p+q}{q}$ of rectangular lattice paths from $(0,0)$ to (p,q). The number $l(p,q)$ is the same as the number of rectangular lattice paths γ' from $(0,-1)$ to $(p,q-1)$ that touch (possibly cross) the diagonal line $y=x$. This follows by shifting paths down one unit, thereby shifting a path γ into a path γ', and this establishes a one-to-one correspondence between the two kinds of paths.

Consider a path γ' from $(0,-1)$ to $(p,q-1)$ that touches the diagonal line $y=x$. Let γ_1' be the subpath of γ' from $(0,-1)$ to the first diagonal point (d,d) touched by γ'. Let γ_2' be the subpath of γ' from (d,d) to $(p,q-1)$. We reflect γ_1' about the line $y=x$ and obtain a path γ_1^* from $(-1,0)$ to (d,d). Following γ_1^* with γ_2, we get a path γ^* from $(-1,0)$ to $(p,q-1)$. This construction is illustrated in Figure 8.7.

Now every rectangular lattice path θ from $(-1,0)$ to $(p,q-1)$ must cross the diagonal line $y=x$, since $(-1,0)$ is above the line and $(p,q-1)$ is below. If we reflect the part of θ that goes from $(-1,0)$ to the first crossing point, we get a path from $(0,-1)$ to $(p,q-1)$ that touches the line $y=x$. This shows that the correspondence γ' to γ^* is a one-to-one correspondence and hence that $l(p,q)$ equals the number of rectangular lattice paths from $(-1,0)$ to $(p,q-1)$. By Theorem 8.5.1, we have

$$l(p,q)=\binom{p+1+q-1}{q-1}=\binom{p+q}{q-1}.$$

Therefore, the number of subdiagonal rectangular lattice paths from $(0,0)$ to (p,q) equals

$$\binom{p+q}{q}-l(p,q)=\binom{p+q}{q}-\binom{p+q}{q-1}=\frac{(p+q)!}{p!q!}-\frac{(p+q)!}{(q-1)!(p+1)!},$$

which simplifies to

$$\frac{p-q+1}{p+1}\binom{p+q}{q}.$$

□

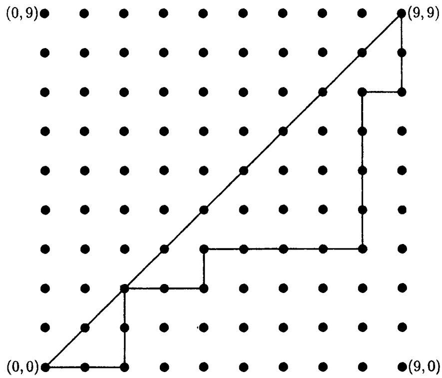


Figure 8.6

We now consider lattice paths where, in addition to horizontal steps $H = (1,0)$ and vertical steps $V = (0,1)$, we allow *diagonal steps* $D = (1,1)$. We call such paths *HVD-lattice paths.* Let p and q be nonnegative integers, and let $K(p,q)$ be the number of HVD-lattice paths from $(0,0)$ to (p,q), and $K(p,q : rD)$ be the number of such paths that use exactly r diagonal steps D. We have $K(p,q : 0D)$ equal to the number of rectangular lattice paths from $(0,0)$ to (p,q); thus, by Theorem 8.5.1,

$$K(p,q : 0D) = \binom{p+q}{p}.$$

We also have $K(p,q : rD) = 0$ if $r > \min\{p,q\}$.

Theorem 8.5.4 *Let $r \leq \min\{p,q\}$. Then*

$$K(p,q : rD) = \binom{p+q-r}{p-r \quad q-r \quad r} = \frac{(p+q-r)!}{(p-r)!(q-r)!r!},$$

and

$$K(p,q) = \sum_{r=0}^{\min\{p,q\}} \frac{(p+q-r)!}{(p-r)!(q-r)!r!}.$$

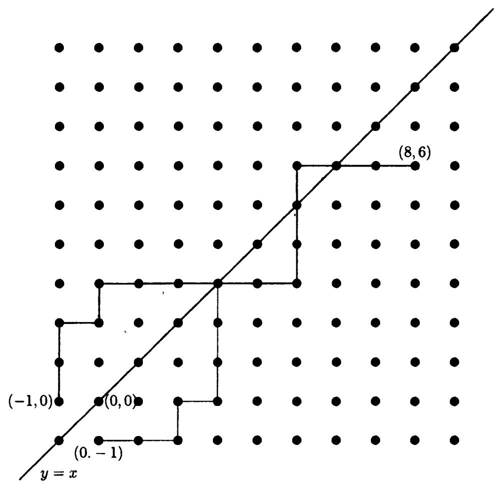


Figure 8.7

Proof. An HVD-lattice path from $(0,0)$ to (p,q) that uses r diagonal steps D must use $p-r$ horizontal steps H and $q-r$ vertical steps V, and is uniquely determined by its sequence of $p-r$ H's, $q-r$ V's, and r D's. Thus, the number of such paths is the number of permutations of the multiset

$$\{(p-r)\cdot H, (q-r)\cdot V, r\cdot D\}.$$

From Chapter 2, we know the number of such permutations to be the multinomial number in the statement of the theorem. If we do not specify the number r of diagonal steps, then by summing $K(p,q:rD)$ from $r=0$ to $r=\min\{p,q\}$, we obtain $K(p,q)$ as given in the theorem. □

Now let $p \geq q$ and let $R(p,q)$ equal the number of subdiagonal HVD-lattice paths from $(0,0)$ to (p,q). Also, let $R(p,q:rD)$ be the number of subdiagonal HVD-lattice paths from $(0,0)$ to (p,q) that use exactly r diagonal steps D. We have

$$R(p,q) = \sum_{r=0}^{q} R(p,q:rD).$$

Theorem 8.5.5 *Let p and q be positive integers with $p \geq q$, and let r be a nonnegative integer with $r \leq q$. Then*

$$\begin{aligned} R(p,q:rD) &= \frac{p-q+1}{p-r+1}\frac{(p+q-r)!}{r!(p-r)!(q-r)!} \\ &= \frac{p-q+1}{p-r+1}\binom{p+q-r}{r \quad (p-r) \quad (q-r)}, \end{aligned}$$

and

$$R(p,q) = \sum_{r=0}^{q} \frac{p-q+1}{p-r+1}\frac{(p+q-r)!}{r!(p-r)!(q-r)!}.$$

Proof. A subdiagonal HVD-lattice path γ from $(0,0)$ to (p,q) with r diagonal steps D becomes a subdiagonal rectangular lattice path π from $(0,0)$ to $(p-r, q-r)$ after removing the r diagonal steps D. Conversely, a subdiagonal rectangular lattice path π from $(0,0)$ to $(p-r,q-r)$ becomes a subdiagonal HVD-lattice path, with r diagonal steps, from $(0,0)$ to (p,q) by inserting r diagonal steps in any of the $p+q-2r+1$ places before, between, and after the horizontal and vertical steps. The number of ways to insert the diagonal steps D in π equals the number of solutions in nonnegative integers of the equation

$$x_1 + x_2 + \cdots + x_{p+q-2r+1} = r,$$

and from Section 3.5, we know this number to be

$$\binom{p+q-2r+1+r-1}{r} = \binom{p+q-r}{r}. \tag{8.28}$$

Thus, to each subdiagonal rectangular lattice path from $(0,0)$ to $(p-r,q-r)$ there correspond a number of subdiagonal HVD-lattice paths from $(0,0)$ to (p,q) with r diagonal steps, and this number is given by (8.28). Therefore,

$$R(p,q:rD) = \binom{p+q-r}{r} R(p-r, q-r:0D).$$

Using Theorem 8.5.3, we get

$$R(p,q:rD) = \binom{p+q-r}{r}\frac{p-q+1}{p-r+1}\binom{p+q-2r}{q-r},$$

which simplifies to

$$\frac{p-q+1}{p-r+1}\frac{(p+q-r)!}{r!(p-r)!(q-r)!} = \frac{p-q+1}{p-r+1}\binom{p+q-r}{r \quad (p-r) \quad (q-r)}.$$

Summing $R(p, q : rD)$ from $r = 0$ to q, we get the formula for $R(p, q)$ given in the theorem. $\square$

Notice that, by taking $r = 0$ in Theorem 8.5.5, we get Theorem 8.5.3.

We now suppose that $p = q = n$. The subdiagonal HVD-lattice paths from $(0, 0)$ to (n, n) are called *Schröder paths.*[21] The *large Schröder number* R_n is the number of Schröder paths from $(0, 0)$ to (n, n). Thus, by Theorem 8.5.5,

$$R_n = R(n, n) = \sum_{r=0}^{n} \frac{1}{n - r + 1} \frac{(2n - r)!}{r!((n - r)!)^2}.$$

The sequence $R_0, R_1, R_2, \ldots, R_n, \ldots$ of large Schröder numbers begins as

$$1, 2, 6, 22, 90, 394, 1806, \ldots.$$

We now turn to the small Schröder numbers, which are defined in terms of constructs called bracketings. Let $n \geq 1$, and let $a_1, a_2, \ldots, a_n$ be a sequence of n symbols. We generalize the idea of a multiplication scheme for $a_1, a_2, \ldots, a_n$ described in Section 8.2 to that of a bracketing of the sequence $a_1, a_2, \ldots, a_n$. For our multiplication schemes, we had a binary operation $\times$ that combined two quantities, and a multiplication scheme was a way to put $n - 1$ sets of parentheses on the sequence $a_1, a_2, \ldots, a_n$, with each set of parentheses corresponding to a multiplication of two quantities. In a bracketing, a set of parentheses can enclose any number of symbols. For clarity, we shall now drop the symbol $\times$ since its use now introduces some ambiguity. Before giving the formal definition of bracketing, we list the bracketings for $n = 1, 2, 3,$ and 4 and, at the same time, introduce some of the simplifications we adopt for purposes of clarity.

Example. If $n = 1$, then there is only one bracketing, namely, a_1. To be precise, we should write this as (a_1) but, also for clarity, we shall remove parentheses around single elements and let the parentheses be implicit. For $n = 2$, there is also only one bracketing, namely, $(a_1 a_2)$, or, for more clarity, $a_1 a_2$. In general, we omit the last set of parentheses corresponding to the final bracketing of the remaining symbols. For $n = 3$, we have three bracketings:

$$a_1 a_2 a_3, (a_1 a_2) a_3, \text{ and } a_1 (a_2 a_3).$$[22]

[21]After Friedrich Wilhelm Karl Ernst Schröder (1841–1902). See R. P. Stanley, Hipparchus, Plutarch, Schröder, and Hough, *American Mathematical Monthly*, 104 (1997), 344–350. Also see L. W. Shapiro and R. A. Sulanke, Bijections for Schröder Numbers, *Mathematics Magazine*, 73 (2000), 369–376. We rely heavily on both of these articles for this section.

[22]Without any of our simplifications, these would be written as $(a_1 \times a_2 \times a_3), ((a_1 \times a_2) \times a_3)$, and $(a_1 \times (a_2 \times a_3))$. The last two are multiplication schemes, since each pair of parentheses in them corresponds to a multiplication of two quantities, but the first is not.

For $n = 4$, we have 11 bracketings:

$$a_1a_2a_3a_4, (a_1a_2)a_3a_4, (a_1a_2a_3)a_4, a_1(a_2a_3)a_4, a_1(a_2a_3a_4), a_1a_2(a_3a_4),$$

and

$$((a_1a_2)a_3)a_4, (a_1(a_2a_3))a_4, a_1((a_2a_3)a_4), a_1(a_2(a_3a_4)), (a_1a_2)(a_3a_4).$$ [23]

□

We now give the formal recursive definition of a *bracketing* of a sequence $a_1, a_2, \ldots, a_n$ Each symbol a_i is itself a bracketing; and any consecutive sequence of two or more bracketings enclosed by a set of parentheses is a bracketing. Thus, in contrast to multiplication schemes in Section 8.2, a pair of parentheses need not correspond to a multiplication of two symbols. Using this definition, we can construct all bracketings of the sequence $a_1, a_2, \ldots, a_n$ by carrying out the following recursive algorithm in all possible ways.

Algorithm to Construct Bracketings

Start with a sequence $a_1, a_2, \ldots, a_n$.

1. Let γ equal $a_1a_2 \ldots a_n$.
2. While γ has at least three symbols, do the following:
   (a) Put a set of parentheses around any number $k \geq 2$ of consecutive symbols, say, $a_ia_{i+1} \cdots a_{i+k-1}$, to form a new symbol $(a_ia_{i+1} \cdots a_{i+k-1})$.
   (b) Replace γ with the expression in which $(a_ia_{i+1} \cdots a_{i+k-1})$ is one symbol.[24]
3. Output the current expression.

A multiplication scheme for $a_1, a_2, \ldots, a_n$ is a *binary bracketing*—that is, a bracketing in which each set of parentheses encloses two symbols.

Example. We give an example of an application of the algorithm. Let $n = 9$ so that we start with $a_1a_2a_3a_4a_5a_6a_7a_8a_9$. We arrive at a bracketing by making the following choices:

$$\begin{aligned} a_1a_2a_3a_4a_5a_6a_7a_8a_9 &\rightarrow a_1a_2a_3(a_4a_5a_6)a_7a_8a_9 \\ &\rightarrow (a_1a_2)a_3(a_4a_5a_6)a_7a_8a_9 \\ &\rightarrow (a_1a_2)a_3((a_4a_5a_6)a_7a_8)a_9 \\ &\rightarrow (a_1a_2)(a_3((a_4a_5a_6)a_7a_8)a_9). \end{aligned}$$

[23]Only the last five are multiplication schemes.

[24]But recall that, if we choose the entire sequence of symbols, we don't put in parentheses. Since $k \geq 2$, we don't put a set of parentheses around one symbol.

This bracketing is not a binary bracketing, since there are sets of parenthesis which enclose more than two symbols; for instance, $(a_4a_5a_6)$ does, and so does $((a_4a_5a_6)a_7a_8)$ (which encloses the three symbols $(a_4a_5a_6)$, a_7, and a_8), and $(a_3((a_4a_5a_6)a_7a_8)a_9)$ (which encloses the three symbols a_3, $((a_4a_5a_6)a_7a_8)$, and a_9) . $\square$

For $n \geq 1$, the *small Schröder number* s_n is defined to be the number of bracketings of a sequence $a_1, a_2, \ldots, a_n$ of n symbols. We have seen that $s_1 = 1$, $s_2 = 1, s_3 = 3$ and $s_4 = 11$. In fact, the sequence $(s_n : n = 1, 2, 3, \ldots)$ begins as

$$1, 1, 3, 11, 45, 197, 903, \ldots.$$

Comparing this with the initial part of the sequence of large Schröder numbers leads us to the tentative conclusion that $R_n = 2s_{n+1}$ for $n \geq 1$ with $R_0 = 1$. We give a proof of this by computing the generating functions for both the small and large Schröder numbers.

Theorem 8.5.6 *The generating function for the sequence* $(s_n : n \geq 1)$ *of small Schröder numbers is*

$$\sum_{n=1}^{\infty} s_n x^n = \frac{1}{4}\left(1 + x - \sqrt{x^2 - 6x + 1}\right).$$

Proof. Let $g(x) = \sum_{n=1}^{\infty} s_n x^n$ be the generating function of the small Schröder numbers. The recursive definition of bracketing implies that

$$\begin{aligned} g(x) &= x + g(x)^2 + g(x)^3 + g(x)^4 + \cdots \\ &= x + g(x)^2(1 + g(x) + g(x)^2 + \cdots) \\ &= x + \frac{g(x)^2}{1 - g(x)}. \end{aligned}$$

This gives

$$(1 - g(x))g(x) = (1 - g(x))x + g(x)^2;$$

hence,

$$2g(x)^2 - (1 + x)g(x) + x = 0.$$

Therefore, $g(x)$ is a solution of the quadratic equation

$$2y^2 - (1 + x)y + x = 0.$$

The two solutions of this quadratic equation are

$$y_1(x) = \frac{(1 + x) + \sqrt{(1 + x)^2 - 8x}}{4}$$

and

$$y_2(x) = \frac{(1+x) - \sqrt{(1+x)^2 - 8x}}{4}.$$

Since $g(0) = 0$, and $y_1(0) = 1/2$ and $y_2(0) = 0$, we have

$$g(x) = y_2(x) = \frac{1 + x - \sqrt{x^2 - 6x + 1}}{4}.$$

□

The generating function $g(x) = \sum_{n=1}^{\infty} s_n x^n$, as evaluated in Theorem 8.5.6, can be used to obtain a recurrence relation for the small Schröder numbers that is useful for computation. We return to the quadratic equation

$$2y^2 - (1+x)y + x = 0 \tag{8.29}$$

that arose in the proof of Theorem 8.5.6. If we differentiate each side of this quadratic equation with respect to x,[25] we get

$$4y\frac{dy}{dx} - y - (1+x)\frac{dy}{dx} + 1 = 0;$$

hence,

$$\begin{aligned} \frac{dy}{dx} &= \frac{y-1}{4y-1-x} \\ &= \frac{(x-3)y - x + 1}{x^2 - 6x + 1}. \end{aligned}$$

The last equality can be routinely verified by cross multiplying and then making use of (8.29). We now have

$$(x^2 - 6x + 1)\frac{dy}{dx} - (x-3)y + x - 1 = 0. \tag{8.30}$$

Substituting $y = g(x) = \sum_{n=1}^{\infty} s_n x^n$ in (8.30), we get, after some simplification,

$$\sum_{n=1}^{\infty}(n-1)s_n x^{n+1} - 3\sum_{n=1}^{\infty}(2n-1)s_n x^n + \sum_{n=1}^{\infty} n s_n x^{n-1} + x - 1 = 0,$$

which can be rewritten as

$$\sum_{n=1}^{\infty}(n-1)s_n x^{n+1} - 3\sum_{n=0}^{\infty}(2n+1)s_{n+1}x^{n+1} +$$

[25] Keep in mind that y is a function of x.

$$\sum_{n=-1}^{\infty}(n+2)s_{n+2}x^{n+1} = -x+1.$$

The coefficient of x^{n+1} in the expression on the left equals 0 for $n \geq 1$, we obtain

$$(n+2)s_{n+2} - 3(2n+1)s_{n+1} + (n-1)s_n = 0, \quad (n \geq 1). \tag{8.31}$$

The recurrence relation (8.31) is a homogeneous linear recurrence relation of order 2 with nonconstant coefficients.

We now return to the large Schröder numbers and, in the next theorem, compute their generating function.

Theorem 8.5.7 *The generating function for the sequence* $(R_n : n \geq 0)$ *of large Schröder numbers is*

$$\sum_{n=0}^{\infty} R_n x^n = \frac{1}{2x}\left(-(x-1) - \sqrt{x^2-6x+1}\right).$$

Proof. Let $h(x) = \sum_{n=0}^{\infty} R_n x^n$ be the generating function for the large Schröder numbers. A subdiagonal HVD-lattice path from $(0,0)$ to (n,n)

(1) is the empty path (if $n = 0$),

(2) starts with a diagonal step D, or

(3) starts with a horizontal step H.

The number of paths of type (2) equals the number of subdiagonal HVD-lattice paths from $(1,1)$ to (n,n) and thus equals R_{n-1}. The paths of type (3) begin with a horizontal step H and then follow a path γ from $(1,0)$ to (n,n) without going above the diagonal line joining $(1,1)$ and (n,n). Since γ ends on the diagonal at the point (n,n), there is a first point (k,k) of γ on the diagonal, where $1 \leq k \leq n$. Since (k,k) is the first point of γ on the diagonal, γ arrives at (k,k) by a vertical step V from the point $(k,k-1)$. The part of γ from $(1,0)$ to $(k,k-1)$ is a lattice path γ_1 that does not go above the diagonal line joining $(1,0)$ to $(k,k-1)$. The part of γ from (k,k) to (n,n) is a lattice path γ_2 that does not go above the diagonal line joining (k,k) to (n,n). There are R_{k-1} choices for γ_1 and R_{n-k} choices for γ_2, and hence the number of lattice paths of type (iii) equals $R_{k-1}R_{n-k}$. Summarizing, we get the recurrence relation

$$R_n = R_{n-1} + \sum_{k=1}^{n} R_{k-1}R_{n-k}, \quad (n \geq 1),$$

or, equivalently,

$$R_n = R_{n-1} + \sum_{k=0}^{n-1} R_k R_{n-1-k}, \quad (n \geq 1), \tag{8.32}$$

where $R_0 = 1$. Thus,

$$x^n R_n = x(x^{n-1}R_{n-1}) + x\left(\sum_{k=0}^{n-1} x^k R_k x^{n-1-k} R_{n-1-k}\right), \quad (n \geq 1).$$

Since $R_0 = 1$, the preceding equation implies that the generating function $h(x)$ of the large Schröder numbers satisfies

$$h(x) = 1 + xh(x) + xh(x)^2.$$

Therefore, $h(x)$ is a solution of the quadratic equation

$$xy^2 + (x-1)y + 1 = 0.$$

The two solutions of this quadratic equation are

$$y_1(x) = \frac{-(x-1) + \sqrt{x^2 - 6x + 1}}{2x}$$

and

$$y_2(x) = \frac{-(x-1) - \sqrt{x^2 - 6x + 1}}{2x}.$$

The first of these cannot be the generating function of the large Schröder numbers as it does not give nonnegative integers. Hence,

$$h(x) = y_2(x) = \frac{1 - x - \sqrt{x^2 - 6x + 1}}{2x}.$$

□

Comparing the generating functions for the large and small Schröder numbers, we obtain the following corollary.

Corollary 8.5.8 *The large and small Schröder numbers are related by*

$$R_n = 2s_{n+1}, \quad (n \geq 1).$$

□

In Sections 7.6 and 8.1, we considered triangulating a convex polygonal region by means of its diagonals which do not intersect in the interior of the region. We showed that the number of such triangularizations of a convex polygonal region with $n+1$ sides equals the number of multiplication schemes for n numbers given in a particular order, with the common value equal to the Catalan number

$$C_{n-1} = \frac{1}{n}\binom{2n-2}{n-1}.$$

Thus, the nth Catalan number C_n equals the number of triangularizations of a convex polygonal region with $n+2$ sides. We conclude this section by showing that bracketings can be given a combinatorial geometric interpretation.

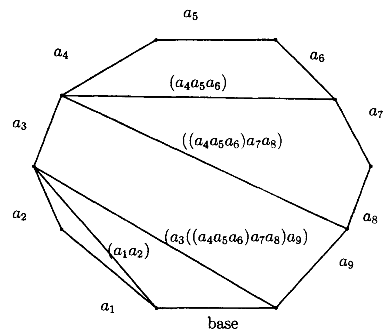


Figure 8.8

Consider a convex polygonal region Π_{n+1} with $n+1$ sides, and the sequence $a_1, a_2, \ldots, a_n$. The base of Π_{n+1} is labeled as base, and the remaining n sides are labeled with $a_1, a_2, \ldots, a_n$, beginning with the side immediately to the left of the base being labeled a_1 and proceeding in order in a clockwise fashion. Bracketings of $a_1, a_2, \ldots, a_n$ are in one-to-one correspondence with *dissections* of Π_{n+1}, where, by a dissection of Π_{n+1}, we mean a partition of Π_{n+1} into regions obtained by inserting diagonals that do not intersect in the interior. In contrast to triangularizations, the regions in the partition of Π_{n+1} are not restricted to be triangles.

We illustrate the correspondence in Figure 8.8, using the example of a bracketing that we constructed with our algorithm:

$$\begin{aligned}
a_1a_2a_3a_4a_5a_6a_7a_8a_9 &\rightarrow a_1a_2a_3(a_4a_5a_6)a_7a_8a_9\\
&\rightarrow (a_1a_2)a_3(a_4a_5a_6)a_7a_8a_9\\
&\rightarrow (a_1a_2)a_3((a_4a_5a_6)a_7a_8)a_9\\
&\rightarrow (a_1a_2)(a_3((a_4a_5a_6)a_7a_8)a_9).
\end{aligned}$$

This correspondence works in general, establishes a one-to-one correspondence between bracketings and dissections, and also proves the next theorem. We adopt the convention that a polygonal region with two sides is a line segment and that it has exactly one dissection (the empty dissection).

Theorem 8.5.9 *Let n be a positive integer. Then the number of dissections of a convex polygonal region of $n+1$ sides equals the small Schröder number s_n.* $\square$

In terms of the polygonal region Π_{n+1}, our algorithm for constructing a bracketing of a sequence of n symbols is both natural and obvious.

Algorithm to construct dissections of Π_{n+1}

Start with the convex polygonal region Π_{n+1}, with the sides labeled as:

$$\text{base}, a_1, a_2, \ldots, a_n,$$

in a clockwise fashion.

1. Let $\Gamma = \Pi_{n+1}$.
   (a) While Γ has three or more sides, insert a diagonal of Γ, thereby partitioning Γ into two parts. (Here we allow the base to be chosen as the diagonal in which case the two parts are Γ and the polygonal region of two sides given by the base.)
   (b) Replace Γ with the part containing the base. (This part will have at least one fewer side and is the base itself if the base was chosen in (a).)
3. Output the full dissected polygonal region Π_{n+1}.

The algorithm comes to an end when the base has been chosen as the diagonal, and Γ is then replaced by the polygonal region of two sides given by the base.

8.6 Exercises

1. Let $2n$ (equally spaced) points on a circle be chosen. Show that the number of ways to join these points in pairs, so that the resulting n line segments do not intersect, equals the nth Catalan number C_n.

2. Prove that the number of 2-by-n arrays

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \end{bmatrix}$$

that can be made from the numbers $1, 2 \ldots, 2n$ such that

$$x_{11} < x_{12} < \cdots < x_{1n},$$

$$x_{21} < x_{22} < \cdots < x_{2n}$$

$$x_{11} < x_{21}, x_{12} < x_{22}, \ldots, x_{1n} < x_{2n},$$

equals the nth Catalan number, C_n.

3. Write out all of the multiplication schemes for four numbers and the triangularization of a convex polygonal region of five sides corresponding to them.

4. Determine the triangularization of a convex polygonal region corresponding to the following multiplication schemes:

   (a) $(a_1 \times (((a_2 \times a_3) \times (a_4 \times a_5)) \times a_6))$

   (b) $(((a_1 \times a_2) \times (a_3 \times (a_4 \times a_5))) \times ((a_6 \times a_7) \times a_8))$

5. $*$ Let m and n be nonnegative integers with $n \geq m$. There are $m+n$ people in line to get into a theater for which admission is 50 cents. Of the $m+n$ people, n have a 50-cent piece and m have a \$1 dollar bill. The box office opens with an empty cash register. Show that the number of ways the people can line up so that change is available when needed is

$$\frac{n-m+1}{n+1}\binom{m+n}{m}.$$

   (The case $m = n$ is the case treated in Section 8.1.)

6. Let the sequence $h_0, h_1, \ldots, h_n, \ldots$ be defined by $h_n = 2n^2 - n + 3$, $(n \geq 0)$. Determine the difference table, and find a formula for $\sum_{k=0}^{n} h_k$.

7. The general term h_n of a sequence is a polynomial in n of degree 3. If the first four entries of the 0th row of its difference table are $1, -1, 3, 10$, determine h_n and a formula for $\sum_{k=0}^{n} h_k$.

8. Find the sum of the fifth powers of the first n positive integers.

9. Prove that the following formula holds for the kth-order differences of a sequence $h_0, h_1, \ldots, h_n, \ldots$:

$$\Delta^k h_n = \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} h_{n+j}.$$

10. If h_n is a polynomial in n of degree m, prove that the constants $c_0, c_1, \ldots, c_m$ such that

$$h_n = c_0\binom{n}{0} + c_1\binom{n}{1} + \cdots + c_m\binom{n}{m}$$

    are uniquely determined. (Cf. Theorem 8.2.2.)

11. Compute the Stirling numbers of the second kind $S(8, k)$, $(k = 0, 1, \ldots, 8)$.

12. Prove that the Stirling numbers of the second kind satisfy the following relations:

    (a) $S(n, 1) = 1, \quad (n \geq 1)$

(b) $S(n,2) = 2^{n-1} - 1, \quad (n \geq 2)$

(c) $S(n,n-1) = \binom{n}{2}, \quad (n \geq 1)$

(d) $S(n,n-2) = \binom{n}{3} + 3\binom{n}{4} \quad (n \geq 2)$

13. Let X be a p-element set and let Y be a k-element set. Prove that the number of functions $f : X \to Y$ which map X *onto* Y equals

$$k!S(p,k) = S^{\#}(p,k).$$

14. $*$ Find and verify a general formula for

$$\sum_{k=0}^{n} k^p$$

involving Stirling numbers of the second kind.

15. The number of partitions of a set of n elements into k distinguishable boxes (some of which may be empty) is k^n. By counting in a different way, prove that

$$k^n = \binom{k}{1}1!S(n,1) + \binom{k}{2}2!S(n,2) + \cdots + \binom{k}{n}n!S(n,n).$$

(If $k > n$, define $S(n,k)$ to be 0.)

16. Compute the Bell number B_8. (Cf. Exercise 11.)

17. Compute the triangle of Stirling numbers of the first kind $s(n,k)$ up to $n = 7$.

18. Write $[n]_k$ as a polynomial in n for $k = 5, 6,$ and 7.

19. Prove that the Stirling numbers of the first kind satisfy the following formulas:

(a) $s(n,1) = (n-1)!, \quad (n \geq 1)$

(b) $s(n,n-1) = \binom{n}{2}, \quad (n \geq 1)$

20. Verify that $[n]_n = n!$, and write $n!$ as a polynomial in n using the Stirling numbers of the first kind. Do this explicitly for $n = 6$.

21. For each integer $n = 1,2,3,4,5$, construct the diagram of the set $\mathcal{P}_n$ of partitions of n, partially ordered by majorization.

22. (a) Calculate the partition number p_6 and construct the diagram of the set $\mathcal{P}_6$, partially ordered by majorization.

    (b) Calculate the partition number p_7 and construct the diagram of the set $\mathcal{P}_7$, partially ordered by majorization.

23. A total order on a finite set has a unique maximal element (a largest element) and a unique minimal element (a smallest element). What are the largest partition and smallest partition in the lexicographic order on $\mathcal{P}_n$ (a total order)?

24. A partial order on a finite set may have many maximal elements and minimal elements. In the set $\mathcal{P}_n$ of partitions of n partially ordered by majorization, prove that there is a unique maximal element and a unique minimal element.

25. Let $t_1, t_2, \ldots, t_m$ be distinct positive integers, and let

$$q_n = q_n(t_1, t_2, \ldots, t_m)$$

equal the number of partitions of n in which all parts are taken from $t_1, t_2, \ldots, t_m$. Define $q_0 = 1$. Show that the generating function for $q_0, q_1, \ldots, q_n, \ldots$ is

$$\prod_{k=1}^{m}(1 - x^{t_k})^{-1}.$$

26. Determine the conjugate of each of the following partitions:

    (a) $12 = 5 + 4 + 2 + 1$

    (b) $15 = 6 + 4 + 3 + 1 + 1$

    (c) $20 = 6 + 6 + 4 + 4$

    (d) $21 = 6 + 5 + 4 + 3 + 2 + 1$

    (e) $29 = 8 + 6 + 6 + 4 + 3 + 2$

27. For each integer $n > 2$, determine a self-conjugate partition of n that has at least two parts.

28. Prove that conjugation reverses the order of majorization; that is, if λ and μ are partitions of n and λ is majorized by μ, then μ^* is majorized by λ^*.

29. Prove that the number of partitions of the positive integer n into parts each of which is at most 2 equals $\lfloor n/2 \rfloor + 1$. (Remark: There is a formula, namely the nearest integer to $\frac{(n+3)^2}{12}$, for the number of partitions of n into parts each of which is at most 3 but it is much more difficult to prove. There is also one for partitions with no part more than 4, but it is even more complicated and difficult to prove.)

30. Prove that the partition function satisfies

$$p_n > p_{n-1} \quad (n \geq 2).$$

31. Evaluate $h_{k-1}^{(k)}$, the number of regions into which k-dimensional space is partitioned by $k-1$ hyperplanes in general position.

32. Use the recurrence relation (8.31) to compute the small Schröder numbers s_8 and s_9.

33. Use the recurrence relation (8.32) to compute the large Schröder numbers R_7 and R_8. Verify that $R_7 = 2s_8$ and $R_8 = 2s_9$, as stated in Corollary 8.5.8.

34. Use the generating function for the large Schröder numbers to compute the first few large Schröder numbers.

35. Use the generating function for the small Schröder numbers to compute the first few small Schröder numbers.

36. Prove that the Catalan number C_n equals the number of lattice paths from $(0,0)$ to $(2n,0)$ using only upsteps $(1,1)$ and downsteps $(1,-1)$ that never go above the horizontal axis (so there are as many upsteps as there are downsteps). (These are sometimes called *Dyck paths.*)

37. * The large Schröder number C_n counts the number of subdiagonal HVD-lattice paths from $(0,0)$ to (n,n). The small Schröder number counts the number of dissections of a convex polygonal region of $n+1$. Since $R_n = 2s_{n+1}$ for $n \geq 1$, there are as many subdiagonal HVD-lattice paths from $(0,0)$ to (n,n) as there are dissections of a convex polygonal region of $n+1$ sides. Find a one-to-one correspondence between these lattice paths and these dissections.