P.35

3

Chapter 2

Generating Functions

Do not pray for tasks equal to your powers. Pray for powers equal to your tasks.

Twenty Sermons, PHILLIPS BROOKS

Generating functions provide an algebraic machinery for solving combinatorial problems. The usual algebraic operations (convolution, especially) facilitate considerably not only the computational aspects but also the thinking processes involved in finding satisfactory solutions. More often than not we remain blissfully unaware of this disinterested service, until trying to reproduce the same by direct calculations (a task usually accompanied by no insignificant mental strain). The main reason for introducing formal power series is the ability to translate key combinatorial operations into algebraic ones that are, in turn, easily and routinely performed within a set (usually an algebra) of generating functions. Generally this is much easier said than done, for it takes great skill to establish such a happy interplay. Yet notable examples exist, and we examine a couple of better known ones in considerable detail.

We begin by introducing the ordinary and exponential generating functions. Upon closely investigating the combinatorial meaning of the operation of convolution in these two well-known cases, we turn to specific generating functions associated with the Stirling and Lah numbers. The latter part of the chapter touches briefly upon the uses of formal power series to recurrence relations and introduces the Bell polynomials, in connection with Faa DiBruno's formula, for explicitly computing the higher order derivatives of a composition of two functions. In ending the chapter we dote upon subjects such as Kirchhoff's tree generating matrix (along with applications to statistical design), partitions of an integer, and a generating function for solutions to Diophantine systems of linear equations in nonnegative integers.

1 THE FORMAL POWER SERIES

2.1

The *generating function* of the sequence (a_n) is the (formal) power series $A(x) = \sum_n a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots$. The summation sign always starts at 0 and extends to infinity in steps of one. By x we understand an indeterminate.

Most of the time we view generating functions as formal power series. Occasionally, however, questions of convergence may arise and the analytic techniques would then come to play an important role. We recall for convenience that two formal power series are equal if (and only if) the coefficients of the corresponding powers of x are equal.

By writing $(a_n) \leftrightarrow A(x)$ we indicate the bijective association between the sequence (a_n) and its generating function $A(x)$. In terms of this association we observe that if $(a_n) \leftrightarrow A(x)$, $(b_n) \leftrightarrow B(x)$, and c is a constant, then

$$\begin{aligned}(a_n + b_n) &\leftrightarrow A(x) + B(x)\\ (ca_n) &\leftrightarrow cA(x)\end{aligned}$$

and, most importantly, multiplication by convolution

$$\left(\sum_{i=0}^{n} a_i b_{n-i}\right) \leftrightarrow A(x)B(x).$$

(The set of generating functions endowed with these operations is said to form an algebra.)

Generating functions A and B are said to be *inverses* of each other if $A(x)B(x) = 1 = B(x)A(x)$. This last relation we sometimes write as $B = A^{-1}$, $B = 1/A$, $A = B^{-1}$, or $A = 1/B$. Note, for example, that $A(x) = 1 - x$ and $B(x) = \sum_n x^n$ are a pair of inverses.

An important operation with power series is that of *composition* (or substitution). By $A \circ B$ we understand the series defined as follows: $(A \circ B)(x) = A(B(x))$. More explicitly still, if $A(x) = \sum_n a_n x^n$ and $B(x) = \sum_n b_n x^n$, then $(A \circ B)(x) = A(B(x)) = \sum_n a_n (B(x))^n$. In order that $A(B(x))$ be a well-defined power series, the original series A and B need be such that the coefficient of each power of x in $A(B(x))$ is obtained as a sum of finitely many terms. [Thus if $A(x) = \sum_n x^n$ and $B(x) = x + x^2$, $A(B(x))$ is well defined, but if $B(x) = 1 + x$, then $A(B(x))$ is not well defined. In the latter case the constant term of $A(B(x))$ involves the summation of infinitely many 1's.] We can see therefore that $A(B(x))$ makes sense essentially under two conditions: when $A(x)$ has infinitely many nonzero coefficients then the constant term in $B(x)$ must be 0, and if $A(x)$ has only finitely many nonzero coefficients [i.e., if $A(x)$ is a polynomial], then $B(x)$ can be arbitrary. Whenever well defined, the series $A \circ B$ is called the composition of A with B (or the substitution of B into A).

We also let the linear operator D (of *formal differentiation*) act upon a generating function A as follows:

$$DA(x) = D\left(\sum_n a_n x^n\right) \underset{\text{def.}}{=} \sum_n (n+1)a_{n+1}x^n.$$

As an example, let $A(x) = 2 - 5x + 3x^2$ and $B(x) = \sum_n (n+1)^{-1}x^n$. The reader may quickly verify that

$$A(x)B(x) = 2 - 4x + \sum_{n=2}^{\infty}(n+5)n^{-1}(n^2-1)^{-1}x^n.$$

Applying the differential operator D to A, B, and AB respectively, we obtain:

$$DA(x) = -5 + 3 \cdot 2x, \quad DB(x) = \sum_n (n+1)(n+2)^{-1}x^n$$

and

$$D(A(x)B(x)) = -4 + \sum_{n=2}^{\infty}(n+5)(n^2-1)^{-1}x^{n-1}.$$

In closing, let us mention that the operator of formal differentiation satisfies the familiar rules of differentiation:

$$\begin{aligned} D(AB) &= (DA)B + A(DB) \\ DA^{-1} &= -A^{-2}DA, \end{aligned}$$

and most importantly, the "chain rule,"

$$D(A \circ B) = ((DA) \circ B)DB.$$

2.2

The *exponential generating function* of the sequence (a_n) is the (formal) power series

$$E(x) = \sum_n a_n \frac{x^n}{n!} = a_0 + a_1 \frac{x}{1!} + a_2 \frac{x^2}{2!} + \cdots + a_n \frac{x^n}{n!} + \cdots .$$

In as much as the exponential generating functions are concerned, if $(a_n) \leftrightarrow E(x)$, $(b_n) \leftrightarrow F(x)$, and c is a constant, then

$$\begin{aligned} (a_n + b_n) &\leftrightarrow E(x) + F(x) \\ (ca_n) &\leftrightarrow cE(x) \end{aligned}$$

and

$$\left(\sum_{i=0}^{n} \binom{n}{i} a_i b_{n-i} \right) \leftrightarrow E(x)F(x).$$

In this case we say that the multiplication of two exponential generating functions corresponds to the binomial convolution of sequences.

As before, we call E and F inverses if $E(x)F(x) = 1 = F(x)E(x)$.

The operator D of formal differentiation acts here as follows:

$$DE(x) = D\left(\sum_n a_n \frac{x^n}{n!} \right) \underset{\text{def.}}{=} \sum_n a_{n+1} \frac{x^n}{n!}.$$

We illustrate the multiplication of two exponential generating functions by a simple example:

$$\begin{aligned} \left(\sum_n 3^n \frac{x^n}{n!} \right) \left(\sum_n \frac{1}{2^n} \frac{x^n}{n!} \right) &= \sum_n \left(\sum_{i=0}^{n} \binom{n}{i} 3^i \frac{1}{2^{n-i}} \right) \frac{x^n}{n!} \\ &= \sum_n \left(\sum_{i=0}^{n} \binom{n}{i} 3^i \left(\frac{1}{2} \right)^{n-i} \right) \frac{x^n}{n!} = \sum_n \left(3 + \frac{1}{2} \right)^n \frac{x^n}{n!} \\ &= \sum_n \left(\frac{7}{2} \right)^n \frac{x^n}{n!}. \end{aligned}$$

The next to the last equality sign is explained by the fact that $\sum_{i=0}^{n}\binom{n}{i}a^i b^{n-i} = (a+b)^n$, where a and b are two entities that commute (such as 3 and $\frac{1}{2}$).

With regard to differentiation,

$$D\left(\sum_n 3^n \frac{x^n}{n!}\right) = \sum_n 3^{n+1}\frac{x^n}{n!}.$$

2.3

The vector space of sequences can be made into an algebra by defining a multiplication of two sequences. We require the rule of multiplication to be "compatible" with the rules of addition and scalar multiplication. Two such rules of multiplication have been described in Sections 2.1 and 2.2. Other rules could be conceived, but one wonders of how much use in combinatorial counting they would be. One well-known multiplication, of interest to number theorists, is as follows:

$$(a_n)(b_n) = \left(\sum_{\substack{d \\ dm=n}} a_d b_m\right)$$

and is called the *Dirichlet convolution.* In this case we attach the formal Dirithlet series $\sum_n (a_n/n^x)$ to the sequence (a_n).

Eulerian generating functions are known to be helpful in enumeration problems over finite vector spaces and with inversion problems in sequences. The *Eulerian series* of the sequence (a_n) is defined as

$$E_q(x) = \sum_n \frac{a_n x^n}{(1-q)(1-q^2)\cdots(1-q^n)}.$$

We briefly discuss these series in Chapter 3.

Let us now make ourselves more aware of what combinatorial operations the generating functions and the exponential generating functions perform for us.

2 THE COMBINATORIAL MEANING OF CONVOLUTION

2.4

In Section 1.2 we established bijective correspondences between the three general problems listed below and showed that they all admit the same numerical solution:

(a) *The number of ways to distribute n indistinguishable balls into m distinguishable boxes is* $\binom{n+m-1}{n}$.

(b) *The number of vectors* $(n_1, n_2, \ldots, n_m)$ *with nonnegative integer entries satisfying*

$$n_1 + n_2 + \cdots + n_m = n$$

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

(c) *The number of ways to select n objects with repetition from m different types of objects is* $\binom{n+m-1}{n}$. (We assume that we have an unlimited supply of objects of each type and that the order of selection of the n objects is irrelevant.)

The three problems just mentioned consociate well to the operation of convolution with generating functions. Specifically, let us explain how we attach combinatorial meaning to the multiplication by convolution of several generating functions with coefficients 0 or 1:

1. *The number of ways of placing n indistinguishable balls into m distinguishable boxes is the coefficient of x^n in*

$$(1 + x + x^2 + \cdots)^m = \left(\sum_k x^k\right)^m = (1-x)^{-m}.$$

Indeed, we can describe the possible contents of our boxes as follows:

$$\begin{array}{ccccc} \text{Box 1} & \text{Box 2} & \text{Box 3} & \cdots & \text{Box } m \\ \hline 1 & 1 & 1 & & 1 \\ x & x & x & & x \\ x^2 & x^2 & x^2 & & x^2 \\ x^3 & x^3 & x^3 & & x^3 \\ \vdots & \vdots & \vdots & & \vdots \end{array} \qquad (*)$$

The symbol x^i beneath box j indicates the fact that we may place i balls in box j. Think of m (the number of boxes) being fixed, but keep n unspecified. With this in mind we can assume that the columns beneath the boxes are of infinite length. How do we then obtain the coefficient of x^n in the product $(1+x+x^2+x^3+\cdots)^m$? We select x^{n_1} from column 1 of $(*)$, x^{n_2} from column 2, $\ldots$, x^{n_m} from column m such that $x^{n_1}x^{n_2}\cdots x^{n_m} = x^n$, and do this in all possible ways. The number of such ways clearly equals the number of vectors $(n_1, n_2, \ldots, n_m)$ satisfying

$$\sum_{i=1}^{m} n_i = n,$$

with $0 \leq n_i$, n_i integers; $1 \leq i \leq m$. By (b) above we conclude that there are precisely $\binom{n+m-1}{n}$ solutions, which is also in agreement with (a), thus proving our statement.

In terms of generating functions, this shows that

$$\left(\sum_k x^k\right)^m = \sum_n \binom{n+m-1}{n} x^n. \tag{2.1}$$

By observing that $(1-x)^{-1} = \sum_n x^n$ we can rewrite relation (2.1) as follows:

$$(1-x)^{-m} = \sum_n \binom{n+m-1}{n} x^n. \tag{2.2}$$

2. *The number of ways of placing n indistinguishable objects into m distinguishable boxes*

with at most r_i objects in box i is the coefficient of x^n in

$$\prod_{i=1}^{m}(1 + x + x^2 + \cdots + x^{r_i}).$$

The contents of our m boxes is now as follows:

Box 1	Box 2	Box 3	$\cdots$	Box m
1	1	1		1
x	x	x		x
x^2	x^2	x^2		x^2
$\vdots$	$\vdots$	$\vdots$		$\vdots$
x^{r_1}	x^{r_2}	x^{r_3}		x^{r_m}

Again, the coefficient of x^n is the number of selections of powers of x (one from each column) such that the sum of these powers is n. To be more precise, the coefficient of x^n is the number of all vectors $(n_1, n_2, \ldots, n_m)$ satisfying

$$\sum_{i=1}^{m} n_i = n,$$

with $0 \leq n_i \leq r_i$, n_i integers; $1 \leq i \leq m$.

Example. At suppertime Mrs. Jones rewards her children, Lorie, Mike, Tammie, and Johnny, for causing only a limited amount of damage to each other during the day. She decides to give them a total of ten identical candies. According to their respective good behavior she chooses to give at most three candies to Lorie, at most four to Mike, at most four to Tammie, and at most one to Johnny. In how many ways can she distribute the candies to the children?

In this problem we make the abstractions as follows:

$$\text{children} \leftrightarrow \text{distinguishable boxes}$$

$$\text{candies} \leftrightarrow \text{indistinguishable balls}$$

The possibilities of assignment are described by

Lorie	Mike	Tammie	Johnny
1	1	1	1
x	x	x	x
x^2	x^2	x^2	
x^3	x^3	x^3	
	x^4	x^4	

The generating function in question is

$$(1+x+x^2+x^3)(1+x+x^2+x^3+x^4)^2(1+x)$$

and the numerical answer we seek will be found in the coefficient of x^{10}. As it seems simple enough to write a computer program that multiplies two formal power series (and, in particular, two polynomials), calculating the coefficient of a power of x can be done expeditiously. Indeed, all it takes to program multiplication by convolution is a DO loop. [The coefficient in question equals, as we saw, the number of solutions (n_1, n_2, n_3, n_4) to

$$n_1 + n_2 + n_3 + n_4 = 10$$

$$0 \le n_1 \le 3$$

$$0 \le n_2, n_3 \le 4$$

$$0 \le n_4 \le 1, \quad n_i \text{ integers.}$$

There are precisely nine such vectors, which we actually list below:

Lorie	Mike	Tammie	Johnny
1	4	4	1
2	3	4	1
2	4	3	1
2	4	4	0
3	2	4	1
3	3	3	1
3	3	4	0
3	4	2	1
3	4	3	0

.]

3. *The number of ways of assigning n indistinguishable balls to m distinguishable boxes such that box j contains at least s_j balls is the coefficient of x^n in* $\prod_{j=1}^{m}(x^{s_j}(1+x+x^2+\cdots)) = x^{\sum s_j}(1-x)^{-m} = \sum_n \binom{n-\sum s_j+m-1}{m-1}x^n$.

The composition of the m boxes is, in this case,

Box 1	Box 2	$\cdots$	Box m
x^{s_1}	x^{s_2}		x^{s_m}
x^{s_1+1}	x^{s_2+1}		x^{s_m+1}
x^{s_1+2}	x^{s_2+2}		x^{s_m+2}
$\vdots$	$\vdots$		$\vdots$

Taking in common factor x^{s_j} from column j we obtain the generating function written above. The process of extracting x^{s_j} in common factor from column j and the writing down of the generating function by multiplying all factors together parallels the combinatorial argument of solving this problem by first leaving s_j balls in box j and then distributing the remaining $n - \sum_{j=1}^{m} s_j$ balls without restrictions to the m boxes. This shows in fact that the coefficient of x^n in the generating function written above is

$$\binom{n - \sum s_j + m - 1}{n - \sum s_j} = \binom{n - \sum s_j + m - 1}{m - 1}.$$

4. *The number of ways to distribute n indistinguishable balls into m distinguishable boxes with box i having the capacity to hold either s_{i1}, or s_{i2}, ..., or s_{ir_i}* (and no other number of) *balls equals the coefficient of x^n in $\prod_{i=1}^{m}(x^{s_{i1}} + x^{s_{i2}} + \cdots + x^{s_{ir_i}})$.*

The composition of the boxes is

Box 1	Box 2	$\cdots$	Box m
$x^{s_{11}}$	$x^{s_{21}}$		$x^{s_{m2}}$
$x^{s_{12}}$	$x^{s_{22}}$		$x^{s_{m2}}$
$\vdots$	$\vdots$		$\vdots$
$x^{s_{1r_1}}$	$x^{s_{2r_2}}$		$x^{s_{mr_m}}$

Placing s_{ij} balls in box j corresponds to selecting the power $x^{s_{ij}}$ in the jth column. Distributing a total of n balls to the m boxes amounts to selecting a vector of powers of x (one from each column), say $(s_{1n_1}, s_{2n_2}, \ldots, s_{mn_m})$, such that $\sum_{i=1}^{m} s_{in_i} = n$. The number of all such distributions of n balls is therefore the coefficient of x^n in the generating function given above. It also equals the number of integer solutions to

$$\sum_{i=1}^{n} s_{in_i} = n$$

with s_{in_i} restricted to belong to $\{s_{i1}, s_{i2}, \ldots, s_{ir_i}\}$, $1 \le i \le m$.

One can geometrically visualize the solutions to these constraints as the points $(y_1, y_2, \ldots, y_m)$, with y_i belonging to the finite set $\{s_{i1}, s_{i2}, \ldots, s_{ir_i}\}$, which are also on the hyperplane $\sum_{i=1}^m y_i = n$.

5. *We conclude this long section with a revision of several useful relations among generating functions. These are*:

(i) $(1-x)^{-1} = \sum_n x^n$.

(ii) $(1-x^{n+1})(1-x)^{-1} = 1 + x + x^2 + \cdots + x^n$.

(iii) $(1-x)^{-m} = \sum_n \binom{n+m-1}{n} x^n$; m *positive integer.*

(iv) $(1+x)^m = \sum_{n=0}^m \binom{m}{n} x^n$; m *positive integer.*

(v) $(x_1 + x_2 + \cdots + x_r)^m = \sum_{(n_1,\ldots,n_r)} m!/(n_1!n_2!\cdots n_r!) x_1^{n_1} x_2^{n_2} \cdots x_r^{n_r}$.

(vi) $\prod_{i=1}^m (\sum_j a_{ij} x^j) = \sum_n (\sum_{\substack{(j_1,\ldots,j_m) \\ \sum_k j_k = n}} a_{1j_1} a_{2j_2} \cdots a_{mj_m}) x^n$.

The contents of (i) and (ii) can be straightforwardly verified by multiplying out. The statement made in (iii) has been established in (2.2).

To understand (iv), write $(1+x)^m$ as

$$\begin{array}{ccccc} 1 & 1 & 1 & \ldots & 1 \\ \mathrm{x} & \mathrm{x} & \mathrm{x} & & \mathrm{x} \end{array} \qquad (m \text{ columns}).$$

A formal term in the expansion of $(1+x)^m$ involves the choice of one entry from each of the m columns. A term containing exactly n x's is obtained by picking x from a subset of n columns, and 1's from the remaining $m-n$ columns. The number of such terms equals the number of subsets with n elements out of the set of m columns, that is, it equals $\binom{m}{n}$. This explains (iv).

The proof of (v) is similar. Write out m columns

$$\begin{array}{cccc} x_1 & x_1 & \cdots & x_1 \\ x_2 & x_2 & & x_2 \\ \vdots & \vdots & & \vdots \\ x_r & x_r & & x_r \end{array}$$

A formal product is obtained by picking an x_i from each column. The coefficient of $x_1^{n_1} x_2^{n_2} \cdots x_r^{n_r}$ is the number of formal products of length m containing n_1 x_1's, n_2 x_2's, …, n_r x_r's. There are $m!/(n_1! n_2! \cdots n_r!)$ such products (see also Section 1.14). This establishes (v).

To realize that (vi) is true, line up m columns of infinite length:

$$\begin{array}{cccc} a_{10} & a_{20} & \cdots & a_{m0} \\ a_{11}x & a_{21}x & & a_{m1}x \\ a_{12}x^2 & a_{22}x^2 & & a_{m2}x^2 \\ a_{13}x^3 & a_{23}x^3 & & a_{m3}x^3 \\ \vdots & \vdots & & \vdots \end{array} .$$

A term involving x^n is obtained by picking $a_{kj_k} x^{j_k}$ from column k $(1 \leq k \leq m)$ and making the product $\prod_{k=1}^m a_{kj_k} x^{j_k}$, with the exponents of x satisfying $\sum_{k=1}^m j_k = n$. The totality of such terms equals

$$\sum_{\substack{(j_1,\ldots,j_m) \\ \sum_{k=1}^m j_k = n}} \prod_{k=1}^m a_{kj_k} x^{j_k},$$

thus explaining the coefficient of x^n.

2.5

We turn our attention now to exponential generating functions. These generating functions are helpful when counting the number of sequences (or words) of length n that can be made with m (possibly repeated) letters and with specified restrictions on the number of occurrences of each letter; such as the number of distinct sequences of length four that can be made with the (distinguishable) letters a, b, c, d, e in which b occurs twice, c at least once, e at most three times, and with no restrictions on the occurrences of a and d.

For convenience we denote the exponential generating function $\sum_n x^n/n!$ by e^x. We invite the reader to observe at once that $(e^x)^m = e^{mx}$. Indeed, the coefficient of $x^n/n!$ in e^{mx} is m^n, while the coefficient of $x^n/n!$ in $(e^x)^m$ is

$$\sum_{\substack{(n_1,\dots,n_m) \\ \sum n_i = m \\ 0 \le n_i,\ \text{integers}}} \frac{n!}{n_1! n_2! \cdots n_m!}.$$

These two expressions count the same thing, however, namely the number of sequences of length n that can be made with m distinguishable letters and with no restrictions on the number of occurrences of each letter. (To be specific, we have n spots to fill with m choices for each spot, and this gives us m^n choices; on the other hand we can sort out the set of sequences by the number of occurrences of each letter, thus obtaining the second expression.)

The mechanism of using exponential generating functions to solve problems in counting is similar to that described in Section 2.4. We present an example that captures all the relevant features of a general case.

Assume at all times that we have available an abundant (and if necessary infinite) supply of replicas of the letters a, b, c, d, e. We want to count *the number of distinct* *sequences of length four containing two b's, at least one c, at most three e's, and with no restrictions on the occurrences of a and d.*

The recipe that leads to the solution is the following: With each distinct letter attach a column in which the powers of x indicate the number of times that letter is allowed to appear in a sequence. Such powers of x are divided by the respective factorials. In this case we have

a	b	c	d	e
1			1	1
$\frac{x}{1!}$		$\frac{x}{1!}$	$\frac{x}{1!}$	$\frac{x}{1!}$
$\frac{x^2}{2!}$	$\frac{x^2}{2!}$	$\frac{x^2}{2!}$	$\frac{x^2}{2!}$	$\frac{x^2}{2!}$
$\frac{x^3}{3!}$		$\frac{x^3}{3!}$	$\frac{x^3}{3!}$	$\frac{x^3}{3!}$
$\frac{x^4}{4!}$		$\frac{x^4}{4!}$	$\frac{x^4}{4!}$	
$\vdots$		$\vdots$	$\vdots$	

The exponential generating function we attach to this problem is (as before) the product of the columns, that is,

$$\begin{aligned}
&\left(\sum_k \frac{x^k}{k!}\right)\frac{x^2}{2!}\left(\sum_{k=1}^{\infty}\frac{x^k}{k!}\right)\left(\sum_k \frac{x^k}{k!}\right)\left(1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}\right)\\
&= e^x\left(\frac{x^2}{2!}\right)(e^x-1)e^x\left(1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}\right)\\
&= e^{2x}(e^x-1)\frac{x^2}{2!}\left(1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}\right).
\end{aligned}$$

The numerical answer we seek is simply the coefficient of $x^4/4!$. If, with the same restrictions, we become interested in the number of sequences of length n, the answer is the coefficient of $x^n/n!$ in the above exponential generating function.

To see that the coefficient of $x^4/4!$ is indeed the answer to our problem one has to observe the following. The act of picking $x^{n_i}/n_i!$ from column i ($1 \le i \le 5$) corresponds to looking at sequences consisting of precisely n_1 a's, n_2 b's, n_3 c's, n_4 d's, and n_5 e's. Taking the product

$$\prod_i \frac{x^{n_i}}{n_i!} = \frac{(\sum_i n_i)!}{\prod_i n_i!} \frac{x^{\sum_i n_i}}{(\sum_i n_i)!}$$

(with $\sum_i n_i = 4$) produces a coefficient of

$$\frac{(\sum_i n_i)!}{\prod_i n_i!} = \frac{4!}{n_1!n_2!n_3!n_4!n_5!}$$

for $x^4/4!$, which equals the number of sequences with precisely n_i copies of each letter. The totality of such pickings, with values of n_i restricted to the exponents of x that appear in column i, leads to the coefficient of $x^4/4!$, which equals, therefore, the number of sequences with occurrences restricted as specified.

Specifically, we have

$$\begin{aligned} 1 \cdot \frac{x^2}{2!} \cdot \frac{x}{1!} \cdot 1 \cdot \frac{x}{1!} + 1 \cdot \frac{x^2}{2!} \cdot \frac{x}{1!} \cdot \frac{x}{1!} \cdot 1 \\ +1 \cdot \frac{x^2}{2!} \cdot \frac{x^2}{2!} \cdot 1 \cdot 1 + \frac{x}{1!} \cdot \frac{x^2}{2!} \cdot \frac{x}{1!} \cdot 1 \cdot 1 \\ = \left(\frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{2}\right) x^4 = 4! \left(\frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{2}\right) \frac{x^4}{4!}. \end{aligned}$$

We thus conclude that there are 42 such sequences.

Let us look at some examples of a more general nature.

Example 1. Find the number of (distinct) sequences of length n formed with m letters ($m \ge n$), with no letter repeated.

The m columns, one for each letter, are:

$$\begin{array}{ccccc} 1 & 1 & 1 & \cdots & 1 \\ x & x & x & & x \end{array}.$$

This gives the exponential generating function $(1+x)^m$. We thus seek the coefficient of $x^n/n!$. And since $(1+x)^m = \sum_{n=0}^{m} \binom{m}{n} x^n$ this shows that $x^n/n!$ has coefficient $m!/(m-n)!(= [m]_n)$, as expected.

Example 2. Find the number of sequences of length n, formed with m letters $(m \leq n)$, in which each letter appears at least once.

The m columns are all the same, namely $x/1!, x^2/2!, x^3/3!, \ldots$. Hence the exponential generating function is $(\sum_{n=1}^{\infty} x^n/n!)^m = (e^x - 1)^m$. The coefficient of $x^n/n!$ turns out to be $m!S_n^m$, where S_n^m is the Stirling number, as we shall see in Section 3.

Example 3. How many sequences of length n can be made with the digits $1, 2, 3, \ldots, m$ such that digit i is not allowed to appear n_{i1} or n_{i2} or $\cdots$ or n_{ir_i} times (these being the only restrictions)?

The ith column in this case consists of the terms of e^x with precisely $x^{n_{ij}}/n_{ij}!$ missing $(1 \leq j \leq r_i)$. We conclude therefore that the exponential generating function in question is

$$\prod_{i=1}^{m} \left(e^x - \sum_{j=1}^{r_i} \frac{x^{n_{ij}}}{n_{ij}!} \right).$$

The numerical answer we seek is the coefficient of $x^n/n!$.

2.6

Having thus shown the computational power of generating functions we address a problem that involves the permutations of the ordered set $1 < 2 < \cdots < m$. If $i < j$ and $\sigma(i) > \sigma(j)$ we say that the permutation σ has an *inversion* at the pair (i, j). Denote by

a_{mk} the number of permutations on $\{1, 2, \ldots, m\}$ with precisely k inversions; $0 \leq k \leq \binom{m}{2}$.

We seek the generating function for a_{mk}.

For a permutation σ and an integer j $(1 \leq j \leq m)$ denote by $\overline{\sigma}(j)$ the cardinality of the set $\{i : 1 \leq i < j \text{ and } \sigma(i) > \sigma(j)\}$. The number of inversions of σ can now be written as $\overline{\sigma}(1) + \overline{\sigma}(2) + \cdots + \overline{\sigma}(m)$. (Note that $\overline{\sigma}(1) = 0$.)

Thus the number of permutations with exactly k inversions is the number of solutions in nonnegative integers to

$$n_1 + n_2 + \cdots + n_m = k$$

with restrictions $0 \leq n_i \leq i - 1$. (For a fixed permutation σ, n_i corresponds to $\overline{\sigma}(i)$.) We know how to interpret the set of such solutions (cf. Section 2.4). Think of m distinguishable boxes (as columns), with column i consisting of $1, x, x^2, \ldots, x^{i-1}$. The generating function that we associate is $\prod_{i=0}^{m}(1 + x + \ldots + x^{i-1})$ and then a_{mk}, being the same as the number of solutions to the constraints mentioned above, equals the coefficient of x^k in this generating function. We conclude, therefore, that

$$\sum_{k=0}^{\binom{m}{2}} a_{mk} x^k = \prod_{i=0}^{m}(1 + x + \ldots + x^{i-1}) = \prod_{i=0}^{m} \left(\frac{1 - x^i}{1 - x} \right).$$

EXERCISES

1. How many ways are there to get a sum of 14 when 4 (distinguishable) dice are rolled?

2. Find the generating function for the number of ways a sum of n can occur when rolling a die an infinite (or at least n) number of times.

3. How many ways are there to collect \$12 from 16 people if each of the first 15 people can give at most \$2 and the last person can give either \$0 or \$1 or \$4?

4. How many ways are there to distribute 20 jelly beans to Mary(G), Larry (B), Sherry (G), Terri (G), and Jerry (B) such that a boy (indicated by B) is given an odd number of jelly beans and a girl is given an even number (0 counts as even).

5. Find the coefficient of x^n in $(1 + x + x^2 + x^3)^m(1 + x)^m$.

6. Find the generating function for the sequence (a_n) if (a) $a_n = n^2$, (b) $a_n = n^3$, (c) $a_n = \binom{n}{2}$, and (d) $a_n = \binom{n}{3}$.

7. In how many ways can ten salespersons be assigned so that two are assigned to district A, three to district B, and five to district C? If five of the salespersons are men and five are women, what is the chance that a random assignment of two salespersons to district A, three to B and five to C will result in segregation of the salespersons by sex? What is the probability that a random assignment will result in at least one female salesperson being assigned to each of the three districts?

8. How many distinct formal words can be made with the letters in the word "abracadabra"?

9. Show that $\sum_k (-1)^k \binom{n}{k}((1+kx)/(l+nk)^k) = 0$, for all x and all positive integers n. What do we obtain by taking $x = 0$, or $x = 1$? [Hint: Write

$$\begin{aligned} 0 &= \left(1 - \frac{1}{1+nx}\right)^n - \left(1 - \frac{1}{1+nx}\right)^n \\ &= \left(1 - \frac{1}{1+nx}\right)^n - \frac{nx}{1+nx}\left(1 - \frac{1}{1+nx}\right)^{n-1}, \end{aligned}$$

expand using the binomial expansion and sort out by $\binom{n}{k}$.]

3 GENERATING FUNCTIONS for STIRLING NUMBERS

Let x and y be indeterminates and denote $\sum_n x^n/n!$ by e^x, $\sum_n(-1)^n x^{n+1}/(n+1)$ by $\ln(1+x)$, and $e^{x\ln y}$ by y^x. We occasionally yield to the temptation of looking at these formal power series as series expansions of analytic functions. While this contemplative attitude is in itself harmless enough, the effective act of assigning numerical values to x and y becomes an unmistakable cause of concern. Questions of convergence immediately arise and they are of crucial importance. It can be shown that both e^x and $\ln(1+x)$ converge for positive values of x. The relations $e^{\ln x} = x = \ln e^x$ are also known to hold and are used freely in what follows. The formal expansion

$$(1+y)^x = \sum_k [x]_k \frac{y^k}{k!}$$

is needed as well; it holds for $|x| < 1$ [here $[x]_k = x(x-1)\cdots(x-k+1)$]. The reader can find these series expansions in most calculus books. We take them for granted here.

2.7

Taking advantage of the new tools just introduced, let us take another look at the Stirling and Bell numbers:

∗ *Compiled beneath are several generating functions for these numbers* (expanding the right-hand side and equating like powers yields many identities):

1. $\sum_n S_n^k y^n/n! = (1/k!)(e^y-1)^k$.

2. $\sum_n s_n^k y^n/n! = (1/k!)(\ln(1+y))^k$.

3. $\sum_k S_n^k x^k = e^{-x} \sum_m m^n x^m/m!$.

4. $\sum_n \sum_k S_n^k x^k y^n/n! = e^x(e^y - 1)$.

5. $\sum_n B_n y^n/n! = e^{e^y-1}$.

6. $\sum_n S_n^k x^{n-k} = (1-x)^{-1}(1-2x)^{-1} \cdots (1-kx)^{-1}$.

7. *The Bell numbers B_n satisfy*

$$\lim_{n\to\infty} \frac{n^{-\frac{1}{2}}(\lambda(n))^{n+\frac{1}{2}} e^{\lambda(n)-n-1}}{B_n} = 1,$$

where $\lambda(n)$ is defined by $\lambda(n) \ln \lambda(n) = n$. (We recall the usual conventions with indices: $S_n^k = 0$ for all $k \geq n$, and $S_n^0 = 0$ for all n.)

Proof. 1. The proof relies on Stirling's formula

$$x^n = \sum_k S_n^k [x]_k,$$

which we proved in (c) of Section 1.7. We proceed as follows:

$$\begin{aligned}
\sum_k \sum_n S_n^k \frac{y^n}{n!}[x]_k &= \sum_n \sum_k S_n^k [x]_k \frac{y^n}{n!} = \sum_n x^n \frac{y^n}{n!} \\
&= \sum_n \frac{(xy)^n}{n!} = e^{xy} = (e^y)^x = (1 + (e^y - 1))^x \\
&= \sum_k \frac{1}{k!}(e^y - 1)^k [x]_k.
\end{aligned}$$

Identifying the coefficients of $[x]_k$ gives

$$\sum_n S_n^k \frac{y^n}{n!} = \frac{1}{k!}(e^y - 1)^k.$$

2. Start out with $[x]_n = \sum_k s_n^k x^k$, a formula that we proved in (c) of Section 1.8. Multiply both sides by $y^n/n!$, sum over n, and use known series expansions to obtain:

$$\begin{aligned}\sum_k \sum_n s_n^k \frac{y^n}{n!} x^k &= \sum_n [x]_n \frac{y^n}{n!} = (1+y)^x = e^{x\ln(1+y)} \\ &= \sum_k \frac{1}{k!} (\ln(1+y))^k x^k.\end{aligned}$$

Identifying the coefficients of x^k yields the result.

3. Observe first that $x^k e^x = \sum_i x^{i+k}/i! = \sum_m [m]_k x^m/m!$, since $[m]_k = 0$ in the first $k-1$ terms. By Stirling's formula, recalling also that $m^n = \sum_k S_n^k [m]_k$, we have

$$\begin{aligned}e^x \sum_k S_n^k x^k &= \sum_k S_n^k x^k e^x = \sum_k S_n^k \sum_m [m]_k \frac{x^m}{m!} \\ &= \sum_m \frac{x^m}{m!} \sum_k S_n^k [m]_k = \sum_m \frac{m^n x^m}{m!}.\end{aligned}$$

If we set $x = 1$, we obtain Dobinski's formula

$$B_n = e^{-1} \sum_m \frac{m^n}{m!}.$$

4. Start with the formula established in **3**, multiply it by $y^n/n!$, and sum. What results is

$$\begin{aligned}\sum_n \sum_k S_n^k x^k \frac{y^n}{n!} &= e^{-x} \sum_m \sum_n \frac{m^n x^m}{m!} \frac{y^n}{n!} = e^{-x} \sum_m \frac{x^m}{m!} \sum_n \frac{(my)^n}{n!} \\ &= e^{-x} \sum_m \frac{x^m}{m!} e^{my} = e^{-x} \sum_m \frac{(xe^y)^m}{m!} = e^{-x} e^{xe^y} = e^{x(e^y-1)}.\end{aligned}$$

5. Recall that $\sum_k S_n^k = B_n$. Set $x = 1$ in **4** to obtain **5**.

6. (Induction on k.) The relation is true for $k = 1$ since it reduces to $1 + x + x^2 + \cdots = 1/(1-x)$. Assume that it holds for $k-1$ and show that it holds for k. Let $f(x) = \sum_{n, n\geq k} S_n^k x^{n-k}$. Then

$$f(x) = \sum_{\substack{n \\ n\geq k}} S_n^k x^{n-k} = \{\text{by the recurrence } S_n^k = S_{n-1}^{k-1} + kS_{n-1}^k\}$$

$$\begin{aligned}
&= \sum_{\substack{n \\ n\geq k}} (S_{n-1}^{k-1} + kS_{n-1}^{k})x^{n-k} \\
&= \sum_{\substack{n \\ n-1\geq k-1}} S_{n-1}^{k-1}x^{(n-1)-(k-1)} + k\sum_{\substack{n \\ n\geq k}} S_{n-1}^{k}x^{n-k} \\
&= \{\text{by induction}\} = \prod_{m=1}^{k-1}(1-mx)^{-1} + k\sum_{\substack{n \\ n\geq k}} S_{n-1}^{k}x^{n-k} \\
&= \prod_{m=1}^{k-1}(1-mx)^{-1} + k(S_{k-1}^{k}x^0 + S_k^k x + S_{k+1}^k x^2 + S_{k+2}^k x^3 + \cdots) \\
&= \prod_{m=1}^{k-1}(1-mx)^{-1} + k\sum_{\substack{n \\ n\geq k}} S_n^k x^{n-k+1} \\
&= \prod_{m=1}^{k-1}(1-mx)^{-1} + kx\sum_{\substack{n \\ n\geq k}} S_n^k x^{n-k} \\
&= \prod_{m=1}^{k-1}(1-mx)^{-1} + kxf(x).
\end{aligned}$$

We can now solve for $f(x)$ and thus obtain the formula we want.

7. The proof of this asymptotic result is somewhat analytic in nature and we omit it to preserve continuity. See reference [10).

2.8

The Stirling numbers occur when relating moments to lower factorial moments. Call $\mathbf{M}_n(f) = \sum_x f(x)x^n$ the nth *moment* of f and $\mathbf{m}_n(f) = \sum_x f(x)[x]_n$ the nth *lower factorial moment of* f. (The sum over x could be an integral as well. The variable x is understood to belong to some subset of the real line Stirling's formulas give us immediately

$$\mathbf{M}_n = \sum_k S_n^k \mathbf{m}_k \quad \text{and} \quad \mathbf{m}_n = \sum_k s_n^k \mathbf{M}_k.$$

We now describe another situation in which the Stirling numbers pop up.

∗ *Let D be the operator of differentiation* (i.e., $D = d/dx$) *and let $\theta = xD$. Then*

$$\theta^n = \sum_{k=0}^{n} S_n^k x^k D^k \quad \left(\textit{and } x^n D^n = \sum_{k=0}^{n} s_n^k \theta^k\right).$$

Proof. Proceed as follows:

$$\begin{aligned}
\theta &= xD = S_1^1 xD \\
\theta^2 &= xD(\theta) = xD(xD) = x(D + xD^2) = xD + x^2D^2 = S_2^1 xD + S_2^2 x^2 D^2 \\
&\vdots \\
\theta^n &= \sum_{k=0}^{n} S_n^k x^k D^k \ \text{(assume this)}.
\end{aligned}$$

Then

$$\begin{aligned}
\theta^{n+1} &= xD(\theta^n) = xd\left(\sum_{k=0}^{n} S_n^k x^k D^k\right) \\
&= x\left(\sum_{k=0}^{n} S_n^k (kx^{k-1}D^k + x^k D^{k+1})\right) \\
&= \sum_{k=0}^{n} S_n^k k x^k D^k + \sum_{k=0}^{n} S_n^k x^{k+1} D^{k+1} \\
&= \sum_{k=0}^{n} S_n^k k x^k D^k + \sum_{k=1}^{n+1} S_n^{k-1} x^k D^k \\
&= \sum_{k=1}^{n+1} (S_n^{k-1} + kS_n^k) x^k D^k \\
&= \sum_{k=0}^{n+1} S_{n+1}^k x^k D^k.
\end{aligned}$$

This ends the proof, by induction.

The second formula, written in parentheses in the statement above, is equivalent to the first through a process of inversion. This process is presented in detail in Chapter 3.

2.9

We discuss here several properties of the Lah numbers L_n^k. A combinatorial interpretation of these numbers was given in Section 1.15, where we labeled

$$L_n^k = (-1)^n \frac{n!}{k!} \binom{n-1}{k-1}.$$

For small values of n and k we have the following table for L_n^k:

$n \backslash k$	1	2	3	4	5
1	-1				
2	2	1		0	
3	-6	-6	-1		
4	24	36	12	1	
5	-120	-240	-120	-20	-1

Define now numbers L_n^k (we show that these are the same as the Lah L_n^k above) by

$$[-x]_n = \sum_{k=1}^{n} L_n^k [x]_k; \quad L_n^k = 0, \quad \text{for } k > n.$$

$*$ *We prove the following*:

1. $[-x]_n = \sum_{k=1}^n L_n^k [x]_k$ *if and only if* $[x]_n = \sum_{k=1}^n L_n^k [-x]_k$.

2. $L_{n+1}^k = -L_n^{k-1} - (n+k)L_n^k$.

3. $\sum_n L_n^k t^n/n! = (1/k!)(-t/(1+t))$.

4. $L_n^k = (-1)^n (n!/k!)\binom{n-1}{k-1}$.

5. $\sum_k \sum_n L_n^k x^k t^n/n! = \exp(-xt/(1+t))$.

6. $L_n^k = \sum_{j=k}^n (-1)^j s_n^j S_j^k$.

Proof. 1. Interchange x and $-x$.

2.

$$\begin{aligned}\sum_{k=1}^{n+1} L_{n+1}^k [x]_k &= [-x]_{n+1} = (-x-n)[-x]_n \\ &= (-x-n) \sum_{k=1}^{n} L_n^k [x]_k = \sum_{k=1}^{n} L_n^k (-x-n)[x]_k\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{n} L_n^k(-(x-k)-(n+k))[x]_k \\
&= \sum_{k=1}^{n}(-L_n^k[x]_{k+1}-(n+k)L_n^k[x]_k).
\end{aligned}
$$

Identifying coefficients of $[x]_k$ gives **2**.

3. Start with $\sum_k L_n^k[x]_k = [-x]_n$. Multiply by $t^n/n!$ and sum:

$$
\begin{aligned}
\sum_k [x]_k \sum_n L_n^k \frac{t^n}{n!} &= \sum_n \sum_k L_n^k [x]_k \frac{t^n}{n!} = \sum_n [-x]_n \frac{t^n}{n!} \\
&= (1+t)^{-x} = \left(\frac{1}{1+t}\right)^x = \left(1-\frac{t}{1+t}\right)^x \\
&= \sum_k \frac{[x]_k(-t/(1+t))^k}{k!},
\end{aligned}
$$

yielding **3**.

4.

$$
\begin{aligned}
\sum_n L_n^k \frac{t^n}{n!} &= \frac{(-t/(1+t))^k}{k!} = \frac{1}{k!}(-t^k(1-t+t^2-t^3+\cdots)^k) \\
&= \sum_n (-1)^n \frac{n!}{k!}\binom{n-1}{k-1}\frac{t^n}{n!}.
\end{aligned}
$$

See (2.2) for an explanation of the last equality sign.

5. Start with $\sum_n L_n^k t^n/n! = (-t/(1+t))^k/k!$, multiply by x^k, and sum

$$
\sum_k \sum_n L_n^k \frac{t^n}{n!} x^k = \sum_k \frac{(-tx/(1+t))^k}{k!} = \exp\left(\frac{-xt}{1+t}\right).
$$

6.

$$
\begin{aligned}
\sum_{k=1}^{n} L_n^k [x]_k &= [-x]_n = \sum_{j=1}^{n}(-1)^j s_n^j x^j \\
&= \sum_{j=1}^{n}(-1)^j s_n^j \sum_{k=1}^{j} s_j^k [x]_k \\
&= \sum_{k=1}^{n}\sum_{j=1}^{n}(-1)^j s_n^j S_j^k [x]_k.
\end{aligned}
$$

$*$ *The Lah numbers occur when expressing the upper factorial moments, defined by* $\overline{\mathbf{m}}_n(f) = \sum_x f(x)[x]^n$, *in terms of the lower factorial moments* $\mathbf{m}_n(f)$, *which we defined earlier in this section.* (*Here* $[x]^n = x(x+1)\cdots(x+n-1)$.) *Specifically,*

$$\overline{\mathbf{m}}_n = \sum_k (-1)^n L_n^k \mathbf{m}_k \quad \textit{and} \quad \mathbf{m}_n = \sum_k (-1)^k L_n^k \overline{\mathbf{m}}_k.$$

Further, in terms of differential operators, if $\theta = x^2 D$ (where D stands for d/dx), *then*

$$\theta^n = \sum_{k=1}^{n} (-1)^n L_n^k x^{n+k} D^k$$

(or, equivalently, $D^n = \sum_{k=1}^{n} (-1)^k L_n^k x^{-n-k} \theta^k$). The proof is similar to the case of $\theta = xD$ involving Stirling numbers.

4 BELL POLYNOMIALS

The object of this section is to bring to attention *an explicit formula by which the higher derivatives of a composition of two functions can be computed.* Partitions of a set, and thus Bell numbers, will enter these calculations in a natural way.

Let $h = f \circ g$ be the composition of f with g, that is, $h(t) = f(g(t))$ where t is an argument. We assume that the functions f, g, and h have derivatives of all orders.

Denote by D_y the operator d/dy of differentiation with respect to y. By D_y^n we indicate the n-fold application of D_y, that is, the nth derivative with respect to y.

We denote as follows:

$$h_n = D_t^n h, \quad f_n = D_g^n f, \quad g_n = D_t^n g.$$

Our aim is to find an explicit formula for h_n *in terms of the* f_k's and g_k's. To begin with, let us look at the first few expressions for h_n:

$$h_1 = f_1 g_1$$

$$
\begin{aligned}
h_2 &= f_2 g_1^2 + f_1 g_2 \\
h_3 &= f_3 g_1^3 + f_2(2g_1 g_2) + f_2 g_1 g_2 + f_1 g_3 \\
&= f_3 g_1 + f_2(3g_1 g_2) + f_1 g_3 \\
h_4 &= f_4 g_1^4 + f_3(6g_2 g_1^2) + f_2(4g_3 g_1 + 3g_2^2) + f_1 g_4 \\
&\vdots
\end{aligned}
$$

Write in general $h_n = \sum_{k=1}^n f_k \alpha_{nk}$. Here the α_{nk}'s are polynomials in g_i's that do not depend upon the choice of f. The h_n's are called *Bell polynomials*. (As is plain to see, these polynomials are linear in the f_k's but highly nonlinear in the g_k's.)

We proceed in establishing the explicit form of the α_{nk}'s and do so in "steps." To this end, define polynomials $\overline{B}_n$ by

$$\overline{B}_n = \sum_{k=1}^n \alpha_{nk}.$$

Step 1. $\overline{B}_n = e^{-g}(D_t^n e^{-g})$.

Indeed, let $f(z) = e^z$ be the exponential series. Then $h = e^g$ and $h_n = D_t^n e^g = \sum_{k=1}^n f_k \alpha_{nk} = \sum_{k=1}^n e^g \alpha_{nk} = e^g \sum_{k=1}^n \alpha_{nk} = e^g \overline{B}_n$.

Step 2. $\overline{B}_{n+1} = \sum_{k=0}^n \binom{n}{k} g_{k+1} \overline{B}_{n-k}$.

Recall that if α_0 and β_0 are functions of t, differentiable any number of times, then

$$
\begin{aligned}
D_t^0 \alpha_0 \beta_0 &= \alpha_0 \beta_0 \quad \text{(ordinary multiplication)} \\
D_t^1 \alpha_0 \beta_0 &= \alpha_1 \beta_0 + \alpha_0 \beta_1 \\
D_t^2 \alpha_0 \beta_0 &= \alpha_2 \beta_0 + \alpha_1 \beta_1 + \alpha_1 \beta_1 + \alpha_0 \beta_2 \\
&= \alpha_2 \beta_0 + 2\alpha_1 \beta_1 + \alpha_0 \beta_2 \\
D_t^3 \alpha_0 \beta_0 &= \alpha_3 \beta_0 + 3\alpha_2 \beta_1 + 3\alpha_1 \beta_2 + \alpha_0 \beta_3
\end{aligned}
$$

$$\vdots$$

$$D_t^n \alpha_0 \beta_0 \quad = \quad \sum_{k=0}^{n} \binom{n}{k} \alpha_k \beta_{n-k}. \quad \text{(Leibnitz's formula)}.$$

This formula is not hard to prove, and it was derived at the end of Section 1.6.

With this at hand,

$$\begin{aligned}
\overline{B}_{n+1} &= e^{-g}(D_t^{n+1} e^g) = e^{-g} D_t^n (g_1 e^g) \\
&= \{\text{let } \alpha_0 = g_1 \text{ and } \beta_0 = e^g\} \\
&= e^{-g} \sum_{k=0}^{n} \binom{n}{k} g_{k+1} D_t^{n-k} e^g \\
&= \sum_{k=0}^{n} \binom{n}{k} g_{k+1} e^{-g} (D_t^{n-k} e^g) \\
&= \sum_{k=0}^{n} \binom{n}{k} g_{k+1} \overline{B}_{n-k}.
\end{aligned}$$

Step 3. $\ln(\sum_n \overline{B}_n x^n/n!) = \sum_n g_{n+1} x^{n+1}/(n+1)!$.

Indeed (formally) differentiating both sides with respect to x we obtain

$$\left(\sum_n \overline{B}_n \frac{x^n}{n!}\right)^{-1} \left(\sum_n \overline{B}_{n+1} \frac{x^n}{n!}\right) = \sum_n g_{n+1} \frac{x^n}{n!}$$

if and only if if and only if

$$\sum_n \overline{B}_{n+1} \frac{x^n}{n!} = \left(\sum_n \overline{B}_n \frac{x^n}{n!}\right) \left(\sum_n g_{n+1} \frac{x^n}{n!}\right)$$

if and only if

$$\overline{B}_{n+1} = \sum_{k=0}^{n} \binom{n}{k} g_{k+1} \overline{B}_{n-k}$$

(which is true by Step 2, above). This shows that the two series in Step 3 are equal, up to a constant term. But the constant term is clearly zero on bott sides. This completes the proof of Step 3.

[We used here the nontrivial but familiar fact that the formal derivative of $\ln y$ (where y is a formal power series in x) equals y^{-1} times the formal derivative of y with respect to x. While true, the verification of this statement is omitted, to preserve continuity. In passing we remind the reader that $\ln(1+y) = \sum_n (-1)^n y^{n+1}/(n+1)$.]

Step 4.

$$\overline{B}_n = \sum_{k=1}^{n} \sum_{\substack{\lambda_i \geq 0 \\ \sum_{i=1}^{k} \lambda_i = k \\ \sum_{i=1}^{k} i\lambda_i = n}} \frac{n!}{(1!)^{\lambda_1} \cdots (k!)^{\lambda_k} (\lambda_1!) \cdots (\lambda_k!)} g_1^{\lambda_1} g_2^{\lambda_2} \cdots g_k^{\lambda_k}$$

(The inner sum is over all partitions of $\{1, 2, \ldots, n\}$ with exactly k classes;

$$\begin{gathered} \lambda_1 \text{ classes of size } 1 \\ \lambda_2 \text{ classes of size } 2 \\ \vdots \\ \lambda_k \text{ classes of size } k). \end{gathered}$$

Indeed, exponentiating both sides of Step 3 we obtain

$$\begin{aligned} \sum_n \overline{B}_n \frac{x^n}{n!} &= \exp\left[\sum_n g_{n+1} \frac{x^{n+1}}{(n+1)!}\right] = \prod_{n=1}^{\infty} \exp\left(g_n \frac{x^n}{n!}\right) \\ &= \prod_{n=1}^{\infty} \left(\sum_{k=0}^{\infty} \frac{(g_n x^n/n!)^k}{k!}\right) = \prod_{n=1}^{\infty} \left(\sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{g_n}{n!}\right)^k x^{nk}\right) \\ &= \left\{\text{write } a_{nk} \text{ (double index) for } \frac{1}{k!}\left(\frac{g_n}{n!}\right)^k\right\} \\ &= \prod_{n=1}^{\infty} \left(\sum_{k=0}^{\infty} a_{nk} x^{nk}\right) \\ &= (a_{10} + a_{11}x^{1\cdot 1} + a_{12}x^{1\cdot 2} + \cdots)(a_{20} + a_{21}x^{2\cdot 1} + a_{22}x^{2\cdot 2} + \cdots) \\ &\quad \cdot (a_{30} + a_{31}x^{3\cdot 1} + a_{32}x^{3\cdot 2} + \cdots) \cdots \\ &= \sum_{n=0}^{\infty} \left(\sum_{\sum_{i=1}^{n} i\lambda_i = n} a_{1\lambda_1} a_{2\lambda_2} \cdots a_{n\lambda_n}\right) x^n \end{aligned}$$

$$= \sum_{n=0}^{\infty} \sum_{\substack{\lambda_i \geq 0 \\ \sum i\lambda_i = n}} \frac{1}{(\lambda_1!)\cdots(\lambda_n!)} \left(\frac{g_1}{1!}\right)^{\lambda_1} \cdots \left(\frac{g_n}{n!}\right)^{\lambda_n} x^n$$

$$= \sum_{n=0}^{\infty} \left[\sum_{k=1}^{n} \sum_{\substack{\sum \lambda_i = k \\ \sum i\lambda_i = n}} \frac{1}{(1!)^{\lambda_1}\cdots(k!)^{\lambda_k}(\lambda_1!)\cdots(\lambda_k!)} g_1^{\lambda_1} g_2^{\lambda_2} \cdots g_k^{\lambda_k} \right] x^n.$$

Equating the coefficients of x^n on both sides explains Step 4.

[*Aside*: The polynomial $\overline{B}_n$ evaluated at $g_1 = 1, g_2 = 1, \ldots, g_n = 1$ becomes the Bell number B_n (this follows immediately from Step 4).]

Step 5.

$$h_n = \sum_{k=1}^{n} f_k \left(\sum_{\substack{\lambda_i \geq 0 \\ \sum \lambda_i = k \\ \sum i\lambda_i = n}} \frac{1}{(1!)^{\lambda_1}\cdots(k!)^{\lambda_k}(\lambda_1!)\cdots(\lambda_k!)} g_1^{\lambda_1} g_2^{\lambda_2} \cdots g_k^{\lambda_k} \right)$$

(i.e., α_{nk} is the inner sum in Step 5). This is *Faa DiBruno's formula.*

To prove this formula denote the inner sum in Step 5 by α^*_{nk}, for convenience. Recall that α_{nk} has been defined by $h_n = \sum_{k=1}^n f_k \alpha_{nk}$ and that the content of Step 4 is (in this notation) $\sum_{k=1}^n \alpha_{nk} = \sum_{k=1}^n \alpha^*_{nk} (= \overline{B}_n)$. Our aim is to prove that $\alpha_{nk} = \alpha^*_{nk}$.

We have the following chain of implications:

$$\begin{aligned} \sum_{k=1}^{n} \alpha_{nk} &= \sum_{k=1}^{n} \alpha^*_{nk} \Rightarrow \sum_{k=1}^{n} (\alpha_{nk} - \alpha^*_{nk}) = 0 \\ &\Rightarrow \alpha_{nk} - \alpha^*_{nk} = 0 \Rightarrow \alpha_{nk} = \alpha^*_{nk}. \end{aligned}$$

The first implication is just rewriting. Let us study the second implication: It is clear that the α^*_{nk}'s are homogeneous polynomials of degree k in the g_i's. We now show that the α_{nk}'s are also homogeneous of degree k. It is easy to verify this statement for small values of n and k. Assume it is so for the α_{nk}'s, for all $1 \leq k \leq n$, and show, by induction, that the $\alpha_{n+1,k}$'s are homogeneous of degree k, $1 \leq k \leq n+1$. Recall that $h_n = \sum_{k=1}^{n} f_k \alpha_{nk}$. The coefficient of f_k in the expression of h_{n+1}, that is, $\alpha_{n+1,k}$, is obtained by differentiating $f_{k-1}\alpha_{n,k-1} + f_k\alpha_{nk}$. That is, $D_t(f_{k-1}\alpha_{n,k-1} + f_k\alpha_{nk}) = f_k g_1 \alpha_{n,k-1} + f_{k-1} D_t \alpha_{n,k-1} + f_{k+1} g_1 \alpha_{nk} + f_k D_t \alpha_{nk}$. We hence have

$$\alpha_{n+1,k} = g_1 \alpha_{n,k-1} + D_t \alpha_{nk}.$$

The right-hand side in this relation has both terms homogeneous of degree k, the first by the inductive assumption, the second using the product rule and induction (on k). Hence the α_{nk}'s are homogeneous polynomials of degree k. The second implication now follows by equating to zero all the *homogeneous components* of the sum. The third implication follows because the monomials of degree k in the g_i's are linearly independent (since the g_i's themselves are, in general). This completes Step 5 and ends the proof of Faa DiBruno's formula.

Bell Polynomials

$$\begin{aligned}
h_1 &= f_1 g_1 \\
h_2 &= f_1 g_2 + f_2 g_1^2 \\
h_3 &= f_1 g_3 + f_2(3g_2 g_1) + f_3 g_1^3 \\
h_4 &= f_1 g_4 + f_2(4g_3 g_1 + 3g_2^2) + f_3(6g_2 g_1^2) + f_4 g_1^4 \\
h_5 &= f_1 g_5 + f_2(5g_4 g_1 + 10g_3 g_2) + f_3(10g_3 g_1^2 + 15g_2^2 g_1) \\
&\quad + f_4(10g_2 g_1^3) + f_5 g_1^5 \\
h_6 &= f_1 g_6 + f_2(6g_5 g_1 + 15g_4 g_2 + 10g_3^2)
\end{aligned}$$

$$
\begin{aligned}
& +f_3(15g_4g_1^2+60g_3g_2g_1+15g_2^3) \\
& +f_4(20g_3g_1^3+45g_2^2g_1^2)+f_5(15g_2g_1^4)+f_6g_1^6 \\
h_7 \quad = \quad & f_1g_7+f_2(7g_6g_1+21g_5g_2+35g_4g_3) \\
& +f_3(21g_5g_1^2+105g_4g_2g_1+70g_3^2g_1+105g_3g_2^2) \\
& +f_4(35g_4g_1^3+210g_3g_2g_1^2+105g_2^3g_1) \\
& +f_5(35g_3g_164+105g_2^2g_1^3)+f_6(21g_2g_1^5)+f_7g_1^7 \\
h_8 \quad = \quad & f_1g_8+f_2(8g_7g_1+28g_6g_2+56g_5g_3+35g_4^2) \\
& +f_3(28g_6g_1^2+168g_5g_2g_1+280g_4g_3g_1+210g_4g_2^2+280g_3^2g_2) \\
& +f_4(56g_5g_1^3+420g_4g_2g_1^2+280g_3^2g_1^2+840g_3g_2^2g_1+105g_2^4) \\
& +f_5(70g_4g_1^4+560g_3g_2g_1^3+420g_2^3g_1^2) \\
& +f_6(56g_3g_1^5+210g_2^2g_1^4)+f_7(28g_2g_1^6)+f_8g_1^8 \\
h_9 \quad = \quad & f_1g_9+f_2(9g_8g_1+36g_7g_2+84g_6g_3+126g_5g_4) \\
& +f_3(36g_7g_1^2+252g_6g_2g_1+504g_5g_3g_1+378g_5g_2^2 \\
& +315g_4^2g_1+1260g_4g_3g_2+280g_3^3) \\
& +f_4(84g_6g_1^3+756g_5g_2g_1^2+1260g_4g_3g_1^2 \\
& +1890g_4g_2^2g_1+2520g_3^2g_2g_1+1260g_3g_2^3) \\
& +f_5(126g_5g_1^4+1260g_4g_2g_1^3 \\
& +840g_3^2g_1^3+3780g_3g_2^2g_1^2+945g_2^4g_1) \\
& +f_6(126g_4g_1^5+1260g_3g_2g_1^4+1260g_2^3g_1^3) \\
& +f_7(84g_3g_1^6+378g_2^2g_1^5)+f_8(36g_2g_1^7)+f_9g_1^9
\end{aligned}
$$

$$\begin{aligned}
h_{10} \quad = \quad & f_1 g_{10} + f_2(10 g_9 g_1 + 45 g_8 g_2 + 120 g_7 g_3 + 210 g_6 g_4 + 126 g_5^2) \\
& + f_3(45 g_8 g_1^2 + 360 g_7 g_2 g_1 + 840 g_6 g_3 g_1 + 630 g_6 g_2^2 \\
& + 1260 g_5 g_4 g_1 + 2520 g_5 g_3 g_2 + 1575 g_4^2 g_2 + 2100 g_4 g_3^2) \\
& + f_4(120 g_7 g_1^3 + 1260 g_6 g_2 g_1^2 + 2520 g_5 g_3 g_1^2 \\
& + 3780 g_5 g_2^2 g_1 + 1575 g_4^2 g_1^2 + 12600 g_4 g_3 g_2 g_1 \\
& + 3150 g_4 g_2^3 + 2800 g_3^3 g_1 + 6300 g_3^2 g_2^2) \\
& + f_5(210 g_6 g_1^4 + 2520 g_5 g_2 g_1^3 + 4200 g_4 g_3 g_1^3 \\
& + 9450 g_4 g_2^2 g_1^2 + 12600 g_3^2 g_2 g_1^2 + 12600 g_3 g_2^3 g_1 + 945 g_2^5) \\
& + f_6(252 g_5 g_1^5 + 3150 g_4 g_2 g_1^4 + 2100 g_3^2 g_1^4 + 12600 g_3 g_2^2 g_1^3 + 4725 g_2^4 g_1^2 \\
& + f_7(210 g_4 g_1^6 + 2520 g_3 g_2 g_1^5 + 3150 g_2^3 g_1^4) \\
& + f_8(120 g_3 g_1^7 + 630 g_2^2 g_1^6) + f_9(45 g_2 g_1^8) + f_{10} g_1^{10}.
\end{aligned}$$

5 RECURRENCE RELATIONS

The general question that we address here is as follows: *From a rule recurrence among the elements of a sequence* (a_n) *determine explicitly that sequence.*

Examples are many. If the recurrence is $a_n = a_{n-1} + n$ with $a_0 = 1$ $(n = 1, 2, 3, \ldots)$, then it easily follows that $a_n = 1 + \binom{n+1}{2}$. On the other hand, if $a_0 = 1$, $a_1 = 1$, and the recurrence relation is $a_n = \sum_{k=1}^{n-1} a_k^2 a_{n-k}$ $(n \geq 2)$, then it is not so easy to determine a_n as a function of n. Indeed, more often than not one will not be able to find a_n explicitly.

Generating functions provide, nonetheless, a powerful technique that leads to complete solutions in many situations. Let us illustrate this by a classic example.

2.10

Mr. Fibonacci just bought a pair of baby rabbits (one of each sex) possessing some remarkable, and perhaps enviable, properties:

They take a month to mature.

When mature, a pair gives birth each month to precisely one new pair (again one of each sex), and with the same remarkable properties.

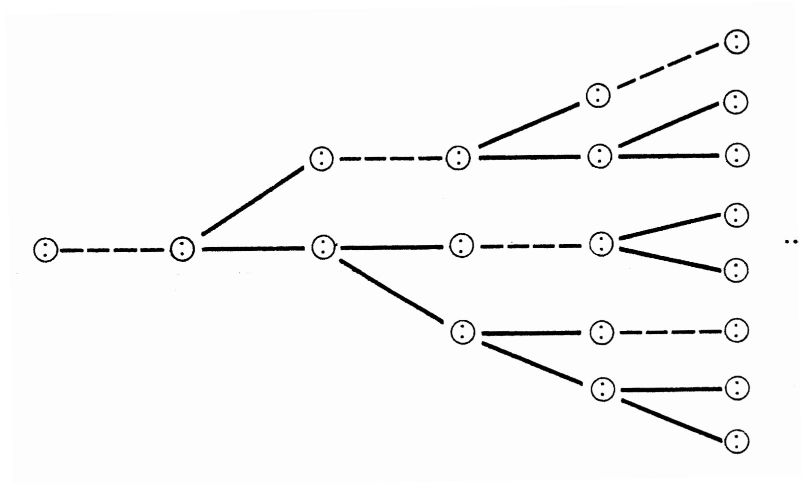
The mating takes place only between the members of a pair born from the same parents.

They live forever!

(Excepting these particulars, the rabbits do resemble in all other respects their more usual mortal counterparts.)

How many pairs of rabbits will Fibonacci have at the beginning of the nth month?

The picture below shows the beginning values of the sequence a_n = the number of pairs of rabbits at the beginning of the nth month ($n \geq 0$). By ----- we indicate the month to mature, and ____ indicates the month of pregnancy. We see from above that $a_0 = 1, a_1 = 1, a_2 = 2, a_3 = 3, a_4 = 5, a_5 = 8, \ldots$.



The sequence (a_n) satisfies in fact the recurrence relation

$$a_{n+2} = a_{n+1} + a_n; \quad n \geq 0.$$

(To see this observe that at stage $n+2$ we have all the a_{n+1} pairs that we had at stage $n+1$ plus the a_n children or grandchildren of the pairs we had at stage n, that is, $a_{n+2} = a_{n+1} + a_n$.)

To find a_n as a function of n only we proceed as follows. Denote by $A(x)$ the generating function of (a_n), that is, $A(x) = \sum_n a_n x^n$. Then

$$a_{n+2} = a_{n+1} + a_n$$

implies

$$a_{n+2}x^{n+2} = a_{n+1}x^{n+2} + a_n x^{n+2}$$

implies

$$\sum_n a_{n+2}x^{n+2} = \sum_n a_{n+1}x^{n+2} + \sum_n a_n x^{n+2}$$

implies

$$A(x) - a_1 x - a_0 = x(A(x) - a_0) + x^2 A(x)$$

implies

$$A(x) - x - 1 = xA(x) - x + x^2 A(x),$$

which leads to

$$A(x) = \frac{1}{1 - x - x^2}.$$

We use this closed form expression of $A(x)$ to find an explicit power series expansion for $A(x)$. Observe first that $1 - x - x^2 = -(a - x)(b - x)$, where $a = \frac{1}{2}(-1 - \sqrt{5})$ and $b = \frac{1}{2}(-1 + \sqrt{5})$. Now

$$\begin{aligned}
A(x) &= \frac{1}{1 - x - x^2} = \frac{-1}{(a - x)(b - x)} \\
&= (a - b)^{-1}\left((a - x)^{-1} - (b - x)^{-1}\right) \\
&= (a - b)^{-1}\left(a^{-1}\left(1 - \frac{x}{a}\right)^{-1} - b^{-1}\left(1 - \frac{x}{b}\right)^{-1}\right) \\
&= (a - b)^{-1}\left(a^{-1}\sum_n \left(\frac{x}{a}\right)^n - b^{-1}\sum_n \left(\frac{x}{b}\right)^n\right) \\
&= \sum_n \left[(a - b)^{-1}(a^{-n-1} - b^{-n-1})\right] x^n.
\end{aligned}$$

Hence $a_n = (a - b)^{-1}(a^{-n-1} - b^{-n-1})$, or

$$a_n = \frac{1}{\sqrt{5}}\left[\left(\frac{2}{-1 + \sqrt{5}}\right)^{n+1} - \left(\frac{2}{-1 - \sqrt{5}}\right)^{n+1}\right], \quad n \geq 0.$$

In general an explicit expression for a_n in terms of n only (although not always desirable) usually gives a more accurate idea of the magnitude of a_n, a fact that the recurrence might not immediately convey. We have thus found how many pairs of rabbits Fibonacci will have at the beginning of the nth month.

Note: If we expand the generating function $A(x) = (1 - (x + x^2))^{-1}$ as the power series $1 + (x + x^2) + (x + x^2)^2 + (x + x^2)^3 + \cdots$ what expression for the Fibonacci numbers do we obtain?

2.11

The case of the Fibonacci sequence, which we just described, is part of a more general class of problems known as linear recurrence relations with constant coefficients.

$*$ *Let* (a_n) *be a sequence satisfying the recurrence relation*

$$c_0 a_n + c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k} = 0; \tag{2.3}$$

$$c_0 = 1; \quad c_k \neq 0; \quad n \geq k$$

with c_i*'s constants* (not depending on n). *Then the generating function of* (a_n) *is of the form*

$$\frac{p(x)}{q(x)} \tag{2.4}$$

where $q(x)$ *is a polynomial of degree* k *with a nonzero constant term and* $p(x)$ *is a polynomial of degree less than* k.

Conversely, given polynomials $p(x)$ *and* $q(x)$ *as in* (2.4), *there exists a sequence* (a_n) *that satisfies a recurrence relation as in* (2.3) *and whose generating function is* $p(x)/q(x)$.

Indeed, suppose (a_n) satisfies (2.3) and has initial values $a_0, a_1, \ldots, a_{k-1}$. Proceed exactly as in the case of the Fibonacci sequence treated in Section 2.10 to obtain $A(x)$, the generating function of (a_n). In fact, $A(x) = p(x)/q(x)$, where $q(x) = \sum_{i=0}^{k} c_i x^i$, and $p(x) = \sum_{j=0}^{k} (\sum_{i=0}^{k-j-1} a_i x^i)$.

Conversely, given $q(x) = b_0 + b_1 x + \cdots + b_k x^k$ with $b_0 \neq 0$, $b_k \neq 0$ and $p(x) = d_0 + d_1 x + \cdots + d_{k-1} x^{k-1}$, using partial fractions and the expansion $1 - y^{-1} = \sum_n y^n$ we can write

$$\frac{p(x)}{g(x)} = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots \tag{2.5}$$

Rewrite (2.5) as follows:

$$\begin{aligned} d_0 + d_1x + \cdots + d_{k-1}x^{k-1} &= (b_0 + b_1x + \cdots + b_kx^k) \\ &\quad \cdot(a_0 + a_1x + a_2x^2 + \cdots). \end{aligned}$$

Identifying coefficients of powers of x on both sides we obtain

$$\begin{aligned} b_0a_0 &= d_0 \\ b_0a_1 + b_1a_0 &= d_1 \\ &\vdots \\ b_0a_{k-1} + b_1a_{k-2} + \cdots + b_{k-1}a_0 &= d_{k-1} \end{aligned} \tag{2.6}$$

and

$$b_0a_n + b_1a_{n-1} + \cdots + b_ka_{n-k} = 0, \quad \text{for } n \geq k.$$

Divide this last relation by b_0 and set $c_j = b_j/b_0$ to obtain the recurrence relation mentioned in (2.3). The initial values $a_0, a_1, \ldots, a_{k-1}$ can be determined from (2.6).

2.12

Merely as an exercise, *consider finding all sequences* (a_n) *that satisfy the recurrence relation*

$$a_{n+1} = 3a_n - 5(n+1) + 7 \cdot 2^n, \quad n \geq 0.$$

The way we proceed is typical of how one uses generating functions to solve problems of this sort.

Let $A(x) = \sum_n a_n x^n$. Then

$$\begin{aligned}
a_{n+1}x^n &= 3a_n x^n - 5(n+1)x^n + 7 \cdot 2^n x^n \\
\sum_n a_{n+1}x^n &= 3\sum_n a_n x^n - 5\sum_n (n+1)x^n + 7\sum_n 2^n x^n \\
x^{-1}(A(x) - a_0) &= 3A(x) - 5(1-x)^{-2} + 7(1-2x)^{-1} \\
(1-3x)A(x) &= a_0 - 5x(1-x)^{-2} + 7x(1-2x)^{-1} \\
A(x) &= \frac{a_0}{1-3x} - 5x\frac{1}{(1-3x)(1-x)^2} + 7x\frac{1}{(1-3x)(1-2x)}.
\end{aligned}$$

We expand $A(x)$ in a power series again, but first we use partial fraaction decompositions as follows:

$$\frac{1}{(1-3x)(1-x)^2} = \frac{A}{1-3x} + \frac{Bx+C}{(1-x)^2},$$

which upon solving for A, B, and C gives $A = \frac{9}{4}$, $B = \frac{3}{4}$, $C = -\frac{5}{4}$. Similarly

$$\frac{1}{(1-3x)(1-2x)} = \frac{3}{1-3x} - \frac{2}{1-2x}.$$

We now proceed

$$\begin{aligned}
A(x) &= \frac{a_0}{1-3x} - 5x\left(\frac{\frac{9}{4}}{1-3x} + \frac{\frac{3}{4}x - \frac{5}{4}}{(1-x)^2}\right) + 7x\left(\frac{3}{1-3x} + \frac{2}{1-2x}\right) \\
A(x) &= a_0\sum_n (3x)^n - \frac{45}{4}x\sum_n (3x)^n - \frac{15}{4}x^2\sum_n (n+1)x^n \\
&\quad + \frac{25}{4}x\sum_n (n+1)x^n + 21x\sum_n (3x)^n - 14x\sum_n (2x)^n.
\end{aligned}$$

Looking at the coefficient of x^n we immediately obtain

$$a_0, a_1 = 3a_0 + 2,$$

and

$$a_{n+2} = \left(3a_0 + \frac{39}{4}\right)3^{n+1} - 7 \cdot 2^{n+2} + \frac{10}{4}(n+1) + \frac{25}{4}, \quad n \geq 0.$$

This sequence does indeed verify the original recurrence.

2.13

Let us count the number of permutations σ *on the set* $1 < 2 < 3 < \cdots < n$ *that satisfy* $\sigma(1) > \sigma(2) < \sigma(3) > \sigma(4) < \cdots$. (*The signs* $>$ *and* $<$ *alternate*). Denote by a_n the number of such permutations.

To begin with, let us look at the initial values of the sequence a_n:

n:	1	2	3	4
σ:	1	21	213 312	2143 3142 3241 4132 4231
a_n:	1	1	2	5

It is well worth observing that the sequence (a_n) satisfies the recurrence

$$a_{n+1} = \sum_{\substack{k=0 \\ (k \text{ even})}} \binom{n}{k} a_k a_{n-k} \tag{2.7}$$

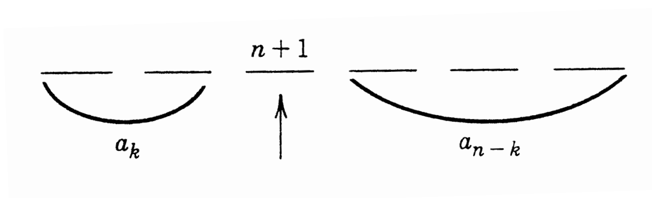
where, for convenience, we define $a_0 = 1$.

We explain this for $n = 5$ and the argument will carry over to any value of n. Take any permutation σ that satisfies $\sigma(1) > \sigma(2) < \sigma(3) > \sigma(4) < \sigma(5)$, say $\sigma = 32514$. Then $n + 1$, in this case 6, can be inserted in all the "even" positions in a to produce permutations on 6 symbols with the same property, that is, 3 2 5 1 4

$\uparrow$ $\uparrow$ $\uparrow$
Possible places to insert 6.

(An "even" position is defined by the fact that there is an even number of symbols at the left of the place where $n+1$ is inserted – note that $n+1$ may not be inserted in "odd" positions.)

For fixed σ the insertion of $n+1$ in position k (k even) produces precisely $\binom{n}{k} a_k a_{n-k}$ new permutations (on $n+1$ symbols).



Indeed, for each selection of digits at the left of $n+1$ there are $a_k a_{n-k}$ possible permutations. And there are $\binom{n}{k}$ possible choices for the digits left of $n+1$.

Summing up we obtain the recurrence in (2.7).

(The case $k=0$ requires, in fact, special attention. Note that 6 3 2 5 1 4 is not of the form we want. But by changing the digits $1 \leftrightarrow n, 2 \leftrightarrow n-1$, etc we produce 6 3 4 1 5 2, which is fine. This being a bijection the recurrence (2.7) remains valid as stated.)

Denote by $A(x) = \sum_n a_n x^n/n!$ the exponential generating function of (a_n). With the substitution $b_n = a_n/n!$ the recurrence relation (2.7) can be written as

$$(n+1)b_{n+1} = \sum_{\substack{k=0 \\ (k \text{ even})}} b_k b_{n-k}.$$

This leads to

$$(n+1)b_{n+1}x^n = (b_k b_{n-k})\, x^n.$$

Summing, we obtain

$$\begin{aligned}\sum_n (n+1)b_{n+1}x^n &= \sum_n \left(\sum_{\substack{k=0 \\ (k \text{ even})}} b_k b_{n-k} \right) x^n \\ &= (b_0 + b_2x^2 + b_4x^4 + b_6x^6 + \cdots) \left(\sum_n b_n x^n \right) \\ &= \frac{1}{2}(A(x) + A(-x))A(x).\end{aligned}$$

Since $\sum_n (n+1)b_{n+1}x^n = DA(x)$ [the formal derivative of $A(x)$] we obtain

$$DA(x) = \frac{1}{2}(A(x) + A(-x))A(x). \tag{2.8}$$

This functional equation, along with the knowledge that the constant term is 1, force a unique solution for $A(x)$. Indeed, (2.8) and the constant term being 1, determine uniquely the coefficient of x, then that of x^2, of x^3, and so on.

If we denote $1-(x^2/2!)+(x^4/4!)-(x^6/6!)\pm\cdots$ by $\cos x$ and $(x/1!)-(x^3/3!)+(x^5/5!)-(x^7/7!)\pm\cdots$ by $\sin x$, a solution (and therefore the solution) to (2.8) is $(\sin x/\cos x)+(1/\cos x)$. If, by analogy to the notation in trigonometry, we further denote $(\sin x/\cos x)$ by $\tan x$ and $(1/\cos x)$ by $\sec x$, the unique solution to (2.8) can be written as

$$A(x) = \tan x + \sec x.$$

We conclude, therefore, that *the exponential generating function for sequence of permutations* (a_n) *defined at the beginning of this paragraph is* $A(x) = \tan x + \sec x$.

[While most of us surely can appreciate a wild guess that works, the claim that $\tan x + \sec x$ is a solution to (2.8) touches undeniably upon the miraculous. Let us sketch a proof that $A'(x) = \frac{1}{2}(A(x)+A(-x))A(x)$ and $A(0) = a_0 = 1$ imply $A(x) = \sec x + \tan x$ (here the prime denotes the derivative).

Let $B(x) = \frac{1}{2}(A(x) + A(-x))$ and $C(x) = \frac{1}{2}(A(x) - A(-x))$. Note that

$$\begin{aligned} B'(x) &= \frac{1}{2}(A'(x) - A'(-x)) = \frac{1}{4}(A(x) + A(-x))(A(x) - A(-x)) \\ &= B(x)C(x) \qquad (2.9) \\ C'(x) &= \frac{1}{2}(A'(x) + A'(-x)) = \frac{1}{4}(A(x) + A(-x))^2 = B(x)^2 \end{aligned}$$

Hence $(B(x)^2 - C(x)^2)' = 2B(x)B(x)C(x) - 2C(x)B(x)^2 = 0$. And since $B(0) = 1$ and $C(0) = 0$ we have

$$B(x)^2 - C(x)^2 = 1. \qquad (2.10)$$

Next note that and

$$\left(\frac{1}{B(x)}\right)' = -\frac{B'(x)}{B(x)^2} = -\frac{C(x)}{B(x)},$$

and

$$\begin{aligned} \left(\frac{1}{B(x)}\right)'' &= -\left(\frac{C(x)}{B(x)}\right)' = -\frac{C'B - CB'}{B^2} = \{\text{by } (2.9)\} \\ &= -\frac{B^3 - BC^2}{B^2} = -\frac{1}{B}(B^2 - C^2) = \{\text{by } (2.10)\} \\ &= -\frac{1}{B(x)}. \end{aligned}$$

Now,

$$\left(\frac{1}{B(x)}\right)'' = -\frac{1}{B(x)} \quad \text{and} \quad -\frac{1}{B(0)} = 1 \quad \text{imply} \quad -\frac{1}{B(x)} = \cos x.$$

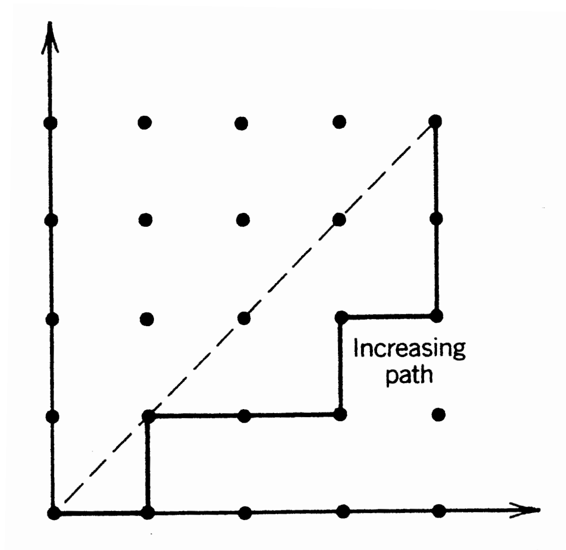
Hence $B(x) = \sec x$, and by (2.10) $C(x) = \pm \tan x$. By (2.9) $C(x) = \tan x$, necessarily. This gives

$$A(x) = B(x) + C(x) = \sec x + \tan x.]$$

EXERCISES

1. Let $c_n = (n+1)^{-1}\binom{2n}{n}$.

   (a) Find the number of increasing lattice paths from (0, 0) to (n, n) that never cross, but may touch, the main diagonal [i.e., the line joining (0, 0) with (n, n)].



Answer: $2c_n$

   (b) How many ways can the product $x_1 x_2 \cdots x_n$ be parenthesized? (Note: we do not allow the order of the x's to change.)

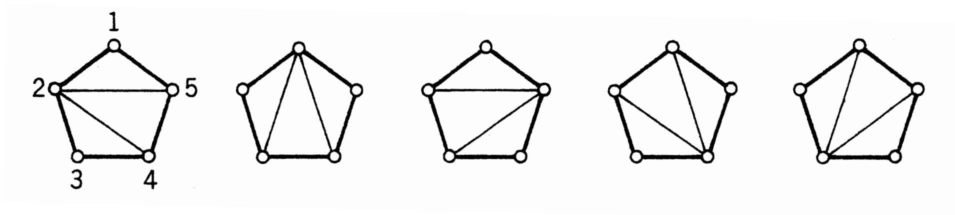
   Example: $n = 4$

$$((x_1x_2)(x_3x_4)), \quad (((x_1x_2)x_3)x_4), \quad ((x_1(x_2x_3))x_4),$$
$$(x_1((x_2x_3)x_4)), \quad (x_1(x_2(x_3x_4))).$$

Answer: c_{n-1}

   (c) Let P_n be the regular n-gon on n labeled vertices. A *diagonal triangulation* of P_n is a triangulation of P_n that involves exactly $n-3$ nonintersecting diagonals of P_n. Find the number of diagonal triangulations of P_n (Euler).

Example:



Answer: c_{n-2}

(d) Given $2n$ people of different heights, in how many ways can these $2n$ people be lined up in two rows of length n each so that everyone in the first row is taller than the corresponding person in the second row?

Answer: c_n

(e) (Application to politics.) In an election candidate A receives a votes and candidate B receives b votes ($a > b$). In how many ways can the ballots be arranged so that when they are counted, one at a time, there are always (strictly) more votes for A than B?

Answer: $((a-b)/(a+b))\binom{a+b}{a}$

(If the election ends in a tie with n votes to each, then the number of sequences in which at no time of the counting is B ahead is $2c_n$.)

(f) Show: $c_n = \sum_{k=0}^{n-1} c_k c_{n-k-1}$, $c_0 = 1$.

(g) Show: $\sum_n c_n x^n = (1 - \sqrt{1-4x})/2x$.

The (c_n)'s are called *Catalan numbers.*

$$c_n: \quad 1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, \ldots$$

2. Let (a_n) be a sequence satisfying the recurrence relation

$$a_n + a_{n-1} - 16a_{n-2} + 20a_{n-3} = 0, \quad n \geq 3$$

with $a_0 = 0, a_1 = 1, a_2 = -1$. Find a_n (as a function of n).

3. Let (a_n) be the Fibonacci sequence (take $a_0 = 0, a_1 = 1, a_2 = 1$ and $a_n = a_{n-1} + a_{n-2}$, $n \geq 3$). Verify that:

   (a) $a_1 + a_2 + \cdots + a_n = a_{n+2} - 1$.

   (b) $a_1 + a_3 + a_5 + \cdots + a_{2n-1} = a_{2n}$.

   (c) $a_2 + a_4 + a_6 + \cdots + a_{2n} = a_{2n+1} - 1$.

   (d) $a_n^3 + a_{n+1}^3 - a_{n-1}^3 = a_{3n}$.

   (e) $\binom{n}{0} + \binom{n-1}{1} + \binom{n-2}{2} + \cdots = a_{n+1}$.

   (f) $a_{n+m} = a_m a_{n+1} + a_{m-1} a_n$. Show also that a_{mn} is a multiple of a_n.

   (g) a_n is $(1/\sqrt{5})((1+\sqrt{5})/2)^n$ rounded off to the nearest integer.

   (h) $a_1 a_2 + a_2 a_3 + \cdots + a_{2n-1} a_{2n} = a_{2n}^2$.

4. Place n points on the circumference of a circle and draw all possible chords through pairs of these points. Assume (at least formally) that no three chords are concurrent. Let a_n be the number of regions formed inside the circle. Find (a_n) and the generating function of (a_n).

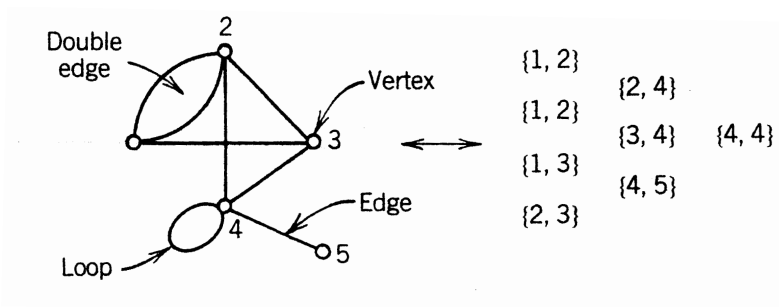
5. Define a_0 to be 1. For $n \geq 1$, let a_n be the number of $n \times n$ symmetric matrices with entries 0 or 1 and row sums equal to 1 (i.e., symmetric permutation matrices). Show that $a_{n+1} = a_n + n a_{n-1}$ and then prove that $\sum_n a_n x^n / n! = \exp(x + \frac{1}{2}x^2)$.

6 THE GENERATING FUNCTION OF LABELED SPANNING TREES

Let us temporarily drift away from generating functions of sequences to present a result in graph theory: the generating function for the spanning trees of a graph.

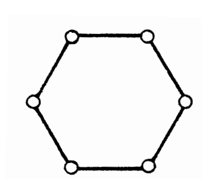
2.14

A *graph* G is a collection of (possibly repeated) subsets of cardinality two (called *edges*) of a finite set of points (called *vertices*). Below is an example of a graph:



In the definition of a graph we also allow the notation $\{4,4\}$ for an edge joining the vertex 4 to itself, which we call a *loop*. All edges, including the multiple ones, are *distinguishable* from each other. A *path* is a collection of edges like this

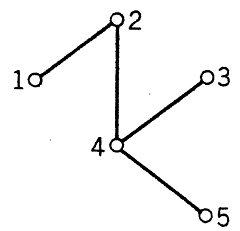
(any length). A *cycle* is a collection of edges like this:



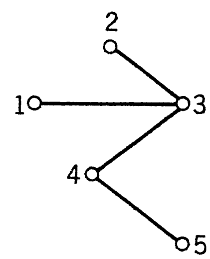
(any length). We call a graph *connected* if any two distinct vertices can be joined by a path. A *tree* is a set of edges containing no cycles. By a *spanning tree* of a graph with n

vertices we understand a set of $n-1$ edges containing no cycles. (Graphs that are not connected have no spanning trees.) A path, cycle, tree, or spanning tree is understood to contain no loops or multiple edges.

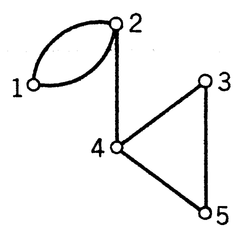
The following two pictures are spanning trees in the graph above:

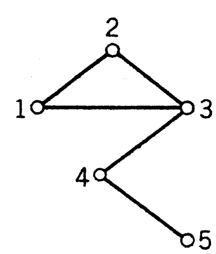


and



These two are not:





Two spanning trees are the *same* if they consist of exactly the same $n-1$ edges. (There are two spanning trees associated with our first picture of a spanning tree above, because there are two *distinguishable* edges $\{1,2\}$ in G.) Given a graph G we address two issues

(i) *How many spanning trees does G have*?

(ii) *Generate a list of all spanning trees of G.*

2.15

To a graph G we associate its *information* (or *Kirchhoff*) matrix $C = (c_{ij})$ (both rows and columns indexed by the vertices of G in the same fixed order) as follows:

$$-c_{ij} \;=\; \text{number of edges between vertices } i \text{ and } j, \quad i \neq j$$

$$c_{ii} \;=\; -\sum_{\substack{j\\ j\neq i}} c_{ij}.$$

Suppose G has n vertices labeled $1, 2, \ldots, n$. The matrix C is then $n \times n$. Denote by $\mathbf{1}$ the column vector with all its entries 1 (and by $\mathbf{1}^t$ its transpose). Properties of the matrix C:

1. $C\mathbf{1} = \mathbf{0}$ (i.e., $\mathbf{1} \in \ker C = \text{kernel of } C$).
2. *If rank* $C = n - 1$, *then all cofactors of* C *are equal and nonzero.*
3. $C \geq 0$ (i.e., $x^t C x \geq 0$, for all vectors x).
4. C *is of rank* $n - 1$ *if and only if* G *is connected.*

Proof. Statement 1 follows from the definition of C. To realize that statement 2 is true denote by C_{ij} the cofactor of c_{ij}. Then

$$(c_{ij})(C_{ij})^t = (\det C)I$$

where $\det C$ stands for the determinant of C and I is the $n \times n$ identity matrix. This equality holds for *any* square matrix. In our case $\det C = 0$ since C is singular, by statement 1. Hence all column vectors of $(C_{ij})^t$ belong to $\ker C = \langle \mathbf{1} \rangle$ (because the rank of C is $n-1$). Thus for fixed i all C_{ij}'s are equal. Similarly (working with transposes) for fixed j, all C_{ij}'s are equal and therefore (C_{ij}) is a multiple of J, the matrix with all entries 1. This proves statement 2. For an edge $\{i, j\}$ in G denote by C^{ij} the Kirchoff matrix of the graph on n vertices and with $\{i, j\}$ the only edge (of multiplicity 1). Then C^{ij} is the $n \times n$ matrix $\begin{bmatrix} & 1 & -1 \\ & & \\ -1 & & 1 \end{bmatrix}$ with 1's in ith and jth diagonal positions, -1 in positions

(i,j) and (j,i), and 0 elsewhere. It is easy to check that $x^tC^{ij}x \geq 0$, for all x. (Note that $C^{ii} = 0$, i.e., the Kirchhoff matrix of a loop is zero.) Then

$$C = \sum_{\substack{\{i,j\} \\ \text{edge of} \\ \text{graph } G}} C^{ij} \quad \text{and} \quad x^tC^{ij}x = \sum x^tC^{ij}x \geq 0,$$

which gives statement 3. (The expression of C as a sum of C^{ij}'s is important because it shows how C gathers "information.") We now prove statement 4. If $A \geq 0$, $B \geq 0$, and $B \leq A$ (notation for $A - B \geq 0$), then the row span of B is included in the row span of A [because $\ker A \subseteq \ker B$ – this is easy to check – and the column (or row) span of B is included in the column (or row) span of A as orthogonal complements of kernels]; keep this in mind. Let G be connected. Then a path γ exists between 1 and any other vertex k. Say the path is $(12\cdots k)$ (without loss). Then the Kirchhoff $n \times n$ matrix of the path is

$$C_\gamma = \begin{bmatrix} 1 & -1 & & & & & \\ -1 & 2 & -1 & & & \mathbf{0} & \\ & -1 & 2 & -1 & & & \\ & & & \ddots & & & \\ & & & -1 & 2 & -1 & \\ & & & & -1 & 1 & \\ & & & & & & \mathbf{0} \end{bmatrix}$$

with the $\mathbf{0}$ in the bottom right-hand corner of dimension $(n-k) \times (n-k)$. Let $e_i = (0\cdots 010\cdots 0)$ with 1 in the ith place. The first row of C_γ is $f_1 = e_1 - e_2$, the sum of first two rows gives $f_2 = e_2 - e_3, \ldots$, the sum of first $k-1$ rows gives $f_{k-1} = e_{k-1} - e_k$. Then $\sum_{i=1}^m f_i = e_1 - e_m$, $2 \leq m \leq k$, are in the row span of C_γ. But for any path γ,

$C = C_\gamma + C_\beta$, where β is the set of edges in G but not in γ, that is, $C_\beta = \sum_{\{i,j\}\in\beta} C^{ij}$. Clearly $C_\gamma \leq C$, and by the above remark $e_1 - e_k$ is also in the row span of C; $2 \leq k \leq n$. These $n-1$ vectors span a subspace of dimension $n-1$. The converse is easy. If G is not connected, then C can be written as

$$C = \begin{bmatrix} C_1 & \mathbf{0} \\ \mathbf{0} & C_2 \end{bmatrix}$$

where C_1 is the Kirchhoff matrix of a connected part (or component) of G. The vectors ($\mathbf{1}$,$\mathbf{0}$) and ($\mathbf{0}$,$\mathbf{1}$) are both in the kernel of C, showing that C can be of rank $n-2$ at the most [$\mathbf{1}$ in ($\mathbf{1}$, $\mathbf{0}$) has $|C_1|$ entries, and $\mathbf{1}$ in ($\mathbf{0}$,$\mathbf{1}$) has $|C_2|$ entries, or coordinates]. This proves statement 4.

2.16

Let G be a graph. We label by the indeterminate x_{ij} the edge between vertices i and j (if there are multiple edges between i and j we use $x_{ij}^{(1)}, x_{ij}^{(2)}, \ldots$, etc.). To each spanning tree of G we associate a monomial of degree $n-1$, the product of all x_{ij}'s, where $\{i,j\}$'s are the $n-1$ edges of the spanning tree.

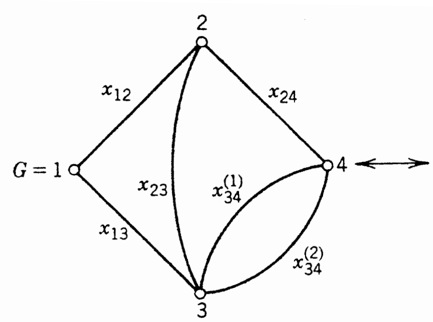
Let $C(G)$ be the (vertex versus vertex) matrix with off diagonal (i,j)th entry $-x_{ij}$ (if multiple edges $-\sum_k x_{ij}^{(k)}$), 0 if there is no edge between i and j, and ith diagonal entry the negative of the sum of the off-diagonal entries in the ith row. (If G has n vertices, then $\overline{C}(G)$ is $n \times n$ with zero row and column sums.)

We now return to the issues considered at the end of Section 2.14, accomplishing (ii) and answering (i).

∗ *Let* G *be a graph with matrix* $\overline{C}(G)$*. Delete a row and* (not necessarily same) *column of* $\overline{C}(G)$*. Denote the resulting matrix by* K*. Let* $\det K$ *be* (the formal expansion of) *the determinant of* K*. Then the monomials in the expansion of* $\det K$ (after cancellations) *are all square free and give a complete list of all spanning trees of* G*.* (Each monomial corresponds uniquely to a spanning tree.) *When setting all* x_{ij}*'s equal to 1 in* $\overline{C}(G)$ $\det K$ *equals* (up to sign) *the number of spanning trees of* G*.*

We call $\det K$ the generating function of the spanning trees of G.

The proof of this result may best be illustrated by an example that captures all the relevant features of a general proof:



$$\begin{bmatrix} x_{11} & -x_{12} & -x_{13} & 0 \\ -x_{12} & x_{22} & -x_{23} & -x_{24} \\ -x_{13} & -x_{23} & x_{33} & -x_{34}^{(1)} - x_{34}^{(2)} \\ 0 & -x_{24} & -x_{34}^{(1)} - x_{34}^{(2)} & x_{44} \end{bmatrix} = \overline{C}(G).$$

[Recall that $x_{ii} = -$(sum of the off-diagonal entries in row i).]

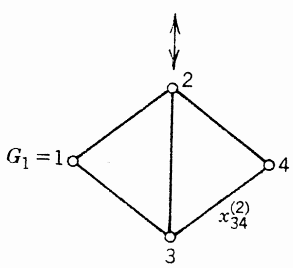
The general idea of the proof is as follows: Select an edge of G (say $x_{34}^{(1)}$). Partition the spanning trees of G into those that do not contain the edge $x_{34}^{(1)}$ and those that do. The first class can be identified with the spanning trees of the graph $G_1 = \{G \text{ without edge } x_{34}^{(1)}\}$,while the second class consists of spanning trees (augmented with edges $x_{34}^{(1)}$) of the graph G_2, obtained from G by shrinking edge $x_{34}^{(1)}$ into a point (thus making vertices 3 and 4 the same vertex and deleting edge $x_{34}^{(1)}$). Both classes defined above involve listing spanning trees in graphs with one edge less than G (G_2 has also one vertex less) and hence we can complete the proof by induction on the number of edges of G.

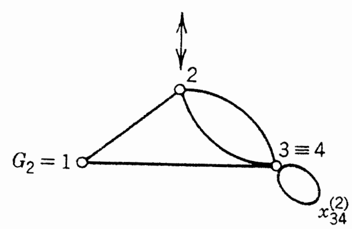
Obtain K by deleting row 4 and column 4 in $\overline{C}(G)$. (The fact that det K is independent of which row or column we delete in $\overline{C}(G)$ to obtain K can be proved as property 2 of matrices C discussed in Section 2.15.) We obtain

$$\det K = \begin{vmatrix} x_{11} & -x_{12} & -x_{13} \\ -x_{12} & x_{22} & -x_{23} \\ -x_{13} & -x_{23} & x_{13}+x_{23}+x_{34}^{(2)}+x_{34}^{(1)} \end{vmatrix}$$

$$= \begin{vmatrix} x_{11} & -x_{12} & -x_{13} \\ -x_{12} & x_{22} & -x_{23} \\ -x_{13} & -x_{23} & x_{13}+x_{23}+x_{34}^{(2)} \end{vmatrix} + \begin{vmatrix} x_{11} & -x_{12} & 0 \\ -x_{12} & x_{22} & 0 \\ 0 & 0 & x_{34}^{(1)} \end{vmatrix}$$

$$\overline{C}(G_1) = \begin{bmatrix} x_{11} & -x_{12} & -x_{13} & 0 \\ -x_{12} & x_{22} & -x_{23} & -x_{24} \\ -x_{13} & -x_{23} & x_{13}+x_{23}+x_{34}^{(2)} & -x_{34}^{(2)} \\ 0 & -x_{24} & -x_{34}^{(2)} & x_{24}+x_{34}^{(2)} \end{bmatrix} \quad \overline{C}(G_2) = \begin{bmatrix} x_{11} & -x_{12} & -x_{13} \\ -x_{12} & x_{22} & -x_{23}-x_{24} \\ -x_{13} & -x_{23}-x_{24} & x_{13}+x_{23}+x_{24} \end{bmatrix}$$





The matrix $\overline{C}(G_1)$ is obtained from $\overline{C}(G)$ by setting $x_{34}^{(1)} = 0$. Add row 4 to row 3 and column 4 to column 3 in $\overline{C}(G)$, then delete row and column 4, to obtain $\overline{C}(G_2)$. [Note that by just looking at G_2 it is not clear whether x_{13} or x_{14} is an edge. But $\overline{C}(G_2)$ clears this up: x_{13} *is* an edge, x_{14} *is not*.]

The first determinant gives the list of trees not containing $x_{34}^{(1)}$ (upon expansion and cancellation). They are

$$x_{12}x_{23}x_{34}^{(2)} + x_{23}x_{34}^{(2)}x_{14} + x_{34}^{(2)}x_{14}x_{12} + x_{14}x_{12}x_{23}$$

$$+x_{12}x_{24}x_{34}^{(2)} + x_{23}x_{24}x_{14} + x_{12}x_{24}x_{23} + x_{14}x_{24}x_{34}^{(2)} \quad (8 \text{ in all}).$$

The second determinant gives the spanning trees containing $x_{34}^{(1)}$:

$$x_{12}x_{13}x_{34}^{(1)} + x_{12}x_{23}x_{34}^{(1)} + x_{12}x_{24}x_{34}^{(1)}$$

$$+x_{13}x_{23}x_{34}^{(1)} + x_{13}x_{24}x_{34}^{(1)} \quad (5 \text{ in all}).$$

Hence G contains 13 spanning trees. Indeed, when all $x_{ij} = 1$ $\det K$ becomes

$$\begin{vmatrix} 2 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 4 \end{vmatrix} = 13.$$

Application to Optimal Statistical Design

The information (or Kirchhoff) matrix C, introduced in Section 2.15, is an important representative of a class of matrices known to statisticians as Fisher information matrices (also known as C-matrices). They capture all the relevant statistical information locked into the actual planning (or design) of an experiment. Without dwelling on the general concerns that surround the planning, we wish to point out (in purely mathematical terms) a specific problem that often arises and that, as yet, has not been brought to a satisfactory solution:

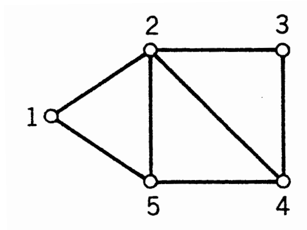
Among all graphs with n vertices and m edges identify those with a maximum number of (labeled) spanning trees.

An understanding of the structure of such graphs translates directly into optimum ways of planning experiments. The resulting design will be called D-optimal by the statistician.

It might not be surprising to find that the Kirchhoff tree generating matrix plays an important part in the solution. For the necessary background in statistics we refer the reader to Chapter 8.

EXERCISES

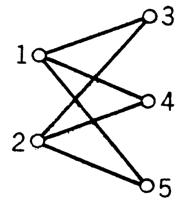
1. How many (labeled) spanning trees does the graph displayed below have?



2. Let $0 = \mu_0(G) \leq \mu_1(G) \leq \cdots \leq \mu_{n-1}(G)$ be the eigenvalues of the Kirchhoff matrix $C(G)$ of a graph G on n vertices. Show that $n^{-1}\prod_{i-1}^{n-1}\mu_i(G)$ = number of labeled spanning trees of G.

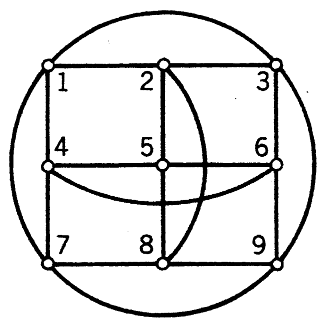
3. A graph is called *simple* if between any two vertices there is at most one edge and no loops are allowed. By K_n we denote a simple graph on n vertices with an edge between any two vertices. We call K_n the *complete graph*; K_n has $\binom{n}{2}$ edges. How many labeled spanning trees does K_n have?

4. Partition $n_1 + n_2 + \cdots + n_m$ vertices into m classes, the ith class containing n_i vertices. Produce a simple graph $K(n_1, n_2, \ldots, n_m)$ by joining each vertex in class i to all vertices outside class i (and to none within class i); do this for all i. The resulting graph is called the *complete multipartite graph* $K(n_1, n_2, \ldots, n_m)$. For example $K(2, 3)$ is



How many labeled spanning trees does $K(n_1, n_2, \ldots, n_m)$ have?

5. Place n^2 vertices into an $n \times n$ square array and join two vertices if and only if they are in the same row or same column. Call the resulting graph S_n. Compute the number of labeled spanning trees of S_n. (S_3 is drawn below.)



6. A graph is called *regular* if each of its vertices has the same degree. The *complementary graph* $\overline{G}$ of a simple graph G is the graph on the same set of vertices as G whose edges are precisely those that are missing in G. For G a regular and simple graph relate the eigenvalues of $C(\overline{G})$ to those of $C(G)$, and (with the help of Exercise 2) obtain a relationship between the number of labeled spanning trees in G and $\overline{G}$.

7. Show that among all graphs on n vertices and e edges (with e sufficiently large) those that have a maximal number of labeled spanning trees must have the degrees of their vertices differ by at most 1 and the number of edges between any two vertices differ by at most 1. [*Hint*: look at $\prod_{i=1}^{n-1}(\mu_i + x)$ for large values of x.]

7 PARTITIONS OF AN INTEGER

We touch only briefly here upon a rich and well-developed subject: that of partitioning an integer.

2.17

The question we raise regards *the number of* (unordered) *ways of writing the number* n *as the sum of exactly* m *positive integers.* Let us call this number $P_m(n)$. More rigorously,

$$\begin{aligned} & |\{(\alpha_1, \alpha_2, \ldots, \alpha_m) : \text{ each } \alpha_i \text{ is a positive} \\ P_m(n) \quad = \quad & \text{integer}, \ \alpha_1 + \alpha_2 + \cdots + \alpha_m = n, \text{ and} \\ & \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_m \geq 1\}|. \end{aligned}$$

The α_i's are called the *parts* of n. Clearly $m \leq n$.

The number of ways of writing n as the sum of 1 integer, as the sum of $n-1$ integers, or as the sum of n integers is unique, so that $P_1(n) = P_n(n) = P_{n-1}(n) = 1$.

We wish to find a pattern, a simple recurrence relation, for $P_m(n)$. *Our first result is the following*:

$$P_1(n) + P_2(n) + \cdots + P_k(n) = P_k(k+n), \quad \text{for } k \leq n.$$

Proof. Let

$$\begin{aligned} P &= \{\text{partitions of } n \text{ into } k \text{ or fewer parts}\} \\ &= \left\{(\alpha_1, \alpha_2, \ldots, \alpha_m, 0, \ldots, 0) \in \{k-\text{tuples}\} \ \sum_i^m \alpha_i = n, m \leq k \right\}. \end{aligned}$$

Define a mapping on P as follows:

$$(\alpha_1, \alpha_2, \ldots, \alpha_m, 0, \ldots, 0) \rightarrow (\alpha_1 + 1, \alpha_2 + 1, \ldots, \alpha_m + 1, 1, \ldots, 1).$$

The image is a k-tuple again, and the number of single 1's is the same as the number of 0's in its preimage.

Note that the image corresponds, in fact, to a partition of $k+n$ into k parts. This mapping is injective, and for each partition of $k+n$ into k parts there is a k-tuple in P that is mapped into it, that is, the mapping is also onto the set of partitions of $n+k$ into k parts. Hence $|P| = |\text{image of } P| = P_k(n+k)$. But also $|P| = P_1(n)+P_2(n)+\cdots+P_k(n)$, from which the recurrence relation follows. This ends the proof.

For small values of m and n we have the following table for $P_m(n)$:

	m					
n	1	2	3	4	5	6
1	1	0	0	0	0	0
2	1	1	0	0	0	0
3	1	1	1	0	0	0
4	1	2	1	1	0	0
5	1	2	2	1	1	0
6	1	3	4	2	1	1

2.18 Ferrer Diagrams

We can also represent a partition by a *Ferrer diagram*, which will be very useful in visualizing many results. Given a partition we represent each part by the appropriate number of dots in a row and place the rows beneath one another. For example, the Ferrer diagram of the partition (6, 4, 3,1) is

$$\begin{array}{llllll l}
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & 6 \text{ dots} \\
\bullet & \bullet & \bullet & \bullet & & & 4 \text{ dots} \\
\bullet & \bullet & \bullet & & & & 3 \text{ dots} \\
\bullet & & & & & & 1 \text{ dot}
\end{array}$$

Given a partition $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ we define a new partition $(\alpha_1^*, \alpha_2^*, \ldots, \alpha_k^*)$, where α_i^* is the number of parts in α that are greater than or equal to i. The new partition α^* is called the *conjugate* of α. For example, if $\alpha = (5, 3, 2)$, then $\alpha^* = (3, 3, 2, 1, 1)$. The simplest and most visual way to construct α^* is by rotating the Ferrer diagram of α about the diagonal. (It is thus clear that $\alpha^{**} = \alpha$.) For example,

$$\alpha = \begin{array}{lllll}
\bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & & \\
\bullet & \bullet & & &
\end{array}
\quad \longleftrightarrow \quad
\begin{array}{lll}
\bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet \\
\bullet & \bullet & \\
\bullet & & \\
\bullet & &
\end{array} = \alpha^*$$

It is also clear from the way α^* is obtained on the Ferrer diagram that $\sum_1^m \alpha_i = \sum_1^k \alpha_i^*$, that is, if α is a partition of n, α^* is also a partition of n. The bijective correspondence between partitions of n and conjugate partitions suggests the following result:

$*$ *The number of partitions of n into k parts is equal to the number of partitions of n into parts the largest of which is k.*

Proof. Let $P = \{(\alpha_1, \ldots, \alpha_k) : \text{ partitions of } n \text{ into } k \text{ parts}\}$. The mapping $(\alpha_1, \ldots, \alpha_k) \rightarrow (\alpha_1^*, \alpha_2^*, \ldots)$ is a bijection, for the conjugate is obtained by a rotation of the Ferrer diagram.

Also, the largest part of $(\alpha_1^*, \alpha_2^*, \ldots)$ does not exceed k.

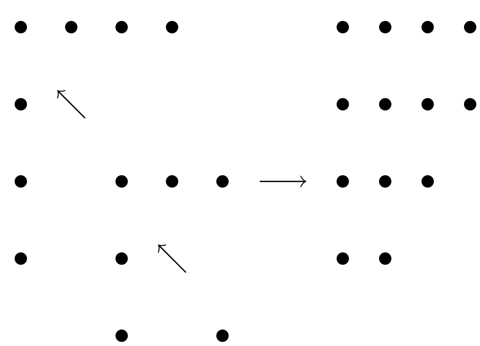
As an easy consequence we have:

$*$ *The number of partitions of n with at most k parts equals the number of partitions of n in which no part exceeds k.*

If $\alpha = \alpha^*$ we call α *self-conjugate*. Note that α is self-conjugate if and only if its Ferrer diagram is symmetric with respect to the diagonal. With this definition we have the following result:

$*$ *The number of self-conjugate partitions of n is equal to the number of partitions of n with all parts unequal and odd.*

Proof. Take each (odd) part of the initial partition, bend in the middle, and reassemble as indicated below:



We thus obtain a self-conjugate partition. This operation produces, in fact, a (visual) bijection from partitions of n with distinct and odd parts to self-conjugate partitions of n.

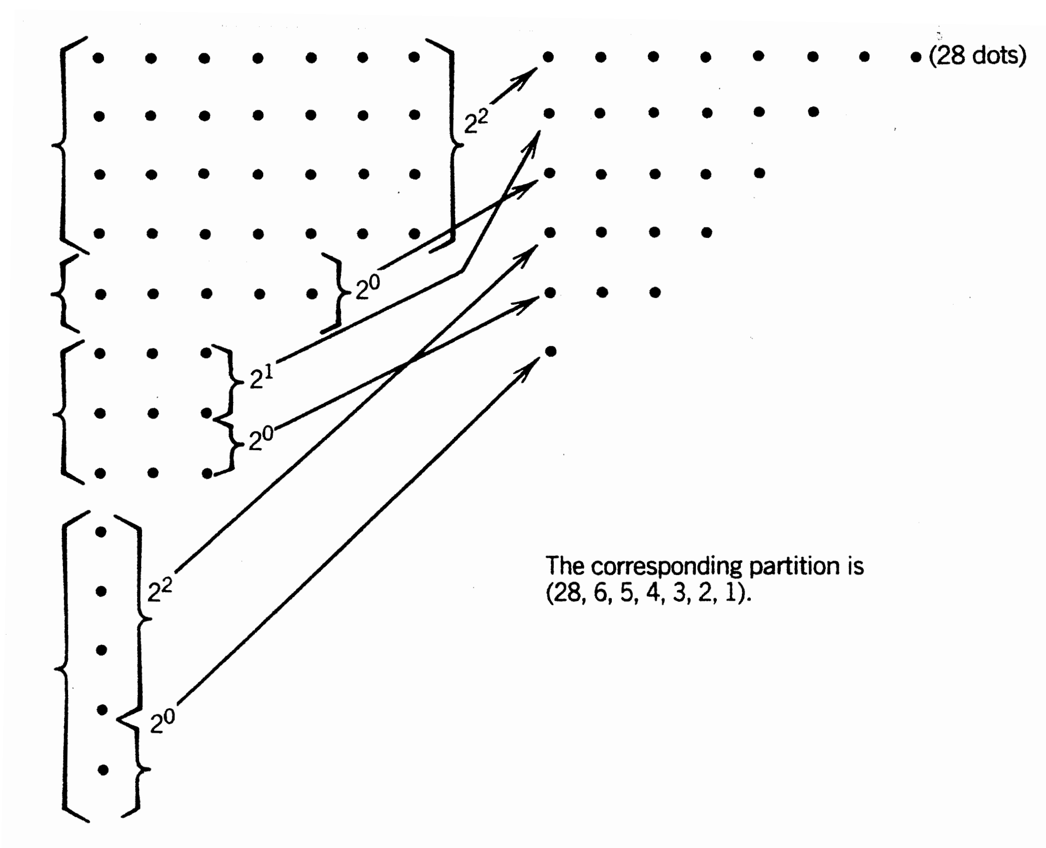
Another transformation of a Ferrer diagram is used in establishing the following result.

∗ *The number of partitions of n into unequal parts is equal to the number of partitions of n into odd parts.*

Proof. Consider a partition of n into odd parts. Write it as $n = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m$ with k_i the multiplicity of α_i (where the α_i's are odd).

We produce a new partition as follows: expand each k_i in binary base, say $k_i = \varepsilon_0 2^0 + \varepsilon_1 2^1 + \cdots + \varepsilon_{r_i} 2^{r_i}$. Group together $\varepsilon_s 2^s$ rows of α_i (attached to k_i), as s ranges between 0 and r_i, and $\varepsilon_s = 0$ or 1. Form the new partition of n with parts $\varepsilon_s 2^s \alpha_i$, as s and i take values in their respective ranges.

As an example, let $\alpha = (7, 7, 7, 7, 5, 3, 3, 3, 1, 1, 1, 1, 1)$, that is, $n = 4{\cdot}7 + 1{\cdot}5 + 3{\cdot}3 + 5{\cdot}1$. Then



Since the α_i's are odd, $2^a\alpha_i \neq 2^b\alpha_j$ for $i \neq j$ or $a \neq b$. Hence the above transformation sends odd partitions of n to unequal partitions, and it is clear that the new partitions are of n, because the total number of dots is preserved.

This transformation can be reversed in a unique way, for given $\beta = (\beta_1, \beta_2, \ldots, \beta_m)$ with β_i's distinct; write each β_i as a product of an odd number and a power of 2. This representation of the β_i's is unique. Therefore, if $\beta_i = 2^{r_i}\alpha_i$ with α_i odd, obtain a new partition α with parts α_i of appropriate multiplicities. Clearly α is a partition of n with odd parts. We have therefore a bijection between unequal partitions of n and odd partitions of n, proving the above statement.

Our next result can be stated as follows:

$*$ *Let $P(n; d, o)$ and $P(n; d, e)$ denote, respectively, the number of partitions of n into an odd/even number of distinct parts. Then*

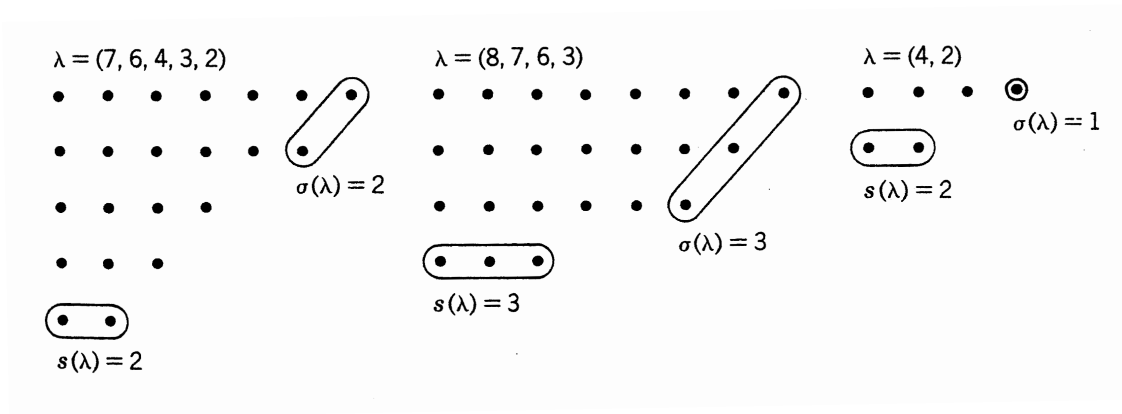
$$P(n; d, e) - P(n; d, o) = \begin{cases} (-1)^m & \text{if } \; n = m(3m+1)/2 \\ 0 & \textit{otherwise.} \end{cases}$$

(This result is known as *Euler's pentagonal theorem.*)

Proof. We initially try to establish a bijective correspondence between the distinct partitions of n into even parts and the distinct partitions of n into odd parts.

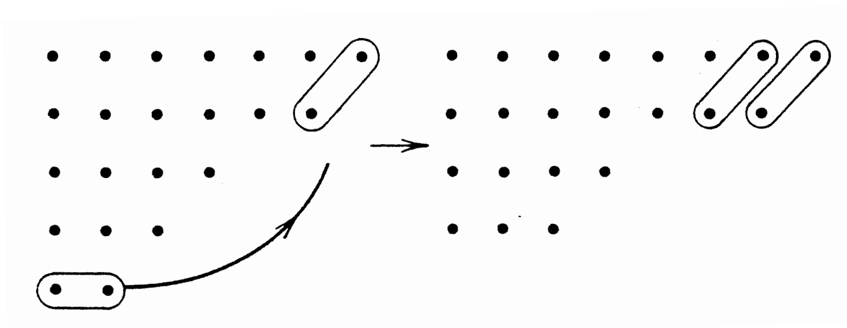
Given a partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_r)$ of n into distinct parts let $s(\lambda) = \lambda_r$, that is, $s(\lambda)$ is the smallest part of λ, and let $\sigma(\lambda)$ be the number of consecutive parts of λ from λ_1 down. [More formally, $\sigma(\lambda) = \max\{j : \lambda_j = \lambda_1 - j + 1\}$.]

Examples.

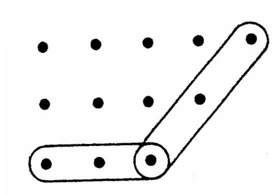


We separate the proof into two cases.

Case 1. $s(\lambda) \leq \sigma(\lambda)$. Add 1 to each of the first $s(\lambda)$ parts of λ and delete the smallest part. Thus $(7, 6, 4, 3, 2) \to (8, 7, 4, 3)$.

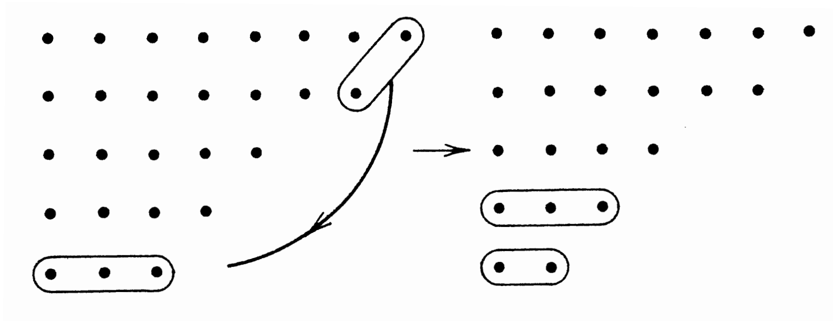


This transformation is always possible, except when the dots enumerated by $s(\lambda)$ and $\sigma(\lambda)$ meet [and $s(\lambda) = \sigma(\lambda)$], for example, if $\lambda = (5, 4, 3)$.

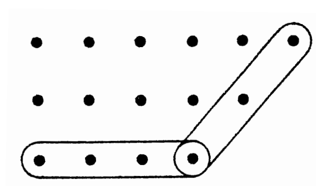


Then, if the initial partition had an odd number of distinct parts, the above transformation does not lead to a partition with an even number of parts. In all other cases, however, the above transformation establishes a bijective map between partitions into distinct, odd parts and partitions into distinct, even parts.

Case 2. $s(\lambda) > \sigma(\lambda)$. Subtract 1 from each of the $\sigma(\lambda)$ largest parts of λ and add a new smallest part of size $\sigma(\lambda)$. Thus $(8, 7, 5, 4, 3) \to (7, 6, 5, 4, 3, 2)$.



This transformation is always possible except when the dots of $\sigma(\lambda)$ and $s(\lambda)$ meet and $s(\lambda) = \sigma(\lambda) + 1$, as in $\lambda = (6, 5, 4)$.



In this case the above transformation will not give a partition into distinct parts but in all other cases it will transform an odd, distinct partition into an even, distinct,partition.

The two exceptional cases depend on the number n, for:

(a) If $s(\lambda) = \sigma(\lambda)$ and the dots in $s(\lambda)$ meet with the dots in $\sigma(\lambda)$, then n is divided into $\sigma(\lambda)$ parts. By writing $m = \sigma(\lambda)$ we conclude that $n = m + (m+1) + \cdots + (m+m-1) = m(3m-1)/2$.

(b) If $s(\lambda) = \sigma(\lambda) + 1$ and the dots in $s(\lambda)$ meet with the dots in $\sigma(\lambda)$, then n is divided into $\sigma(\lambda)$ parts. Hence, if $m = \sigma(\lambda)$, then $n = (m+1) + (m+2) + \cdots + (m+1+m-1) = m(3m+1)/2$.

Therefore, if $n \neq m(3m \pm 1)/2$ for some positive integer m, then Case 1 and Case 2 establish a bijective mapping from partitions of n into an odd number of distinct parts to partitions of n into an even number of distinct parts. For such integers $P(n; d, o) = P(n; d, e)$.

We now investigate the exceptional cases (a) and (b) mentioned above. Let $n = m(3m-1)/2$ for some odd m, $m \geq 1$. For this n only the exceptional situation described in Case (a) can occur, and this exceptional situation involves only the one partition mentioned in Case (a). For this sole partition the bijective transformation fails. The "extra" partition explains why for m odd, and n = m(3m - 1)/2, we have

$$P(n; d, o) = P(n; d, e) + 1.$$

Similar arguments will explain the result for even m and, in Case b, for $n = m(3m + 1)/2$. This ends our proof.

2.19

A lot of results about partitions can be obtained by means of generating functions. Let us look at some of these:

1. $F(x) = (1-x^a)^{-1}(1-x^b)^{-1}(1-x^c)^{-1}\cdots$ *is the generating function of* $P(n; \{a, b, c, \ldots\})$, *the number of ways of writing* n *as the sum of integers from the set* $\{a, b, c, \ldots\}$ *with repetitions allowed.*

Proof. Consider the coefficient of x^n in the series expansion of $F(x)$: $(1-x^a)^{-1}(1-x^b)^{-1}(1-x^c)^{-1}\cdots = (1+x^a+x^{2a}+\cdots+x^{ka}+\cdots)(1+x^b+\cdots+x^{kb}+\cdots)(1+x^c+\cdots)\cdots$.

If the term x^n is formed from the product of $x^{k_1 a}, x^{k_2 b}, x^{k_3 c}, \ldots$ then

$$n = \underbrace{a+\cdots+a}_{k_1 \text{ times}} + \underbrace{b+\cdots+b}_{k_2 \text{ times}} + \underbrace{c+\cdots+c}_{k_3 \text{ times}} + \cdots.$$

Hence the term x^n arises exactly as often as n can be written as the sum of a's, b's, c's, $\ldots$. The coefficient of x^n is therefore $P(n;\{a,b,c,\ldots\})$.

Immediate consequences of the above observation are:

1.1. *The generating function for $P(n)$, the number of ways of writing n as the sum of positive integers, is*

$$F(x) = (1-x)^{-1}(1-x^2)^{-1}(1-x^3)^{-1}\cdots(1-x^k)^{-1}\cdots.$$

1.2. *The generating function for $P(n;\{\text{odd integers}\})$ is*

$$(1-x^1)^{-1}(1-x^3)^{-1}(1-x^5)^{-1}\cdots(1-x^{2k+1})^{-1}\cdots.$$

1.3. *The generating function for $P(n;\{1,2,\ldots,k\})$ is*

$$(1-x)^{-1}(1-x^2)^{-1}\cdots(1-x^k)^{-1}.$$

1.4. *We have*

$$\sum_n P_m(n)x^n = x^m(1-x)^{-1}(1-x^2)^{-1}\cdots(1-x^m)^{-1}.$$

Proof. We prove 1.4. As we just saw $(1-x)^{-1}(1-x^2)^{-1}\cdots(1-x^m)^{-1} = \sum_n P(n;\{1,2,\ldots,m\})x^n$. Multiplying by x^m we obtain $x^m(1-x)^{-1}(1-x^2)^{-1}\cdots(1-x^m)^{-1} = \sum_m P(n;\{1,2,\ldots,m\})x^{n+m} = \sum_{n=m}^{\infty} P(n-m;\{1,2,\ldots,m\})x^n = \sum_{n=m}^{\infty}(\sum_{k=1}^{\infty} P_k(n-m))x^n = \sum_{n=m}^{\infty} P_m(n)x^n$, as claimed. The last two signs of equality are explained by the first two results proved in this section. This proves 1.4.

Our next result is the following:

2. $F(x) = (1+x^a)(1+x^b)(1+x^c)\cdots$ *is the generating function of* $P(n;d,(a,b,c,\ldots\}),$

the number of ways of writing n as a sum using the distinct numbers $a, b, c, \ldots$ at most once each.

Proof. To form x^n we can choose either 1 or x^a from the first factor and there is no option for choosing x^a again. The same is true for $x^b, x^c, \ldots$. Hence $n = \varepsilon_a a + \varepsilon_b b + \varepsilon_c c + \cdots$, where $\varepsilon_k = 1$ or 0. We can see that x^n arises as often as n can be written in the above way; hence the coefficient of x^n is $P(n; d, \{a, b, c, \ldots\})$. This ends our proof.

Immediate consequences are:

2.1. *The generating function of $P(n; d)$, the number of ways of writing n as the sum of distinct integers, is*

$$(1+x)(1+x^2)(1+x^3)\cdots.$$

2.2. *The generating function of $P(n; d, \{\text{odd integers}\})$ is*

$$(1+x)(1+x^3)(1+x^5)\cdots.$$

2.3. *The generating function of $P(n; d, (2^k : k = 0, 1, 2, \ldots\})$ is $\prod_{k=0}^{\infty}(1+x^{2^k})$.*

In Section 2.18 we proved the equality of $P(n; d)$ and $P(n; \{\text{odd integers}\})$ using Ferrer diagrams. Relying on generating functions we can prove this as follows:

$$\begin{aligned}
\sum_n P(n; d)x^n &= (1+x)(1+x^2)(1+x^3)\cdots \\
&= \frac{(1-x)(1+x)(1-x^2)(1+x^2)(1-x^3)(1+x^3)\cdots}{(1-x)(1-x^2)(1-x^3)\cdots} \\
&= \frac{(1-x^2)(1-x^4)(1-x^6)(1-x^8)\cdots}{(1-x)(1-x^2)(1-x^3)(1-x^4)\cdots} \\
&= \frac{1}{(1-x)(1-x^3)(1-x^5)\cdots} \\
&= \sum_n P(n; \{\text{odd integers}\})x^n.
\end{aligned}$$

REMARK. We all know that a positive integer has a unique expression in base 2. It is somewhat amusing to see how this follows from easy work with generating functions.

$$\begin{aligned}
&\sum_n P(n; d, \{2^k : k = 0, 1, 2, \ldots\})x^n \\
&\qquad = (1 + x)(1 + x^2)(1 + x^4)(1 + x^8)\cdots \\
&\qquad = \left(\frac{1 - x^2}{1 - x}\right)\left(\frac{1 - x^4}{1 - x^2}\right)\left(\frac{1 - x^8}{1 - x^4}\right)\left(\frac{1 - x^{16}}{1 - x^8}\right)\cdots \\
&\qquad = \frac{1}{1 - x} = \sum_n x^n.
\end{aligned}$$

Hence $P(n; d, \{2^k : k = 0, 1, 2, \ldots\}) = 1$, for all n.

Let us close this section with a series expansion version of *Euler's pentagonal theorem*:

$$\prod_{n=1}^{m}(1 - x^n) = 1 + \sum_{m=1}^{\infty}(-1)^m(x^{m(3m-1)/2} + x^{m(3m+1)/2}).$$

Proof.

$$\begin{aligned}
&\sum_{m=1}^{\infty}(-1)^m(x^{m(3m-1)/2} + x^{m(3m+1)/2}) \\
&\qquad = (-1)^m x^{m(3m\pm1)/2} \\
&\qquad = \sum_{n=1}^{\infty} x^n \begin{cases} (-1)^m & \text{if } n = m(3m \pm 1)/2 \\ 0 & \text{otherwise} \end{cases} \\
&\qquad = \{\text{by the result in Section 2.18}\} \\
&\qquad = \sum_{n=1}^{\infty}(P(n; d, e) - P(n; d, o))x^n.
\end{aligned}$$

We need to show that

$$1 + \sum_{n=1}^{\infty}(P(n; d, e) - P(n; d, o))x^n = \prod_{n=1}^{\infty}(1 - x^n).$$

Let us look at the coefficient of x^n in $\prod_{n=1}^{\infty}(1 - x^n)$. Since x^n can be formed as a product of $(-x)^{k_1}(-x^2)^{k_2}\cdots(-x^n)^{k_n}$, where $k_i = 0$ or 1, we have $x^n = (-1)^{k_1+\cdots+k_n}x^{k_1+2k_2+\cdots+nk_n}$.

The coefficient of x^n is therefore

$$\sum_{(k_1,\ldots,k_n)} (-1)^{k_1+k_2+\cdots+k_n},$$

where each n-tuple corresponds to a partition of n into distinct integers as $n = k_1 + 2k_2 + 3k_3 + \cdots + nk_n$ ($k_i = 0$ or 1). Note that $k_1 + k_2 + \cdots + k_n$ gives us the number of parts of the partition of n. Hence $(-1)^{k_1+\cdots+k_n}$ if the partition has an even number of parts and -1 if it has an odd number parts. This observation leads us to conclude that the coefficient of x^n in $\prod_{n=1}^{\infty}(1-x^n)$ is

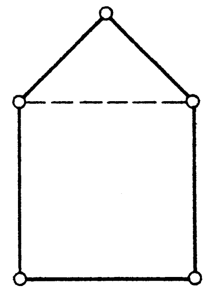
$$\sum_{(k_1,\ldots,k_n)} (-1)^{k_1+\cdots+k_n} = P(n,d,e) - P(n;d,o).$$

The constant term is clearly 1 on both sides. We have therefore proved that

$$\begin{aligned}\prod_{n=1}^{\infty}(1-x^n) &= 1 + \sum_{n=1}^{\infty}(P(n;d,e) - P(n;d,o))x^n \\ &= 1 + \sum_{m=1}^{\infty}(-1)^m(x^{m(3m-1)/2} + x^{m(3m+1)/2}).\end{aligned}$$

Brief Note on Terminology

The word pentagonal has been mentioned more than once in connection with Euler's result. Numbers of the form $m(3m \pm 1)/2$ are called pentagonal. These are exceptional integers for which the number of distinct even partitions does not equal the number of distinct odd ones. We call them so because each can be written as the sum of a square and a "triangular" number, thus producing the geometric effect of a pentagon (or of a house), as displayed below:



Indeed, $m(3m \pm 1)/2 = s + t$, where $s = m^2$ and $t = m(m \pm 1)/2$.

EXERCISES

1. The number of noncongruent triangles with circumference $2n$ and integer sides is equal to $P_3(n)$. Prove this.

2. A partition of the number n is called *perfect* if every integer from 1 to $(n-1)$ can be written in a unique way as the total of a subset of the parts of this partition. Prove that the number of perfect partitions of n is the same as the number of ways of factoring $n+1$, where the order of the factors counts and factors of 1 are not counted. When will the trivial partition $n = 1 + 1 + \cdots + 1$ be the only solution?

3. Find a generating function for the number of integer solutions of $n = 2x + 3y + 7z$ with:

   (a) $x, y, z \geq 0$.

   (b) $0 \leq z \leq 2 \leq y \leq 8 \leq x$.

4. Find a generating function for the number of ways of making n cents change in

pennies, nickels, dimes, and quarters.

5. Show with generating functions that every positive integer can be written as a sum of distinct powers of 10, that is, it has a unique decimal expansion.

6. Prove the identity

$$\frac{1}{1-x} = (1+x+x^2+\cdots+x^9)(1+x^{10}+x^{20}+\cdots+x^{90})$$
$$\cdot(1+x^{100}+x^{200}+\cdots+x^{900})\cdots.$$

7. Show that the number of partitions of the integer $2r+k$ into exactly $r+k$ parts is the same for any nonnegative integer k.

8. Show that the number of partitions of n into at most two parts is $[n/2]+1$, with $[x]$ denoting the integral part of x.

9. Prove that the number of partitions of n in which only odd parts may be repeated equals the number of partitions of n in which no part appears more than three times.

10. Prove that the number of partitions of n with unique smallest part (i.e., the smallest part occurs only once) and largest part at most twice the smallest part equals the number of partitions of n in which the largest part is odd and the smallest part is larger than half the largest part.

8 A GENERATING FUNCTION FOR SOLUTIONS OF DIOPHANTINE SYSTEMS IN NONNEGATIVE INTEGERS

The title, pretty well describes our intentions with regard to the contents of this section. Consider

$$\sum_{j=1}^{n} a_{ij} x_j = b_i; \quad i = 1, 2, \ldots, m \tag{2.11}$$

where a_{ij} and b_i are nonnegative integers. *We investigate the solutions to the Diophantine system* (2.11) *in nonnegative integers.*

Write

$$x = (x_1, \ldots, x_n), \quad s = (s_1, \ldots, s_n)$$

$$b = (b_1, \ldots, b_m), \quad t = (t_1, \ldots, t_m)$$

$$s^x = \prod_{j=1}^{n} s_j^{x_j}, \quad t^b = \prod_{i=1}^{m} t_i^{b_i}.$$

The notation $x \geq 0$ or $b \geq 0$ means that the respective components are nonnegative (and, in this case, also integral).

Assume that each column of the $m \times n$ matrix (a_{ij}) has a nonzero entry. The nonnegativity of the entities involved insures then at most a finite number of solutions to system (2.11).

For $x \geq 0$ and $b \geq 0$ set

$$N_x(b) = \begin{cases} 1 & \text{if } x \text{ is a solution of (2.11)} \\ 0 & \text{otherwise} \end{cases}$$

and let $N(b)$ be the number of solutions to (2.11).

∗ *We assert that*

$$\sum_{\substack{x\geq 0\\ b\geq 0}} N_x(b) s^x t^b = \prod_{j=1}^{n} (1 - s_j t_1^{a_{1j}} t_2^{a_{2j}} \cdots t_m^{a_{mj}})^{-1}$$

and

$$\sum_{b\geq 0} N(b) t^b = \prod_{j=1}^{n} (1 - t_1^{a_{1j}} t_2^{a_{2j}} \cdots t_m^{a_{mj}})^{-1}.$$

The proof rests upon routine expansions:

$$\begin{aligned}
\sum_{\substack{x\geq 0\\ b\geq 0}} N_x(b) s^x t^b &= \sum_{x\geq 0} s_1^{x_1} \cdots s_n^{x_n} t_1^{\sum_j a_{1j}x_j} \cdots t_m^{\sum_j a_{mj}x_j} \\
&= \prod_{j=1}^{n} \left(\sum_{x_j=0}^{\infty} s_j^{x_j} t_1^{a_{1j}x_j} \cdots t_m^{a_{mj}x_j} \right) \\
&= \prod_{j=1}^{n} \sum_{x_j=0}^{\infty} (s_j t_1^{a_{1j}} \cdots t_m^{a_{mj}})^{x_j} \\
&= \prod_{j=1}^{n} (1 - s_j t_1^{a_{1j}} \cdots t_m^{a_{mj}})^{-1}.
\end{aligned}$$

The second formula is explained similarly. This explains the assertion.

Further, by writing our first formula as

$$\left[\prod_{j=1}^{n} (1 - s_j t_1^{a_{1j}} \cdots t_m^{a_{mj}}) \right] \left[\sum_{\substack{x\geq 0\\ b\geq 0}} N_x(b) s^x t^b \right] = 1$$

and equating the coefficients of $s^x t^b$ on both sides we obtain *recursive formulas* for the $N_x(b)$'s. The same can be done to the second formula to obtain recurrences for the $N(b)$'s.

In particular, the reader may wish to investigate in detail the Diophantine system

$$\left.\begin{aligned}
x_0 + x_1 + x_2 + \cdots + x_n &= b_1 \\
x_1 + 2x_2 + \cdots + nx_n &= b_2
\end{aligned}\right\}.$$

It leads to the so-called *Gaussian polynomials* which we discuss in Section 6 of Chapter 3 – see, in particular, Exercise 5 of that section.

9 HISTORICAL NOTE

What we have seen in this chapter is by and large classical material on generating functions. Much of the first two sections introduce the (formal) power series and explain the combinatorial meaning of multiplication by convolution. Of the results in Section 3 those regarding Stirling numbers rely fundamentally on Stirling's formulas, introduced in Sections 1.7(c) and 1.8(c) of Chapter l. The Lah numbers, and their analogous behavior to those of Stirling, were only relatively recently noticed by Ivo Lah [8] of the University of Belgrade, Yugoslavia. Though less fundamental in nature than the numbers of Stirling, we meet them again in connection with inversion formulas.

Faa DiBruno observed the pattern of the higher order derivative of a composition of two functions in terms of (what we now call) Bell polynomials; this result can be found in [1]. We only briefly discussed recurrence relations and only those aspects that call for immediate use of generating functions. The contents of Section 2.13 are based upon a paper of D. André of 1879 [6].

Enumerating labeled spanning trees of a graph, as we did, was (implicitly) noted by Kirchhoff in his classic paper [5] on electrical networks of which the famous Kirchhoff laws of current form the main topic. That the Kirchhoff matrix coincides with the Fisher information matrix in the setting of statistical designs (with blocking in one direction – see [7]) is an unexpected connection with possibly interesting ramifications. We discuss these shared aspects in Chapter 8, the chapter on statistical design. The contents of Section 8 are of recent origin and appear only as part of a more substantial work on Fuchsian groups [9] by R. S. Kulkarni.

The pentagonal theorem (hereinafter written as theorem P) dates back almost to the very beginnings of the work with generating functions [4]. It had preoccupied Euler a good deal over the span of at least a decade. In 1740, while expanding $\prod_n (1 - x^n)$, Euler observed the pattern of -1's and 1's that arises in connection with the pentagonal numbers. The reader may be entertained by how Euler relates this:

> Theorem P is of such a nature that we can be assured of its truth without giving it a perfect demonstration. Nevertheless, I will present evidence for it of such a character that it might be regarded as almost equivalent to a rigorous demonstration.

We are then informed that he has compared coefficients of up to the 40th power of x and that they all follow the proposed pattern.

> I have long searched in vain for a rigorous demonstration of theorem P, and I have proposed the same question to some of my friends with whose ability in these matters I am familiar but all have agreed with me on the truth of theorem P without being able to unearth any clue of a demonstration. Thus it will be a known truth, but not yet demonstrated . . . . And since I must admit that I am not in a position to give it a rigorous demonstration, I will justify it by a sufficiently large number of examples . . . . I think these examples are sufficient to discourage anyone from imagining that it is by pure chance that my rule is in agreement with the truth . . . . If one still doubts that the law is precisely that one which I have indicated, I will give some examples with larger numbers.

Here he tells how he took the trouble to examine the coefficients of x^{101} and x^{301} and how they came out to be just what he had expected.

> These examples which I have just developed undoubtedly will dispel any qualms which we might have had about the truth of theorem P.

Euler did succeed in proving the pentagonal theorem in 1750. The passages above were extracted from Pòlya's work mentioned also as reference [4].

Of the texts available that treat similar material we recommend [1], [2], and [3].

10 REFERENCES

1. J. Riordan, *An Introduction to Combinatorial Analysis*, Wiley, New York, 1958.

2. G. E. Andrews, *The Theory of Partitions*, Addison-Wesley, Reading, MA, 1976.

3. C. Berge, *Graphs and Hypergraphs*, North-Holland, Amsterdam, 1973.

4. *Leonhardi Euleri Opera Omnia*, Ser. 1, Vol. 2, 1915, pp. 241-253 (see G. Pó1ya, *Collected Works, Vol.* 4, The MIT Press, Cambridge, MA, 1984, pp. 186-187.)

5. G. Kirchhoff, Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen verteilung galvanischer Ströme gerführt wird, *Ann. Phys. Chem.*, **72**, 497-508 (1847) (English translation in *Trans. Inst. Radio Engrs.*, **CT-5**, 4-7, March, 1958).

6. D. André, Developpement de *sec* x et de *tang* x, *C. R. Acad. Sci. Paris*, **88** (965-967 (1879).

7. J. C. Kiefer, Optimum experimental designs, *J. Royal Stat. Soc.* (B), **21** 272-304 (1959).

8. I. Lah, Eine neue Art von Zahlen, ihre Eigenschaften and Anwendung in der mathematischen Statistik, *Mitteilungsbl. Math. Statist.*, **7** 203-212 (1955).

9. R. S. Kulkarni, An extension of a theorem of Kurosh and applications to Fuchsian groups, *Mich. Math. Journal*, **30** 259-272 (1983).

10. L. Lovász, *Combinatorial Problems and Exercises*, North Holland, Amsterdam, 1979.