2

Sieve Methods

2.1 Inclusion-Exclusion

Roughly speaking, a "sieve method" in enumerative combinatorics is a method for determining the cardinality of a set S that begins with a larger set and somehow subtracts off or cancels out unwanted elements. Sieve methods have two basic variations: (1) We can first approximate our answer with an overcount, and then subtract off an overcounted approximation of our original error, and so on, until after finitely many steps we have "converged" to the correct answer. This method is the combinatorial essence of the Principle of Inclusion-Exclusion, to which this section and the next four are devoted. (2) The elements of the larger set can be weighted in a natural combinatorial way so that the unwanted elements cancel out, leaving only the original set S. We discuss this technique in Sections 2.6 and 2.7.

The Principle of Inclusion-Exclusion is one of the fundamental tools of enumerative combinatorics. Abstractly, the Principle of Inclusion-Exclusion amounts to nothing more than computing the inverse of a certain matrix. As such, it is simply a minor result in linear algebra. The beauty of the principle lies not in the result itself, but rather in its wide applicability. We will give several example of problems that can be solved by Inclusion-Exclusion, some in a rather subtle way. First we state the principle in its purest form.

2.1.1 Theorem. *Let S be an n-set. Let V be the 2^n-dimensional vector space (over some field K) of all functions $f : 2^S \to K$. Let $\phi : V \to V$ be the linear transformation defined by*

$$\phi f(T) = \sum_{Y \supseteq T} f(Y), \text{ for all } T \subseteq S. \tag{2.1}$$

Then ϕ^{-1} exists and is given by

$$\phi^{-1} f(T) = \sum_{Y \supseteq T} (-1)^{\#(Y-T)} f(Y), \text{ for all } T \subseteq S. \tag{2.2}$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

Proof. Define $\psi : V \to V$ by $\psi f(T) = \sum_{Y \supseteq T} (-1)^{\#(Y-T)} f(Y)$. Then (composing functions right to left)

$$\begin{aligned}\phi\psi f(T) &= \sum_{Y \supseteq T} (-1)^{\#(Y-T)} \phi f(Y) \\ &= \sum_{Y \supseteq T} (-1)^{\#(Y-T)} \sum_{Z \supseteq Y} f(Z) \\ &= \sum_{Z \supseteq T} \left(\sum_{Z \supseteq Y \supseteq T} (-1)^{\#(Y-T)} \right) f(Z).\end{aligned}$$

Setting $m = \#(Z - T)$, we have

$$\sum_{\substack{Z \supseteq Y \supseteq T \\ (Z,T \text{ fixed})}} (-1)^{\#(Y-T)} = \sum_{i=0}^{m} (-1)^i \binom{m}{i} = \delta_{0m},$$

the latter equality by putting $x = -1$ in equation (1.18) or by Exercise 1.3(f), so $\phi\psi f(T) = f(T)$. Hence, $\phi\psi f = f$, so $\psi = \phi^{-1}$. □

The following is the usual combinatorial situation involving Theorem 2.1.1. We think of S as being a set of properties that the elements of some given set A of objects may or may not have. For any subset T of S, let $f_=(T)$ be the number of objects in A that have *exactly* the properties in T (so they fail to have the properties in $\overline{T} = S - T$). More generally, if $w : A \to K$ is any weight function on A with values in a field (or abelian group) K, then one could set $f_=(T) = \sum_x w(x)$, where x ranges over all objects in A having exactly the properties in T. Let $f_\geq(T)$ be the number of objects in A that have *at least* the properties in T. Clearly then,

$$f_\geq(T) = \sum_{Y \supseteq T} f_=(Y). \tag{2.3}$$

Hence by Theorem 2.1.1,

$$f_=(T) = \sum_{Y \supseteq T} (-1)^{\#(Y-T)} f_\geq(Y). \tag{2.4}$$

In particular, the number of objects having *none* of the properties in S is given by

$$f_=(\emptyset) = \sum_{Y \supseteq T} (-1)^{\#Y} f_\geq(Y), \tag{2.5}$$

where Y ranges over all subsets S. In typical applications of the Principle of Inclusion-Exclusion, it will be relatively easy to compute $f_\geq(Y)$ for $Y \subseteq S$, so equation (2.4) will yield a formula for $f_=(T)$.

In equation (2.4) one thinks of $f_\geq(T)$ (the term indexed by $Y = T$) as being a first approximation to $f_=(T)$. We then subtract

$$\sum_{\substack{Y \supseteq T \\ \#(Y-T)=1}} f_\geq(Y),$$

to get a better approximation. Next we add back in

$$\sum_{\substack{Y \supseteq T \\ \#(Y-T)=2}} f_{\geq}(Y),$$

and so on, until finally reaching the explicit formula (2.4). This reasoning explains the terminology "Inclusion-Exclusion."

Perhaps the most standard formulation of the Principle of Inclusion-Exclusion is one that dispenses with the set S of properties per se, and just considers subsets of A. Thus, let $A_1, \dots, A_n$ be subsets of a finite set A. For each subset T of $[n]$, let

$$A_T = \bigcap_{i \in T} A_i$$

(with $A_\emptyset = A$), and for $0 \leq k \leq n$ set

$$S_k = \sum_{\#T=k} \#A_T, \tag{2.6}$$

the sum of the cardinalities, or more generally the weighted cardinalities

$$w(A_T) = \sum_{x \in A_T} w(x),$$

of all k-tuple intersections of the A_i's. Think of A_i as defining a property P_i by the condition that $x \in A$ satisfies P_i if and only if $x \in A_i$. Then A_T is just the set of objects in A that have at least the properties in T, so by (2.5) the number $\#(\overline{A}_1 \cap \cdots \cap \overline{A}_n)$ of elements of A lying in *none* of the A_i's is given by

$$\#(\overline{A}_1 \cap \cdots \cap \overline{A}_n) = S_0 - S_1 + S_2 - \cdots + (-1)^n S_n, \tag{2.7}$$

where $S_0 = A_\emptyset = \#A$.

The Principle of Inclusion-Exclusion and its various reformulations can be dualized by interchanging $\cap$ and $\cup$, $\subseteq$ and $\supseteq$, and so on, throughout. The dual form of Theorem 2.1.1 states that if

$$\widetilde{\phi} f(T) = \sum_{Y \subseteq T} f(Y), \quad \text{for all } T \subseteq S,$$

then $\widetilde{\phi}^{-1}$ exists and is given by

$$\widetilde{\phi}^{-1} f(T) = \sum_{Y \subseteq T} (-1)^{\#(T-Y)} f(Y), \quad \text{for all } T \subseteq S.$$

Similarly, if we let $f_{\leq}(T)$ be the (weighted) number of objects of A having *at most* the properties in T, then

$$f_{\leq}(T) = \sum_{Y \subseteq T} f_{=}(Y),$$

$$f_{=}(T) = \sum_{Y \subseteq T} (-1)^{\#(T-Y)} f_{\leq}(Y). \tag{2.8}$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

A common special case of the Principle of Inclusion-Exclusion occurs when the function f $=$satisfies $f = (T) = f = (T')$ whenever $\#T = \#T'$. Thus also $f \geq (T)$ depends only on $\#T$, and we set $a(n-i) = f = (T)$ and $b(n-i) = f_{\geq}(T)$ whenever $\#T = i$. (Caveat. In many problems the set A of objects and S of properties will depend on a parameter p, and the functions $a(i)$ and $b(i)$ may depend on p. Thus, for example, $a(0)$ and $b(0)$ are the number of objects having *all* the properties, and this number may certainly depend on p. Proposition 2.2.2 is devoted to the situation when $a(i)$ and $b(i)$ are independent of p.) We thus obtain from equations (2.3) and (2.4) the equivalence of the formulas

$$b(m) = \sum_{i=0}^{m} \binom{m}{i} a(i), \quad 0 \leq m \leq n, \tag{2.9}$$

$$a(m) = \sum_{i=0}^{m} \binom{m}{i} (-1)^{m-i} b(i), \quad 0 \leq m \leq n. \tag{2.10}$$

In other words, the inverse of the $(n+1) \times (n+1)$ matrix whose (i, j)-entry $(0 \leq i, j \leq n)$ is $\binom{j}{i}$ has (i, j)-entry $(-1)^{j-i}\binom{j}{i}$. For instance,

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Of course, we may let n approach ∞ so that (2.9) and (2.10) are equivalent for $n = \infty$.

Note that in language of the calculus of finite differences (see equation (1.98)), (2.10) can be rewritten as

$$a(m) = \Delta^m b(0), \quad 0 \leq m \leq n.$$

2.2 Examples and Special Cases

The canonical example of the use of the Principle of Inclusion-Exclusion is the following.

2.2.1 Example. (the "derangement problem" or "*problème des rencontres*"). How many permutations $w \in \mathfrak{S}_n$ have no fixed points, that is, $w(i) \neq i$ for all $i \in [n]$? Such a permutation is called a *derangement*. Call this number $D(n)$. Thus, $D(0) = 1, D(1) = 0, D(2) = 1, D(3) = 2$. Think of the condition $w(i) = i$ as the ith property of w. Now the number of permutations with *at least* the set $T \subseteq [n]$ of points fixed is $f_{\geq}(T) = b(n-i) = (n-i)!$, where $\#T = i$ (since we fix the elements of T and permute the remaining $n - i$ elements arbitrarily). Hence by (2.10), the number

$f_=(\emptyset) = a(n) = D(n)$ of permutations with *no* fixed points is

$$D(n) = \sum_{i=0}^{n} \binom{n}{i} (-1)^{n-i} i!. \tag{2.11}$$

This last expression may be rewritten

$$D(n) = n! \left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!} \right). \tag{2.12}$$

Since $0.36787944\cdots = e^{-1} = \sum_{j\geq 0}(-1)^j/j!$, it is clear from (2.12) that $n!/e$ is a good approximation to $D(n)$, and indeed it is not difficult to show that $D(n)$ is the nearest integer to $n!/e$. It also follows immediately from (2.12) that for $n \geq 1$,

$$D(n) = nD(n-1) + (-1)^n, \tag{2.13}$$

$$D(n) = (n-1)(D(n-1) + D(n-2)). \tag{2.14}$$

While it is easy to give a direct combinatorial proof of equation (2.14), considerably more work is necessary to prove (2.13) combinatorially. (See Exercise 2.8.) In terms of generating functions, we have that

$$\sum_{n\geq 0} D(n) \frac{x^n}{n!} = \frac{e^{-x}}{1-x}.$$

The function $b(i) = i!$ has a very special property—it depends only on i, not on n. Equivalently, the number of permutations $w \in \mathfrak{S}_n$ that have at most the set $T \subseteq [n]$ of points *unfixed* depends only on $\#T$, not on n. This means that equation (2.11) can be rewritten in the language of the calculus of finite differences (see equation (1.98)) as

$$D(n) = \Delta^n x!\big|_{x=0},$$

which is abbreviated $\Delta^n 0!$. Since the number $b(i)$ of permutations in $\mathfrak{S}_n$ that have at most some specified i-set of points unfixed depends only on i, the same is true of the number $a(i)$ of permutations in $\mathfrak{S}_n$ that have exactly some specified i set of points unfixed. It is clear combinatorially that $a(i) = D(i)$, and this fact is also evident from equations (2.10) and (2.11).

Let us state formally the general result that follows from the preceding considerations.

2.2.2 Proposition. *For each $n \in \mathbb{N}$, let B_n be a (finite) set, and let S_n be a set of n properties that elements of B_n may or may not have. Suppose that for every $T \subseteq S_n$, the number of $x \in B_n$ that* lack at most *the properties in T (i.e., that have at least the properties in $S - T$) depends only on $\#T$, not on n. Let $b(n) = \#B_n$, and let $a(n)$ be the number of objects $x \in B_n$ that have* none *of the properties in S_n. Then $a(n) = \Delta^n b(0)$.*

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

2.2.3 Example. Let us consider an example to which the previous proposition does *not* apply. Let $h(n)$ be the number of permutations of the multiset $M_n = \{1^2, 2^2, \dots, n^2\}$ with no two consecutive terms equal. Thus, $h(0) = 1$, $h(1) = 0$, and $h(2) = 2$ (corresponding to the permutations 1212 and 2121). Let P_i, for $1 \le i \le n$, be the property that the permutation w of M_n has two consecutive i's. Hence we seek $f_=(\emptyset) = h(n)$. It is clear by symmetry that for fixed n, $f_\ge(T)$ depends only on $i = \#T$, so write $g(i) = f_\ge(T)$. Clearly $g(i)$ is equal to the number of permutations w of the multiset $\{1, 2, \dots, i, (i+1)^2, \dots, n^2\}$ (replace any $j \ge i$ appearing in w by two consecutive j's), so

$$g(i) = (2n-i)!2^{-(n-i)},$$

a special case of equation (1.23). Note that $b(i) := g(n-i) = (n+i)!2^{-i}$ is not a function of i alone, so that Proposition 2.2.2 is indeed inapplicable. However, we do get from (2.10) that

$$h(n) = \sum_{i=0}^{n} \binom{n}{i} (-1)^{n-i} (n+i)!2^{-i} = \Delta^n (n+i)!2^{-i}\big|_{i=0}.$$

Here the function $(n+i)!2^{-i}$ to which Δ^n is applied depends on n.

We turn next to an example for which the final answer can be represented by a determinant.

2.2.4 Example. Recall that in Chapter 1 (Section 1.4) we defined the *descent set* $D(w)$ of a permutation $w = a_1 a_2 \cdots a_n$ of $[n]$ by $D(w) = \{i : a_i > a_{i+1}\}$. Our object here is to obtain an expression for the quantity $\beta_n(S)$, the number of permutations $w \in \mathfrak{S}_n$ with descent set S. Let $\alpha_n(S)$ be the number of permutations whose descent set is *contained* in S, as in equation (1.31). Thus (as pointed out in equation (1.31))

$$\alpha_n(S) = \sum_{T \subseteq S} \beta_n(T).$$

It was stated in equation (1.34) and follows from (2.8) that

$$\beta_n(S) = \sum_{T \subseteq S} (-1)^{\#(S-T)} \alpha_n(T).$$

Recall also that if the elements of S are given by $1 \le s_1 < s_2 < \cdots < s_k \le n-1$, then by Proposition 1.4.1 we have

$$\alpha_n(S) = \binom{n}{s_1, s_2 - s_1, s_3 - s_2, \dots, n - s_k}.$$

Therefore,

$$\beta_n(S) = \sum_{1 \le i_1 < i_2 < \cdots < i_j \le k} (-1)^{k-j} \binom{n}{s_{i_1}, s_{i_2} - s_{i_1}, \dots, n - s_{i_j}}. \tag{2.15}$$

We can write (2.15) in an alternative form as follows. Let f be any function defined on $[0,k+1]\times[0,k+1]$ satisfying $f(i,i)=1$ and $f(i,j)=0$ if $i>j$. Then the terms in the sum

$$A_k = \sum_{1\le i_1<i_2<\cdots<i_j\le k} (-1)^{k-j} f(0,i_1)f(i_1,i_2)\cdots f(i_j,k+1)$$

are just the nonzero terms in the expansion of the $(k+1)\times(k+1)$ determinant with (i,j)-entry $f(i,j+1)$, $(i,j)\in[0,k]\times[0,k]$. Hence, if we set $f(i,j)=1/(s_j-s_i)!$ (with $s_0=0$, $s_{k+1}=n$), we obtain from (2.15) that

$$\beta_n(S) = n!\det[1/(s_{j+1}-s_i)!], \tag{2.16}$$

$(i,j)\in[0,k]\times[0,k]$. For instance, if $n=8$ and $S=\{1,5\}$, then

$$\beta_n(S) = 8!\begin{vmatrix} \frac{1}{1!} & \frac{1}{5!} & \frac{1}{8!} \\ 1 & \frac{1}{4!} & \frac{1}{7!} \\ 0 & 1 & \frac{1}{3!} \end{vmatrix} = 217.$$

By an elementary manipulation (whose details are left to the reader), equation (2.16) can also be written in the form

$$\beta_n(S) = \det\left[\binom{n-s_i}{s_{j+1}-s_i}\right], \tag{2.17}$$

where $(i,j)\in[0,k]\times[0,k]$ as before.

2.2.5 Example. We can obtain a q-analogue of the previous example with very little extra work. We seek some statistic $s(w)$ of permutations $w\in\mathfrak{S}_n$ such that

$$\sum_{\substack{w\in\mathfrak{S}_n \\ D(w)\subseteq S}} q^{s(w)} = \binom{\boldsymbol{n}}{\boldsymbol{s_1,s_2-s_1,\dots,n-s_k}}, \tag{2.18}$$

where the elements of S are $1\le s_1<s_2<\cdots<s_k\le n-1$ as above. We will then automatically obtain q-analogues of equations (2.15), (2.16), and (2.17). We claim that (2.18) holds when $s(w)=\operatorname{inv}(w)$, the number of inversions of w. To see this, set $t_1=s_1$, $t_2=s_2-s_1,\dots,t_{k+1}=n-s_k$. Let $M=\{1^{t_1},\dots,(k+1)^{t_{k+1}}\}$. Recall from Proposition 1.7.1 that

$$\sum_{u\in\mathfrak{S}(M)} q^{\operatorname{inv}(u)} = \binom{\boldsymbol{n}}{\boldsymbol{t_1,t_2,\dots,t_{k+1}}}. \tag{2.19}$$

Now given $u\in\mathfrak{S}(M)$, let $v\in\mathfrak{S}_n$ be the standardization of u as defined after the second proof of Proposition 1.7.1, so $\operatorname{inv}(u)=\operatorname{inv}(v)$. We call v a *shuffle* of the sets $[1,s_1]$, $[s_1+1,s_2],\dots,[s_k+1,n]$. Now set $w=v^{-1}$. It is easy to see that v is a shuffle of $[1, s_1]$, $[s_1+1, s_2], \ldots, [s_k+1, n]$ if and only if $D(w) \subseteq \{s_1, s_2, \ldots, s_k\}$. Since $\mathrm{inv}(v) = \mathrm{inv}(w)$ by Proposition 1.3.14, we obtain

$$\sum_{\substack{w \in \mathfrak{S}_n \\ D(w) \subseteq S}} q^{\mathrm{inv}(w)} = \binom{\boldsymbol{n}}{\boldsymbol{s_1, s_2 - s_1, \ldots, n - s_k}}, \tag{2.20}$$

as desired.

Thus, set

$$\beta_n(S, q) = \sum_{\substack{w \in \mathfrak{S}_n \\ D(w) = S}} q^{\mathrm{inv}(w)}.$$

By simply mimicking the reasoning of Example 2.2.4, we obtain

$$\begin{aligned} \beta_n(S, q) &= \boldsymbol{(n)!} \det \left[1/\boldsymbol{(s_{j+1} - s_i)!} \right]_0^k \\ &= \det \left[\binom{\boldsymbol{n - s_i}}{\boldsymbol{s_{j+1} - s_i}} \right]_0^k. \end{aligned} \tag{2.21}$$

For instance, if $n = 8$ and $S = \{1, 5\}$, then

$$\beta_n(S, q) = \boldsymbol{(8)!} \begin{vmatrix} \dfrac{1}{\boldsymbol{(1)!}} & \dfrac{1}{\boldsymbol{(5)!}} & \dfrac{1}{\boldsymbol{(8)!}} \\ 1 & \dfrac{1}{\boldsymbol{(4)!}} & \dfrac{1}{\boldsymbol{(7)!}} \\ 0 & 1 & \dfrac{1}{\boldsymbol{(3)!}} \end{vmatrix}$$

$$\begin{aligned} &= q^2 + 3q^3 + 6q^4 + 9q^5 + 13q^6 + 17q^7 + 21q^8 + 23q^9 \\ &\quad + 24q^{10} + 23q^{11} + 21q^{12} + 18q^{13} + 14q^{14} + 10q^{15} \\ &\quad + 7q^{16} + 4q^{17} + 2q^{18} + q^{19}. \end{aligned}$$

If we analyze the reason why we obtained a determinant in the previous two examples, then we get the following result.

2.2.6 Proposition. *Let $S = \{P_1, \ldots, P_n\}$ be a set of properties, and let $T = \{P_{s_1}, \ldots, P_{s_k}\} \subseteq S$, where $1 \le s_1 < \cdots < s_k \le n$. Suppose that $f_{\le}(T)$ has the form*

$$f_{\le}(T) = h(n) e(s_0, s_1) e(s_1, s_2) \cdots e(s_k, s_{k+1})$$

for certain functions h on $\mathbb{N}$ and e on $\mathbb{N} \times \mathbb{N}$, where we set $s_0 = 0$, $s_{k+1} = n$, $e(i, i) = 1$, and $e(i, j) = 0$ if $j < i$. Then

$$f_{=}(T) = h(n) \det \left[e(s_i, s_{j+1}) \right]_0^k.$$

2.3 Permutations with Restricted Position

The derangement problem asks for the number of permutations $w \in \mathfrak{S}_n$ where for each i, certain values of $w(i)$ are disallowed (namely, we disallow $w(i) = i$). We

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

now consider a general theory of such permutations. It is traditionally described using terminology from the game of chess. Let $B \subseteq [n] \times [n]$, called a *board*. If $w \in \mathfrak{S}_n$, then define the graph $G(w)$ of w by

$$G(w) = \{(i, w(i)) : i \in [n]\}.$$

Now define

$$\begin{aligned} N_j &= \#\{w \in \mathfrak{S}_n : j = \#(B \cap G(w)\}, \\ r_k &= \text{number of } k\text{-subsets of } B \text{ such that no two} \\ &\qquad \text{elements have a common coordinate} \\ &= \text{number of ways to place } k \text{ nonattacking rooks on } B. \end{aligned}$$

Also define the *rook polynomial* $r_B(x)$ of the board B by

$$r_B(x) = \sum_k r_k x^k.$$

We may identify $w \in \mathfrak{S}_n$ with the placement of n nonattacking rooks on the squares $(i, w(i))$ of $[n] \times [n]$. Thus N_j is the number of ways to place n nonatttacking rooks on $[n] \times [n]$ such that exactly j of these rooks lie in B. For instance, if $n = 4$ and $B = \{(1,1),(2,2),(3,3),(3,4),(4,4)\}$, then $N_0 = 6$, $N_1 = 9$, $N_2 = 7$, $N_3 = 1$, $N_4 = 1$, $r_0 = 1$, $r_1 = 5$, $r_2 = 8$, $r_3 = 5$, $r_4 = 1$. Our object is to describe the numbers N_j, and especially N_0, in terms of the numbers r_k. To this end, define the polynomial

$$N_n(x) = \sum_j N_j x^j.$$

2.3.1 Theorem. *We have*

$$N_n(x) = \sum_{k=0}^{n} r_k (n-k)!(x-1)^k. \tag{2.22}$$

In particular,

$$N_0 = N_n(0) = \sum_{k=0}^{n} (-1)^k r_k (n-k)!. \tag{2.23}$$

First proof. Let C_k be the number of pairs (w, C), where $w \in \mathfrak{S}_n$ and C is a k-element subset of $B \cap G(w)$. For each j, choose w in N_j ways so that $j = \#(B \cap G(w))$, and then choose C is $\binom{j}{k}$ ways. Hence, $C_k = \sum_j \binom{j}{k} N_j$. On the other hand, we first could choose C in r_k ways and then "extend" to w in $(n-k)!$ ways. Hence, $C_k = r_k(n-k)!$. Therefore,

$$\sum_j \binom{j}{k} N_j = r_k (n-k)!,$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

or equivalently (multiplying by y^k and summing on k),

$$\sum_j (y+1)^j N_j = \sum_k r_k (n-k)! y^k.$$

Putting $y = x-1$ yields the desired formula.

Second proof. It suffices to assume $x \in \mathbb{P}$. The left-hand side of equation (2.22) counts the number of ways to place n nonattacking rooks on $[n]\times[n]$ and labeling each rook on B with an element of $[x]$. On the other hand, such a configuration can be obtained by placing k nonattacking rooks on B, labeling each of them with an element of $\{2,\dots,x\}$, placing $n-k$ additional nonattacking rooks on $[n]\times[n]$ in $(n-k)!$ ways, and labeling the new rooks on B with 1. This argument establishes the desired bijection. □

The two proofs of Theorem 2.3.1 provide another illustration of the principle enunciated in Chapter 1 (third proof of Proposition 1.3.7) about the two combinatorial methods for showing that two polynomials are identical. It is certainly also possible to prove (2.23) by a direct application of Inclusion-Exclusion, generalizing Example 2.2.1. Such a proof would not be considered combinatorial since we have not explicitly constructed a bijection between two sets (but see Section 2.6 for a method of making such a proof combinatorial). The two proofs we have given may be regarded as "semicombinatorial," since they yield by direct bijections formulas involving parameters y and x, respectively; and we then obtain (2.23) by setting $y=-1$ and $x=0$, respectively. In general, a semicombinatorial proof of (2.5) can easily be given by first showing combinatorially that

$$\sum_X f_=(X) x^{\#X} = \sum_Y f_\geq(Y)(x-1)^{\#Y} \tag{2.24}$$

or

$$\sum_X f_=(X)(y+1)^{\#X} = \sum_Y f_\geq(Y) y^{\#Y}, \tag{2.25}$$

and then setting $x=0$ and $y=-1$, respectively.

As a example of Theorem 2.3.1, take $B=\{(1,1),(2,2),(3,3),(3,4),(4,4)\}$ as earlier. Then

$$\begin{aligned} N_4(x) &= 4! + 5\cdot 3!(x-1) + 8\cdot 2!(x-1)^2 + 5\cdot 1!(x-1)^3 + (x-1)^4 \\ &= x^4 + x^3 + 7x^2 + 9x + 6. \end{aligned}$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

2.3.2 Example (derangements revisted). Take $B = \{(1,1),(2,2),\dots,(n,n)\}$. We want to compute $N_0 = D(n)$. Clearly, $r_k = \binom{n}{k}$, so

$$\begin{aligned} N_n(x) &= \sum_{k=0}^{n} \binom{n}{k} (n-k)!(x-1)^k \\ &= \sum_{k=0}^{n} \frac{n!}{k!}(x-1)^k \\ \Rightarrow N_0 &= \sum_{k=0}^{n} (-1)^k \frac{n!}{k!}. \end{aligned}$$

2.3.3 Example. (Problème des ménages). This famous problem is equivalent to asking for the number $M(n)$ of permutations $w \in \mathfrak{S}_n$ such that $w(i) \not\equiv i, i+1 \pmod{n}$ for all $i \in [n]$. In other words, we seek N_0 for the board

$$B = \{(1,1),(2,2),\dots,(n,n),(1,2),(2,3),\dots,(n-1,n),(n,1)\}.$$

By looking at a picture of B, we see that r_k is equal to the number of ways to choose k points, no two consecutive, from a collection of $2n$ points arranged in a circle.

2.3.4 Lemma. *The number of ways to choose k points, no two consecutive, from a collection of m points arranged in a circle is* $\frac{m}{m-k}\binom{m-k}{k}$.

First Proof. Let $f(m,k)$ be the desired number; and let $g(m,k)$ be the number of ways to choose k nonconsecutive points from m points arranged in a circle, next coloring the k points red, and then coloring one of the non-red points blue. Clearly $g(m,k) = (m-k)f(m,k)$. But we can also compute $g(m,k)$ as follows. First, color a point blue in m ways. We now need to color k points red, no two consecutive, from a *linear* array of $m-1$ points. One way to proceed is as follows. (See also Exercise 1.34.) Place $m-1-k$ uncolored points on a line, and insert k red points into the $m-k$ spaces between the uncolored points (counting the beginning and end) in $\binom{m-k}{k}$ ways. Hence, $g(m,k) = m\binom{m-k}{k}$, so $f(m,k) = \frac{m}{m-k}\binom{m-k}{k}$. □

This proof is based on a general principle of passing from "circular" to "linear" arrays. We will discuss this principle further in Chapter 4 (see Proposition 4.7.13).

Second Proof. Label the points $1,2,\dots,m$ in clockwise order. We wish to color k of them red, no two consecutive. First, we count the number of ways when 1 isn't colored red. Place $m-k$ uncolored points on a circle, label one of these 1, and insert k red points into the $m-k$ spaces between the uncolored points in $\binom{m-k}{k}$ ways. On the other hand, if 1 is to be colored red, then place $m-k-1$ points on the circle, color one of these points red and label it 1, and then insert in $\binom{m-k-1}{k-1}$

ways $k-1$ red points into the $m-k-1$ allowed spaces. Hence,

$$f(m,k) = \binom{m-k}{k} + \binom{m-k-1}{k-1} = \frac{m}{m-k}\binom{m-k}{k}.$$

□

2.3.5 Corollary. *The polynomial* $N_n(x)$ *for the board* $B = \{(i,i),(i,i+1)\,(\mathrm{mod}\, n) : 1 \le i \le n\}$ *is given by*

$$N_n(x) = \sum_{k=0}^{n} \frac{2n}{2n-k}\binom{2n-k}{k}(n-k)!(x-1)^k.$$

In particular, the number N_0 *of permutations* $w \in \mathfrak{S}_n$ *such that* $w(i) \ne i, i+1\,(\mathrm{mod}\, n)$ *for* $1 \le i \le n$ *is given by*

$$N_0 = \sum_{k=0}^{n} \frac{2n}{2n-k}\binom{2n-k}{k}(n-k)!(-1)^k.$$

Corollary 2.3.5 suggest the following question. Let $1 \le k \le n$, and let $B_{n,k}$ denote the board

$$B_{n,k} = \{(i,i),(i,i+1),\dots,(i,i+k-1)\,(\mathrm{mod}\, n) : 1 \le i \le n\}.$$

Find the rook polynomial $R_{n,k}(x) = \sum_i r_i(n,k)x^i$ of $B_{n,k}$. Thus, by equation (2.23) the number $f(n,k)$ of permutations $w \in \mathfrak{S}_n$ satisfying $w(i) \not\equiv i, i+1,\dots,i+k-1\,(\mathrm{mod}\, n)$ is given by

$$f(n,k) = \sum_{i=0}^{n}(-1)^i r_i(n,k)(n-i)!.$$

Such permutations are called *k-discordant*. For instance, 1-discordant permutations are just derangements. When $k > 2$ there is no simple explicit expression for $r_i(n,k)$ as there was for $k = 1,2$. However, we shall see in Example 4.7.19 that there exists a polynomial $Q_k(x,y) \in \mathbb{Z}[x,y]$ such that

$$\sum_n R_{n,k}(x)y^n = \frac{-y\dfrac{\partial}{\partial y}Q_k(x,y)}{Q_k(x,y)},$$

provided that $R_{n,k}(x)$ is suitably interpreted when $n < k$. For instance,

$$\begin{aligned}
Q_1(x,y) &= 1-(1+x)y,\\
Q_2(x,y) &= (1-(1+2x)y+x^2y^2)(1-xy),\\
Q_3(x,y) &= (1-(1+2x)y-xy^2+x^3y^3)(1-xy).
\end{aligned}$$

2.4 Ferrers Boards

Given a particular board or class of boards B, we can ask whether the rook numbers r_i have any special properties of interest. Here we will discuss a class of boards called *Ferrers boards*. Given integers $0 \le b_1 \le \cdots \le b_m$, the Ferrers board of shape $(b_1, \ldots, b_m)$ is defined by

$$B = \{(i,j) : 1 \le i \le m,\ 1 \le j \le b_i\},$$

where we are using ordinary cartesian coordinates so the $(1,1)$ square is at the bottom left. The board B depends (up to translation) only on the *positive* b_i's. However, it will prove to be a technical convenience to allow $b_i = 0$. Note that B is just a reflection and rotation of the Young diagram of the partition $\lambda = (b_m, \ldots, b_1)$.

2.4.1 Theorem. *Let $\sum r_k x^k$ be the rook polynomial of the Ferrers board B of shape $(b_1, \ldots, b_m)$. Set $s_i = b_i - i + 1$. Then*

$$\sum_k r_k \cdot (x)_{m-k} = \prod_{i=1}^{m} (x + s_i).$$

Proof. Let $x \in \mathbb{N}$, and let B' be the Ferrers board of shape $(b_1 + x, \ldots, b_m + x)$. Regard $B' = B \cup C$, where C is an $x \times m$ rectangle placed below B. See Figure 2.1 for the case $(b_1, b_2, b_3) = (1,2,4)$. We count $r_m(B')$ in two ways.

1. Place k rooks on B in r_k ways, then $m-k$ rooks on C in $(x)_{m-k}$ ways, to get

$$r_m(B') = \sum_k r_k \cdot (x)_{m-k}.$$

2. Place a rook in the first column of B' in $x + b_1 = x - s_1$ ways, then a rook in the second column in $x + b_2 - 1 = x + s_2$ ways, and so on, to get

$$f_m(B') = \prod_{i=1}^{m} (x + s_i).$$

This completes the proof. □

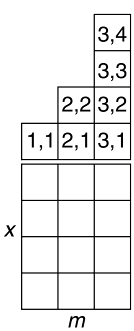

Figure 2.1 The Ferrers board of shape $(1,2,4)$ with a rectangle C underneath.

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

2.4.2 Corollary. *Let B be the triangular board (or* staircase*) of shape* $(0,1,2,\ldots,m-1)$. *Then* $r_k = S(m,m-k)$.

Proof. We have that each $s_i = 0$. Hence, by Theorem 2.4.1,

$$x^m = \sum r_k \cdot (x)_{m-k}.$$

It follows from equation (1.94d) that $r_k = S(m,m-k)$. □

A combinatorial proof of Corollary 2.4.2 is clearly desirable. We wish to associate a partition of $[m]$ into $m-k$ blocks with a placement of k nonattacking rooks on $B = \{(i,j) : 1 \le i \le m,\ 1 \le j < i\}$. If a rook occupies (i,j), then define i and j to be in the same block of the partition. It is easy to check that this procedure yields the desired correspondence. □

2.4.3 Corollary. *Two Ferrers boards, each with m columns (allowing empty columns), have the same rook polynomial if and only if their multisets of the numbers s_i are the same.*

Corollary 2.4.3 suggests asking for the number of Ferrers boards with a rook polynomial equal to that of a given board B.

2.4.4 Theorem. *Let $0 \le c_1 \le \cdots \le c_m$, and let $f(c_1,\ldots,c_m)$ be the number of Ferrers boards with no empty columns and having the same rook polynomial as the Ferrers board of shape $(c_1,\ldots,c_m)$. Add enough intial 0's to $c_1,\ldots,c_m$ to get a shape $(b_1,\ldots,b_t) = (0,0,\ldots,0,c_1,\ldots,c_m)$ such that if $s_i = b_i - i + 1$, then $s_1 = 0$ and $s_i < 0$ for $2 \le i \le t$. Suppose that a_i of the s_j's are equal to $-i$, so $\sum_{i\ge 1} a_i = t-1$. Then*

$$f(c_1,\ldots,c_m) = \binom{a_1+a_2-1}{a_2}\binom{a_2+a_3-1}{a_3}\binom{a_3+a_4-1}{a_4}\cdots.$$

Proof (sketch). By Corollary 2.4.3, we seek the number of permutations $d_1d_2\cdots d_{t-1}$ of the multiset $\{1^{a_1},2^{a_2},\ldots\}$ such that $0 \ge d_1 - 1 \ge d_2 - 2 \ge \cdots \ge d_{t-1} - t + 1$. Equivalently, $d_1 = 1$ and d_i must be followed by a number $d_{i+1} \le d_i + 1$. Place the a_1 1's down in a line. The a_2 2's may be placed arbitrarily in the a_1 spaces following each 1 in $\left(\!\binom{a_1}{a_2}\!\right) = \binom{a_1+a_2-1}{a_2}$ ways. Now the a_3's may be placed arbitrarily in the a_2 spaces following each 2 in $\left(\!\binom{a_2}{a_3}\!\right) = \binom{a_2+a_3-1}{a_3}$ ways, and so on, completing the proof. □

For instance, there are no other Ferrers boards with the same rook polynomial as the triangular board $(0,1,\ldots,n-1)$, while there are 3^{n-1} Ferrers boards with the same rook polynomial as the $n \times n$ chessboard $[n]\times[n]$.

If in the proof of Theorem 2.4.4 we want all the columns of our Ferrers board to have distinct lengths, then we must arrange the multiset $\{1^{a_1},2^{a_2},\ldots\}$ to first strictly increase from 1 to its maximum in unit steps and then to be non-increasing. Hence, we obtain the following result.

2.4.5 Corollary. *Let B be a Ferrers board. Then there is a unique Ferrers board whose columns have distinct (nonzero) lengths and that has the same rook polynomial as B.*

For instance, the unique "increasing" Ferrers board with the same rook polynomial as $[n] \times [n]$ has shape $(1, 3, 5, \ldots, 2n-1)$.

2.5 V-Partitions and Unimodal Sequences

We now give an example of a sieve process that cannot be derived (except in a very contrived way) using the Principle of Inclusion-Exclusion. By a *unimodal sequence of weight n* (also called an *n-stack*), we mean a $\mathbb{P}$-sequence $d_1 d_2 \cdots d_m$ such that

a. $\sum d_i = n$.
b. For some j, we have $d_1 \leq d_2 \leq \cdots \leq d_j \geq d_{j+1} \geq \cdots \geq d_m$.

Many interesting combinatorial sequences turn out to be unimodal. (See Exercise 1.50 for some examples.) In this section, we shall be concerned not with any particular sequence, but rather with counting the total number $u(n)$ of unimodal sequences of weight n. By convention we set $u(0) = 0$. For instance, $u(5) = 15$, since all 16 compositions of 5 are unimodal except 212. Now set

$$\begin{aligned} U(q) &= \sum_{n \geq 0} u(n) q^n \\ &= q + 2q^2 + 4q^3 + 8q^4 + 15q^5 + 27q^6 + 47q^7 + 79q^8 + \cdots . \end{aligned}$$

Our object is to find a nice expression for $U(q)$. Write $[j]! = (1-q)(1-q^2)\cdots(1-q^j)$. It is easy to see that the number of unimodal sequences of weight n with largest term k is the coefficient of q^n in $q^k/[k-1]![k]!$. Hence,

$$U(q) = \sum_{k \geq 1} \frac{q^k}{[k-1]![k]!} \tag{2.26}$$

This is analogous to the formula

$$\sum_{n \geq 0} p(n) q^n = \sum_{k \geq 0} \frac{q^k}{[k]!},$$

where $p(n)$ is the number of partitions of n. (Put $x = 1$ in equation (1.82).) What we want, however, is an analogue of equation (1.77), which states that

$$\sum_{n \geq 0} p(n) q^n = \prod_{i \geq 1} (1 - q^i)^{-1}.$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

It turns out to be easier to work with objects slightly different from unimodal sequences, and then relate them to unimodal sequences at the end. We define a *V-partition of n* to be an $\mathbb{N}$-array

$$\left[\begin{array}{cccc} & a_1 & a_2 & \cdots \\ c & & & \\ & b_1 & b_2 & \cdots \end{array}\right] \tag{2.27}$$

such that $c+\sum a_i+\sum b_i = n$, $c \geq a_1 \geq a_2 \geq \cdots$ and $c \geq b_1 \geq b_2 \geq \cdots$. Hence, a V-partition may be regarded as a unimodal sequence "rooted" at one of its largest parts. Let $v(n)$ be the number of V-partitions of n, with $v(0)=1$. Thus, for instance $v(4)=12$, since there is one way of rooting 4, one way for 13, one for 31, two for 22, one for 211, one for 112, and four for 1111. Set

$$\begin{aligned} V(q) &= \sum_{n\geq 0} v(n)q^n \\ &= 1+q+3q^2+6q^3+12q^4+21q^5+38q^6+63q^7+106q^8+\cdots. \end{aligned}$$

Analogously to (2.26), we have

$$V(q) = \sum_{k\geq 0} \frac{q^k}{[k]!^2},$$

but as before we want a product formula for $V(q)$.

Let V_n be the set of all V-partitions of n, and let D_n be the set of all *double partitions* of n, that is, $\mathbb{N}$-arrays

$$\left[\begin{array}{ccc} a_1 & a_2 & \cdots \\ b_1 & b_2 & \cdots \end{array}\right] \tag{2.28}$$

such that $\sum a_i + \sum b_i = n$, $a_1 \geq a_2 \geq \cdots$ and $b_1 \geq b_2 \geq \cdots$. If $d(n)\#D_n$, then clearly

$$\sum_{n\geq 0} d(n)q^n = \prod_{i\geq 1}(1-q^i)^{-2}. \tag{2.29}$$

Now define $\Gamma_1 : D_n \to V_n$ by

$$\Gamma_1\left[\begin{array}{ccc} a_1 & a_2 & \cdots \\ b_1 & b_2 & \cdots \end{array}\right] = \begin{cases} \left[\begin{array}{cccc} & a_2 & a_3 & \cdots \\ a_1 & & & \\ & b_1 & b_2 & \cdots \end{array}\right], & \text{if } a_1 \geq b_1, \\ \left[\begin{array}{cccc} & a_1 & a_2 & \cdots \\ b_1 & & & \\ & b_2 & b_3 & \cdots \end{array}\right], & \text{if } b_1 > a_1. \end{cases}$$

Clearly, Γ_1 is surjective, but it is not injective. Every V-partition in the set

$$V_n^1 = \left\{ \left[\begin{array}{cccc} & a_1 & a_2 & \cdots \\ c & & & \\ & b_1 & b_2 & \cdots \end{array}\right] \in V_n : c > a_1 \right\}$$

appears twice as a value of Γ_1, so

$$\#V_n = \#D_n - \#V_n^1.$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

Next define $\Gamma_2\colon D_{n-1} \to V_n^1$ by

$$\Gamma_2 \begin{bmatrix} a_1 & a_2 & \cdots \\ b_1 & b_2 & \cdots \end{bmatrix} = \begin{cases} \begin{bmatrix} a_1+1 & \begin{matrix} a_2 & a_3 & \cdots \\ b_1 & b_2 & \cdots \end{matrix} \end{bmatrix}, & \text{if } a_1+1 \geq b_1, \\ \begin{bmatrix} b_1 & \begin{matrix} a_1+1 & a_2 & \cdots \\ b_2 & b_3 & \cdots \end{matrix} \end{bmatrix}, & \text{if } b_1 > a_1+1. \end{cases}$$

Again Γ_2 is surjective, but every V-partition in the set

$$V_n^2 = \left\{ \begin{bmatrix} c & \begin{matrix} a_1 & a_2 & \cdots \\ b_1 & b_2 & \cdots \end{matrix} \end{bmatrix} \in V_n \,:\, c > a_1 > a_2 \right\}$$

appears twice as a value of Γ_2. Hence, $\#V_n^1 = \#D_{n-1} - \#V_n^2$, so

$$\#V_n = \#D_n - \#D_{n-1} + \#V_n^2.$$

Next define $\Gamma_3\colon D_{n-3} \to V_n^2$ by

$$\Gamma_3 \begin{bmatrix} a_1 & a_2 & \cdots \\ b_1 & b_2 & \cdots \end{bmatrix} = \begin{cases} \begin{bmatrix} a_1+2 & \begin{matrix} a_2+1 & a_3 & a_4 & \cdots \\ b_1 & b_2 & b_3 & \cdots \end{matrix} \end{bmatrix}, & \text{if } a_1+2 \geq b_1, \\ \begin{bmatrix} b_1 & \begin{matrix} a_1+2 & a_2+1 & a_3 & \cdots \\ b_2 & b_3 & b_4 & \cdots \end{matrix} \end{bmatrix}, & \text{if } b_1 > a_1+2. \end{cases}$$

We obtain

$$\#V_n = \#D_n - \#D_{n-1} + \#D_{n-3} - \#V_n^3,$$

where

$$V_n^3 = \left\{ \begin{bmatrix} c & \begin{matrix} a_1 & a_2 & \cdots \\ b_1 & b_2 & \cdots \end{matrix} \end{bmatrix} \in V_n \,:\, c > a_1 > a_2 > a_3 \right\}.$$

Continuing this process, we obtain maps $\Gamma_i\colon D_{n-\binom{i}{2}} \to V_n^{i-1}$. The process stops when $\binom{i}{2} > n$, so we obtain the sieve-theoretic formula

$$v(n) = d(n) - d(n-1) + d(n-3) - d(n-6) + \cdots,$$

where we set $d(m) = 0$ for $m < 0$. Thus, using equation (2.29), we obtain the following result.

2.5.1 Proposition. *We have*

$$V(q) = \left(\sum_{n \geq 0} (-1)^n q^{\binom{n+1}{2}} \right) \prod_{i \geq 1} (1-q^i)^{-2}.$$

We can now obtain an expression for $U(q)$ using the following result.

2.5.2 Proposition. *We have*

$$U(q) + V(q) = \prod_{i \geq 1} (1-q^i)^{-2}.$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

Proof. Let U_n be the set of all unimodal sequences of weight n. We need to find a bijection $D_n \to U_n \cup V_n$. Such a bijection is given by

$$\begin{bmatrix} a_1 & a_2 & \cdots \\ b_1 & b_2 & \cdots \end{bmatrix} \mapsto \begin{cases} \begin{bmatrix} a_1 & \begin{matrix} a_2 & a_3 & \cdots \\ b_1 & b_2 & \cdots \end{matrix} \end{bmatrix}, & \text{if } a_1 \geq b_1 \\ \cdots a_2\, a_1\, b_1\, b_2 \cdots, & \text{if } b_1 > a_1 \end{cases}$$

□

2.5.3 Corollary. *We have*

$$U(q) = \left(\sum_{n \geq 1} (-1)^{n-1} q^{\binom{n+1}{2}} \right) \prod_{i \geq 1} (1 - q^i)^{-2}.$$

2.6 Involutions

Recall now the viewpoint of Section 1.1 that the best way to determine that two finite sets have the same cardinality is to exhibit a bijection between them. We will show how to apply this principle to the identity (2.5). (The seemingly more general (2.4) is done exactly the same way.) As it stands this identity does not assert that two sets have the same cardinality. Therefore, we rearrange terms so that all signs are positive. Thus, we wish to prove the identity

$$f_=(\emptyset) + \sum_{\#Y \text{ odd}} f_\geq(Y) = \sum_{\#Y \text{ even}} f_\geq(Y), \tag{2.30}$$

where $f_=(Y)$ (respectively, $f_\geq(Y)$) denotes the number of objects in a set A having exactly (respectively, at least) the properties in $T \subseteq S$. The left-hand side of (2.30) is the cardinality of the set $M \cup N$, where M is the set of objects x having none of the properties in S, and N is the set of ordered triples (x, Y, Z), where $x \in A$ has exactly the properties $Z \supseteq Y$ with $\#Y$ odd. The right-hand side of (2.30) is the cardinality of the set N' of ordered triples (x', Y', Z'), where $x' \in A$ has exactly the properties $Z' \supseteq Y'$ with $\#Y'$ even. Totally order the set S of properties, and define $\sigma : M \cup N \to N'$ as follows:

$$\sigma(x) = (x, \emptyset, \emptyset), \quad \text{if } x \in M$$

$$\sigma(x, Y, Z) = \begin{cases} (x, Y - i, Z), & \text{if } (x, Y, Z) \in N \\ & \text{and } \min Y = \min Z = i, \\ & \\ (x, Y \cup i, Z), & \text{if } (x, Y, Z) \in N \\ & \text{and } \min Z = i < \min Y. \end{cases}$$

It is easily seen that σ is a bijection with inverse

$$\sigma^{-1}(x,Y,Z) = \begin{cases} x \in M, & \text{if } Y = Z = \emptyset, \\ (x, Y-i, Z) \in N, & \text{if } Y \neq \emptyset \\ & \text{and } \min Y = \min Z = i, \\ (x, Y \cup i, Z) \in N, & \text{if } Z \neq \emptyset \text{ and} \\ & \min Z = i < \min Y \\ & \text{(where we set } \min Y = \infty \text{ if } Y = \emptyset). \end{cases}$$

This construction yields the desired bijective proof of (2.30).

Note that if in the definition of σ^{-1} we identify $x \in M$ with $(x,\emptyset,\emptyset) \in N'$ (so $\sigma^{-1}(x,\emptyset,\emptyset) = (x,\emptyset,\emptyset)$), then $\sigma \cup \sigma^{-1}$ is a function $\tau\colon N \cup N' \to N \cup N'$ satisfying: (a) τ is an *involution*; that is, $\tau^2 = \mathrm{id}$; (b) the fixed points of τ are the triples $(x,\emptyset,\emptyset)$, so are in one-to-one correspondence with M; and (c) if (x,Y,Z) is not a fixed point of τ and we set $\tau(x,Y,Z) = (x,Y',Z')$, then

$$(-1)^{\#Y} + (-1)^{\#Y'} = 0.$$

Thus, the involution τ selects terms from the right-hand side of (2.5) (or rather, terms from the right-hand side of (2.5) after each $f_{\geq}(Y)$ is written as a sum (2.3)) that add up to the left-hand side, and then τ cancels out the remaining terms.

We can put the preceding discussion in the following general context. Suppose that the finite set X is written as a disjoint union $X^+ \cup X^-$ of two subsets X^+ and X^-, called the "positive" and "negative" parts of X, respectively. Let τ be an involution on X that satisfies:

a. If $\tau(x) = y$ and $x \neq y$, then exactly one of x, y belongs to X^+ (so the other belongs to X^-).
b. If $\tau(x) = x$ then $x \in X^+$.

If we define a weight function w on X by

$$w(x) = \begin{cases} 1, & x \in X^+, \\ -1, & x \in X^-, \end{cases}$$

then clearly

$$\#\mathrm{Fix}(\tau) = \sum_{x \in X} w(x), \tag{2.31}$$

where $\mathrm{Fix}(\tau)$ denotes the fixed point set of τ. Just as in the previous paragraph, the involution τ has selected terms from the right-hand side of (2.31) which add up to the left-hand side, and has cancelled the remaining terms.

We now consider a more complicated situation. Suppose that we have another set $\widetilde{X}$ that is also expressed as a disjoint union $\widetilde{X} = \widetilde{X}^+ \cup \widetilde{X}^-$, and an involution $\tilde{\tau}$ on $\widetilde{X}$ satisfying (a) and (b) Suppose that we also are given a sign-preserving bijection $f\colon X \to \widetilde{X}$, that is, $f(X^+) = \widetilde{X}^+$ and $f(X^-) = \widetilde{X}^-$. Clearly, then

$\#\mathrm{Fix}(\tau) = \#\mathrm{Fix}(\tilde{\tau})$, since $\#\mathrm{Fix}(\tau) = \#X^+ - \#X^-$ and $\#\mathrm{Fix}(\tilde{\tau}) = \#\widetilde{X}^+ - \#\widetilde{X}^-$. We wish to construct in a canonical way a bijection g between $\mathrm{Fix}(\tau)$ and $\mathrm{Fix}(\tilde{\tau})$. This construction is known as the *involution principle* and is a powerful technique for converting noncombinatorial proofs into combinatorial ones.

The bijection $g\colon \mathrm{Fix}(\tau) \to \mathrm{Fix}(\tilde{\tau})$ is defined as follows. Let $x \in \mathrm{Fix}(\tau)$. It is easily seen, since X is finite, that there is a nonnegative integer n for which

$$f(\tau f^{-1}\tilde{\tau} f)^n(x) \in \mathrm{Fix}(\tilde{\tau}). \tag{2.32}$$

Define $g(x)$ to be $f(\tau f^{-1}\tilde{\tau} f)^n(x)$ where n is the *least* nonnegative integer for which (2.32) holds.

We leave it to the reader to verify rigorously that g is a bijection from $\mathrm{Fix}(\tau)$ to $\mathrm{Fix}(\tilde{\tau})$. There is, however, a nice geometric way to visualize the situation. Represent the elements of X and $\widetilde{X}$ as vertices of a graph Γ. Draw an undirected edge between two distinct vertices x and y if (1) $x, y \in X$ and $\tau(x) = y$; or (2) $x, y \in \widetilde{X}$ and $\tilde{\tau}(x) = y$; or (3) $x \in X$, $y \in \widetilde{X}$, and $f(x) = y$. Every component of Γ will then be either a cycle disjoint from $\mathrm{Fix}(\tau)$ and $\mathrm{Fix}(\tilde{\tau})$, or a path with one endpoint z in $\mathrm{Fix}(\tau)$ and the other endpoint $\tilde{z}$ in $\mathrm{Fix}(\tilde{\tau})$. Then g is defined by $g(z) = \tilde{z}$. See Figure 2.2.

There is a variation of the involution principle that is concerned with "sieve-equivalence." We will mention only the simplest case here; see Exercise 2.36 for further development. Suppose that X and $\widetilde{X}$ are (disjoint) finite sets. Let $Y \subseteq X$ and $\widetilde{Y} \subseteq \widetilde{X}$, and suppose that we are given bijections $f\colon X \to \widetilde{X}$ and $g\colon Y \to \widetilde{Y}$. Hence $\#(X - Y) = \#(\widetilde{X} - \widetilde{Y})$, and we wish to construct an explicit bijection h between $X - Y$ and $\widetilde{X} - \widetilde{Y}$. Pick $x \in X - Y$. As in equation (2.32), there will be a nonnegative integer n for which

$$f(g^{-1} f)^n(x) \in \widetilde{X} - \widetilde{Y}. \tag{2.33}$$

In this case, n is unique since if $x \in \widetilde{X} - \widetilde{Y}$, then $g^{-1}(y)$ is undefined. Define $h(x)$ to be $f(g^{-1} f)^n(x)$ where n satisfies (2.33). One easily checks that $h\colon X - Y \to \widetilde{X} - \widetilde{Y}$ is a bijection.

Let us consider a simple example of the bijection $h\colon X - Y \to \widetilde{X} - \widetilde{Y}$.

2.6.1 Example. Let Y be the set of all permutations $w \in \mathfrak{S}_n$ that fix 1, that is, $w(1) = 1$. Let $\widetilde{Y}$ be the set of all permutations $w \in \mathfrak{S}_n$ with exactly one cycle.

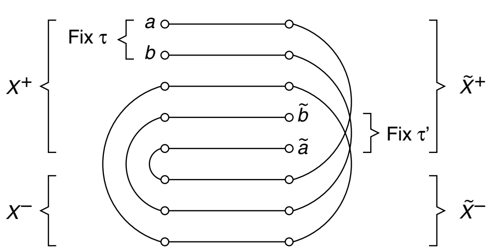

Figure 2.2 An illustration of the involution principle.

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

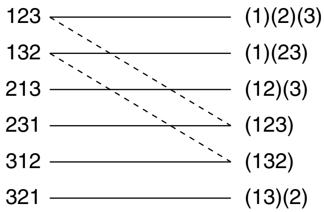

Figure 2.3 The bijection $h\colon \mathfrak{S}_n - Y \to \mathfrak{S}_n - \widetilde{Y}$.

Thus, $\#Y = \#\widetilde{Y} = (n-1)!$, so

$$\#(\mathfrak{S}_n - Y) = \#(\mathfrak{S}_n - \widetilde{Y}) = n! - (n-1)!.$$

It may not be readily apparent, however, how to construct a bijection h between $\mathfrak{S}_n - Y$ and $\mathfrak{S}_n - \widetilde{Y}$. On the other hand, it is easy to construct a bijection g between Y and $\widetilde{Y}$; namely, if $w = 1a_2 \cdots a_n \in Y$ (where w is written as a word, i.e., $w(i) = a_i$), then set $g(w) = (1, a_2, \dots, a_n)$ (written as a cycle). Of course we choose the bijection $f\colon \mathfrak{S}_n \to \mathfrak{S}_n$ to be the identity. Then equation (2.33) defines the bijection $h\colon \mathfrak{S}_n - Y \to \mathfrak{S}_n - \widetilde{Y}$. For example, when $n = 3$ we depict f by solid lines and g by broken lines in Figure 2.3. Hence (writing permutations in the domain as words and in the range as products of cycles),

$$\begin{aligned} h(213) &= (12)(3) \\ h(231) &= (1)(2)(3) \\ h(312) &= (1)(23) \\ h(321) &= (13)(2). \end{aligned}$$

It is natural to ask here (and in other uses of the involution and related principles) whether there is a more direct description of h. In this example there is little difficulty because Y and $\widetilde{Y}$ are *disjoint* subsets (when $n \geq 2$) of the same set $\mathfrak{S}_n$. This special situation yields

$$h(w) = \begin{cases} w, & \text{if } w \notin \widetilde{Y}, \\ g^{-1}(w), & \text{if } w \in \widetilde{Y}. \end{cases} \tag{2.34}$$

2.7 Determinants

In Proposition 2.2.6 we saw that a determinant $\det[a_{ij}]_0^n$, with $a_{ij} = 0$ if $j < i-1$, can be interpreted combinatorially using the Principle of Inclusion-Exclusion. In this section we will consider the combinatorial significance of arbitrary determinants, by setting up a combinatorial problem in which the right-hand side of equation (2.31) is the expansion of a determinant.

We will consider lattice paths $L = (v_0, v_1, \dots, v_k)$ in $\mathbb{N}^2$, as defined in Section 1.2, with steps $v_i - v_{i-1} = (1,0)$ or $(0,-1)$. We picture L by drawing an edge between v_{i-1} and v_i, $1 \leq i \leq k$. For instance, the lattice path

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

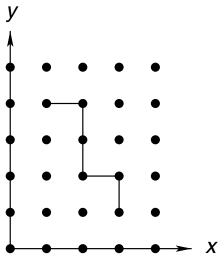

Figure 2.4 A lattice path in $\mathbb{N}^2$.

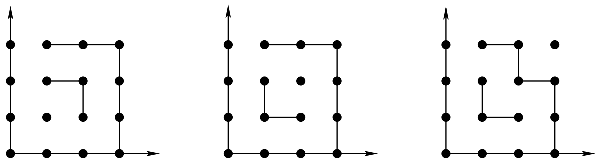

Figure 2.5 Three non-intersecting 2-paths.

$((1,4),(2,4),(2,3),(2,2),(3,2),(3,1))$ is drawn in Figure 2.4. An *n-path* is an n-tuple $\boldsymbol{L} = (L_1,\ldots,L_n)$ of lattice paths. Let $\alpha,\beta,\gamma,\delta \in \mathbb{N}^n$. Then $\boldsymbol{L}$ is of *type* $(\alpha,\beta,\gamma,\delta)$ if L_i goes from (β_i,γ_i) to (α_i,δ_i). (Clearly then, $\alpha_i \geq \beta_i$ and $\gamma_i \geq \delta_i$.) The n-path $\boldsymbol{L}$ is *intersecting* if for some $i \neq j$ L_i and L_j have a point in common; otherwise, $\boldsymbol{L}$ is *nonintersecting*. Define the *weight* of a horizontal step from (i,j) to $(i+1,j)$ to be the indeterminate x_j, and the weight $\Lambda(\boldsymbol{L})$ of $\boldsymbol{L}$ to be the product of the weights of its horizontal steps. For instance, the path in Figure 2.4 has weight x_2x_4.

If $\alpha = (\alpha_1,\ldots,\alpha_n) \in \mathbb{N}^n$ and $w \in \mathfrak{S}_n$, then let $w(\alpha) = (\alpha_{w(1)},\ldots,\alpha_{w(n)})$. Let $\mathcal{A} = \mathcal{A}(\alpha,\beta,\gamma,\delta)$ be the set of all n-paths of type $(\alpha,\beta,\gamma,\delta)$, and let $A = A(\alpha,\beta,\gamma,\delta)$ be the sum of their weights. Consider a path from (β_i,γ_i) to (α_i,δ_i). Let $m = \alpha_i - \beta_i$. For each j satisfying $1 \leq j \leq m$, there is exactly one horizontal step of the form $(j-1+\beta_i,k_j) \to (j+\beta_i,k_j)$. The numbers $k_1,\ldots,k_m$ can be chosen arbitrarily provided

$$\gamma_i \geq k_1 \geq k_2 \geq \cdots \geq k_m \geq \delta_i. \tag{2.35}$$

Hence, if we define

$$h(m;\gamma_i,\delta_i) = \sum x_{k_1}x_{k_2}\cdots x_{k_m},$$

summed over all integer sequences (2.35), then

$$A(\alpha,\beta,\gamma,\delta) = \prod_{i=1}^{n} h(\alpha_i - \beta_i;\gamma_i,\delta_i). \tag{2.36}$$

(In the terminology of Section 7.4, $h(m;\gamma_i,\delta_i)$ is the complete homogeneous symmetric function $h_m(x_{\delta_i},x_{\delta_i+1},\ldots,x_{\gamma_i}.)$

Now let $\mathcal{B} = \mathcal{B}(\alpha,\beta,\gamma,\delta)$ be the set of all nonintersecting n-paths of type $(\alpha,\beta,\gamma,\delta)$, and let $B = B(\alpha,\beta,\gamma,\delta)$ be the sum of their weights. For instance, let $\alpha = (2,3)$, $\beta = (1,1)$, $\gamma = (2,3)$, $\delta = (1,0)$. Then $B(\alpha,\beta,\gamma,\delta) = x_2x_3^2 + x_1x_3^2 + x_1x_2x_3$, corresponding to the nonintersecting 2-paths shown in Figure 2.5.

2.7.1 Theorem. *Let $\alpha,\beta,\gamma,\delta \in \mathbb{N}^n$ such that for $w \in \mathfrak{S}_n$, $\mathcal{B}(w(\alpha),\beta,\gamma,w(\delta))$ is empty unless w is the identity permutation. (For example, this condition occurs if $\alpha_i < \alpha_{i+1}$, $\beta_i < \beta_{i+1}$, $\gamma_i \le \gamma_{i+1}$, and $\delta_i \le \delta_{i+1}$ for $1 \le i \le n-1$.) Then*

$$B(\alpha,\beta,\gamma,\delta) = \det[h(\alpha_j - \beta_i;\gamma_i,\delta_j)]_1^n, \tag{2.37}$$

where we set $h(\alpha_j - \beta_i;\gamma_i,\delta_j) = 0$ whenever there are no sequences (2.35).

Proof. When we expand the right-hand side of equation (2.37), we obtain

$$\sum_{w\in\mathfrak{S}_n} (\operatorname{sgn} w) A(w(\alpha),\beta,\gamma,w(\delta)). \tag{2.38}$$

Let $\mathcal{A}_w = \mathcal{A}(w(\alpha),\beta,\gamma,w(\delta))$. We will construct a bijection $\boldsymbol{L} \to \boldsymbol{L}^*$ from $\left(\bigcup_{w\in\mathfrak{S}_n} \mathcal{A}_w\right) - \mathcal{B}$ to itself satisfying:

a. $\boldsymbol{L}^{**} = \boldsymbol{L}$; that is, * is an involution.
b. $\Lambda(\boldsymbol{L}^*) = \Lambda(\boldsymbol{L})$, that is, * is weight-preserving.
c. If $\boldsymbol{L} \in \mathcal{A}_u$ and $\boldsymbol{L}^* \in \mathcal{A}_v$ then $\operatorname{sgn} u = -\operatorname{sgn} v$.

Then by grouping together terms of (2.38) corresponding to pairs $(\boldsymbol{L},\boldsymbol{L}^*)$ of intersecting n-paths, we see that all terms cancel except for those producing the desired result $B(\alpha,\beta,\gamma,\delta)$.

To construct the involution *, let $\boldsymbol{L}$ be an intersecting n-path. We need to single out some canonically defined pair (L_i,L_j) of paths from $\boldsymbol{L}$ that intersect, and then some canonically defined intersection point (x,y) of these paths. One of many ways to do this is the following. Let i be the least integer for which L_i and L_k intersect for some $k \ne i$, and let x be the least integer such that L_i intersects some L_k with $k > i$ at a point (x,y), and then of all such k let j be the minimum. Construct L_i^* by following L_i to its first intersection point $v = (x,y)$ with L_j, and then following L_j to the end. Construct L_j^* similarly by following L_j to v and then L_i to the end. For $k \ne i,j$ let $L_k^* = L_k$.

Property (a) follows since the triple (i,j,v) can be obtained from $\boldsymbol{L}^*$ by the same rule that $\boldsymbol{L}^*$ is obtained from $\boldsymbol{L}$. Property (b) is immediate since the totality of single steps in $\boldsymbol{L}$ and $\boldsymbol{L}^*$ is identical. Finally, v is obtained from u by multiplication by the transposition (i,j), so (c) follows. □

Theorem 2.7.1 has important applications in the theory of symmetric functions (see the first proof of Theorem 7.16.1), but let us be content here with a simple example of its use.

2.7.2 Example. Let $r,s \in \mathbb{N}$ and let S be a subset of $[0,r] \times [0,s]$. How many lattice paths are there between $(0,r)$ and $(s,0)$ that don't intersect S? Call this number $f(r,s,S)$. Let $S = \{(a_1,b_1),\ldots,(a_k,b_k)\}$, and set

$$\alpha = (s,a_1,\ldots,a_k), \quad \beta = (0,a_1,\ldots,a_k),$$

$$\gamma = (r,b_1,\ldots,b_k), \quad \delta = (0,b_1,\ldots,b_k).$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

Then $f(r,s,S) = B(\alpha,\beta,\gamma,\delta)$, where we set each weight $x_m = 1$. Now

$$h(\alpha_j - \beta_i;\gamma_i,\delta_j)\big|_{xm=1} = \binom{\alpha_j + \gamma_i - \beta_i - \delta_j}{\alpha_j - \beta_i}.$$

Hence by Theorem 2.7.1,

$$f(r,s,S) = \begin{vmatrix} \binom{r+s}{r} & \binom{r+a_1-b_1}{a_1} & \cdots & \binom{r+a_k-b_k}{a_k} \\ \binom{s-a_1+b_1}{s-a_1} & 1 & \cdots & \binom{a_k-b_k-a_1+b_1}{a_k-a_1} \\ & \vdots & & \\ \binom{s-a_k+b_k}{s-a_k} & \binom{a_1-b_1-a_k+b_k}{a_1-a_k} & \cdots & 1 \end{vmatrix},$$

where we set $\binom{i}{j} = 0$ if $j < 0$ or $i - j < 0$. When we expand this determinant we obtain a formula for $f(r,s,S)$ that can also be deduced directly from the Principle of Inclusion-Exclusion. Indeed, by a suitable permutation of rows and columns the above expression for $f(r,s,S)$ becomes a special case of Proposition 2.2.6. (In its full generality, however, Theorem 2.7.1 cannot be deduced from Proposition 2.2.6; indeed, the determinant (2.37) will in general have no zero entries.)

Notes

As P. Stein says in his valuable monograph [1.71], the Principle of Inclusion-Exclusion "is doubtless very old; its origin is probably untraceable." An extensive list of references is given by Takács [2.22], and exact citations for results listed here without reference may be found there. In probabilistic form, the Principle of Inclusion-Exclusion can be traced back to A. de Moivre and less clearly to J. Bernoulli, and is sometimes referred to as "Poincaré's theorem." The first statement in combinatorial terms may be due to C. P. da Silva and is sometimes attributed to Sylvester.

Example 2.2.1 (the derangement problem) was first solved by P. R. de Montmort (in probabilistic terms) and later independently investigated by Euler.

Example 2.2.4 (enumeration of permutations by descent set) was first obtained by MacMahon [1.55, vol. 1, p. 190] and has been rediscovered several times since. Example 2.2.5 first appears in Stanley [2.21]. The problème des ménages (or *menage problem*) (Example 2.3.3) was suggested by Tait to Cayley and Muir, but they did not reach a definitive answer. The problem was independently considered by Lucas and solved by him in a rather unsatisfactory form. The elegant formula given in Corollary 2.3.5 is due to Touchard. For references to more recent work see Comtet [2.3, p. 185] and Dutka [2.4]. The theory of rook polynomials in general is due to Kaplansky and Riordan [2.13]; see Riordan [2.17, Chs. 7–8]. Ferrers boards were first considered by D. Foata and M.-P. Schützenberger [2.5] and developed further by Goldman, Joichi, and White [2.8]–[2.11]. The proof given here of Theorem 2.4.4 was suggested by P. Leroux. There have been many further developments in the area of rook theory; see for instance Sjöstrand [2.18] and the references given there. The results of Section 2.5 first appeared in Stanley [2.19, Ch. IV.3] and were restated in [2.20, §23].

The involution principle was first stated by Garsia and Milne [2.6], where it was used to give a long-sought-for combinatorial proof of the Rogers-Ramanujan identities. (See Pak [1.62, §7] for more information.) For further discussion of the involution principle, sieve equivalence, and related results, see Cohen [2.2], Gordon [2.12], and Wilf [2.23]. The combinatorial proof of the Principle of Inclusion-Exclusion given in Section 2.6 appears implicitly in Remmel [2.16] and is made more explicit in Zeilberger [2.24]. Theorem 2.7.1 and its proof are anticipated by Chaundy [2.1], Karlin and McGregor [2.14], and Lindström [2.15], though the first explicit statement appears in a paper of Gessel and Viennot [2.7]. It was independently rediscovered several times since the paper of Gessel and Viennot. Our presentation closely follows that of Gessel and Viennot.

Bibliography

[1] T. W. Chaundy, Partition-generating functions, *Quart. J. Math. (Oxford)* **2** (1931), 234–240.

[2] D. I. A. Cohen, PIE-sums: A combinatorial tool for partition theory, *J. Combinatorial Theory, Ser. A* **31** (1981), 223–236.

[3] L. Comtet, *Advanced Combinatorics*, Reidel, Boston, 1974.

[4] J. Dutka, On the problème des ménages, *Math. Intelligencer* **8** (1986), no. 3, 18–25 and 33.

[5] D. Foata and M.-P. Schützenberger, On the rook polynomials of Ferrers relations, *Colloquia Mathematica Societatis Jano Bolyai, 4*, Combinatorial Theory and Its Applications, vol. 2 (P. Erdős, A. Renyi, and V. Sós, eds.), North-Holland, Amsterdam, 1970, pp. 413–436.

[6] A. M. Garsia and S. C. Milne, A Rogers-Ramanujan bijection, *J. Combinatorial Theory, Ser. A* **31** (1981), 289–339.

[7] I. M. Gessel and X. G. Viennot, Binomial determinants, paths, and hook-length formulas, *Advances in Math.* **58** (1985), 300–321.

[8] J. Goldman, J. Joichi, and D. White, Rook theory I: Rook equivalence of Ferrers boards, *Proc. Amer. Math. Soc.* **52** (1975), 485–492.

[9] J. Goldman, J. Joichi, and D. White, Rook polynomials, Möbius inversion, and the umbral calculus, *J. Combinatorial Theory, Ser. A* **21** (1976), 230–239.

[10] J. Goldman, J. Joichi, and D. White, Rook theory IV: Orthogonal sequences of rook polynomials, *Studies in Applied Math.* **56** (1977), 267–272.

[11] J. Goldman, J. Joichi, and D. White, Rook theory III: Rook polynomials and the chromatic structure of graphs, *J. Combinatorial Theory Ser. B* **25** (1978), 135–142.

[12] B. Gordon, Sieve-equivalence and explicit bijections, *J. Combinatorial Theory Ser. A* **34** (1983), 90–93.

[13] I. Kaplansky and J. Riordan, The problem of the rooks and its applications, *Duke Math. J.* **13** (1946), 259–268.

[14] S. Karlin and G. McGregor, Coincidence probabilities, *Pacific J. Math.* **9** (1959), 1141–1164.

[15] B. Lindström, On the vector representation of induced matroids, *Bull. London Math. Soc.* **5** (1973), 85–90.

[16] J. Remmel, Bijective proofs of some classical partition identities, *J. Combinatorial Theory Ser. A* **33** (1982), 273–286.
[17] J. Riordan, *An Introduction to Combinatorial Analysis*, Wiley, New York, 1958.
[18] J. Sjöstrand, Bruhat intervals as rooks on skew Ferrers boards, *J. Combinatorial Theory Ser. A* **114** (2007), 1182–1198.
[19] R. Stanley, *Ordered structures and partitions*, Ph.D. thesis, Harvard University, 1971.
[20] R. Stanley, Ordered structures and partitions, *Mem. Amer. Math. Soc.* **119** (1972), iii+104 pages.
[21] R. Stanley, Binomial posets, Möbius inversion, and permutation enumeration, *J. Combinatorial Theory* **20** (1976), 336–356.
[22] L. Takács, On the method of inclusion and exclusion, *J. Amer. Stat. Soc.* **62** (1967), 102–113.
[23] H. S. Wilf, Sieve-equivalence in generalized partition theory, *J. Combinatorial Theory, Ser. A* **34** (1983), 80–89.
[24] D. Zeilberger, Garsia and Milne's bijective proof of the inclusion-exclusion principle, *Discrete Math.* **51** (1984), 109–110.

Exercises for Chapter 2

1. [3] Explain why the Principle of Inclusion-Exclusion has the numerical value

$$8.53973422267356706546\cdots.$$

2. [2–]* Give a bijective proof of equation (2.24) or (2.25), that is,

$$\sum_X f_=(X)x^{\#X} = \sum_Y f_\geq(Y)(x-1)^{\#Y}$$

or

$$\sum_X f_=(X)(y+1)^{\#X} = \sum_Y f_\geq(Y)y^{\#Y}.$$

3. [2] Let $S=\{P_1,\ldots,P_n\}$ be a set of properties, and let f_k (respectively, $f_{\geq k}$) denote the number of objects in a finite set A that have *exactly* k (respectively, *at least* k of the properties). Show that

$$f_k = \sum_{i=k}^{n}(-1)^{i-k}\binom{i}{k}g_i, \tag{2.39}$$

and

$$f_{\geq k} = \sum_{i=k}^{n}(-1)^{i-k}\binom{i-1}{k-1}g_i, \tag{2.40}$$

where

$$g_i = \sum_{\substack{T\subseteq S\\ \#T=i}} f_\geq(T).$$

4. a. [2] Let $A_1,\ldots,A_n$ be subsets of a finite set A, and define S_k, $0\leq k\leq n$, by (2.6). Show that

$$S_k - S_{k+1} + \cdots + (-1)^{n-k}S_n \geq 0,\ \ 0\leq k\leq n. \tag{2.41}$$

b. [2+] Find necessary and sufficient conditions on a vector $(S_0,S_1,\ldots,S_n)\in\mathbb{N}^{n+1}$ so that there exist subsets $A_1,\ldots,A_n$ of a finite set A satisfying (2.6).

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

5. **a.** [2] Let

$$0 \to V_n \xrightarrow{\partial_n} V_{n-1} \xrightarrow{\partial_{n-1}} \cdots \xrightarrow{\partial_1} V_0 \xrightarrow{\partial_0} W \to 0 \tag{2.42}$$

be an exact sequence of finite-dimensional vector spaces over some field; that is, the ∂_j's are linear transformations satisfying $\operatorname{im}\partial_{j+1} = \ker\partial_j$ (with ∂_n injective and ∂_0 surjective). Show that

$$\dim W = \sum_{i=0}^{n}(-1)^i \dim V_i. \tag{2.43}$$

b. [2] Show that for $0 \le j \le n$,

$$\operatorname{rank}\partial_j = \sum_{i=j}^{n}(-1)^{i-j}\dim V_i, \tag{2.44}$$

so in particular the quantity on the right-hand side is nonnegative.

c. [2] Suppose that we are given only that equation (2.42) is a *complex*; that is, $\partial_j\partial_{j+1} = 0$ for $0 \le j \le n-1$, or equivalently $\operatorname{im}\partial_{j+1} \subseteq \ker\partial_j$. Show that if equation (2.44) holds for $0 \le j \le n$, then (2.42) is exact.

d. [2+] Let $A_1,\dots,A_n$ be subsets of a finite set A, and for $T \subseteq [n]$ set $A_T = \bigcap_{i\in T} A_i$. In particular, $A_\emptyset = A$. Let V_T be the vector space (over some field) with a basis consisting of all symbols $[a,T]$ where $a \in A_T$. Set $V_j = \bigoplus_{\#T=j} V_T$, and define for $1 \le i \le n$ linear transformations $\partial_j\colon V_j \to V_{j-1}$ by

$$\partial_j[a,T] = \sum_{i=1}^{j}(-1)^{i-1}[a, T - t_i], \tag{2.45}$$

where the elements of T are $t_1 < \cdots < t_j$. Also, define W to be the vector space with basis $\{[a] : a \in \bar{A}_1 \cap \cdots \cap \bar{A}_n\}$, and define $\partial_0\colon V_0 \to W$ by

$$\partial_0[a,\emptyset] = \begin{cases} [a], & \text{if } a \in \bar{A}_1 \cap \cdots \cap \bar{A}_n, \\ 0, & \text{otherwise.} \end{cases}$$

(Here $\bar{A}_i = A - A_i$.) Show that (2.42) is an exact sequence.

e. [1+] Deduce equation (2.7) from (a) and (d).

f. [1+] Deduce Exercise 2.4(a) from (b) and (d).

6. In this exercise, we consider a *multiset* generalization of the Principle of Inclusion-Exclusion.

a. [2] Let N be a finite multiset, say $N = \{x_1^{a_1},\dots,x_k^{a_k}\}$. For each $1 \le r \le k$ and $1 \le i \le a_r$, let P_{ir} be some property that each of the elements of a set A may or may not have, with the condition that if $1 \le i \le j \le a_r$ then any object with property P_{jr} also has P_{ir}. (For instance, if A is a set of integers, then P_{ir} could be the property of being divisible by r^i.) For every submultiset $M \subseteq N$, let $f_=(M)$ be the number of objects in A with *exactly* the properties in M; in other words, if $M = \{x_1^{b_1},\dots,x_k^{b_k}\}$, then $f_=(M)$ counts those objects in A that have property $P_{b_r,r}$ but fail to have $P_{b_r+1,r}$ for $1 \le r \le k$. Similarly, define $f_\ge(M)$ so

$$f_\ge(M) = \sum_{Y\supseteq M} f_=(Y). \tag{2.46}$$

Show that

$$f_=(M) = \sum_{\substack{Y\supseteq M\\ Y-M \text{ is a set}}} (-1)^{\#(Y-M)} f_\ge(Y). \tag{2.47}$$

Dually, if

$$f_{\leq}(M) = \sum_{Y \subseteq M} f_{=}(Y), \tag{2.48}$$

then

$$f_{=}(M) = \sum_{\substack{Y \subseteq M \\ M-Y \text{ is a set}}} (-1)^{\#(M-Y)} f_{\leq}(Y). \tag{2.49}$$

b. [2] Suppose that we encode the multiset $N = \{x_1^{a_1}, \dots, x_k^{a_k}\}$ by the integer $n = p_1^{a_1} \cdots p_k^{a_k}$, where $p_1, \dots, p_k$ are distinct primes. Thus, submultisets M of N correspond to (positive) divisors d of n. What do equations (2.48) and (2.49) become in this setting?

7. [2] Fix a prime power q. Prove equation (1.103), namely, the number $\beta(n)$ of monic irreducible polynomials of degree n over the field $\mathbb{F}_q$ is given by

$$\beta(n) = \frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}.$$

(Use Exercise 2.6(b).)

8. **a.** [3–] Give a direct combinatorial proof of equation (2.13); that is,

$$D(n) = nD(n-1) + (-1)^n.$$

b. [2] Let $\mathcal{E}(n)$ denote the set of permutations $w \in \mathfrak{S}_n$ whose first ascent is in an even position (where we always count n as an ascent). For instance, $\mathcal{E}(3) = \{213, 312\}$, and $\mathcal{E}(4) = \{2134, 2143, 3124, 3142, 3241, 4123, 4132, 4231, 4321\}$. Set $E(n) = \#\mathcal{E}(n)$. Show that $E(n) = nE(n-1) + (-1)^n$. Hence (since $E(1) = D(1) = 0$), we have $E(n) = D(n)$.

c. [2+] Give a bijection between the permutations being counted by $E(n)$ and the derangements of $[n]$.

9. [2–] Prove the formula $\Delta^k 0^d = k!S(d,k)$ of Proposition 1.9.2(c) (equivalent to equation (1.94a)) using the Principle of Inclusion-Exclusion.

10. **a.** [1+]* How many functions $f : [n] \to [n]$ have no fixed points?

b. [2] Let $E(n)$ be the number obtained in (a). Show that $\lim_{n\to\infty} E(n)/n! = 1/e$, the same as $\lim_{n\to\infty} D(n)/n!$ (Example 2.2.1). Which of $D(n)/n!$ and $E(n)/n!$ gives the better approximation to $1/e$?

11. [3–] Let $a_1, \dots, a_k$ be positive integers with $\sum a_i = n$. Let $S = \{a_1, a_1 + a_2, \dots, a_1 + a_2 + \cdots + a_{k-1}\}$. Show that the number of derangements in $\mathfrak{S}_n$ with descent set S is the coefficient of $x_1^{a_1} \cdots x_k^{a_k}$ in the expansion of

$$\frac{1}{(1+x_1)\cdots(1+x_k)(1-x_1-\cdots-x_k)}.$$

12. [2+] Let $\alpha = (\alpha_1, \dots, \alpha_k) \in \mathbb{N}^k$, and let M_α be the multiset $\{1^{\alpha_1}, \dots, k^{\alpha_k}\}$. A *derangement* of M_α is a permutation $a_1 a_2 \cdots a_n$ (where $n = \sum \alpha_i$) of M_α that disagrees in every position with the permutation we get by listing the elements of M in weakly increasing order. For instance, the multiset $\{1, 2^2, 3\}$ has the two derangements 2132 and 2312.

Let $D(\alpha)$ denote the number of derangments of M_α. Show that

$$\sum_{\alpha\in\mathbb{N}^k} D(\alpha)x^\alpha = \frac{1}{(1+x_1)\cdots(1+x_k)\left(1-\frac{x_1}{1+x_1}-\cdots-\frac{x_k}{1+x_k}\right)}$$
$$= \frac{1}{1-\sum_S(\#S-1)\prod_{i\in S}x_i},$$

where S ranges over all nonempty subsets of $[n]$.

13. Let $w = a_1a_2\cdots a_n \in \mathfrak{S}_n$. The *connectivity set* $C(w)$ of w is defined by

$$C(w) = \{i : a_j < a_k \text{ for all } j \le i < k\} \subseteq [n-1].$$

In other words, $i \in C(w)$ if $\{a_1,\dots,a_i\} = [i]$. For instance, $C(2314675) = \{3,4\}$. (Exercise 1.128(a) deals with the enumeration of permutations $w \in \mathfrak{S}_n$ satisfying $C(w) = \emptyset$.)

a. [2] If $S = \{i_1,\dots,i_k\}_< \subset [n-1]$, then let

$$\eta(S) = i_1!(i_2-i_1)!\cdots(i_k-i_{k-1})!(n-i_k)!.$$

Hence by Proposition 1.4.1 we have $\alpha(S) = n!/\eta(S)$, the number of permutations $w \in \mathfrak{S}_n$ with descent set $D(w) \subseteq S$. Show that

$$\#\{w \in \mathfrak{S}_n : S \subseteq C(w)\} = \eta(S).$$

b. [2+] Given $S,T \subseteq [n-1]$, let $\overline{S} = [n-1] - S$, and define

$$X_{ST} = \#\{w \in \mathfrak{S}_n : C(w) = \overline{S},\ D(w) = T\},$$

$$Z_{ST} = \#\{w \in \mathfrak{S}_n : \overline{S} \subseteq C(w),\ T \subseteq D(w)\},$$
$$= \sum_{\substack{S'\supseteq S\\ T'\supseteq T}} X_{S'T'}.$$

For instance, for $n = 4$, we have the following table of X_{ST}.

$S\backslash T$	$\emptyset$	1	2	3	12	13	23	123
$\emptyset$	1							
1	0	1						
2	0	0	1					
3	0	0	0	1				
12	0	1	1	0	1			
13	0	0	0	0	0	1		
23	0	0	1	1	0	0	1	
123	0	1	2	1	2	4	2	1

Show that

$$Z_{ST} = \begin{cases} \eta(\overline{S})/\eta(\overline{T}), & \text{if } S \supseteq T, \\ 0, & \text{otherwise.} \end{cases} \tag{2.50}$$

c. [2–] Let $M = (M_{ST})$ be the matrix whose rows and columns are indexed by subsets $S,T \subseteq [n-1]$ (taken in some order), with

$$M_{ST} = \begin{cases} 1, & \text{if } S \supseteq T, \\ 0, & \text{otherwise.} \end{cases}$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

Let $D = (D_{ST})$ be the diagonal matrix with $D_{SS} = \eta(\overline{S})$. Let $Z = (Z_{ST})$, that is, the matrix whose (S,T)-entry is Z_{ST}. Show that equation (2.50) can be restated as follows:

$$Z = DMD^{-1}.$$

Similarly show that if $X = (X_{ST})$, then

$$MXM = Z.$$

d. [1+] For an invertible matrix $A = (A_{ST})$, write A^{-1}_{ST} for the (S,T)-entry of the inverse matrix A^{-1}. Show that the Principle of Inclusion-Exclusion (Theorem 2.1.1) is equivalent to

$$M^{-1}_{ST} = (-1)^{\#S+\#T} M_{ST}.$$

e. [2–] Define the matrix $Y = (Y_{ST})$ by

$$Y_{ST} = \#\{w \in \mathfrak{S}_n : \overline{S} \subseteq C(w),\ T = D(w)\}.$$

Show that $Y = MX = ZM^{-1}$.

f. [2+] Show that the matrices Z, Y, X have the following inverses:

$$\begin{aligned}
Z^{-1}_{ST} &= (-1)^{\#S+\#T} Z_{ST},\\
Y^{-1}_{ST} &= (-1)^{\#S+\#T} \#\{w \in \mathfrak{S}_n : \overline{S} = C(w),\ T \subseteq D(w)\},\\
X^{-1}_{ST} &= (-1)^{\#S+\#T} X_{ST}.
\end{aligned}$$

14. a. [2+]* Let $A_k(n)$ denote the number of k-element antichains in the boolean algebra B_n (i.e., the number of subsets S of $2^{[n]}$ such that no element of S is a subset of another). Show that

$$\begin{aligned}
A_1(n) &= 2^n,\\
A_2(n) &= \frac{1}{2}\left(4^n - 2\cdot 3^n + 2^n\right),\\
A_3(n) &= \frac{1}{6}\left(8^n - 6\cdot 6^n + 6\cdot 5^n + 3\cdot 4^n - 6\cdot 3^n + 2\cdot 2^n\right),\\
A_4(n) &= \frac{1}{24}\left(16^n - 12\cdot 12^n + 24\cdot 10^n + 4\cdot 9^n - 18\cdot 8^n\right.\\
&\qquad \left. +6\cdot 7^n - 36\cdot 6^n + 11\cdot 4^n - 22\cdot 3^n + 6\cdot 2^n\right).
\end{aligned}$$

b. [2+]* Show that for fixed $k \in \mathbb{P}$ there exist integers $a_{k,2}, a_{k,3}, \dots, a_{k,2^k}$ such that

$$A_k(n) = \frac{1}{k!}\sum_{i=2}^{2^k} a_{k,i}\, i^n.$$

Show in particular that $a_{k,2^k} = 1$, $a_{k,i} = 0$ if $3\cdot 2^{k-2} < i < 2^k$, and $a_{k,3\cdot 2^{k-2}} = k(k-1)$.

15. a. [2–] Given a permutation $w \in \mathfrak{S}_3$, let P_w denote the corresponding permutation matrix; that is, the (i,j)-entry of P_w is equal to $\delta_{w(i),j}$. Let α_w, where $w \in \mathfrak{S}_3$, be integers satisfying $\sum_w \alpha_w P_w = 0$. Show that

$$\alpha_{123} = \alpha_{231} = \alpha_{312} = -\alpha_{132} = -\alpha_{213} = -\alpha_{321}.$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

b. [2] Let $H_n(r)$ denote the number of $n \times n$ $\mathbb{N}$-matrices A for which every row and column sums to r. Assume the theorem that A is a sum of permutation matrices is known. Deduce from this result (for the case $n = 3$) and (a) that

$$H_3(r) = \binom{r+5}{5} - \binom{r+2}{5}. \tag{2.51}$$

c. [3–] Give a direct combinatorial proof that

$$H_3(r) = \binom{r+4}{4} + \binom{r+3}{4} + \binom{r+2}{4}.$$

16. [2] Fix $k \geq 1$. How many permutations of $[n]$ have no cycle of length k? If $f_k(n)$ denotes this number, then compute $\lim_{n\to\infty} f_k(n)/n!$.

17. a. [2] Let $f_2(n)$ be the number of permutations of the integers modulo n that consist of a single cycle $(a_1, a_2, \ldots, a_n)$ and for which $a_i + 1 \not\equiv a_{i+1} \pmod{n}$ for all i (with $a_{n+1} = a_1$). For example, for $n = 4$, there is one such permutation; namely, $(1,4,3,2)$. Set $f_2(0) = 1$ and $f_2(1) = 0$. Use the Principle of Inclusion-Exclusion to find a formula for $f_2(n)$.

b. [1+] Write the answer to (a) in the form $\Delta^n g(0)$ for some function g.

c. [2–] Find the generating function $\sum_{n\geq 0} f_2(n) x^n/n!$.

d. [2–] Express the derangment number $D(n)$ in terms of the numbers $f_2(k)$.

e. [2–] Show that

$$\lim_{n\to\infty} \frac{f_2(n)}{(n-1)!} = \frac{1}{e}.$$

f. [3–] Generalize (e) to show that $f_2(n)$ has the asymptotic expansion

$$\frac{f_2(n)}{(n-1)!} \sim \frac{1}{e}\left(1 - \frac{1}{n} + \frac{1}{n^3} + \frac{1}{n^4} - \frac{2}{n^5} - \frac{9}{n^6} + \cdots + \frac{a_i}{n^i} + \cdots\right), \tag{2.52}$$

where $\sum_{i\geq 0} a_i x^i/i! = \exp(1 - e^x)$. By definition, equation (2.52) means that for any $k \in \mathbb{N}$,

$$\lim_{n\to\infty} n^k \left[\frac{f_2(n)}{(n-1)!} - \frac{1}{e}\sum_{i=0}^{k} \frac{a_i}{n^i}\right] = 0.$$

18. [3] Let $k \geq 2$. Let $f_k(n)$ be the number of cycles as in Exercise 2.17 such that for no i do we have

$$w(i+j) \equiv w(i) + j \pmod{n}, \quad \text{for all } j = 1, 2, \ldots, k-1,$$

where the argument $i + j$ is taken modulo n. Use the Principle of Inclusion-Exclusion to show that

$$\frac{f_3(n)}{(n-1)!} = 1 - \frac{1}{n} - \frac{3}{2}\frac{1}{n^2} - \frac{14}{3}\frac{1}{n^3} + O(n^{-4}),$$

$$\frac{f_4(n)}{(n-1)!} = 1 - \frac{1}{n^2} - \frac{5}{n^3} - \frac{29}{2}\frac{1}{n^4} + O(n^{-5}),$$

$$\frac{f_k(n)}{(n-1)!} = 1 - \frac{1}{n^{k-2}} - \frac{(k-2)(k+1)}{2}\frac{1}{n^{k-1}}$$

$$-\frac{k(k+1)(3k^2 - 5k - 10}{24}\frac{1}{n^k} + O(n^{-k-1}),$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

for fixed $k \geq 5$.
In particular, for fixed $k \geq 3$ we have $\lim_{n\to\infty} f_k(n)/(n-1)! = 1$.

19. [2] Suppose that $2n$ persons are sitting in a circle. In how many ways can they form n pairs if no two adjacent persons can form a pair? Express your answer as a finite sum.

20. [2] Call two permutations of the $2n$-element set $S = \{a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n\}$ *equivalent* if one can be obtained from the other by interchanges of *consecutive* elements of the form $a_i b_i$ or $b_i a_i$. For example, $a_2b_3a_3b_2a_1b_1$ is equivalent to itself and to $a_2a_3b_3b_2a_1b_1$, $a_2b_3a_3b_2b_1a_1$, and $a_2a_3b_3b_2b_1a_1$. How many equivalence classes are there?

21. **a.** [2+]* Given numbers (or elements of a commutative ring with 1) a_i for $i \in \mathbb{Z}$, with $a_i = 0$ for $i < 0$ and $a_0 = 1$, let $f(k) = \det[a_{j-i+1}]_1^k$. In particular, $f(0) = 1$. Show that

$$\sum_{k\geq 0} f(k)x^k = \frac{1}{1 - a_1x + a_2x^2 - \cdots}.$$

b. [2] Suppose that in (a) we drop the condition $a_0 = 1$, say $a_0 = \alpha$. Deduce from (a) that

$$\sum_{k\geq 0} f(k)x^k = \frac{1}{1 + \sum_{i\geq 1}(-1)^i\alpha^{i-1}a_ix^i}.$$

c. [2+] Suppose that in (b) we let the first row of the matrix be arbitrary, that is, let $M_k = (m_{ij})_1^k$ be the $k \times k$ matrix defined by

$$\begin{aligned} m_{1j} &= b_j, \\ m_{ij} &= a_{j-i+1},\ i \geq 2, \end{aligned}$$

where $a_0 = \alpha$ and $a_i = 0$ for $i < 0$. Let $g(k) = \det M_k$. Show that

$$\sum_{k\geq 1} g(k)x^k = \frac{\sum_{j\geq 1}(-1)^{j-1}\alpha^{j-1}b_jx^j}{1 + \sum_{i\geq 1}(-1)^i\alpha^{i-1}a_ix^i}.$$

d. [2] Fix $0 < a \leq d$. Let $\beta(k) = \beta_{a+kd}(a, a+d, a+2d, ..., a+(k-1)d)$. Deduce from equation (2.16) that

$$\sum_{k\geq 0} \beta(k)\frac{x^k}{(a+kd)!} = \frac{\displaystyle\sum_{j\geq 0}(-1)^j\frac{x^j}{(a+jd)!}}{\displaystyle\sum_{i\geq 0}(-1)^i\frac{x^i}{(id)!}}.$$

Give a q-analogue based on Example 2.2.5.

e. [2]* Suppose that in Proposition 2.2.6 the function $e(i,j)$ has the form

$$e(i,j) = \alpha_{j-i}$$

for certain numbers α_k, with $\alpha_0 = 1$ and $\alpha_k = 0$ for $k < 0$. Show that $f_=(S)$ is equal to the coefficient of x^{n+1} in the power series

$$h(n)(1 - \alpha_1x + \alpha_2x^2 - \alpha_3x^3 + \cdots)^{-1}.$$

22. **a.** [2+] Let E_{2n} denote the number of alternating permutations $w \in \mathfrak{S}_{2n}$. Thus by Proposition 1.6.1, we have

$$\left(\sum_{n\geq 0} E_{2n}\frac{x^{2n}}{(2n)!}\right)\left(1-\frac{x^2}{2!}+\frac{x^4}{4!}-\cdots\right)=1.$$

Equating coefficients of $x^{2n}/(2n)!$ on both sides gives

$$E_{2n}=\binom{2n}{2}E_{2n-2}-\binom{2n}{4}E_{2n-4}+\binom{2n}{6}E_{2n-6}-\cdots. \tag{2.53}$$

Give a sieve-theoretic proof of equation (2.53).

b. [2+] State and prove a similar result for E_{2n+1}.

23. **a.** [2+] Give a sieve-theoretic proof of Exercise 1.61(c), that is, if $f(n)$ is the number of permutations $w \in \mathfrak{S}_n$ with no proper double descents, then

$$\sum_{n\geq 0} f(n)\frac{x^n}{n!}=\frac{1}{\sum_{j\geq 0}\left(\frac{x^{3j}}{(3j)!}-\frac{x^{3j+1}}{(3j+1)!}\right)}.$$

b. [2+]* Generalize (a) as follows. Let $f_r(n)$ be the number of permutations $w \in S_n$ with no r consecutive descents (where n is not considered a descent). Give a sieve-theoretic proof that

$$\sum_{n\geq 0} f_r(n)\frac{x^n}{n!}=\frac{1}{\sum_{j\geq 0}\left(\frac{x^{(r+1)j}}{((r+1)j)!}-\frac{x^{(r+1)j+1}}{((r+1)j+1)!}\right)}.$$

24. [2+]* Fix $j,k\geq 1$. For $n\geq 0$ let $f(n)$ be the number of integer sequences $a_1,a_2,\dots,a_n$ such that $1\leq a_i\leq k$ for $1\leq i\leq n$, and $a_i\geq a_{i-1}-j$ for $2\leq i\leq n$. Give a sieve-theoretic proof that

$$F(x):=\sum_{n\geq 0} f(n)x^n=\frac{1}{\sum_{i\geq 0}(-1)^i\binom{k-j(i-1)}{i}x^i}.$$

(Note that the denominator is actually a finite sum.)

25. **a.** [2]* Let $f_i(m,n)$ be the number of $m\times n$ matrices of 0's and 1's with at least one 1 in every row and column, and with a total of i 1's. Use the Principle of Inclusion-Exclusion to show that

$$\sum_i f_i(m,n)t^i=\sum_{k=0}^{n}(-1)^k\binom{n}{k}((1+t)^{n-k}-1)^m. \tag{2.54}$$

b. [2]* Show that

$$\sum_{m,n\geq 0}\sum_{i\geq 0}\sum_{i\geq 0} f_i(m,n)y^i\frac{x^m y^n}{m!n!}=e^{-x-y}\sum_{i\geq 0}\sum_{j\geq 0}(1+t)^{ij}\frac{x^i y^j}{i!j!}.$$

Note that this formula, unlike equation (2.54), exhibits the symmetry between m and n.

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

26. [2+]* Let $\pi \in \Pi_n$, the set of partitions of $[n]$. Let $S(\pi,r)$ denote the number of $\sigma \in \Pi_n$ such that $|\sigma| = r$ and $\#(A \cap B) \le 1$ for all $A \in \pi$ and $B \in \sigma$. (This last condition is equivalent to $\pi \wedge \sigma = \hat{0}$ in the lattice structure on Π_n defined in Example 3.10.4.) Show that

$$S(\pi,r) = \frac{1}{r!}\sum_{i=0}^{r}\binom{r}{i}(-1)^{r-i}\prod_{A\in\pi}(i)_{\#A}$$

$$= \frac{1}{r!}\Delta^r \prod_{A\in\pi}(n)_{\#A}|_{n=0}.$$

27. a. [3–] Let F be a forest, with $\ell = \ell(F)$ components, on the vertex set $[n]$. We say that F is *rooted* if we specify a root vertex for each connected component of F. Thus, if $c_1,\dots,c_\ell$ are the number of vertices of the components of F (so $\sum c_i = n$), then the number $p(F)$ of ways to root F is $c_1c_2\cdots c_\ell$. Show that the number of k-component rooted forests on $[n]$ that contain F is equal to

$$p(F)\binom{\ell-1}{\ell-k}n^{\ell-k}.$$

b. [2+] Given any graph G on $[n]$ with no multiple edges, define the polynomial

$$P(G,x) = \sum_F x^{\ell(F)-1}, \tag{2.55}$$

summed over all rooted forests F on $[n]$ contained in G. Let $\overline{G}$ denote the complement of G; that is, $\{i,j\} \in \binom{[n]}{2}$ is an edge of G if and only if $\{i,j\}$ is not an edge of $\overline{G}$. Use (a) and the Principle of Inclusion-Exclusion to show that

$$P(\overline{G},x) = (-1)^{n-1}P(G,-x-n). \tag{2.56}$$

In particular, the number $c(\overline{G})$ of spanning trees of $\overline{G}$ (i.e., subgraphs of $\overline{G}$ that are trees and that use all the vertices of $\overline{G}$) is given by

$$c(\overline{G}) = (-1)^{n-1}P(G,-n)/n. \tag{2.57}$$

c. [2] The *complete graph* K_n has vertex set $[n]$ and an edge between any two distinct vertices (so $\binom{n}{2}$ edges in all). The *complete bipartite graph* $K_{r,s}$ has vertex set $A \cup B$, where A and B are disjoint with $\#A = r$ and $\#B = s$, and with one edge between each vertex of A and each vertex of B (so rs edges in all). Use (b) to find the number of spanning trees of K_n and $K_{r,s}$.

28. [3] Let $r \ge 1$. An *r-stemmed V-partition* of n is an array

$$\begin{bmatrix} & & & & b_1 & b_2 & b_3 & \cdots \\ a_1 & a_2 & \cdots & a_r & & & & \\ & & & & c_1 & c_2 & c_3 & \cdots \end{bmatrix}$$

of nonnegative integers satisfying $a_1 \ge a_2 \ge \cdots \ge a_r \ge b_1 \ge b_2 \ge b_3 \ge \cdots$, $a_r \ge c_1 \ge c_2 \ge c_3 \ge \cdots$, and $\sum a_i + \sum b_i + \sum c_i = n$. Hence, a 1-stemmed V-partition is just a V-partition. Let $v_r(n)$ denote the number of r-stemmed V-partitions of n. Show that

$$\sum_{n\ge 0} v_r(n)x^n = \frac{p_r(x)T(x) - q_r(x)}{(1-x)(1-x^2)\cdots(1-x^{r-1})\prod_{i\ge 1}(1-x^i)^2},$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

where

$$\begin{aligned}
&p_1(x)=1,\ p_2(x)=2,\ q_1(x)=0,\ q_2(x)=1,\\
&p_r(x)=2p_{r-1}(x)+(x^{r-2}-1)p_{r-2}(x),\ r>2,\\
&q_r(x)=2q_{r-1}(x)+(x^{r-2}-1)q_{r-2}(x),\ r>2,\\
&T(x)=\sum_{i\geq 0}(-1)^i x^{\binom{i+1}{2}}.
\end{aligned}$$

29. **a.** [2]* A *concave composition* of n is a nonnegative integer sequence $a_1>a_2>\cdots>a_r=b_r<b_{r-1}<\cdots<b_1$ such that $\sum(a_i+b_i)=n$. For instance, the eight concave compositions of 6 are 33, 5001, 4002, 3003, 2112, 2004, 1005, and 210012. Let $f(n)$ denote the number of concave partitions of n. Give a combinatorial proof that $f(n)$ is even for $n\geq 1$.

b. [5–] Set

$$F(q)=\sum_{n\geq 0}f(n)q^n=1+2q^2+2q^3+4q^4+4q^5+8q^6+\cdots.$$

Give an Inclusion-Exclusion proof, analogous to the proof of Proposition 2.5.1, that

$$F(q)=\frac{1-\sum_{n\geq 1}q^{n(3n-1)/2}(1-q^n)}{(1-q)(1-q^2)(1-q^3)\cdots}.$$

30. [3] Give a sieve-theoretic proof of the Pentagonal Number Formula (Proposition 1.8.7), namely,

$$\frac{1+\sum_{n\geq 1}(-1)^n[x^{n(3n-1)/2}+x^{n(3n+1)/2}]}{\prod_{i\geq 1}(1-x^i)}=1.$$

Your sieve should start with all partitions of $n\geq 0$ and sieve out all but the empty partition of 0.

31. [3–] Give cancellation proofs, similar to our proof of the Pentagonal Number Formula (Proposition 1.8.7), of the two identities of Exercise 1.91(c), namely,

$$\prod_{k\geq 1}\frac{1-q^k}{1+q^k}=\sum_{n\in\mathbb{Z}}(-1)^n q^{n^2},$$

$$\prod_{k\geq 1}\frac{1-q^{2k}}{1-q^{2k-1}}=\sum_{n\geq 0}q^{\binom{n+1}{2}}.$$

32. [3–] Give a cancellation proof of the identity

$$\sum_{k=0}^{n}(-1)^k\binom{\boldsymbol{n}}{\boldsymbol{k}}=\begin{cases}(1-q)(1-q^3)\cdots(1-q^{n-1}), & n \text{ even},\\ 0, & n \text{ odd}.\end{cases}$$

33. [2–] Deduce from equation (2.21) that

$$\det\left[\binom{\boldsymbol{n-i}}{\boldsymbol{j-i+1}}\right]_0^{n-1}=q^{\binom{n}{2}}. \qquad (2.58)$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

34. A *tournament* T on the vertex set $[n]$ is a directed graph on $[n]$ with no loops such that each pair of distinct vertices is joined by exactly one directed edge. The *weight* $w(e)$ of a directed edge e from i to j (denoted $i \to j$) is defined to be x_j if $i < j$ and $-x_j$ if $i > j$. The weight of T is defined to be $w(T) = \prod_e w(e)$, where e ranges over all edges of T.

a. [2–] Show that

$$\sum_T w(T) = \prod_{1 \le i < j \le n} (x_j - x_i), \tag{2.59}$$

where the sum is over all $2^{\binom{n}{2}}$ tournaments on $[n]$.

b. [2–] The tournament T is *transitive* if there is a permutation $z \in \mathfrak{S}_n$ for which $z(i) < z(j)$ if and only if $i \to j$. Show that a nontransitive tournament contains a 3-cycle (i.e., a triple (t, u, v) of vertices for which $t \to u \to v \to t$).

c. [1+] If T and T' are tournaments on $[n]$ then write $T \leftrightarrow T'$ if T' can be obtained from T by reversing a 3-cycle; that is, replacing the edges $t \to u, u \to v, v \to t$ with $u \to t$, $v \to u$, $t \to v$, and leaving all other edges unchanged. Show that $w(T') = -w(T)$.

d. [2] Show that if $T \leftrightarrow T'$ then T and T' have the same number of 3-cycles.

e. [2+] Deduce from (a)–(d) that

$$\det \left[x_i^{j-1} \right]_1^n = \prod_{1 \le i < j \le n} (x_j - x_i),$$

by canceling out all terms in the left-hand side of (2.59) except those corresponding to transitive T.

35. a. [2] Let $f(x_1, \ldots, x_n)$ be a homogeneous polynomial of degree n over a field K. Show that

$$[x_1 x_2 \cdots x_n] f(x_1, \ldots, x_n) = \sum_{(\epsilon_1, \ldots, \epsilon_n) \in \{0,1\}^n} (-1)^{n - \sum \epsilon_i} f(\epsilon_1, \ldots, \epsilon_n). \tag{2.60}$$

(Regard each ϵ_i in the exponent of -1 as an integer and in the argument of f as an element of K.)

b. [2] Let $A = (a_{ij})$ be an $n \times n$ matrix. The *permanent* of A is defined by

$$\operatorname{per}(A) = \sum_{w \in \mathfrak{S}_n} a_{1,w(1)} a_{2,w(2)} \cdots a_{n,w(n)}.$$

In other words, the formula for $\operatorname{per}(A)$ is the same as the expansion of $\det(A)$ but with all signs positive. Show that

$$\operatorname{per}(A) = \sum_{S \subseteq [n]} (-1)^{n - \#S} \prod_{i=1}^{n} \sum_{j \in S} a_{ij}. \tag{2.61}$$

36. [3–] Let $A_1, \ldots, A_n$ be subsets of a finite set A, and $B_1, \ldots, B_n$ subsets of a finite set B. For each subset S of $[n]$, let $A_S = \bigcap_{i \in S} A_i$ and $B_S = \bigcap_{i \in S} B_i$. Given bijections $f_S \colon A_S \to B_S$ for each $S \subseteq [n]$, construct an explicit bijection $h \colon A - \bigcup_{i=1}^n A_i \to B - \bigcup_{i=1}^n B_i$. Your definition of h should depend only on the f_S's, and not on some ordering of the elements of A or on the labeling of the subsets $A_1, \ldots, A_n$ and $B_1, \ldots, B_n$.

37. [3–]* Given $a, b \in \mathbb{P}$ with $a < b$, let $C(b-a)$ denote the number of lattice paths in $\mathbb{Z}^2$ from $(2a, 0)$ to $(2b, 0)$ with steps $(1,1)$ or $(1,-1)$ that never pass below the x-axis. (It follows from Corollary 6.2.3(iv) that $C(b-a)$ is the Catalan number $\frac{1}{b-a+1}\binom{2(b-a)}{b-a}$, but this fact is irrelevant here.) Now given $\{a_1, a_2, \dots, a_{2n}\}_< \subset \mathbb{Z}$, let $C(a_1, a_2, \dots, a_{2n})$ denote the number of ways to connect the points $(2a_1, 0)$, $(2a_2, 0)$,…, $(2a_{2n}, 0)$ with n pairwise disjoint lattice paths $L_1, \dots, L_n$ of the type just described. (Thus, each L_i connects some $(2a_j, 0)$ to some $(2a_k, 0)$, $j \neq k$. If $i \neq j$ then L_i and L_j do not intersect, including endpoints, so each $(2a_i, 0)$ is an endpoint of exactly one L_i.)
Now given a triangular array $A = (a_{ij})$ with $1 \le i < j \le 2n$, define the *pfaffian* of A by

$$\mathrm{Pf}(A) = \sum \varepsilon(i_1, j_1, \dots, i_n, j_n) a_{i_1 j_1} \cdots a_{i_n j_n},$$

where the summation is over all partitions $\{\{i_1, j_1\}_<, \dots, \{i_n, j_n\}_<\}$ of $[2n]$ into 2-element blocks, and where $\varepsilon(i_1, j_1, \dots, i_n, j_n)$ denotes the sign of the permutation (written in two-line form)

$$\begin{pmatrix} 1 & 2 & \cdots & 2n-1 & 2n \\ i_1 & j_1 & \cdots & i_n & j_n \end{pmatrix}.$$

(It is easy to see that $\varepsilon(i_1, j_1, \dots, i_n, j_n)$ does not depend on the order of the n blocks.) Give a proof analogous to that of Theorem 2.7.1 of the formula

$$C(a_1, a_2, \dots, a_{2n}) = \mathrm{Pf}(C(a_j - a_i)).$$

For instance,

$$\begin{aligned} C(0,3,5,6) &= \mathrm{Pf}\begin{vmatrix} C(3) & C(5) & C(6) \\ & C(2) & C(3) \\ & & C(1) \end{vmatrix} \\ &= \mathrm{Pf}\begin{vmatrix} 5 & 42 & 132 \\ & 2 & 5 \\ & & 1 \end{vmatrix} \\ &= 5 \cdot 1 + 132 \cdot 2 - 42 \cdot 5 \\ &= 59. \end{aligned}$$

Solutions to Exercises

1. We have

$$\begin{aligned} \textbf{P}\text{rinciple of } \textbf{I}\text{nclusion-}\textbf{E}\text{xclusion} &= \text{PIE} \\ &= \pi e \\ &= (3.141592653\cdots)(2.718281828\cdots) \\ &= 8.53973422267356706546\cdots. \end{aligned}$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

3. We have

$$\begin{aligned}\sum_{i=k}^{n}(-1)^{i-k}\binom{i}{k}g_i &= \sum_{i=k}^{n}(-1)^{i-k}\binom{i}{k}\sum_{T\subseteq S\#T=i} f_{\geq}(T)\\ &= \sum_{i=k}^{n}(-1)^{i-k}\binom{i}{k}\sum_{\substack{T\subseteq R\subseteq S\\ \#T=i}} f_{=}(R)\\ &= \sum_{R\subseteq S} f_{=}(R)\sum_{T\subseteq R}(-1)^{\#T-k}\binom{\#T}{k}.\end{aligned}$$

If $\#R=r$ then the inner sum is equal to

$$\sum_{j=0}^{r}(-1)^{j-k}\binom{r}{j}\binom{j}{k}=\binom{r}{k}\sum_{j=0}^{r}(-1)^{j-k}\binom{r-k}{r-j}=\delta_{kr},$$

and the proof of equation (2.39) follows. The sum (2.40) is evaluated similarly. An extensive bibliography appears in Takács [2.22].

4. **a.** If we regard A_i as the set of elements having property P_i, then

$$A_T=f_{\geq}(T)=\sum_{Y\supseteq T} f_{=}(Y).$$

Hence,

$$\begin{aligned}S_k-S_{k-1}+\cdots+(-1)^{n-k}S_n &= \sum_{\#T\geq k}(-1)^{\#T-k}f_{\geq}(T)\\ &= \sum_{\#T\geq k}\sum_{Y\supseteq T}(-1)^{\#T-k}f_{=}(Y)\\ &= \sum_{\#Y\geq k} f_{=}(Y)\sum_{\substack{T\subseteq Y\\ \#T\geq k}}(-1)^{\#T-k}\\ &= \sum_{\#Y\geq k} f_{=}(Y)\sum_{i=k}^{\#Y}(-1)^{i-k}\binom{\#Y}{i}.\end{aligned}$$

It is easy to see that $\sum_{i=k}^{m}(-1)^{i-k}\binom{m}{i}=\binom{m-1}{k-1}\geq 0$. Since $f_{=}(Y)\geq 0$, equation (2.41) follows.

Setting

$$S=f_{=}(\emptyset)=\#(\bar{A}_1\cap\cdots\cap\bar{A}_n)=S_0-S_1+\cdots+(-1)^n S_n,$$

the inequality (2.41) can be rewritten

$$\begin{aligned}&S\geq 0\\ &S\leq S_0\\ &S\geq S_0-S_1\\ &S\leq S_0-S_1+S_2\\ &\quad\vdots\end{aligned}$$

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

In other words, the partial sums $S_0 - S_1 + \cdots + (-1)^k S_k$ successively overcount and undercount the value of S. In this form, equation (2.41) is due to Carlo Bonferroni (1892–1960), *Pubblic. Ist. Sup. Sc. Ec. Comm. Firenze* **8** (1936), 1–62. These inequalities sometimes make it possible to estimate S accurately when not all the S_i's can be computed explicitly.

b. *Answer.* $\sum_{i=1}^{k}(-1)^{i-k}\binom{i}{k}S_i \geq 0,\quad 0 \leq k \leq n.$

5. **a.** The most straightforward proof is by induction on n, the case $n = 0$ being trivial (since when $n = 0$ exactness implies that $W \cong V_0$). The details are omitted.

b. The sequence

$$0 \to V_n \xrightarrow{\partial_n} V_{n-1} \xrightarrow{\partial_{n-1}} \cdots \xrightarrow{\partial_{j+1}} V_j \xrightarrow{\partial_j} \operatorname{im}\partial_j \to 0$$

is exact. But $\dim(\operatorname{im}\partial_j) = \operatorname{rank}\partial_j$, so the proof follows from (a).

c. By equation (2.44), we have $\dim V_j = \operatorname{rank}\partial_j + \operatorname{rank}\partial_{j+1}$. On the other hand, $\operatorname{rank}\partial_{j+1} = \dim(\operatorname{im}\partial_{j+1})$ and $\operatorname{rank}\partial_j = \dim V_j - \dim(\ker\partial_j)$, so $\dim(\operatorname{im}\partial_{j+1}) = \dim(\ker\partial_j)$. Since $\operatorname{im}\partial_{j+1} \subseteq \ker\partial_j$, the proof follows.

d. For fixed $a \in A$, let V_T^a be the span of the symbols $[a,T]$ if $a \in A_T$; otherwise, $V_T^a = 0$. Let $V_j^a = \bigoplus_{\#T=j} V_T^a$, and let W^a be the span of the single element $[a]$ if $a \in \bar{A}_1 \cap \cdots \cap \bar{A}_n$; otherwise, $W^a = 0$. Then $\partial_j\colon V_j^a \to V_{j-1}^a$, $j \geq 1$, and $\partial_0\colon V_0^a \to W^a$. (Thus, the sequence (2.42) is the *direct sum* of such sequences for fixed a.) From this discussion, it follows that we may assume $A = \{a\}$.

Clearly, ∂_0 is surjective, so exactness holds at W. It is straightforward to check that $\partial_j\partial_{j+1} = 0$, so (2.42) is a complex. Since $A = \{a\}$, we have $\dim V_j = \binom{n}{j}$ and $\sum_{i=j}^{n}(-1)^{i-j}\dim V_i = \binom{n-1}{j-1}$. There are several ways to show that $\operatorname{rank}\partial_j = \binom{n-1}{j-1}$, so the proof follows from (c).

There are many other proofs, whose accessibility depends on background. For instance, the complex (2.42) in the case at hand (with $A = \{a\}$) is the tensor product of the complexes $\mathcal{C}_i\colon 0 \to U_i \xrightarrow{\partial_0} W \to 0$, where U_i is spanned by $[a,\{t_i\}]$. Clearly, each $\mathcal{C}_i$ is exact; hence, so is (2.42). (The definition (2.45) was not plucked out of the air; it is a *Koszul relation*, and (2.42) (with $A = \{a\}$) is a *Koszul complex*. See almost any textbook on homological algebra for further information.).

e,f. Follows from $\dim V_T = \#A_T$, whence $\dim V_j = S_j$.

6. **a.** Straightforward generalization of Theorem 2.1.1.

b. We obtain the classical Möbius inversion formula (see Example 3.8.4). More specifically, let D_n denote the set of all divisors of n, and let $f,g\colon D_n \to K$. Equations (2.48) and (2.49) then assert that the following two formulas are equivalent:

$$g(m) = \sum_{d|m} f(d), \text{ for all } m|n,$$

$$f(m) = \sum_{d|m} \mu(m/d)g(d), \text{ for all } m|n. \tag{2.62}$$

7. Each element $\alpha \in \mathbb{F}_{q^n}$ generates a subfield $\mathbb{F}_q(\alpha)$ of order q^d for some $d|n$. Thus, α is a zero of a unique monic irreducible polynomial $f_\alpha(x)$ of degree d over $\mathbb{F}_q$. Every

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

such polynomial has d distinct zeros, all belonging to $\mathbb{F}_{q^n}$. Hence,

$$q^n = \sum_{d|n} d\beta(d).$$

Möbius inversion (see equation (2.62)) gives

$$n\beta(n) = \sum_{d|n} \mu(n/d)q^d = \sum_{d|n} \mu(d)q^{n/d}.$$

8. **a.** See J. B. Remmel, *Europ. J. Combinatorics* **4** (1983), 371–374, and H. S. Wilf, *Math. Mag.* **57** (1984), 37–40.

b. Note that the last entry a_n of a permutation $w = a_1 \cdots a_n \in \mathfrak{S}_n$ has no effect on the location of the first ascent unless $w = n, n-1, \ldots, 1$, in which case the contribution to $E(n)$ is $(-1)^n$. See J. Désarménien, *Sem. Lotharingien de Combinatoire* (electronic) **8** (1983), B08b; formerly *Publ. I.R.M.A. Strasbourg*, 229/S-08, 1984, pp. 11–16. For a generalization, see J. Désarménien and M. L. Wachs, *Sem. Lotharingien de Combinatoire* (electronic) **19** (1988), B19a; formerly *Publ. I.R.M.A. Strasbourg*, 361/S-19, 1988, pp. 13–21. See also Vol. II, Exercise 7.65 for a related result dealing with symmetric functions.

NOTE. We can also see that $E(n) = D(n)$ by noting that the number of permutations $w \in \mathfrak{S}_n$, whose first ascent is in position k, is $n!\left(\frac{1}{k!} - \frac{1}{(k+1)!}\right)$ for $0 \le k < n$ and is $1/n!$ for $k = n$, and then comparing with equation (2.12). Moreover, in an unpublished paper at

⟨`http://people.brandeis.edu/~gessel/homepage/papers/color.pdf`⟩,

Gessel gives an elegant bijective proof in terms of "hook factorizations" of permutations that

$$\sum_{n \ge 0} E(n)\frac{x^n}{n!} = \frac{1}{1 - \sum_{k \ge 2}(k-1)\frac{x^k}{k!}} = \frac{e^{-x}}{1-x}.$$

c. Write a derangement w as a product of cycles. Arrange these cycles in decreasing order of their smallest element. Within each cycle, put the smallest element in the *second* position. Then erase the parentheses, obtaining another permutation w'. For instance, let $w = 974382651 = (85)(43)(627)(91)$; then $w' = 854362791$. It is not hard to check that the map $w \mapsto w'$ is a bijection from derangements in $\mathfrak{S}_n$ to $\mathcal{E}(n)$. This bijection is due to J. Désarménien, ibid.

I am grateful to Ira Gessel for providing most of the information for this exercise.

9. We interpret $k!S(d,k)$ as the number of surjective functions $f : [d] \to [k]$. Let A be the set of all functions $f : [d] \to [k]$, and for $i \in [k]$ let P_i be the property that $i \notin \operatorname{im} f$. A function $f \in A$ lacks at most the properties $T \subseteq S \subseteq \{P_1, \ldots, P_k\}$ if and only if $\operatorname{im} f \subseteq \{i : P_i \in T\}$; hence, the number of such f is i^d, where $\#T = i$. The proof follows from Proposition 2.2.2.

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

10. b. We have

$$\frac{D(n)}{n!}-\frac{1}{e}=\left(1-\frac{1}{1!}+\frac{1}{2!}-\cdots+(-1)^n\frac{1}{n!}\right)$$
$$-\left(1-\frac{1}{1!}+\frac{1}{2!}-\cdots+(-1)^n\frac{1}{n!}+(-1)^{n+1}\frac{1}{(n+1)!}+\cdots\right)$$
$$=\frac{(-1)^n}{(n+1)!}+\cdots,$$

while

$$\left(1-\frac{1}{n}\right)^n-\frac{1}{e}=e^{n\log(1-\frac{1}{n})}-\frac{1}{e}$$
$$=e^{n(-\frac{1}{n}-\frac{1}{2n^2}+\cdots)}-\frac{1}{e}$$
$$=-\frac{1}{2ne}+\cdots.$$

Hence, $D(n)/n!$ is a *much* better approximation to $1/e$ than $E(n)/n!$.

11. This result was proved by G.-N. Han and G. Xin, *J. Combinatorial Theory Ser. A* **116** (2009), 449–459 (Theorems 1 and 9), using the theory of symmetric functions. A bijective proof was given by N. Eriksen, R. Freij, and J. Wästlund, *Electronic J. Combinatorics* **16**(1) (2009), #R32 (Theorem 2.1).

12. The Inclusion-Exclusion formula (2.11) for $D(n)$ generalizes straightforwardly to

$$D(\alpha)=\sum_{\beta_1=0}^{\alpha_1}\cdots\sum_{\beta_k=0}^{\alpha_k}\binom{\alpha_1}{\beta_1}\cdots\binom{\alpha_k}{\beta_k}(-1)^{\beta_1+\cdots+\beta_k}\binom{\sum(\alpha_i-\beta_i)}{\alpha_1-\beta_1,\ldots,\alpha_k-\beta_k}.$$

Let $\gamma_i=\alpha_i-\beta_i$. We get

$$\sum_\alpha D(\alpha)x^\alpha=\sum_{\beta,\gamma}\binom{\beta_1+\gamma_1}{\beta_1}\cdots\binom{\beta_k+\gamma_k}{\beta_k}(-1)^{\sum\beta_i}\binom{\gamma_1+\cdots+\gamma_k}{\gamma_1,\ldots,\gamma_k}x^{\sum(\beta_i+\gamma_i)}$$
$$=\sum_\gamma\binom{\gamma_1+\cdots+\gamma_k}{\gamma_1,\ldots,\gamma_k}x^{\sum\gamma_i}\sum_\beta\binom{\beta_1+\gamma_1}{\beta_1}\cdots\binom{\beta_k+\gamma_k}{\beta_k}(-1)^{\sum\beta_i}x^{\sum\beta_i}$$
$$=\sum_\gamma\binom{\gamma_1+\cdots+\gamma_k}{\gamma_1,\ldots,\gamma_k}x^{\sum\gamma_i}(1+x_1)^{-\gamma_1-1}\cdots(1+x_k)^{-\gamma_k-1}$$
$$=\frac{1}{(1+x_1)\cdots(1+x_k)}\sum_{n\geq 0}\left(\frac{x_1}{1+x_1}+\cdots+\frac{x_k}{1+x_k}\right)^n$$
$$=\frac{1}{(1+x_1)\cdots(1+x_k)\left(1-\frac{x_1}{1+x_1}-\cdots-\frac{x_k}{1+x_k}\right)}.$$

This result appears as Exercise 4.5.5 in Goulden and Jackson [3.32], as a special case of the more general Exercise 4.5.4.

13. These results appear in R. Stanley, *J. Integer Sequences* **8** (2005), article 05.3.8. This paper also gives an extension to multisets and a q-analogue. For a generalization to arbitrary Coxeter groups, see N. Bergeron, C. Hohlweg, and M. Zabrocki, *J. Algebra* **303** (2006), 831–846, and M. Marietti, *Europ. J. Combinatorics* **29** (2008), 1555–1562.

15. a. The result follows easily after checking that any five of the matrices P_w are linearly independent.

b. Let A be a 3×3 $\mathbb{N}$-matrix for which every row and column sums to r. It is given that we can write

$$A = \sum_{w \in \mathfrak{S}_3} \alpha_w P_w, \tag{2.63}$$

where $\alpha_w \in \mathbb{N}$ and $\sum \alpha_w = r$. By Section 1.2, the number of ways to choose $\alpha_w \in \mathbb{N}$ such that $\sum \alpha_w = r$ is $\binom{r+5}{5}$. By (a), the representation (2.63) is unique provided at least one of $\alpha_{213}, \alpha_{132}, \alpha_{321}$ is 0. The number of ways to choose $\alpha_{123}, \alpha_{231}, \alpha_{312} \in \mathbb{N}$ and $\alpha_{213}, \alpha_{132}, \alpha_{321} \in \mathbb{P}$ such that $\sum \alpha_w = r$ is equal to the number of weak compositions of $r-3$ into six parts; that is, $\binom{r+2}{5}$. Hence, $H_3(r) = \binom{r+5}{5} - \binom{r+2}{5}$. Equation (2.51) appears in §407 of MacMahon [1.55], essentially with the previous proof. To evaluate $H_4(r)$ by a similar technique would be completely impractical, although it can be shown using the Hilbert syzygy theorem that such a computation could be done in principle. See R. Stanley, *Duke Math. J.* **40** (1973), 607–632. For a different approach toward evaluating $H_n(r)$ for any n, see Proposition 4.6.2. The theorem mentioned in the statement of (b) is called the Birkhoff–von Neumann theorem and is proved for general n in Lemma 4.6.1.

c. One can check that every matrix being counted can be represented in exactly one way in one of the forms

$$\begin{bmatrix} a+e & b+d & c \\ c+d & a & b+e \\ b & c+e & a+d \end{bmatrix}, \begin{bmatrix} a & b+d & c+e+1 \\ c+d & a+e+1 & b \\ b+e+1 & c & a+d \end{bmatrix},$$

$$\begin{bmatrix} a+d+1 & b & c+e+1 \\ c & a+e+1 & b+d+1 \\ b+e+1 & c+d+1 & a \end{bmatrix},$$

where $a,b,c,d,e \in \mathbb{N}$, from which the proof is immediate. The idea behind this proof is to associate an indeterminate x_w to each $w \in \mathfrak{S}_3$, and then to use the identity

$$1 - x_{132}x_{213}x_{321} = (1 - x_{321}) + x_{321}(1 - x_{132}) + x_{321}x_{132}(1 - x_{213}).$$

Details are left to the reader. For yet another way to obtain $H_3(r)$, see M. Bóna, *Math. Mag.* **70** (1997), 201–203.

16. *Answer.*

$$f_k(n) = \sum_{i=0}^{\lfloor n/k \rfloor} (-1)^i \frac{n!}{i!k^i},$$

$$\lim_{n\to\infty} \frac{f_k(n)}{n!} = \sum_{i \geq 0} \frac{(-1)^i}{i!k^i} = e^{-1/k}.$$

17. a. $\sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i-1)!$, provided we define $(-1)! = 1$.

b. $g(n) = (n-1)!$, with $g(0) = 1$.

c. $e^{-x}(1 - \log(1-x))$.

d. $D(n) = f_2(n) + f_2(n+1)$.

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

This problem goes back to W. A. Whitworth, *Choice and Chance*, 5th ed. (and presumably earlier editions), Stechert, New York, 1934 (Prop. 34 and Ex. 217). For further information and references, including solutions to (e) and (f), see R. Stanley, *JPL Space Programs Summary 37–40*, vol. 4 (1966), 208–214, and S. M. Tanny, *J. Combinatorial Theory* **21** (1976), 196–202.

18. See R. Stanley, ibid.

19. *Answer.* $\sum_{k=0}^{n}(-1)^k\binom{2n-k}{k}(2n-2k-1)!!$, where $(2m-1)!! = 1\cdot 3\cdot 5\cdots(2m-1)$.

20. Call a permutation *standard* if b_i is not immediately followed by a_i for $1\le i\le n$. Clearly, each equivalence class contains exactly one standard permutation. A straightforward use of Inclusion-Exclusion shows that the number of standard permutations is equal to

$$\sum_{i=0}^{n}\binom{n}{i}(-1)^i(2n-i)! = \Delta^n(n+i)!\big|_{i=0}.$$

21. **b.** Let $A_k = [a_{j-i+1}]_1^k$. Let D be the diagonal matrix $\mathrm{diag}(\alpha,\alpha^2,\dots,\alpha^k)$. Then $D^{-1}AD = [\alpha^{j-i}a_{j-i+1}]$. Since $\det A = \det D^{-1}AD$, the proof follows from (a).

c. If we remove the first row and ith column from M, then we obtain a matrix $M_i = \begin{bmatrix} B & C \\ 0 & A_{n-k}\end{bmatrix}$, where B is an upper triangular $(i-1)\times(i-1)$ matrix with α's on the diagonal. Hence, when we expand $\det M$ along the first row, we get

$$\det M = b_1(\det A_{k-1}) - \alpha b_2(\det A_{k-2}) + \cdots + (-1)^{k-1}\alpha^{k-1}b_k(\det A_0).$$

The proof follows.

d. For some alternative approaches and results related to this item, see Proposition 1.6.1, equation (1.59), Exercise 2.22, and equation (3.98).

22. **a.** Let S_k be the set of permutations $w = a_1a_2\cdots a_{2n}\in\mathfrak{S}_{2n}$ satisfying

$$a_1 > a_2 < a_3 > a_4 < \cdots > a_{2n-2k},\;\; a_{2n-2k+1} > a_{2n-2k+2} > \cdots > a_{2n},$$

and let T_k be those permutations in S_k that also satisfy $a_{2n-2k} > a_{2n-2k+1}$. Hence, $S_1 - T_1$ consists of all alternating permutations in $\mathfrak{S}_n$. Moreover, $T_i = S_{i+1} - T_{i+1}$. Hence,

$$E_n = \#(S_1-T_1) = \#S_1 - \#(S_2-T_2) = \cdots = \#S_1 - \#S_2 + \#S_3 - \cdots.$$

A permutation in S_k is obtained by choosing $a_{2n-2k+1}, a_{2n-2k+2},\dots,a_{2n}$ in $\binom{2n}{2k}$ ways and then $a_1,a_2,\dots,a_{2n-2k}$ in $E_{2(n-k)}$ ways. Hence, $\#S_k = \binom{2n}{2k}E_{2(n-k)}$, and the proof follows.

b. The recurrence is

$$E_{2n+1} = \binom{2n+1}{2}E_{2n-1} - \binom{2n+1}{4}E_{2n-3} + \binom{2n+1}{6}E_{2n-5} - \cdots + (-1)^n,$$

proved similarly to (a) but with the additional complication of accounting for the term $(-1)^n$.

23. **a.** The argument is analogous to that of the previous exercise. Let S_k be the set of those permutations $a_1a_2\cdots a_n\in\mathfrak{S}_n$ such that $a_1a_2\cdots a_{n-k}$ has no proper double descents and $a_{n-k+1} > a_{n-k+1} > \cdots > a_n$. Let T_k consist of those permutations in S_k that also satisfy $a_{n-k-1} > a_{n-k} > a_{n-k+1}$. Let U_k consist of those permutations in S_k that also satisfy $a_{n-k} > a_{n-k+1}$. Then $T_k = S_{k+2} - U_{k+2}$, $U_k = S_{k+1} - T_{k+1}$, and $S_0 = S_1 - T_1$. Hence,

$$\begin{aligned} f(n) = \#S_0 &= \#(S_1 - T_1) = \#S_1 - \#(S_3 - U_3) \\ &= \#S_1 - \#S_3 + \#(S_4 - T_4) = \#S_1 - \#S_3 + \#S_4 - \#(S_6 - U_6), \end{aligned}$$

and so on. Since $\#S_k = \binom{n}{k} f(n-k)$, the proof follows. This result (with a different proof) appears in F. N. David and D. E. Barton, *Combinatorial Chance*, Hafner, New York, 1962, pp. 156–157. See also I. M. Gessel, Ph.D. thesis, M.I.T. (Example 3, page 51), and I. P. Goulden and D. M. Jackson, *Combinatorial Enumeration*, John Wiley & Sons, New York, 1983; reprinted by Dover, Mineola, N.Y., 2004 (Exercise 5.2.17).

27. a. Follows easily from Proposition 5.3.2.

b. Let $f_k(\overline{G})$ denote the coefficient of x^{k-1} in $P(\overline{G}, x)$; that is, $f_k(\overline{G})$ is equal to the number of k-component rooted forests F of $\overline{G}$. By the Principle of Inclusion-Exclusion,

$$f_k(\overline{G}) = \sum_F (-1)^{n-\ell(F)} g_k(F),$$

where F ranges over all spanning forests of G, and where $g_k(F)$ denotes the number of k-component rooted forests on $[n]$ that contain F. (Note that $n - \ell(F)$ is equal to the number of edges of F.) By (a), $g_k(F) = p(F)\binom{\ell-1}{\ell-k} n^{\ell-k}$, where $\ell = \ell(F)$. Hence,

$$f_k(\overline{G}) = \sum_F (-1)^{n-\ell} p(F) \binom{\ell-1}{\ell-k} n^{\ell-k}. \tag{2.64}$$

On the other hand, from equation (2.56) the coefficient of x^{k-1} in $(-1)^{n-1} P(G, -x-n)$ is equal to

$$(-1)^{n-1} \sum_F (-1)^{\ell-1} p(F) \binom{\ell-1}{k-1} n^{\ell-1-(k-1)}, \tag{2.65}$$

again summed over all spanning forests F of G, with $\ell = \ell(F)$. Since equations (2.64) and (2.65) agree, the result follows.

Equation (2.57) (essentially the case $x = 0$ of (2.56)) is implicit in H. N. V. Temperley, *Proc. Phys. Soc.* **83** (1984), 3–16. See also Theorem 6.2 of J. W. Moon, *Counting Labelled Trees*, Canadian Mathematical Monographs, no. 1, 1970. The general case (2.56) is due to S. D. Bedrosian, *J. Franklin Inst.* **227** (1964), 313–326. A subsequent proof of (2.56) using matrix techniques is due to A. K. Kelmans. See equation (2.19) in D. M. Cvetković, M. Doob, and H. Sachs, *Spectra of Graphs*, 2nd ed., Johann Ambrosius Barth Verlag, Heidelberg, 1995. A simple proof of (2.56) and additional references appear in J. W. Moon and S. D. Bedrosian, *J. Franklin Inst.* **316** (1983), 187–190.

Equation (2.56) may be regarded as a "reciprocity theorem" for rooted trees. It can be used, in conjunction with the obvious fact $P(G+H, x) = xP(G,x)P(H,x)$ (where $G+H$ denotes the disjoint union of G and H) to unify and simplify many known results involving the enumeration of spanning trees and forests. Part (c) illustrates this technique.

c. We have

$$\begin{aligned}
P(K_1,x) = 1 &\Rightarrow P(nK_1,x) = x^{n-1}\\
&\Rightarrow P(K_n,x) = (x+n)^{n-1} \text{ (so } c(K_n) = n^{n-2})\\
&\Rightarrow P(K_r+K_s,x) = x(x+r)^{r-1}(x+s)^{s-1}\\
&\Rightarrow P(K_{r,s},x) = (x+r+s)(x+s)^{r-1}(x+r)^{s-1}\\
&\Rightarrow c(K_{r,s}) = s^{r-1}r^{s-1}.
\end{aligned}$$

28. This result appeared in R. Stanley [2.19, Ch. 5.3] and was stated without proof in [2.20, Prop. 23.8].

29. b. A generating function proof was given by G. E. Andrews, Concave compositions,

⟨http://www.math.psu.edu/andrews/pdf/277.pdf⟩.

30. See G. W. E. Andrews, in *The Theory of Arithmetic Functions* (A. A. Gioia and D. L. Goldsmith, eds.), Lecture Notes in Math., no. 251, Springer, Berlin, 1972, pp. 1–20. See also Chapter 9 in [1.2].

31. These identities are due to Gauss. See I. Pak [1.62, §5.5].

32. This identity is due to Gauss. A cancellation proof was given by W. Y. C. Chen, Q.-H. Hou and A. Lascoux, *J. Combinatorial Theory Ser. A* **102** (2003), 309–320, where several other proofs are also cited.

33. Let $S = \{1,2,\dots,n-1\}$ in (2.21). There is a unique $w \in \mathfrak{S}_n$ with $D(w) = S$, namely, $w = n, n-1, \dots, 1$, and then $\mathrm{inv}(w) = \binom{n}{2}$. Hence, $\beta_n(S,q) = q^{\binom{n}{2}}$. On the other hand, the right-hand side of (2.21) becomes the left-hand side of (2.58), and the proof follows.

34. This exercise is due to I. M. Gessel, *J. Graph Theory* **3** (1979), 305–307. Part (d) was first shown by M. G. Kendall and B. Babington Smith, *Biometrika* **33** (1940), 239–251. The crucial point in (e) is the following. Let G be the graph whose vertices are the tournaments T on $[n]$ and whose edges consist of pairs T, T' with $T \leftrightarrow T'$. Then from (c) and (d), we deduce that G is bipartite and that every connected component of G is regular, so the connected component containing the vertex T consists of a certain number of tournaments of weight $w(T)$ and an equal number of weight $-w(T)$.

Some far-reaching generalizations appear in D. Zeilberger and D. M. Bressoud, *Discrete Math.* **54** (1985), 201–224 (reprinted in *Discrete Math.* **306** (2006), 1039–1059); D. M. Bressoud, *Europ. J. Combinatorics* **8** (1987), 245–255; and R. M. Calderbank and P. J. Hanlon, *J. Combinatorial Theory Ser. A* **41** (1986), 228–245. The first of these references gives a solution to Exercise 1.19(c).

35. a. By linearity, it suffices to assume that f is a monomial of degree n. If the support of f (set of variables occurring in f) is S, then

$$f(\epsilon_1,\dots,\epsilon_n) = \begin{cases} 1, & \epsilon_i = 1 \text{ for all } x_i \in S,\\ 0, & \text{otherwise.}\end{cases}$$

Hence,

$$\begin{aligned}
\sum_{(\epsilon_1,\dots,\epsilon_n)\in\{0,1\}^n} (-1)^{n-\sum\epsilon_i} f(\epsilon_1,\dots,\epsilon_n) &= \prod_{x_i \notin S}(1-1)\\
&= \begin{cases} 1, & f = x_1x_2\cdots x_n,\\ 0, & \text{otherwise,}\end{cases}
\end{aligned}$$

and the proof follows.

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press

b. Note that

$$\operatorname{per}(A) = [x_1x_2\cdots x_n]\prod_{i=1}^{n}(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n),$$

and use (a). Equation (2.61) is due to H. J. Ryser, *Combinatorial Mathematics*, Math. Assoc. of America, 1963 (Chap. 2, Cor. 4.2). For further information on permanents, see H. Minc, *Permanents*, Encyclopedia of Mathematics and Its Applications, Vol. 6, Addison-Wesley, Reading, Mass., 1978; reprinted by Cambridge University Press, 1984.

36. See B. Gordon [2.12].

https://doi.org/10.1017/CBO9781139058520.003 Published online by Cambridge University Press