

1

What Is Enumerative Combinatorics?

1.1 How to Count

The basic problem of enumerative combinatorics is that of counting the number of elements of a finite set. Usually we are given an infinite collection of finite sets S_i where i ranges over some index set I (such as the nonnegative integers \mathbb{N}), and we wish to count the number $f(i)$ of elements in each S_i “simultaneously.” Immediate philosophical difficulties arise. What does it mean to “count” the number of elements of S_i ? There is no definitive answer to this question. Only through experience does one develop an idea of what is meant by a “determination” of a counting function $f(i)$. The counting function $f(i)$ can be given in several standard ways:

1. The most satisfactory form of $f(i)$ is a completely explicit closed formula involving only well-known functions, and free from summation symbols. Only in rare cases will such a formula exist. As formulas for $f(i)$ become more complicated, our willingness to accept them as “determinations” of $f(i)$ decreases. Consider the following examples.

1.1.1 Example. For each $n \in \mathbb{N}$, let $f(n)$ be the number of subsets of the set $[n] = \{1, 2, \dots, n\}$. Then $f(n) = 2^n$, and no one will quarrel about this being a satisfactory formula for $f(n)$.

1.1.2 Example. Suppose n men give their n hats to a hat-check person. Let $f(n)$ be the number of ways that the hats can be given back to the men, each man receiving one hat, so that no man receives his own hat. For instance, $f(1) = 0$, $f(2) = 1$, $f(3) = 2$. We will see in Chapter 2 (Example 2.2.1) that

$$f(n) = n! \sum_{i=0}^n \frac{(-1)^i}{i!}. \quad (1.1)$$

This formula for $f(n)$ is not as elegant as the formula in Example 1.1.1, but for lack of a simpler answer we are willing to accept (1.1) as a satisfactory formula. It certainly has the virtue of making it easy (in a sense that can be made precise) to compute the values $f(n)$. Moreover, once the derivation of (1.1) is understood

(using the Principle of Inclusion–Exclusion), every term of (1.1) has an easily understood combinatorial meaning. This enables us to “understand” (1.1) intuitively, so our willingness to accept it is enhanced. We also remark that it follows easily from (1.1) that $f(n)$ is the nearest integer to $n!/e$. This is certainly a simple explicit formula, but it has the disadvantage of being “noncombinatorial”; that is, dividing by e and rounding off to the nearest integer has no direct combinatorial significance.

1.1.3 Example. Let $f(n)$ be the number of $n \times n$ matrices \mathbf{M} of 0’s and 1’s such that every row and column of \mathbf{M} has three 1’s. For example, $f(0) = 1$, $f(1) = f(2) = 0$, $f(3) = 1$. The most explicit formula known at present for $f(n)$ is

$$f(n) = 6^{-n} n!^2 \sum \frac{(-1)^\beta (\beta + 3\gamma)! 2^\alpha 3^\beta}{\alpha! \beta! \gamma!^2 6^\gamma}, \quad (1.2)$$

where the sum ranges over all $(n+2)(n+1)/2$ solutions to $\alpha + \beta + \gamma = n$ in nonnegative integers. This formula gives very little insight into the behavior of $f(n)$, but it does allow one to compute $f(n)$ much faster than if only the combinatorial definition of $f(n)$ were used. Hence with some reluctance we accept (1.2) as a “determination” of $f(n)$. Of course, if someone were later to prove that $f(n) = (n-1)(n-2)/2$ (rather unlikely), then our enthusiasm for (1.2) would be considerably diminished.

1.1.4 Example. There are actually formulas in the literature (“nameless here for evermore”) for certain counting functions $f(n)$ whose evaluation requires listing all (or almost all) of the $f(n)$ objects being counted! Such a “formula” is completely worthless.

2. A recurrence for $f(i)$ may be given in terms of previously calculated $f(j)$ ’s, thereby giving a simple procedure for calculating $f(i)$ for any desired $i \in I$. For instance, let $f(n)$ be the number of subsets of $[n]$ that do not contain two consecutive integers. For example, for $n = 4$ we have the subsets \emptyset , $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\{1,3\}$, $\{1,4\}$, $\{2,4\}$, so $f(4) = 8$. It is easily seen that $f(n) = f(n-1) + f(n-2)$ for $n \geq 2$. This makes it trivial, for example, to compute $f(20) = 17711$. On the other hand, it can be shown (see Section 4.1 for the underlying theory) that

$$f(n) = \frac{1}{\sqrt{5}} \left(\tau^{n+2} - \bar{\tau}^{n+2} \right),$$

where $\tau = \frac{1}{2}(1 + \sqrt{5})$, $\bar{\tau} = \frac{1}{2}(1 - \sqrt{5})$. This is an explicit answer, but because it involves irrational numbers, it is a matter of opinion (which may depend on the context) whether it is a better answer than the recurrence $f(n) = f(n-1) + f(n-2)$.

3. An algorithm may be given for computing $f(i)$. This method of determining f subsumes the previous two, as well as method 5, which follows. Any counting function likely to arise in practice can be computed from an algorithm, so the acceptability of this method will depend on the elegance and performance of the algorithm. In general, we would like the time that it takes the algorithm to compute $f(i)$ to be “substantially less” than $f(i)$ itself. Otherwise, we are accomplishing little more than a brute force listing of the objects counted by $f(i)$. It would take us too far afield to discuss the profound contributions that computer science has made to the problem of analyzing, constructing, and evaluating algorithms. We will be concerned almost exclusively with enumerative problems that admit solutions that are more concrete than an algorithm.
4. An estimate may be given for $f(i)$. If $I = \mathbb{N}$, this estimate frequently takes the form of an *asymptotic formula* $f(n) \sim g(n)$, where $g(n)$ is a “familiar function.” The notation $f(n) \sim g(n)$ means that $\lim_{n \rightarrow \infty} f(n)/g(n) = 1$. For instance, let $f(n)$ be the function of Example 1.1.3. It can be shown that

$$f(n) \sim e^{-2} 36^{-n} (3n)!.$$

For many purposes this estimate is superior to the “explicit” formula (1.2).

5. The most useful but most difficult to understand method for evaluating $f(i)$ is to give its *generating function*. We will not develop in this chapter a rigorous abstract theory of generating functions, but will instead content ourselves with an informal discussion and some examples. Informally, a generating function is an “object” that represents a counting function $f(i)$. Usually this object is a *formal power series*. The two most common types of generating functions are *ordinary* generating functions and *exponential* generating functions. If $I = \mathbb{N}$, then the ordinary generating function of $f(n)$ is the formal power series

$$\sum_{n \geq 0} f(n)x^n,$$

while the exponential generating function of $f(n)$ is the formal power series

$$\sum_{n \geq 0} f(n) \frac{x^n}{n!}.$$

(If $I = \mathbb{P}$, the positive integers, then these sums begin at $n = 1$.) These power series are called “formal” because we are not concerned with letting x take on particular values, and we ignore questions of convergence and divergence. The term x^n or $x^n/n!$ merely marks the place where $f(n)$ is written.

If $F(x) = \sum_{n \geq 0} a_n x^n$, then we call a_n the *coefficient* of x^n in $F(x)$, and write

$$a_n = [x^n]F(x).$$

Similarly, if $F(x) = \sum_{n \geq 0} a_n x^n/n!$, then we write

$$a_n = n![x^n]F(x).$$

In the same way, we can deal with generating functions of several variables, such as

$$\sum_{l \geq 0} \sum_{m \geq 0} \sum_{n \geq 0} f(l, m, n) \frac{x^l y^m z^n}{n!}$$

(which may be considered as “ordinary” in the indices l, m and “exponential” in n), or even of infinitely many variables. In this latter case every term should involve only finitely many of the variables. A simple generating function in infinitely many variables is $x_1 + x_2 + x_3 + \cdots$.

Why bother with generating functions if they are merely another way of writing a counting function? The answer is that we can perform various natural operations on generating functions that have a combinatorial significance. For instance, we can add two generating functions, say in one variable with $I = \mathbb{N}$, by the rule

$$\left(\sum_{n \geq 0} a_n x^n \right) + \left(\sum_{n \geq 0} b_n x^n \right) = \sum_{n \geq 0} (a_n + b_n) x^n$$

or

$$\left(\sum_{n \geq 0} a_n \frac{x^n}{n!} \right) + \left(\sum_{n \geq 0} b_n \frac{x^n}{n!} \right) = \sum_{n \geq 0} (a_n + b_n) \frac{x^n}{n!}.$$

Similarly, we can multiply generating functions according to the rule

$$\left(\sum_{n \geq 0} a_n x^n \right) \left(\sum_{n \geq 0} b_n x^n \right) = \sum_{n \geq 0} c_n x^n,$$

where $c_n = \sum_{i=0}^n a_i b_{n-i}$, or

$$\left(\sum_{n \geq 0} a_n \frac{x^n}{n!} \right) \left(\sum_{n \geq 0} b_n \frac{x^n}{n!} \right) = \sum_{n \geq 0} d_n \frac{x^n}{n!},$$

where $d_n = \sum_{i=0}^n \binom{n}{i} a_i b_{n-i}$, with $\binom{n}{i} = n! / i!(n-i)!$. Note that these operations are just what we would obtain by treating generating functions as if they obeyed the ordinary laws of algebra, such as $x^i x^j = x^{i+j}$. These operations coincide with the addition and multiplication of functions when the power series converge for appropriate values of x , and they obey such familiar laws of algebra as associativity and commutativity of addition and multiplication, distributivity of multiplication over addition, and cancellation of multiplication (i.e., if $F(x)G(x) = F(x)H(x)$ and $F(x) \neq 0$, then $G(x) = H(x)$). In fact, the set of all formal power series $\sum_{n \geq 0} a_n x^n$ with complex coefficients a_n (or more generally, coefficients in any integral domain R , where integral domains are assumed to be commutative with a multiplicative identity 1) forms a (commutative) integral domain under the operations just defined. This integral domain is denoted $\mathbb{C}[[x]]$ (or more generally, $R[[x]]$). Actually, $\mathbb{C}[[x]]$, or more generally $K[[x]]$ when K is a field, is a very

special type of integral domain. For readers with some familiarity with algebra, we remark that $\mathbb{C}[[x]]$ is a principal ideal domain and therefore a unique factorization domain. In fact, every ideal of $\mathbb{C}[[x]]$ has the form (x^n) for some $n \geq 0$. From the viewpoint of commutative algebra, $\mathbb{C}[[x]]$ is a one-dimensional complete regular local ring. Moreover, the operation $[x^n] : \mathbb{C}[[x]] \rightarrow \mathbb{C}$ of taking the coefficient of x^n (and similarly $[x^n/n!]$) is a linear functional on $\mathbb{C}[[x]]$. These general algebraic considerations will not concern us here; rather we will discuss from an elementary viewpoint the properties of $\mathbb{C}[[x]]$ that will be useful to us.

There is an obvious extension of the ring $\mathbb{C}[[x]]$ to formal power series in m variables x_1, \dots, x_m . The set of all such power series with complex coefficients is denoted $\mathbb{C}[[x_1, \dots, x_m]]$ and forms a unique factorization domain (though not a principal ideal domain for $m \geq 2$).

It is primarily through experience that the combinatorial significance of the algebraic operations of $\mathbb{C}[[x]]$ or $\mathbb{C}[[x_1, \dots, x_m]]$ is understood, as well as the problems of whether to use ordinary or exponential generating functions (or various other kinds discussed in later chapters). In Section 3.18 we will explain to some extent the combinatorial significance of these operations, but even then experience is indispensable.

If $F(x)$ and $G(x)$ are elements of $\mathbb{C}[[x]]$ satisfying $F(x)G(x) = 1$, then we (naturally) write $G(x) = F(x)^{-1}$. (Here 1 is short for $1 + 0x + 0x^2 + \dots$.) It is easy to see that $F(x)^{-1}$ exists (in which case it is unique) if and only if $a_0 \neq 0$, where $F(x) = \sum_{n \geq 0} a_n x^n$. One commonly writes “symbolically” $a_0 = F(0)$, even though $F(x)$ is not considered to be a function of x . If $F(0) \neq 0$ and $F(x)G(x) = H(x)$, then $G(x) = F(x)^{-1}H(x)$, which we also write as $G(x) = H(x)/F(x)$. More generally, the operation $^{-1}$ satisfies all the familiar laws of algebra, provided it is only applied to power series $F(x)$ satisfying $F(0) \neq 0$. For instance, $(F(x)G(x))^{-1} = F(x)^{-1}G(x)^{-1}$, $(F(x)^{-1})^{-1} = F(x)$, and so on. Similar results hold for $\mathbb{C}[[x_1, \dots, x_m]]$.

1.1.5 Example. Let $(\sum_{n \geq 0} \alpha^n x^n)(1 - \alpha x) = \sum_{n \geq 0} c_n x^n$, where α is nonzero complex number. (We could also take α to be an indeterminate, in which case we should extend the coefficient field to $\mathbb{C}(\alpha)$, the field of rational functions over \mathbb{C} in the variable α .) Then by definition of power series multiplication,

$$c_n = \begin{cases} 1, & n = 0 \\ \alpha^n - \alpha(\alpha^{n-1}) = 0, & n \geq 1. \end{cases}$$

Hence, $\sum_{n \geq 0} \alpha^n x^n = (1 - \alpha x)^{-1}$, which can also be written

$$\sum_{n \geq 0} \alpha^n x^n = \frac{1}{1 - \alpha x}.$$

This formula comes as no surprise; it is simply the formula (in a formal setting) for summing a geometric series.

Example 1.1.5 provides a simple illustration of the general principle that, informally speaking, if we have an identity involving power series that is valid when the power series are regarded as functions (so that the variables are sufficiently small complex numbers), then this identity continues to remain valid when regarded as an identity among formal power series, *provided* the operations defined in the formulas are well defined for formal power series. It would be unnecessarily pedantic for us to state a precise form of this principle here, since the reader should have little trouble justifying in any particular case the formal validity of our manipulations with power series. We will give several examples throughout this section to illustrate this contention.

1.1.6 Example. The identity

$$\left(\sum_{n \geq 0} \frac{x^n}{n!} \right) \left(\sum_{n \geq 0} (-1)^n \frac{x^n}{n!} \right) = 1 \quad (1.3)$$

is valid at the function-theoretic level (it states that $e^x e^{-x} = 1$) and is well defined as a statement involving formal power series. Hence, (1.3) is a valid formal power series identity. In other words (equating coefficients of $x^n/n!$ on both sides of (1.3)), we have

$$\sum_{k=0}^n (-1)^k \binom{n}{k} = \delta_{0n}. \quad (1.4)$$

To justify this identity directly from (1.3), we may reason as follows. Both sides of (1.3) converge for all $x \in \mathbb{C}$, so we have

$$\sum_{n \geq 0} \left(\sum_{k=0}^n (-1)^k \binom{n}{k} \right) \frac{x^n}{n!} = 1, \quad \text{for all } x \in \mathbb{C}.$$

But if two power series in x represent the same function $f(x)$ in a neighborhood of 0, then these two power series must agree term-by-term, by a standard elementary result concerning power series. Hence, (1.4) follows.

1.1.7 Example. The identity

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

is valid at the function-theoretic level (it states that $e^{x+1} = e \cdot e^x$) but does not make sense as a statement involving formal power series. There is no *formal* procedure for writing $\sum_{n \geq 0} (x+1)^n/n!$ as a member of $\mathbb{C}[[x]]$. For instance, the constant term of $\sum_{n \geq 0} (x+1)^n/n!$ is $\sum_{n \geq 0} 1/n!$, whose interpretation as a member of $\mathbb{C}[[x]]$ involves the consideration of convergence.

Although the expression $\sum_{n \geq 0} (x+1)^n/n!$ does not make sense *formally*, there are nevertheless certain infinite processes that can be carried out formally in $\mathbb{C}[[x]]$. (These concepts extend straightforwardly to $\mathbb{C}[[x_1, \dots, x_m]]$, but for simplicity we consider only $\mathbb{C}[[x]]$.) To define these processes, we need to put some additional structure on $\mathbb{C}[[x]]$ —namely, the notion of *convergence*. From an algebraic standpoint, the definition of convergence is inherent in the statement that $\mathbb{C}[[x]]$ is *complete* in a certain standard topology that can be put on $\mathbb{C}[[x]]$. However, we will assume no knowledge of topology on the part of the reader and will instead give a self-contained, elementary treatment of convergence.

If $F_1(x), F_2(x), \dots$ is a sequence of formal power series, and if $F(x) = \sum_{n \geq 0} a_n x^n$ is another formal power series, we say by definition that $F_i(x)$ *converges* to $F(x)$ as $i \rightarrow \infty$, written $F_i(x) \rightarrow F(x)$ or $\lim_{i \rightarrow \infty} F_i(x) = F(x)$, provided that for all $n \geq 0$ there is a number $\delta(n)$ such that the coefficient of x^n in $F_i(x)$ is a_n whenever $i \geq \delta(n)$. In other words, for every n the sequence

$$[x^n]F_1(x), [x^n]F_2(x), \dots$$

of complex numbers eventually becomes constant (or *stabilizes*) with value a_n . An equivalent definition of convergence is the following. Define the *degree* of a nonzero formal power series $F(x) = \sum_{n \geq 0} a_n x^n$, denoted $\deg F(x)$, to be the least integer n such that $a_n \neq 0$. Note that $\deg F(x)G(x) = \deg F(x) + \deg G(x)$. Then $F_i(x)$ converges if and only if $\lim_{i \rightarrow \infty} \deg(F_{i+1}(x) - F_i(x)) = \infty$, and $F_i(x)$ converges to $F(x)$ if and only if $\lim_{i \rightarrow \infty} \deg(F(x) - F_i(x)) = \infty$.

We now say that an infinite sum $\sum_{j \geq 0} F_j(x)$ has the value $F(x)$ provided that $\sum_{j=0}^i F_j(x) \rightarrow F(x)$. A similar definition is made for the infinite product $\prod_{j \geq 1} F_j(x)$. To avoid unimportant technicalities we assume that, in any infinite product $\prod_{j \geq 1} F_j(x)$, each factor $F_j(x)$ satisfies $F_j(0) = 1$.

For instance, let $F_j(x) = a_j x^j$. Then for $i \geq n$, the coefficient of x^n in $\sum_{j=0}^i F_j(x)$ is a_n . Hence $\sum_{j \geq 0} F_j(x)$ is just the power series $\sum_{n \geq 0} a_n x^n$. Thus, we can think of the formal power series $\sum_{n \geq 0} a_n x^n$ as actually being the “sum” of its individual terms. The proofs of the following two elementary results are left to the reader.

1.1.8 Proposition. The infinite series $\sum_{j \geq 0} F_j(x)$ converges if and only if

$$\lim_{j \rightarrow \infty} \deg F_j(x) = \infty.$$

1.1.9 Proposition. The infinite product $\prod_{j \geq 1} (1 + G_j(x))$, where $G_j(0) = 0$, converges if and only if $\lim_{j \rightarrow \infty} \deg G_j(x) = \infty$.

It is essential to realize that in evaluating a convergent series $\sum_{j \geq 0} F_j(x)$ (or similarly a product $\prod_{j \geq 1} F_j(x)$), the coefficient of x^n for any given n can be

computed using only *finite* processes. For if j is sufficiently large, say $j > \delta(n)$, then $\deg F_j(x) > n$, so that

$$[x^n] \sum_{j \geq 0} F_j(x) = [x^n] \sum_{j=0}^{\delta(n)} F_j(x).$$

The latter expression involves only a *finite* sum.

The most important combinatorial application of the notion of convergence is to the idea of power series composition. If $F(x) = \sum_{n \geq 0} a_n x^n$ and $G(x)$ are formal power series with $G(0) = 0$, define the *composition* $F(G(x))$ to be the infinite sum $\sum_{n \geq 0} a_n G(x)^n$. Since $\deg G(x)^n = n \cdot \deg G(x) \geq n$, we see by Proposition 1.1.8 that $F(G(x))$ is well defined as a *formal* power series. We also see why an expression such as e^{1+x} does not make sense formally; namely, the infinite series $\sum_{n \geq 0} (1+x)^n / n!$ does not converge in accordance with the preceding definition. On the other hand, an expression like $e^{e^x - 1}$ makes good sense formally, since it has the form $F(G(x))$ where $F(x) = \sum_{n \geq 0} x^n / n!$ and $G(x) = \sum_{n \geq 1} x^n / n!$.

1.1.10 Example. If $F(x) \in \mathbb{C}[[x]]$ satisfies $F(0) = 0$, then we can *define* for any $\lambda \in \mathbb{C}$ the formal power series

$$(1 + F(x))^\lambda = \sum_{n \geq 0} \binom{\lambda}{n} F(x)^n, \quad (1.5)$$

where $\binom{\lambda}{n} = \lambda(\lambda-1) \cdots (\lambda-n+1)/n!$. In fact, we may regard λ as an indeterminate and take (1.5) as the definition of $(1 + F(x))^\lambda$ as an element of $\mathbb{C}[[x, \lambda]]$ (or of $\mathbb{C}[\lambda][[x]]$; that is, the coefficient of x^n in $(1 + F(x))^\lambda$ is a certain *polynomial* in λ). All the expected properties of exponentiation are indeed valid, such as

$$(1 + F(x))^{\lambda+\mu} = (1 + F(x))^\lambda (1 + F(x))^\mu,$$

regarded as an identity in the ring $\mathbb{C}[[x, \lambda, \mu]]$, or in the ring $\mathbb{C}[[x]]$ where one takes $\lambda, \mu \in \mathbb{C}$.

If $F(x) = \sum_{n \geq 0} a_n x^n$, define the *formal derivative* $F'(x)$ (also denoted $\frac{dF}{dx}$ or $DF(x)$) to be the formal power series

$$F'(x) = \sum_{n \geq 0} n a_n x^{n-1} = \sum_{n \geq 0} (n+1) a_{n+1} x^n.$$

It is easy to check that all the familiar laws of differentiation that are well defined formally continue to be valid for formal power series. In particular,

$$\begin{aligned} (F + G)' &= F' + G', \\ (FG)' &= F'G + FG', \\ F(G(x))' &= G'(x)F'(G(x)). \end{aligned}$$

We thus have a theory of *formal calculus* for formal power series. The usefulness of this theory will become apparent in subsequent examples. We first give an example of the use of the formal calculus that should shed some additional light on the validity of manipulating formal power series $F(x)$ as if they were actual functions of x .

1.1.11 Example. Suppose $F(0) = 1$, and let $G(x)$ be the power series (easily seen to be unique) satisfying

$$G'(x) = F'(x)/F(x), \quad G(0) = 0. \quad (1.6)$$

From the function-theoretic viewpoint we can “solve” (1.6) to obtain $F(x) = \exp G(x)$, where by definition

$$\exp G(x) = \sum_{n \geq 0} \frac{G(x)^n}{n!}.$$

Since $G(0) = 0$ everything is well defined formally, so (1.6) should remain equivalent to $F(x) = \exp G(x)$ even if the power series for $F(x)$ converges only at $x = 0$. How can this assertion be justified without actually proving a combinatorial identity? Let $F(x) = 1 + \sum_{n \geq 1} a_n x^n$. From (1.6) we can compute explicitly $G(x) = \sum_{n \geq 1} b_n x^n$, and it is quickly seen that each b_n is a *polynomial* in finitely many of the a_i 's. It then follows that if $\exp G(x) = 1 + \sum_{n \geq 1} c_n x^n$, then each c_n will also be a polynomial in finitely many of the a_i 's, say $c_n = p_n(a_1, a_2, \dots, a_m)$, where m depends on n . Now we know that $F(x) = \exp G(x)$ provided $1 + \sum_{n \geq 1} a_n x^n$ converges. If two Taylor series convergent in some neighborhood of the origin represent the same function, then their coefficients coincide. Hence $a_n = p_n(a_1, a_2, \dots, a_m)$ provided $1 + \sum_{n \geq 1} a_n x^n$ converges. Thus, the two polynomials a_n and $p_n(a_1, \dots, a_m)$ agree in some neighborhood of the origin of \mathbb{C}^m , so they must be equal. (It is a simple result that if two complex polynomials in m variables agree in some open subset of \mathbb{C}^m , then they are identical.) Since $a_n = p_n(a_1, a_2, \dots, a_m)$ as polynomials, the identity $F(x) = \exp G(x)$ continues to remain valid for *formal* power series.

There is an alternative method for justifying the formal solution $F(x) = \exp G(x)$ to (1.6), which may appeal to topologically inclined readers. Given $G(x)$ with $G(0) = 0$, define $F(x) = \exp G(x)$ and consider a map $\phi : \mathbb{C}[[x]] \rightarrow \mathbb{C}[[x]]$ defined by $\phi(G(x)) = G'(x) - \frac{F'(x)}{F(x)}$. One easily verifies the following: (a) if G converges in some neighborhood of 0, then $\phi(G(x)) = 0$; (b) the set \mathcal{G} of all power series $G(x) \in \mathbb{C}[[x]]$ that converge in some neighborhood of 0 is dense in $\mathbb{C}[[x]]$, in the topology defined earlier (in fact, the set $\mathbb{C}[x]$ of polynomials is dense); and (c) the function ϕ is continuous in the topology defined earlier. From this it follows that $\phi(G(x)) = 0$ for all $G(x) \in \mathbb{C}[[x]]$ with $G(0) = 0$.

We now present various illustrations in the manipulation of generating functions. Throughout we will be making heavy use of the principle that formal power series can be treated as if they were functions.

1.1.12 Example. Find a simple expression for the generating function $F(x) = \sum_{n \geq 0} a_n x^n$, where $a_0 = a_1 = 1$, $a_n = a_{n-1} + a_{n-2}$ if $n \geq 2$. We have

$$\begin{aligned} F(x) &= \sum_{n \geq 0} a_n x^n = 1 + x + \sum_{n \geq 2} a_n x^n \\ &= 1 + x + \sum_{n \geq 2} (a_{n-1} + a_{n-2}) x^n \\ &= 1 + x + x \sum_{n \geq 2} a_{n-1} x^{n-1} + x^2 \sum_{n \geq 2} a_{n-2} x^{n-2} \\ &= 1 + x + x(F(x) - 1) + x^2 F(x). \end{aligned}$$

Solving for $F(x)$ yields $F(x) = 1/(1 - x - x^2)$. The number a_n is just the *Fibonacci number* F_{n+1} . For some combinatorial properties of Fibonacci numbers, see Exercises 1.35–1.42. For the general theory of rational generating functions and linear recurrences with constant coefficients illustrated in the present example, see Section 4.1.

1.1.13 Example. Find a simple expression for the generating function $F(x) = \sum_{n \geq 0} a_n x^n / n!$, where $a_0 = 1$,

$$a_{n+1} = a_n + n a_{n-1}, \quad n \geq 0. \quad (1.7)$$

(Note that if $n = 0$ we get $a_1 = a_0 + 0 \cdot a_{-1}$, so the value of a_{-1} is irrelevant.) Multiply the recurrence (1.7) by $x^n / n!$ and sum on $n \geq 0$. We get

$$\begin{aligned} \sum_{n \geq 0} a_{n+1} \frac{x^n}{n!} &= \sum_{n \geq 0} a_n \frac{x^n}{n!} + \sum_{n \geq 0} n a_{n-1} \frac{x^n}{n!} \\ &= \sum_{n \geq 0} a_n \frac{x^n}{n!} + \sum_{n \geq 1} a_{n-1} \frac{x^n}{(n-1)!}. \end{aligned}$$

The left-hand side is just $F'(x)$, while the right-hand side is $F(x) + xF(x)$. Hence, $F'(x) = (1+x)F(x)$. The unique solution to this differential equation satisfying $F(0) = 1$ is $F(x) = \exp(x + \frac{1}{2}x^2)$. (As shown in Example 1.1.11, solving this differential equation is a purely formal procedure.) For the combinatorial significance of the numbers a_n , see equation (5.32).

NOTE. With the benefit of hindsight we wrote the recurrence $a_{n+1} = a_n + n a_{n-1}$ with indexing that makes the computation simplest. If for instance we had written $a_n = a_{n-1} + (n-1)a_{n-2}$, then the computation would be more complicated (though still quite tractable). In converting recurrences to generating function identities, it can be worthwhile to consider how best to index the recurrence.

1.1.14 Example. Let $\mu(n)$ be the Möbius function of number theory; that is, $\mu(1) = 1$, $\mu(n) = 0$ if n is divisible by the square of an integer greater than one,

and $\mu(n) = (-1)^r$ if n is the product of r distinct primes. Find a simple expression for the power series

$$F(x) = \prod_{n \geq 1} (1 - x^n)^{-\mu(n)/n}. \quad (1.8)$$

First let us make sure that $F(x)$ is well defined as a formal power series. We have by Example 1.1.10 that

$$(1 - x^n)^{-\mu(n)/n} = \sum_{i \geq 0} \binom{-\mu(n)/n}{i} (-1)^i x^{in}.$$

Note that $(1 - x^n)^{-\mu(n)/n} = 1 + H(x)$, where $\deg H(x) = n$. Hence, by Proposition 1.1.9 the infinite product (1.8) converges, so $F(x)$ is well defined. Now apply \log to (1.8). In other words, form $\log F(x)$, where

$$\log(1 + x) = \sum_{n \geq 1} (-1)^{n-1} \frac{x^n}{n},$$

the power series expansion for the natural logarithm at $x = 0$. We obtain

$$\begin{aligned} \log F(x) &= \log \prod_{n \geq 1} (1 - x^n)^{-\mu(n)/n} \\ &= - \sum_{n \geq 1} \log(1 - x^n)^{\mu(n)/n} \\ &= - \sum_{n \geq 1} \frac{\mu(n)}{n} \log(1 - x^n) \\ &= - \sum_{n \geq 1} \frac{\mu(n)}{n} \sum_{i \geq 1} \left(-\frac{x^{in}}{i} \right). \end{aligned}$$

The coefficient of x^m in the preceding power series is

$$\frac{1}{m} \sum_{d|m} \mu(d),$$

where the sum is over all positive integers d dividing m . It is well known that

$$\frac{1}{m} \sum_{d|m} \mu(d) = \begin{cases} 1, & m = 1 \\ 0, & \text{otherwise.} \end{cases}$$

Hence, $\log F(x) = x$, so $F(x) = e^x$. Note that the derivation of this miraculous formula involved only *formal* manipulations.

1.1.15 Example. Find the unique sequence $a_0 = 1, a_1, a_2, \dots$ of real numbers satisfying

$$\sum_{k=0}^n a_k a_{n-k} = 1 \quad (1.9)$$

for all $n \in \mathbb{N}$. The trick is to recognize the left-hand side of (1.9) as the coefficient of x^n in $(\sum_{n \geq 0} a_n x^n)^2$. Letting $F(x) = \sum_{n \geq 0} a_n x^n$, we then have

$$F(x)^2 = \sum_{n \geq 0} x^n = \frac{1}{1-x}.$$

Hence,

$$F(x) = (1-x)^{-1/2} = \sum_{n \geq 0} \binom{-1/2}{n} (-1)^n x^n,$$

so

$$\begin{aligned} a_n &= (-1)^n \binom{-1/2}{n} \\ &= (-1)^n \frac{(-\frac{1}{2})(-\frac{3}{2})(-\frac{5}{2}) \cdots (-\frac{2n-1}{2})}{n!} \\ &= \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2^n n!}. \end{aligned}$$

Note that a_n can also be rewritten as $4^{-n} \binom{2n}{n}$. The identity

$$\binom{2n}{n} = (-1)^n 4^n \binom{-1/2}{n} \quad (1.10)$$

can be useful for problems involving $\binom{2n}{n}$.

Now that we have discussed the manipulation of formal power series, the question arises as to the advantages of using generating functions to represent a counting function $f(n)$. Why, for instance, should a formula such as

$$\sum_{n \geq 0} f(n) \frac{x^n}{n!} = \exp\left(x + \frac{x^2}{2}\right) \quad (1.11)$$

be regarded as a “determination” of $f(n)$? Basically, the answer is that there are many standard, routine techniques for extracting information from generating functions. Generating functions are frequently the most concise and efficient way of presenting information about their coefficients. For instance, from (1.11) an experienced enumerative combinatorialist can tell at a glance the following:

1. A simple recurrence for $f(n)$ can be found by differentiation. Namely, we obtain

$$\sum_{n \geq 1} f(n) \frac{x^{n-1}}{(n-1)!} = (1+x)e^{x+x^2/2} = (1+x) \sum_{n \geq 0} f(n) \frac{x^n}{n!}.$$

Equating coefficients of $x^n/n!$ yields

$$f(n+1) = f(n) + nf(n-1), \quad n \geq 1.$$

Note that in Example 1.1.13 we went in the opposite direction (i.e., we obtained the generating function from the recurrence, a less straightforward procedure).

2. An explicit formula for $f(n)$ can be obtained from $e^{x+(x^2/2)} = e^x e^{x^2/2}$. Namely,

$$\begin{aligned} \sum_{n \geq 0} f(n) \frac{x^n}{n!} &= e^x e^{x^2/2} = \left(\sum_{n \geq 0} \frac{x^n}{n!} \right) \left(\sum_{n \geq 0} \frac{x^{2n}}{2^n n!} \right) \\ &= \left(\sum_{n \geq 0} \frac{x^n}{n!} \right) \left(\sum_{n \geq 0} \frac{(2n)!}{2^n n!} \frac{x^{2n}}{(2n)!} \right), \end{aligned}$$

so that

$$f(n) = \sum_{\substack{i \geq 0 \\ i \text{ even}}} \binom{n}{i} \frac{i!}{2^{i/2}(i/2)!} = \sum_{j \geq 0} \binom{n}{2j} \frac{(2j)!}{2^j j!}.$$

3. Regarded as a function of a complex variable, $\exp(x + \frac{x^2}{2})$ is a nicely behaved entire function, so that standard techniques from the theory of asymptotic analysis can be used to estimate $f(n)$. As a first approximation, it is routine (for someone sufficiently versed in complex variable theory) to obtain the asymptotic formula

$$f(n) \sim \frac{1}{\sqrt{2}} n^{n/2} e^{-\frac{n}{2} + \sqrt{n} - \frac{1}{4}}. \quad (1.12)$$

No other method of describing $f(n)$ makes it so easy to determine these fundamental properties. Many other properties of $f(n)$ can also be easily obtained from the generating function; for instance, we leave to the reader the problem of evaluating, essentially by inspection of (1.11), the sum

$$\sum_{i=0}^n (-1)^{n-i} \binom{n}{i} f(i) \quad (1.13)$$

(see Exercise 1.7). Therefore, we are ready to accept the generating function $\exp(x + \frac{x^2}{2})$ as a satisfactory determination of $f(n)$.

This completes our discussion of generating functions and more generally the problem of giving a satisfactory description of a counting function $f(n)$. We now turn to the question of what is the best way to *prove* that a counting function has some given description. In accordance with the principle from other branches of mathematics that it is better to exhibit an explicit isomorphism between two objects than merely prove that they are isomorphic, we adopt the general principle that it is better to exhibit an explicit one-to-one correspondence (bijection) between two finite sets than merely to prove that they have the same number of elements. A proof that shows that a certain set S has a certain number m of elements by constructing an explicit bijection between S and some other set that is known to have m elements is called a *combinatorial proof* or *bijective proof*. The precise border between combinatorial and noncombinatorial proofs is rather hazy, and certain arguments that to an inexperienced enumerator will appear noncombinatorial will

be recognized by a more facile counter as combinatorial, primarily because he or she is aware of certain standard techniques for converting apparently noncombinatorial arguments into combinatorial ones. Such subtleties will not concern us here, and we now give some clear-cut examples of the distinction between combinatorial and noncombinatorial proofs. We use the notation $\#S$ or $|S|$ for the cardinality (number of elements) of the finite set S .

1.1.16 Example. Let n and k be fixed positive integers. How many sequences (X_1, X_2, \dots, X_k) are there of subsets of the set $[n] = \{1, 2, \dots, n\}$ such that $X_1 \cap X_2 \cap \dots \cap X_k = \emptyset$? Let $f(k, n)$ be this number. If we were not particularly inspired we could perhaps argue as follows. Suppose $X_1 \cap X_2 \cap \dots \cap X_{k-1} = T$, where $\#T = i$. If $Y_j = X_j - T$, then $Y_1 \cap \dots \cap Y_{k-1} = \emptyset$ and $Y_j \subseteq [n] - T$. Hence, there are $f(k-1, n-i)$ sequences (X_1, \dots, X_{k-1}) such that $X_1 \cap X_2 \cap \dots \cap X_{k-1} = T$. For each such sequence, X_k can be any of the 2^{n-i} subsets of $[n] - T$. As is probably familiar to most readers and will be discussed later, there are $\binom{n}{i} = n!/i!(n-i)!$ i -element subsets T of $[n]$. Hence,

$$f(k, n) = \sum_{i=0}^n \binom{n}{i} 2^{n-i} f(k-1, n-i). \quad (1.14)$$

Let $F_k(x) = \sum_{n \geq 0} f(k, n) x^n / n!$. Then (1.14) is equivalent to

$$F_k(x) = e^x F_{k-1}(2x).$$

Clearly $F_1(x) = e^x$. It follows easily that

$$\begin{aligned} F_k(x) &= \exp(x + 2x + 4x + \dots + 2^{k-1}x) \\ &= \exp((2^k - 1)x) \\ &= \sum_{n \geq 0} (2^k - 1)^n \frac{x^n}{n!}. \end{aligned}$$

Hence, $f(k, n) = (2^k - 1)^n$. This argument is a flagrant example of a noncombinatorial proof. The resulting answer is extremely simple despite the contortions involved to obtain it, and it cries out for a better understanding. In fact, $(2^k - 1)^n$ is clearly the number of n -tuples (Z_1, Z_2, \dots, Z_n) , where each Z_i is a subset of $[k]$ not equal to $[k]$. Can we find a bijection θ between the set S_{kn} of all $(X_1, \dots, X_k) \subseteq [n]^k$ such that $X_1 \cap \dots \cap X_k = \emptyset$, and the set T_{kn} of all (Z_1, \dots, Z_n) where $[k] \neq Z_i \subseteq [k]$? Given an element (Z_1, \dots, Z_n) of T_{kn} , define (X_1, \dots, X_k) by the condition that $i \in X_j$ if and only if $j \in Z_i$. This rule is just a precise way of saying the following: The element 1 can appear in any of the X_i 's except all of them, so there are $2^k - 1$ choices for which of the X_i 's contain 1; similarly there are $2^k - 1$ choices for which of the X_i 's contain 2, 3, \dots , n , so there are $(2^k - 1)^n$ choices in all. Thus, the crucial point of the problem is that the different elements of $[n]$ behave *independently*, so we end up with a simple product. We leave to the reader the (rather

dull) task of rigorously verifying that θ is a bijection, but this fact should be intuitively clear. The usual way to show that θ is a bijection is to construct explicitly a map $\phi : T_{kn} \rightarrow S_{kn}$, and then to show that $\phi = \theta^{-1}$; for example, by showing that $\phi\theta(X) = X$ and that θ is surjective. CAVEAT. Any proof that θ is bijective must not use a priori the fact that $\#S_{kn} = \#T_{kn}$!

Not only is the preceding combinatorial proof much shorter than our previous proof, but it also makes the reason for the simple answer completely transparent. It is often the case, as occurred here, that the first proof to come to mind turns out to be laborious and inelegant, but that the final answer suggests a simpler combinatorial proof.

1.1.17 Example. Verify the identity

$$\sum_{i=0}^n \binom{a}{i} \binom{b}{n-i} = \binom{a+b}{n}, \quad (1.15)$$

where a, b , and n are nonnegative integers. A noncombinatorial proof would run as follows. The left-hand side of (1.15) is the coefficient of x^n in the power series (polynomial) $(\sum_{i \geq 0} \binom{a}{i} x^i) (\sum_{j \geq 0} \binom{b}{j} x^j)$. But by the binomial theorem,

$$\begin{aligned} \left(\sum_{i \geq 0} \binom{a}{i} x^i \right) \left(\sum_{j \geq 0} \binom{b}{j} x^j \right) &= (1+x)^a (1+x)^b \\ &= (1+x)^{a+b} \\ &= \sum_{n \geq 0} \binom{a+b}{n} x^n, \end{aligned}$$

so the proof follows. A combinatorial proof runs as follows. The right-hand side of (1.15) is the number of n -element subsets X of $[a+b]$. Suppose X intersects $[a]$ in i elements. There are $\binom{a}{i}$ choices for $X \cap [a]$, and $\binom{b}{n-i}$ choices for the remaining $n-i$ elements $X \cap \{a+1, a+2, \dots, a+b\}$. Thus, there are $\binom{a}{i} \binom{b}{n-i}$ ways that $X \cap [a]$ can have i elements, and summing over i gives the total number $\binom{a+b}{n}$ of n -element subsets of $[a+b]$.

There are many examples in the literature of finite sets that are known to have the same number of elements but for which no combinatorial proof of this fact is known. Some of these will appear as exercises throughout this book.

1.2 Sets and Multisets

We have (finally!) completed our description of the solution of an enumerative problem, and we are now ready to delve into some actual problems. Let us begin with the basic problem of counting subsets of a set. Let $S = \{x_1, x_2, \dots, x_n\}$ be an

n -element set, or n -set for short. Let 2^S denote the set of all subsets of S , and let $\{0, 1\}^n = \{(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) : \varepsilon_i = 0 \text{ or } 1\}$. Since there are two possible values for each ε_i , we have $\#\{0, 1\}^n = 2^n$. Define a map $\theta : 2^S \rightarrow \{0, 1\}^n$ by $\theta(T) = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$, where

$$\varepsilon_i = \begin{cases} 1, & \text{if } x_i \in T \\ 0, & \text{if } x_i \notin T. \end{cases}$$

For example, if $n = 5$ and $T = \{x_2, x_4, x_5\}$, then $\theta(T) = (0, 1, 0, 1, 1)$. Most readers will realize that $\theta(T)$ is just the *characteristic vector* of T . It is easily seen that θ is a bijection, so that we have given a combinatorial proof that $\#2^S = 2^n$. Of course, there are many alternative proofs of this simple result, and many of these proofs could be regarded as combinatorial.

Now define $\binom{S}{k}$ (sometimes denoted $S^{(k)}$ or otherwise, and read “ S choose k ”) to be the set of all k -element subsets (or k -subsets) of S , and define $\binom{n}{k} = \#\binom{S}{k}$, read “ n choose k ” (ignore our previous use of the symbol $\binom{n}{k}$) and called a *binomial coefficient*. Our goal is to prove the formula

$$\binom{n}{k} = \frac{n(n-1) \cdots (n-k+1)}{k!}. \quad (1.16)$$

Note that if $0 \leq k \leq n$ then the right-hand side of equation (1.16) can be rewritten $n!/k!(n-k)!$. The right-hand side of (1.16) can be used to define $\binom{n}{k}$ for any complex number (or indeterminate) n , provided $k \in \mathbb{N}$. The numerator $n(n-1) \cdots (n-k+1)$ of (1.16) is read “ n lower factorial k ” and is denoted $(n)_k$. CAVEAT. Many mathematicians, especially those in the theory of special functions, use the notation $(n)_k = n(n+1) \cdots (n+k-1)$.

We would like to give a bijective proof of (1.16), but the factor $k!$ in the denominator makes it difficult to give a “simple” interpretation of the right-hand side. Therefore, we use the standard technique of clearing the denominator. To this end we count in two ways the number $N(n, k)$ of ways of choosing a k -subset T of S and then linearly ordering the elements of T . We can pick T in $\binom{n}{k}$ ways, then pick an element of T in k ways to be first in the ordering, then pick another element in $k-1$ ways to be second, and so on. Thus,

$$N(n, k) = \binom{n}{k} k!.$$

On the other hand, we could pick any element of S in n ways to be first in the ordering, then another element in $n-1$ ways to be second, on so on, down to any remaining element in $n-k+1$ ways to be k th. Thus,

$$N(n, k) = n(n-1) \cdots (n-k+1).$$

We have therefore given a combinatorial proof that

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

and hence of equation (1.16).

A generating function approach to binomial coefficients can be given as follows. Regard x_1, \dots, x_n as independent indeterminates. It is an immediate consequence of the process of multiplication (one could also give a rigorous proof by induction) that

$$(1 + x_1)(1 + x_2) \cdots (1 + x_n) = \sum_{T \subseteq S} \prod_{x_i \in T} x_i. \quad (1.17)$$

If we put each $x_i = x$, then we obtain

$$(1 + x)^n = \sum_{T \subseteq S} \prod_{x_i \in T} x = \sum_{T \subseteq S} x^{\#T} = \sum_{k \geq 0} \binom{n}{k} x^k, \quad (1.18)$$

since the term x^k appears exactly $\binom{n}{k}$ times in the sum $\sum_{T \subseteq S} x^{\#T}$. This reasoning is an instance of the simple but useful observation that if \mathcal{S} is a collection of finite sets such that \mathcal{S} contains exactly $f(n)$ sets with n elements, then

$$\sum_{S \in \mathcal{S}} x^{\#S} = \sum_{n \geq 0} f(n) x^n.$$

Somewhat more generally, if $g : \mathbb{N} \rightarrow \mathbb{C}$ is any function, then

$$\sum_{S \in \mathcal{S}} g(\#S) x^{\#S} = \sum_{n \geq 0} g(n) f(n) x^n.$$

Equation (1.18) is such a simple result (the binomial theorem for the exponent $n \in \mathbb{N}$) that it is hardly necessary to obtain first the more refined (1.17). However, it is often easier in dealing with generating functions to work with the most number of variables (indeterminates) possible and then specialize. Often the more refined formula will be more transparent, and its various specializations will be automatically unified.

Various identities involving binomial coefficients follow easily from the identity $(1 + x)^n = \sum_{k \geq 0} \binom{n}{k} x^k$, and the reader will find it instructive to find combinatorial proofs of them. (See Exercise 1.3 for further examples of binomial coefficient identities.) For instance, put $x = 1$ to obtain $2^n = \sum_{k \geq 0} \binom{n}{k}$; put $x = -1$ to obtain $0 = \sum_{k \geq 0} (-1)^k \binom{n}{k}$ if $n > 0$; differentiate and put $x = 1$ to obtain $n2^{n-1} = \sum_{k \geq 0} k \binom{n}{k}$, and so on.

There is a close connection between subsets of a set and compositions of a nonnegative integer. A *composition* of n can be thought of as an expression of n as an *ordered* sum of integers. More precisely, a composition of n is a sequence $\alpha = (a_1, \dots, a_k)$ of positive integers satisfying $\sum a_i = n$. For instance, there are eight compositions of 4; namely,

$$\begin{array}{ll} 1 + 1 + 1 + 1 & 3 + 1 \\ 2 + 1 + 1 & 1 + 3 \\ 1 + 2 + 1 & 2 + 2 \\ 1 + 1 + 2 & 4. \end{array}$$

If exactly k summands appear in a composition α , then we say that α has k parts, and we call α a k -composition. If $\alpha = (a_1, a_2, \dots, a_k)$ is a k -composition of n , then define a $(k-1)$ -subset S_α of $[n-1]$ by

$$S_\alpha = \{a_1, a_1 + a_2, \dots, a_1 + a_2 + \dots + a_{k-1}\}.$$

The correspondence $\alpha \mapsto S_\alpha$ gives a bijection between all k -compositions of n and $(k-1)$ -subsets of $[n-1]$. Hence, there are $\binom{n-1}{k-1}$ k -compositions of n and 2^{n-1} compositions of $n > 0$. The inverse bijection $S_\alpha \mapsto \alpha$ is often represented schematically by drawing n dots in a row and drawing vertical bars between $k-1$ of the $n-1$ spaces separating the dots. This procedure divides the dots into k linearly ordered (from left-to-right) “compartments” whose number of elements is a k -composition of n . For instance, the compartments

$$\cdot | \cdot \cdot | \cdot | \cdot | \cdot \cdot \cdot | \cdot \cdot \quad (1.19)$$

correspond to the 6-composition $(1, 2, 1, 1, 3, 2)$ of 10. The diagram (1.19) illustrates another very general principle related to bijective proofs—it is often efficacious to represent the objects being counted geometrically.

A problem closely related to compositions is that of counting the number $N(n, k)$ of solutions to $x_1 + x_2 + \dots + x_k = n$ in *nonnegative* integers. Such a solution is called a *weak composition* of n into k parts, or a *weak k -composition* of n . (A solution in *positive* integers is simply a k -composition of n .) If we put $y_i = x_i + 1$, then $N(n, k)$ is the number of solutions in positive integers to $y_1 + y_2 + \dots + y_k = n + k$, that is, the number of k -compositions of $n + k$. Hence, $N(n, k) = \binom{n+k-1}{k-1}$. A further variant is the enumeration of \mathbb{N} -solutions (that is, solutions where each variable lies in \mathbb{N}) to $x_1 + x_2 + \dots + x_k \leq n$. Again we use a standard technique, namely, introducing a *slack variable* y to convert the inequality $x_1 + x_2 + \dots + x_k \leq n$ to the equality $x_1 + x_2 + \dots + x_k + y = n$. An \mathbb{N} -solution to this equation is a weak $(k+1)$ -composition of n , so the number $N(n, k+1)$ of such solutions is $\binom{n+(k+1)-1}{k} = \binom{n+k}{k}$.

A k -subset T of an n -set S is sometimes called a *k -combination of S without repetitions*. This suggests the problem of counting the number of k -combinations of S *with repetitions*; that is, we choose k elements of S , disregarding order and allowing repeated elements. Denote this number by $\left(\binom{n}{k}\right)$, which could be read “ n multichoose k .” For instance, if $S = \{1, 2, 3\}$, then the combinations counted by $\left(\binom{3}{2}\right)$ are 11, 22, 33, 12, 13, 23. Hence, $\left(\binom{3}{2}\right) = 6$. An equivalent but more precise treatment of combinations with repetitions can be made by introducing the concept of a *multiset*. Intuitively, a multiset is a set with repeated elements; for instance, $\{1, 1, 2, 5, 5, 5\}$. More precisely, a *finite multiset* M on a set S is a pair (S, ν) , where ν is a function $\nu : S \rightarrow \mathbb{N}$ such that $\sum_{x \in S} \nu(x) < \infty$. One regards $\nu(x)$ as the number of repetitions of x . The integer $\sum_{x \in S} \nu(x)$ is called the *cardinality*, *size*, or *number of elements* of M and is denoted $|M|$, $\#M$, or $\text{card } M$. If $S = \{x_1, \dots, x_n\}$ and $\nu(x_i) = a_i$, then we call a_i the *multiplicity* of x_i in M and

write $M = \{x_1^{a_1}, \dots, x_n^{a_n}\}$. If $\#M = k$, then we call M a k -multiset. The set of all k -multisets on S is denoted $\left(\left(\begin{smallmatrix} S \\ k \end{smallmatrix}\right)\right)$. If $M' = (S, v')$ is another multiset on S , then we say that M' is a *submultiset* of M if $v'(x) \leq v(x)$ for all $x \in S$. The number of submultisets of M is $\prod_{x \in S} (v(x) + 1)$, since for each $x \in S$ there are $v(x) + 1$ possible values of $v'(x)$. It is now clear that a k -combination of S with repetition is simply a multiset on S with k elements.

Although the reader may be unaware of it, we have already evaluated the number $\left(\left(\begin{smallmatrix} n \\ k \end{smallmatrix}\right)\right)$. If $S = \{y_1, \dots, y_n\}$ and we set $x_i = v(y_i)$, then we see that $\left(\left(\begin{smallmatrix} n \\ k \end{smallmatrix}\right)\right)$ is the number of solutions in nonnegative integers to $x_1 + x_2 + \dots + x_n = k$, which we have seen is $\binom{n+k-1}{n-1} = \binom{n+k-1}{k}$.

There are two elegant direct combinatorial proofs that $\left(\left(\begin{smallmatrix} n \\ k \end{smallmatrix}\right)\right) = \binom{n+k-1}{k}$. For the first, let $1 \leq a_1 < a_2 < \dots < a_k \leq n+k-1$ be a k -subset of $[n+k-1]$. Let $b_i = a_i - i + 1$. Then, $\{b_1, b_2, \dots, b_k\}$ is a k -multiset on $[n]$. Conversely, given a k -multiset $1 \leq b_1 \leq b_2 \leq \dots \leq b_k \leq n$ on $[n]$, then defining $a_i = b_i + i - 1$ we see that $\{a_1, a_2, \dots, a_k\}$ is a k -subset of $[n+k-1]$. Hence, we have defined a bijection between $\left(\left(\begin{smallmatrix} [n] \\ k \end{smallmatrix}\right)\right)$ and $\binom{[n+k-1]}{k}$, as desired. This proof illustrates the technique of *compression*, where we convert a strictly increasing sequence to a weakly increasing sequence.

Our second direct proof that $\left(\left(\begin{smallmatrix} n \\ k \end{smallmatrix}\right)\right) = \binom{n+k-1}{k}$ is a “geometric” (or “balls into boxes” or “stars and bars”) proof, analogous to the preceding proof that there are $\binom{n-1}{k-1}$ k -compositions of n . There are $\binom{n+k-1}{k}$ sequences consisting of k dots and $n-1$ vertical bars. An example of such a sequence for $k=5$ and $n=7$ is given by

$$|| \cdot \cdot | \cdot ||| \cdot \cdot$$

The $n-1$ bars divide the k dots into n compartments. Let the number of dots in the i th compartment be $v(i)$. In this way the diagrams correspond to k -multisets on $[n]$, so $\left(\left(\begin{smallmatrix} n \\ k \end{smallmatrix}\right)\right) = \binom{n+k-1}{k}$. For the preceding example, the multiset is $\{3, 3, 4, 7, 7\}$.

The generating function approach to multisets is instructive. In exact analogy to our treatment of subsets of a set $S = \{x_1, \dots, x_n\}$, we have

$$(1 + x_1 + x_1^2 + \dots)(1 + x_2 + x_2^2 + \dots) \cdots (1 + x_n + x_n^2 + \dots) = \sum_{M=(S,v)} \prod_{x_i \in S} x_i^{v(x_i)},$$

where the sum is over all finite multisets M on S . Put each $x_i = x$. We get

$$\begin{aligned} (1 + x + x^2 + \dots)^n &= \sum_{M=(S,v)} x^{v(x_1) + \dots + v(x_n)} \\ &= \sum_{M=(S,v)} x^{\#M} \\ &= \sum_{k \geq 0} \left(\left(\begin{smallmatrix} n \\ k \end{smallmatrix}\right)\right) x^k. \end{aligned}$$

But

$$(1 + x + x^2 + \cdots)^n = (1 - x)^{-n} = \sum_{k \geq 0} \binom{-n}{k} (-1)^k x^k, \quad (1.20)$$

so $\left(\binom{n}{k}\right) = (-1)^k \binom{-n}{k} = \binom{n+k-1}{k}$. The elegant formula

$$\left(\binom{n}{k}\right) = (-1)^k \binom{-n}{k} \quad (1.21)$$

is no accident; it is the simplest instance of a *combinatorial reciprocity theorem*. A partially ordered set generalization appears in Section 3.15.3, while a more general theory of such results is given in Chapter 4.

The binomial coefficient $\binom{n}{k}$ may be interpreted in the following manner. Each element of an n -set S is placed into one of two categories, with k elements in Category 1 and $n - k$ elements in Category 2. (The elements of Category 1 form a k -subset T of S .) This suggests a generalization allowing more than two categories. Let (a_1, a_2, \dots, a_m) be a sequence of nonnegative integers summing to n , and suppose that we have m categories C_1, \dots, C_m . Let $\binom{n}{a_1, a_2, \dots, a_m}$ denote the number of ways of assigning each element of an n -set S to one of the categories C_1, \dots, C_m so that exactly a_i elements are assigned to C_i . The notation is somewhat at variance with the notation for binomial coefficients (the case $m = 2$), but no confusion should result when we write $\binom{n}{k}$ instead of $\binom{n}{k, n-k}$. The number $\binom{n}{a_1, a_2, \dots, a_m}$ is called a *multinomial coefficient*. It is customary to regard the elements of S as being n distinguishable balls and the categories as being m distinguishable boxes. Then $\binom{n}{a_1, a_2, \dots, a_m}$ is the number of ways to place the balls into the boxes so that the i th box contains a_i balls.

The multinomial coefficient can also be interpreted in terms of “permutations of a multiset.” If S is an n -set, then a *permutation* w of S can be defined as a linear ordering w_1, w_2, \dots, w_n of the elements of S . Think of w as a *word* $w_1 w_2 \cdots w_n$ in the alphabet S . If $S = \{x_1, \dots, x_n\}$, then such a word corresponds to the bijection $w : S \rightarrow S$ given by $w(x_i) = w_i$, so that a permutation of S may also be regarded as a bijection $S \rightarrow S$. Many interesting combinatorics are based on these two different ways of representing permutations; a good example is the second proof of Proposition 5.3.2.

We write \mathfrak{S}_S for the set of permutations of S . If $S = [n]$, then we write \mathfrak{S}_n for $\mathfrak{S}_{[n]}$. Since we choose w_1 in n ways, then w_2 in $n - 1$ ways, and so on, we clearly have $\#\mathfrak{S}_S = n!$. In an analogous manner, we can define a permutation w of a multiset M of cardinality n to be a linear ordering w_1, w_2, \dots, w_n of the “elements” of M ; that is, if $M = (S, \nu)$ then the element $x \in S$ appears exactly $\nu(x)$ times in the permutation. Again, we think of w as a word $w_1 w_2 \cdots w_n$. For instance, there are 12 permutations of the multiset $\{1, 1, 2, 3\}$; namely, 1123, 1132, 1213, 1312, 1231, 1321, 2113, 2131, 2311, 3112, 3121, 3211. Let \mathfrak{S}_M denote the set of permutations of M . If $M = \{x_1^{a_1}, \dots, x_m^{a_m}\}$ and $\#M = n$, then it is

clear that

$$\#\mathfrak{S}_M = \binom{n}{a_1, a_2, \dots, a_m}. \quad (1.22)$$

Indeed, if x_i appears in position j of the permutation, then we put the element j of $[n]$ into Category i .

Our results on binomial coefficients extend straightforwardly to multinomial coefficients. In particular, we have

$$\binom{n}{a_1, a_2, \dots, a_m} = \frac{n!}{a_1! a_2! \cdots a_m!}. \quad (1.23)$$

Among the many ways to prove this result, we can place a_1 elements of S into Category 1 in $\binom{n}{a_1}$ ways, then a_2 of the remaining $n - a_1$ elements of $[n]$ into Category 2 in $\binom{n-a_1}{a_2}$ ways, and so on, yielding

$$\begin{aligned} \binom{n}{a_1, a_2, \dots, a_m} &= \binom{n}{a_1} \binom{n-a_1}{a_2} \cdots \binom{n-a_1-\cdots-a_{m-1}}{a_m} \\ &= \frac{n!}{a_1! a_2! \cdots a_m!}. \end{aligned} \quad (1.24)$$

Equation (1.24) is often a useful device for reducing problems on multinomial coefficients to binomial coefficients. We leave to the reader the (easy) multinomial analogue (known as the *multinomial theorem*) of equation (1.18), namely,

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{a_1 + \cdots + a_m = n} \binom{n}{a_1, a_2, \dots, a_m} x_1^{a_1} \cdots x_m^{a_m},$$

where the sum ranges over all $(a_1, \dots, a_m) \in \mathbb{N}^m$ satisfying $a_1 + \cdots + a_m = n$. Note that $\binom{n}{1, 1, \dots, 1} = n!$, the number of permutations of an n -element set.

Binomials and multinomial coefficients have an important geometric interpretation in terms of lattice paths. Let S be a subset of \mathbb{Z}^d . More generally, we could replace \mathbb{Z}^d by any lattice (discrete subgroup of full rank) in \mathbb{R}^d , but for simplicity we consider only \mathbb{Z}^d . A *lattice path* L in \mathbb{Z}^d of length k with steps in S is a sequence $v_0, v_1, \dots, v_k \in \mathbb{Z}^d$ such that each consecutive difference $v_i - v_{i-1}$ lies in S . We say that L *starts at* v_0 and *ends at* v_k , or more simply that L *goes from* v_0 *to* v_k . Figure 1.1 shows the six lattice paths in \mathbb{Z}^2 from $(0, 0)$ to $(2, 2)$ with steps $(1, 0)$ and $(0, 1)$.

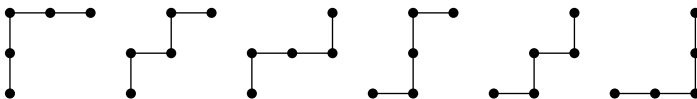


Figure 1.1 Six lattice paths.

1.2.1 Proposition. Let $v = (a_1, \dots, a_d) \in \mathbb{N}^d$, and let e_i denote the i th unit coordinate vector in \mathbb{Z}^d . The number of lattice paths in \mathbb{Z}^d from the origin $(0, 0, \dots, 0)$ to v with steps e_1, \dots, e_d is given by the multinomial coefficient $\binom{a_1 + \dots + a_d}{a_1, \dots, a_d}$.

Proof. Let v_0, v_1, \dots, v_k be a lattice path being counted. Then the sequence $v_1 - v_0, v_2 - v_1, \dots, v_k - v_{k-1}$ is simply a sequence consisting of a_i e_i 's in some order. The proof follows from equation (1.22). \square

Proposition 1.2.1 is the most basic result in the vast subject of *lattice path enumeration*. Further results in this area will appear throughout this book.

1.3 Cycles and Inversions

Permutations of sets and multisets are among the richest objects in enumerative combinatorics. A basic reason for this fact is the wide variety of ways to *represent* a permutation combinatorially. We have already seen that we can represent a set permutation either as a *word* or a *function*. In fact, for any set S , the function $w : [n] \rightarrow S$ given by $w(i) = w_i$ corresponds to the word $w_1 w_2 \dots w_n$. Several additional representations will arise in Section 1.5. Many of the basic results derived here will play an important role in later analysis of more complicated objects related to permutations.

A second reason for the richness of the theory of permutations is the wide variety of interesting “statistics” of permutations. In the broadest sense, a statistic on some class \mathcal{C} of combinatorial objects is just a function $f : \mathcal{C} \rightarrow S$, where S is any set (often taken to be \mathbb{N}). We want $f(x)$ to capture some combinatorially interesting feature of x . For instance, if x is a (finite) set, then $f(x)$ could be its number of elements. We can think of f as *refining* the enumeration of objects in \mathcal{C} . For instance, if \mathcal{C} consists of all subsets of an n -set S and $f(x) = \#x$, then f refines the number 2^n of subsets of S into a sum $2^n = \sum_k \binom{n}{k}$, where $\binom{n}{k}$ is the number of subsets of S with k elements. In this section and the next two, we will discuss a number of different statistics on permutations.

Cycle Structure

If we regard a set permutation w as a bijection $w : S \rightarrow S$, then it is natural to consider for each $x \in S$ the sequence $x, w(x), w^2(x), \dots$. Eventually (since w is a bijection and S is assumed finite) we must return to x . Thus for some unique $\ell \geq 1$, we have that $w^\ell(x) = x$ and that the elements $x, w(x), \dots, w^{\ell-1}(x)$ are distinct. We call the sequence $(x, w(x), \dots, w^{\ell-1}(x))$ a *cycle* of w of length ℓ . The cycles $(x, w(x), \dots, w^{\ell-1}(x))$ and $(w^i(x), w^{i+1}(x), \dots, w^{\ell-1}(x), x, \dots, w^{i-1}(x))$ are considered the same. Every element of S then appears in a unique cycle of w , and we may regard w as a disjoint union or *product* of its distinct cycles C_1, \dots, C_k , written $w = C_1 \dots C_k$. For instance, if $w : [7] \rightarrow [7]$ is defined by $w(1) = 4$, $w(2) = 2$, $w(3) = 7$, $w(4) = 1$, $w(5) = 3$, $w(6) = 6$, $w(7) = 5$ (or $w = 4271365$ as a word), then $w = (14)(2)(375)(6)$. Of course this representation of w in disjoint cycle notation is not unique; we also have for instance $w = (753)(14)(6)(2)$.

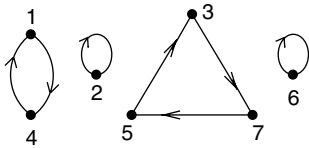


Figure 1.2 The digraph of the permutation $(14)(2)(375)(6)$.

A geometric or graphical representation of a permutation w is often useful. A finite *directed graph* or *digraph* D is a triple (V, E, ϕ) , where $V = \{x_1, \dots, x_n\}$ is a set of *vertices*, E is a finite set of (directed) *edges* or *arcs*, and ϕ is a map from E to $V \times V$. If ϕ is injective, then we call D a *simple* digraph, and we can think of E as a subset of $V \times V$. If e is an edge with $\phi(e) = (x, y)$, then we represent e as an arrow directed from x to y . If w is permutation of the set S , then define the *digraph* D_w of w to be the directed graph with vertex set S and edge set $\{(x, y) : w(x) = y\}$. In other words, for every vertex x , there is an edge from x to $w(x)$. Digraphs of permutations are characterized by the property that every vertex has one edge pointing out and one pointing in. The disjoint cycle decomposition of a permutation of a finite set guarantees that D_w will be a disjoint union of directed cycles. For instance, Figure 1.2 shows the digraph of the permutation $w = (14)(2)(375)(6)$.

We noted earlier that the disjoint cycle notation of a permutation is not unique. We can define a *standard representation* by requiring that (a) each cycle is written with its largest element first, and (b) the cycles are written in increasing order of their largest element. Thus, the standard form of the permutation $w = (14)(2)(375)(6)$ is $(2)(41)(6)(753)$. Define \hat{w} to be the word (or permutation) obtained from w by writing it in standard form and erasing the parentheses. For example, with $w = (2)(41)(6)(753)$, we have $\hat{w} = 2416753$. Now observe that we can uniquely recover w from \hat{w} by inserting a left parenthesis in $\hat{w} = a_1 a_2 \dots a_n$ preceding every *left-to-right maximum* or *record* (also called *outstanding element*); that is, an element a_i such that $a_i > a_j$ for every $j < i$. Then insert a right parenthesis where appropriate; that is, before every internal left parenthesis and at the end. Thus, the map $w \mapsto \hat{w}$ is a *bijection* from \mathfrak{S}_n to itself, known as the *fundamental bijection*. Let us sum up this information as a proposition.

1.3.1 Proposition. *a. The map $\mathfrak{S}_n \xrightarrow{\hat{\cdot}} \mathfrak{S}_n$ defined above is a bijection.
b. If $w \in \mathfrak{S}_n$ has k cycles, then \hat{w} has k left-to-right maxima.*

If $w \in \mathfrak{S}_S$ where $\#S = n$, then let $c_i = c_i(w)$ be the number of cycles of w of length i . Note that $n = \sum i c_i$. Define the *type* of w , denoted $\text{type}(w)$, to be the sequence (c_1, \dots, c_n) . The total number of cycles of w is denoted $c(w)$, so $c(w) = c_1(w) + \dots + c_n(w)$.

1.3.2 Proposition. *The number of permutations $w \in \mathfrak{S}_S$ of type (c_1, \dots, c_n) is equal to $n! / 1^{c_1} c_1! 2^{c_2} c_2! \dots n^{c_n} c_n!$.*

Proof. Let $w = w_1 w_2 \cdots w_n$ be any permutation of S . Parenthesize the word w so that the first c_1 cycles have length 1, the next c_2 have length 2, and so on. For instance, if $(c_1, \dots, c_9) = (1, 2, 0, 1, 0, 0, 0, 0, 0)$ and $w = 427619583$, then we obtain $(4)(27)(61)(9583)$. In general, we obtain the disjoint cycle decomposition of a permutation w' of type (c_1, \dots, c_n) . Hence, we have defined a map $\Phi : \mathfrak{S}_S \rightarrow \mathfrak{S}_S^c$, where \mathfrak{S}_S^c is the set of all $u \in \mathfrak{S}_S$ of type $c = (c_1, \dots, c_n)$. Given $u \in \mathfrak{S}_S^c$, we claim that there are $1^{c_1} c_1! 2^{c_2} c_2! \cdots n^{c_n} c_n!$ ways to write it in disjoint cycle notation so that the cycle lengths are weakly increasing from left to right. Namely, order the cycles of length i in $c_i!$ ways, and choose the first elements of these cycles in i^{c_i} ways. These choices are all independent, so the claim is proved. Hence for each $u \in \mathfrak{S}_S^c$, we have $\#\Phi^{-1}(u) = 1^{c_1} c_1! 2^{c_2} c_2! \cdots n^{c_n} c_n!$, and the proof follows since $\#\mathfrak{S}_S = n!$. \square

NOTE. The proof of Proposition 1.3.2 can easily be converted into a bijective proof of the identity

$$n! = 1^{c_1} c_1! 2^{c_2} c_2! \cdots n^{c_n} c_n! (\#\mathfrak{S}_S^c),$$

analogous to our bijective proof of equation (1.16).

Proposition 1.3.2 has an elegant and useful formulation in terms of generating functions. Suppose that $w \in \mathfrak{S}_n$ has type (c_1, \dots, c_n) . Write

$$t^{\text{type}(w)} = t_1^{c_1} t_2^{c_2} \cdots t_n^{c_n},$$

and define the *cycle indicator* or *cycle index* of \mathfrak{S}_n to be the polynomial

$$Z_n = Z_n(t_1, \dots, t_n) = \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} t^{\text{type}(w)}. \quad (1.25)$$

(Set $Z_0 = 1$.) For instance,

$$Z_1 = t_1,$$

$$Z_2 = \frac{1}{2}(t_1^2 + t_2),$$

$$Z_3 = \frac{1}{6}(t_1^3 + 3t_1 t_2 + 2t_3),$$

$$Z_4 = \frac{1}{24}(t_1^4 + 6t_1^2 t_2 + 8t_1 t_3 + 3t_2^2 + 6t_4).$$

1.3.3 Theorem. *We have*

$$\sum_{n \geq 0} Z_n x^n = \exp \left(t_1 x + t_2 \frac{x^2}{2} + t_3 \frac{x^3}{3} + \cdots \right). \quad (1.26)$$

Proof. We give a naive computational proof. For a more conceptual proof, see Example 5.2.10. Let us expand the right-hand side of equation (1.26):

$$\begin{aligned}\exp\left(\sum_{i \geq 1} t_i \frac{x^i}{i}\right) &= \prod_{i \geq 1} \exp\left(t_i \frac{x^i}{i}\right) \\ &= \prod_{i \geq 1} \sum_{j \geq 0} t_i^j \frac{x^{ij}}{i^j j!}.\end{aligned}\quad (1.27)$$

Hence, the coefficient of $t_1^{c_1} \cdots t_n^{c_n} x^n$ is equal to 0 unless $\sum i c_i = n$, in which case it is equal to

$$\frac{1}{1^{c_1} c_1! 2^{c_2} c_2! \cdots} = \frac{1}{n!} \frac{n!}{1^{c_1} c_1! 2^{c_2} c_2! \cdots}.$$

Comparing with Proposition 1.3.2 completes the proof. \square

Let us give two simple examples of the use of Theorem 1.3.3. For some additional examples, see Exercises 5.10 and 5.11. A more general theory of cycle indicators based on symmetric functions is given in Section 7.24. Write $F(t; x) = F(t_1, t_2, \dots; x)$ for the right-hand side of equation (1.26).

1.3.4 Example. Let $e_6(n)$ be the number of permutations $w \in \mathfrak{S}_n$ satisfying $w^6 = 1$. A permutation w satisfies $w^6 = 1$ if and only if all its cycles have length 1, 2, 3 or 6. Hence,

$$e_6(n) = n! Z_n(t_i = 1 \text{ if } i|6, t_i = 0 \text{ otherwise}).$$

There follows

$$\begin{aligned}\sum_{n \geq 0} e_6(n) \frac{x^n}{n!} &= F(t_i = 1 \text{ if } i|6, t_i = 0 \text{ otherwise}) \\ &= \exp\left(x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^6}{6}\right).\end{aligned}$$

For the obvious generalization to permutations w satisfying $w^r = 1$, see equation (5.31).

1.3.5 Example. Let $E_k(n)$ denote the expected number of k -cycles in a permutation $w \in \mathfrak{S}_n$. It is understood that the expectation is taken with respect to the uniform distribution on \mathfrak{S}_n , so

$$E_k(n) = \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} c_k(w),$$

where $c_k(w)$ denotes the number of k -cycles in w . Now note that from the definition (1.25) of Z_n we have

$$E_k(n) = \frac{\partial}{\partial t_k} Z_n(t_1, \dots, t_n) |_{t_i=1}.$$

Hence,

$$\begin{aligned}
 \sum_{n \geq 0} E_k(n) x^n &= \frac{\partial}{\partial t_k} \exp \left(t_1 x + t_2 \frac{x^2}{2} + t_3 \frac{x^3}{3} + \cdots \right) \Big|_{t_i=1} \\
 &= \frac{x^k}{k} \exp \left(x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots \right) \\
 &= \frac{x^k}{k} \exp \log(1-x)^{-1} \\
 &= \frac{x^k}{k} \frac{1}{1-x} \\
 &= \frac{x^k}{k} \sum_{n \geq 0} x^n.
 \end{aligned}$$

It follows that $E_k(n) = 1/k$ for $n \geq k$. Can the reader think of a simple explanation (Exercise 1.120)?

Now define $c(n, k)$ to be the number of permutations $w \in \mathfrak{S}_n$ with exactly k cycles. The number $s(n, k) := (-1)^{n-k} c(n, k)$ is known as a *Stirling number of the first kind*, and $c(n, k)$ is called a *signless Stirling number of the first kind*.

1.3.6 Lemma. *The numbers $c(n, k)$ satisfy the recurrence*

$$c(n, k) = (n-1)c(n-1, k) + c(n-1, k-1), \quad n, k \geq 1,$$

with the initial conditions $c(n, k) = 0$ if $n < k$ or $k = 0$, except $c(0, 0) = 1$.

Proof. Choose a permutation $w \in \mathfrak{S}_{n-1}$ with k cycles. We can insert the symbol n after any of the numbers $1, 2, \dots, n-1$ in the disjoint cycle decomposition of w in $n-1$ ways, yielding the disjoint cycle decomposition of a permutation $w' \in \mathfrak{S}_n$ with k cycles for which n appears in a cycle of length at least 2. Hence, there are $(n-1)c(n-1, k)$ permutations $w' \in \mathfrak{S}_n$ with k cycles for which $w'(n) \neq n$.

On the other hand, if we choose a permutation $w \in \mathfrak{S}_{n-1}$ with $k-1$ cycles, we can extend it to a permutation $w' \in \mathfrak{S}_n$ with k cycles satisfying $w'(n) = n$ by defining

$$w'(i) = \begin{cases} w(i), & \text{if } i \in [n-1] \\ n, & \text{if } i = n. \end{cases}$$

Thus there are $c(n-1, k-1)$ permutations $w' \in \mathfrak{S}_n$ with k cycles for which $w'(n) = n$, and the proof follows. \square

Most of the elementary properties of the numbers $c(n, k)$ can be established using Lemma 1.3.6 together with mathematical induction. However, combinatorial proofs are to be preferred whenever possible. An illuminating illustration of the various techniques available to prove elementary combinatorial identities is provided by the next result.

1.3.7 Proposition. Let t be an indeterminate and fix $n \geq 0$. Then

$$\sum_{k=0}^n c(n, k) t^k = t(t+1)(t+2) \cdots (t+n-1). \quad (1.28)$$

First Proof. This proof may be regarded as “semi-combinatorial” since it is based directly on Lemma 1.3.6, which had a combinatorial proof. Let

$$F_n(t) = t(t+1) \cdots (t+n-1) = \sum_{k=0}^n b(n, k) t^k.$$

Clearly $b(n, k) = 0$ if $n = 0$ or $k = 0$, except $b(0, 0) = 1$ (an empty product is equal to 1). Moreover, since

$$\begin{aligned} F_n(t) &= (t+n-1)F_{n-1}(t) \\ &= \sum_{k=1}^n b(n-1, k-1) t^k + (n-1) \sum_{k=0}^{n-1} b(n-1, k) t^k, \end{aligned}$$

there follows $b(n, k) = (n-1)b(n-1, k) + b(n-1, k-1)$. Hence $b(n, k)$ satisfies the same recurrence and initial conditions as $c(n, k)$, so they agree. \square

Second Proof. Our next proof is a straightforward argument using generating functions. In terms of the cycle indicator Z_n , we have

$$\sum_{k=0}^n c(n, k) t^k = n! Z_n(t, t, \dots).$$

Hence substituting $t_i = t$ in equation (1.26) gives

$$\begin{aligned} \sum_{n \geq 0} \sum_{k=0}^n c(n, k) t^k \frac{x^n}{n!} &= \exp t \left(x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots \right) \\ &= \exp t (\log(1-x)^{-1}) \\ &= (1-x)^{-t} \\ &= \sum_{n \geq 0} (-1)^n \binom{-t}{n} x^n \\ &= \sum_{n \geq 0} t(t+1) \cdots (t+n-1) \frac{x^n}{n!}, \end{aligned}$$

and the proof follows from taking coefficient of $x^n/n!$. \square

Third Proof. The coefficient of t^k in $F_n(t)$ is

$$\sum_{1 \leq a_1 < a_2 < \cdots < a_{n-k} \leq n-1} a_1 a_2 \cdots a_{n-k}, \quad (1.29)$$

where the sum is over all $\binom{n-1}{n-k}$ $(n-k)$ -subsets $\{a_1, \dots, a_{n-k}\}$ of $[n-1]$. (Though irrelevant here, it is interesting to note that this sum is just the $(n-k)$ th elementary symmetric function of $1, 2, \dots, n-1$.) Clearly (1.29) counts the number of pairs (S, f) , where $S \in \binom{[n-1]}{n-k}$ and $f: S \rightarrow [n-1]$ satisfies $f(i) \leq i$. Thus, we seek a bijection $\phi: \Omega \rightarrow \mathfrak{S}_{n,k}$ between the set Ω of all such pairs (S, f) , and the set $\mathfrak{S}_{n,k}$ of $w \in \mathfrak{S}_n$ with k cycles.

Given $(S, f) \in \Omega$ where $S = \{a_1, \dots, a_{n-k}\} \subset [n-1]$, define $T = \{j \in [n] : n-j \notin S\}$. Let the elements of $[n] - T$ be $b_1 > b_2 > \dots > b_{n-k}$. Define $w = \phi(S, f)$ to be that permutation that when written in standard form satisfies: (i) the first (= greatest) elements of the cycles of w are the elements of T , and (ii) for $i \in [n-k]$ the number of elements of w preceding b_i and larger than b_i is $f(a_i)$. We leave it to the reader to verify that this construction yields the desired bijection. \square

1.3.8 Example. Suppose that in the preceding proof $n=9, k=4, S=\{1, 3, 4, 6, 8\}$, $f(1)=1, f(3)=2, f(4)=1, f(6)=3, f(8)=6$. Then $T=\{2, 4, 7, 9\}$, $[9]-T=\{1, 3, 5, 6, 8\}$, and $w=(2)(4)(753)(9168)$.

Fourth Proof of Proposition 1.3.7. There are two basic ways of giving a combinatorial proof that two polynomials are equal: (i) showing that their coefficients are equal and (ii) showing that they agree for sufficiently many values of their variable(s). We have already established Proposition 1.3.7 by the first technique; here we apply the second. If two polynomials in a single variable t (over the complex numbers, say) agree for all $t \in \mathbb{P}$, then they agree as polynomials. Thus, it suffices to establish (1.28) for all $t \in \mathbb{P}$.

Let $t \in \mathbb{P}$, and let $C(w)$ denote the set of cycles of $w \in \mathfrak{S}_n$. The left-hand side of (1.28) counts all pairs (w, f) , where $w \in \mathfrak{S}_n$ and $f: C(w) \rightarrow [t]$. The right-hand side counts integer sequences (a_1, a_2, \dots, a_n) where $0 \leq a_i \leq t+n-i-1$. (There are historical reasons for this restriction of a_i , rather than, say, $1 \leq a_i \leq t+i-1$.) Given such a sequence (a_1, a_2, \dots, a_n) , the following simple algorithm may be used to define (w, f) . First, write down the number n and regard it as starting a cycle C_1 of w . Let $f(C_1) = a_n + 1$. Assuming $n, n-1, \dots, n-i+1$ have been inserted into the disjoint cycle notation for w , we now have two possibilities:

- i. $0 \leq a_{n-i} \leq t-1$. Then start a new cycle C_j with the element $n-i$ to the left of the previously inserted elements, and set $f(C_j) = a_{n-i} + 1$.
- ii. $a_{n-i} = t+k$ where $0 \leq k \leq i-1$. Then insert $n-i$ into an old cycle so that it is not the leftmost element of any cycle, and so that it appears to the right of $k+1$ of the numbers previously inserted.

This procedure establishes the desired bijection. \square

1.3.9 Example. Suppose $n = 9$, $t = 4$, and $(a_1, \dots, a_9) = (4, 8, 5, 0, 7, 5, 2, 4, 1)$. Then w is built up as follows:

(9)
 (98)
 (7)(98)
 (7)(968)
 (7)(9685)
 (4)(7)(9685)
 (4)(73)(9685)
 (4)(73)(96285)
 (41)(73)(96285).

Moreover, $f(96285) = 2$, $f(73) = 3$, $f(41) = 1$.

Note that if we set $t = 1$ in the preceding proof, we obtain a combinatorial proof of the following result.

1.3.10 Proposition. Let $n, k \in \mathbb{P}$. The number of integer sequences (a_1, \dots, a_n) such that $0 \leq a_i \leq n - i$ and exactly k values of a_i equal 0 is $c(n, k)$.

Note that because of Proposition 1.3.1, we obtain “for free” the enumeration of permutations by left-to-right maxima.

1.3.11 Corollary. The number of $w \in \mathfrak{S}_n$ with k left-to-right maxima is $c(n, k)$.

Corollary 1.3.11 illustrates one benefit of having different ways of representing the same object (here a permutation) – different enumerative problems involving the object turn out to be equivalent.

Inversions

The fourth proof of Proposition 1.3.7 (in the case $t = 1$) associated a permutation $w \in \mathfrak{S}_n$ with an integer sequence (a_1, \dots, a_n) , $0 \leq a_i \leq n - i$. There is a different method for accomplishing this which is perhaps more natural. Given such a vector (a_1, \dots, a_n) , assume that $n, n - 1, \dots, n - i + 1$ have been inserted into w , expressed this time as a *word* (rather than a product of cycles). Then insert $n - i$ so that it has a_{n-i} elements to its left. For example, if $(a_1, \dots, a_9) = (1, 5, 2, 0, 4, 2, 0, 1, 0)$, then w is built up as follows:

9
 98
 798
 7968
 79685
 479685
 4739685
 47396285
 417396285.

Clearly a_i is the number of entries j of w to the left of i satisfying $j > i$. A pair (w_i, w_j) is called an *inversion* of the permutation $w = w_1 w_2 \cdots w_n$ if $i < j$ and $w_i > w_j$. The sequence $I(w) = (a_1, \dots, a_n)$ is called the *inversion table* of w . The preceding algorithm for constructing w from its inversion table $I(w)$ establishes the following result.

1.3.12 Proposition. *Let*

$$\mathcal{T}_n = \{(a_1, \dots, a_n) : 0 \leq a_i \leq n - i\} = [0, n-1] \times [0, n-2] \times \cdots \times [0, 0].$$

The map $I : \mathfrak{S}_n \rightarrow \mathcal{T}_n$ that sends each permutation to its inversion table is a bijection.

Therefore, the inversion table $I(w)$ is yet another way to represent a permutation w . Let us also mention that the *code* of a permutation w is defined by $\text{code}(w) = I(w^{-1})$. Equivalently, if $w = w_1 \cdots w_n$ and $\text{code}(w) = (c_1, \dots, c_n)$, then c_i is equal to the number of elements w_j to the right of w_i (i.e., $i < j$) such that $w_i > w_j$. The question of whether to use $I(w)$ or $\text{code}(w)$ depends on the problem at hand and is clearly only a matter of convenience. Often it makes no difference which is used, such as in obtaining the next corollary.

1.3.13 Corollary. *Let $\text{inv}(w)$ denote the number of inversions of the permutation $w \in \mathfrak{S}_n$. Then*

$$\sum_{w \in \mathfrak{S}_n} q^{\text{inv}(w)} = (1+q)(1+q+q^2) \cdots (1+q+q^2+\cdots+q^{n-1}). \quad (1.30)$$

Proof. If $I(w) = (a_1, \dots, a_n)$ then $\text{inv}(w) = a_1 + \cdots + a_n$. Hence,

$$\begin{aligned} \sum_{w \in \mathfrak{S}_n} q^{\text{inv}(w)} &= \sum_{a_1=0}^{n-1} \sum_{a_2=0}^{n-2} \cdots \sum_{a_n=0}^0 q^{a_1+a_2+\cdots+a_n} \\ &= \left(\sum_{a_1=0}^{n-1} q^{a_1} \right) \left(\sum_{a_2=0}^{n-2} q^{a_2} \right) \cdots \left(\sum_{a_n=0}^0 q^{a_n} \right), \end{aligned}$$

as desired. \square

The polynomial $(1+q)(1+q+q^2) \cdots (1+q+\cdots+q^{n-1})$ is called “the q -analogue of $n!$ ” and is denoted $(n)!$. Moreover, we denote the polynomial $1+q+\cdots+q^{n-1} = (1-q^n)/(1-q)$ by (n) and call it “the q -analogue of n ,” so that

$$(n)! = (1)(2) \cdots (n).$$

In general, a q -analogue of a mathematical object is an object depending on the variable q that “reduces to” (an admittedly vague term) the original object when we set $q = 1$. To be a “satisfactory” q -analogue more is required, but there is no precise definition of what is meant by “satisfactory.” Certainly one desirable property is that the original object concerns finite sets, while the q -analogue can

be interpreted in terms of subspaces of finite-dimensional vector spaces over the finite field \mathbb{F}_q . For instance, $n!$ is the number of sequences $\emptyset = S_0 \subset S_1 \subset \cdots \subset S_n = [n]$ of subsets of $[n]$. (The symbol \subset denotes strict inclusion, so $\#S_i = i$.) Similarly if q is a prime power then $(n)!$ is the number of sequences $0 = V_0 \subset V_1 \subset \cdots \subset V_n = \mathbb{F}_q^n$ of subspaces of the n -dimensional vector space \mathbb{F}_q^n over \mathbb{F}_q (so $\dim V_i = i$). For this reason $(n)!$ is regarded as a satisfactory q -analogue of $n!$. We can also regard an i -dimensional vector space over \mathbb{F}_q as the q -analogue of an i -element set. Many more instances of q -analogues will appear throughout this book, especially in Section 1.10. The theory of binomial posets developed in Section 3.18 gives a partial explanation for the existence of certain classes of q -analogues including $(n)!$.

We conclude this section with a simple but important property of the statistic inv .

1.3.14 Proposition. *For any $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$, we have $\text{inv}(w) = \text{inv}(w^{-1})$.*

Proof. The pair (i, j) is an inversion of w if and only if (w_j, w_i) is an inversion of w^{-1} . \square

1.4 Descents

In addition to cycle type and inversion table, there is one other fundamental statistic associated with a permutation $w \in \mathfrak{S}_n$. If $w = w_1 w_2 \cdots w_n$ and $1 \leq i \leq n-1$, then i is a *descent* of w if $w_i > w_{i+1}$, while i is an *ascent* if $w_i < w_{i+1}$. (Sometimes it is desirable to also define n to be a descent, but we will adhere to the previous definition.) Define the *descent set* $D(w)$ of w by

$$D(w) = \{i : w_i > w_{i+1}\} \subseteq [n-1].$$

If $S \subseteq [n-1]$, then denote by $\alpha(S)$ (or $\alpha_n(S)$ if necessary) the number of permutations $w \in \mathfrak{S}_n$ whose descent set is contained in S , and by $\beta(S)$ (or $\beta_n(S)$) the number whose descent set is equal to S . In symbols,

$$\alpha(S) = \#\{w \in \mathfrak{S}_n : D(w) \subseteq S\}, \quad (1.31)$$

$$\beta(S) = \#\{w \in \mathfrak{S}_n : D(w) = S\}. \quad (1.32)$$

Clearly,

$$\alpha(S) = \sum_{T \subseteq S} \beta(T). \quad (1.33)$$

As explained in Example 2.2.4, we can invert this relationship to obtain

$$\beta(S) = \sum_{T \subseteq S} (-1)^{\#(S-T)} \alpha(T). \quad (1.34)$$

1.4.1 Proposition. *Let $S = \{s_1, \dots, s_k\}_< \subseteq [n-1]$. Then*

$$\alpha(S) = \binom{n}{s_1, s_2 - s_1, s_3 - s_2, \dots, n - s_k}. \quad (1.35)$$

Proof. To obtain a permutation $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$ satisfying $D(w) \subseteq S$, first choose $w_1 < w_2 < \cdots < w_{s_1}$ in $\binom{n}{s_1}$ ways. Then choose $w_{s_1+1} < w_{s_1+2} < \cdots < w_{s_2}$ in $\binom{n-s_1}{s_2-s_1}$ ways, and so on. We therefore obtain

$$\begin{aligned}\alpha(S) &= \binom{n}{s_1} \binom{n-s_1}{s_2-s_1} \binom{n-s_2}{s_3-s_2} \cdots \binom{n-s_k}{n-s_k} \\ &= \binom{n}{s_1, s_2-s_1, s_3-s_2, \dots, n-s_k},\end{aligned}$$

as desired. \square

1.4.2 Example. Let $n \geq 9$. Then

$$\begin{aligned}\beta_n(3, 8) &= \alpha_n(3, 8) - \alpha_n(3) - \alpha_n(8) + \alpha_n(\emptyset) \\ &= \binom{n}{3, 5, n-8} - \binom{n}{3} - \binom{n}{8} + 1.\end{aligned}$$

Two closely related descent sets are of special combinatorial interest. We say that a permutation $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$ (or more generally any sequence of distinct numbers) is *alternating* (or *zigzag* or *down-up*) if $w_1 > w_2 < w_3 > w_4 < \cdots$. Equivalently, $D(w) = \{1, 3, 5, \dots\} \cap [n-1]$. The alternating permutations in \mathfrak{S}_4 are 2143, 3142, 3241, 4132, 4231. Similarly, w is *reverse alternating* (or *up-down*) if $w_1 < w_2 > w_3 < w_4 > \cdots$. Equivalently, $D(w) = \{2, 4, 6, \dots\} \cap [n-1]$. The reverse alternating permutations in \mathfrak{S}_4 are 1324, 1423, 2314, 2413, 3412. The number of alternating permutations $w \in \mathfrak{S}_n$ is denoted E_n (with $E_0 = 1$) and is called an *Euler number*. (Originally, $(-1)^n E_{2n}$ was called an Euler number.) Since w is alternating if and only if $n+1-w_1, n+1-w_2, \dots, n+1-w_n$ is reverse alternating, it follows that E_n is also the number of reverse alternating permutations in \mathfrak{S}_n . We will develop some properties of alternating permutations and Euler numbers in various subsequent sections, especially Section 1.6.

NOTE. Some mathematicians define alternating permutations to be our reverse alternating permutations, while others define them to be permutations which are either alternating or reverse alternating according to our definition.

For the remainder of this section, we discuss some additional permutation statistics based on the descent set. The first of these is the *number of descents* of w , denoted $\text{des}(w)$. Thus, $\text{des}(w) = \#D(w)$. Let

$$\begin{aligned}A_d(x) &= \sum_{w \in \mathfrak{S}_d} x^{1+\text{des}(w)} \\ &= \sum_{k=1}^d A(d, k) x^k.\end{aligned}\tag{1.36}$$

Hence $A(d, k)$ is the number of permutations $w \in \mathfrak{S}_d$ with exactly $k - 1$ descents. The polynomial $A_d(x)$ is called an *Eulerian polynomial*, while $A(d, k)$ is an *Eulerian number*. We set $A(0, k) = \delta_{0k}$. The first few Eulerian polynomials are

$$\begin{aligned} A_0(x) &= 1 \\ A_1(x) &= x \\ A_2(x) &= x + x^2 \\ A_3(x) &= x + 4x^2 + x^3 \\ A_4(x) &= x + 11x^2 + 11x^3 + x^4 \\ A_5(x) &= x + 26x^2 + 66x^3 + 26x^4 + x^5 \\ A_6(x) &= x + 57x^2 + 302x^3 + 302x^4 + 57x^5 + x^6 \\ A_7(x) &= x + 120x^2 + 1191x^3 + 2416x^4 + 1191x^5 + 120x^6 + x^7 \\ A_8(x) &= x + 247x^2 + 4293x^3 + 15619x^4 + 15619x^5 + 4293x^6 \\ &\quad + 247x^7 + x^8. \end{aligned}$$

The bijection $w \mapsto \widehat{w}$ of Proposition 1.3.1 yields an interesting alternative description of the Eulerian numbers. Suppose that

$$w = (a_1, a_2, \dots, a_{i_1})(a_{i_1+1}, a_{i_1+2}, \dots, a_{i_2}) \cdots (a_{i_{k-1}+1}, a_{i_{k-1}+2}, \dots, a_d)$$

is a permutation in \mathfrak{S}_d written in standard form. Thus, $a_1, a_{i_1+1}, \dots, a_{i_{k-1}+1}$ are the largest elements of their cycles, and $a_1 < a_{i_1+1} < \cdots < a_{i_{k-1}+1}$. It follows that if $w(a_i) \neq a_{i+1}$, then $a_i < a_{i+1}$. Hence, $a_i < a_{i+1}$ or $i = d$ if and only if $w(a_i) \geq a_i$, so that

$$d - \text{des}(\widehat{w}) = \#\{i \in [d] : w(i) \geq i\}.$$

A number i for which $w(i) \geq i$ is called a *weak excedance* of w , while a number i for which $w(i) > i$ is called an *excedance* of w . One easily sees that a permutation $w = w_1 w_2 \cdots w_d$ has k weak excedances if and only if the permutation $u_1 u_2 \cdots u_d$ defined by $u_i = d + 1 - w_{d-i+1}$ has $d - k$ excedances. Moreover, w has $d - 1 - j$ descents if and only if $w_d w_{d-1} \cdots w_1$ has j descents. We therefore obtain the following result.

1.4.3 Proposition. *The number of permutations $w \in \mathfrak{S}_d$ with k excedances, as well as the number with $k + 1$ weak excedances, is equal to the Eulerian number $A(d, k + 1)$.*

The next result gives a fundamental property of Eulerian polynomials related to generating functions.

1.4.4 Proposition. *For every $d \geq 0$, we have*

$$\sum_{m \geq 0} m^d x^m = \frac{A_d(x)}{(1-x)^{d+1}}. \quad (1.37)$$

Proof. The proof is by induction on d . Since $\sum_{m \geq 0} x^m = 1/(1-x)$, the assertion is true for $d = 0$. Now assume that equation (1.37) holds for some $d \geq 0$. Differentiate with respect to x and multiply by x to obtain

$$\sum_{m \geq 0} m^{d+1} x^m = \frac{x(1-x)A'_d(x) + (d+1)x A_d(x)}{(1-x)^{d+2}}. \quad (1.38)$$

Hence, it suffices to show that

$$A_{d+1}(x) = x(1-x)A'_d(x) + (d+1)x A_d(x).$$

Taking coefficients of x^k on both sides and simplifying yields

$$A(d+1, k) = kA(d, k) + (d-k+2)A(d, k-1). \quad (1.39)$$

The left-hand side of equation (1.39) counts permutations in \mathfrak{S}_{d+1} with $k-1$ descents. We can obtain such a permutation uniquely in one of two ways. For the first way, choose a permutation $w = w_1 \cdots w_d \in \mathfrak{S}_d$ with $k-1$ descents, and insert $d+1$ after w_i if $i \in D(w)$, or insert $d+1$ at the end. There are k ways to insert $d+1$, so we obtain by this method $kA(d, k)$ permutations in \mathfrak{S}_{d+1} with $k-1$ descents. For the second way, choose $w = w_1 \cdots w_d \in \mathfrak{S}_d$ with $k-2$ descents, and insert $d+1$ after w_i if $i \notin D(w)$, or insert $d+1$ at the beginning. There are $d-k+2$ ways to insert $d+1$, so we obtain a further $(d-k+2)A(d, k-1)$ permutations in \mathfrak{S}_{d+1} with $k-1$ descents. We have verified that the recurrence (1.39) holds, so the proof follows by induction. \square

The appearance of the expression m^d in equation (1.37) suggests that there might be a more conceptual proof involving functions $f: [d] \rightarrow [m]$. We give such a proof at the end of this section.

We can also give a formula for the exponential generating function of the Eulerian polynomials themselves. For this purpose, define $A_0(x) = 1$.

1.4.5 Proposition. *We have*

$$\sum_{d \geq 0} A_d(x) \frac{t^d}{d!} = \frac{1-x}{1-xe^{(1-x)t}}. \quad (1.40)$$

Proof. Perhaps the simplest proof at this point is to multiply equation (1.37) by $t^d/d!$ and sum on $d \geq 0$. We get (using the convention $0^0 = 1$, which is often “correct” in enumerative combinatorics)

$$\begin{aligned} \sum_{d \geq 0} \frac{A_d(x)}{(1-x)^{d+1}} \frac{t^d}{d!} &= \sum_{d \geq 0} \sum_{m \geq 0} m^d x^m \frac{t^d}{d!} \\ &= \sum_{m \geq 0} x^m e^{mt} \\ &= \frac{1}{1-xe^t}. \end{aligned}$$

Now multiply both sides by $1 - x$ and substitute $(1 - x)t$ for t to complete the proof. (A more conceptual proof will be given in Section 3.19.) \square

A further interesting statistic associated with the descent set $D(w)$ is the *major index* (originally called the *greater index*), denoted $\text{maj}(w)$ (originally $\iota(w)$) and defined to be the sum of the elements of $D(w)$:

$$\text{maj}(w) = \sum_{i \in D(w)} i.$$

We next give a bijective proof of the remarkable result that inv and maj are *equidistributed*, that is, for any k ,

$$\#\{w \in \mathfrak{S}_n : \text{inv}(w) = k\} = \#\{w \in \mathfrak{S}_n : \text{maj}(w) = k\}. \quad (1.41)$$

Note that in terms of generating functions, equation (1.41) takes the form

$$\sum_{w \in \mathfrak{S}_n} q^{\text{inv}(w)} = \sum_{w \in \mathfrak{S}_n} q^{\text{maj}(w)}.$$

1.4.6 Proposition. *We have*

$$\sum_{w \in \mathfrak{S}_n} q^{\text{maj}(w)} = (n)!. \quad (1.42)$$

Proof. We will recursively define a bijection $\varphi : \mathfrak{S}_n \rightarrow \mathfrak{S}_n$ as follows. Let $w = w_1 \cdots w_n \in \mathfrak{S}_n$. We will define words (or sequences) $\gamma_1, \dots, \gamma_n$, where γ_k is a permutation of $\{w_1, \dots, w_k\}$.

First, let $\gamma_1 = w_1$. Assume that γ_k has been defined for some $1 \leq k < n$. If the last letter of γ_k (which turns out to be w_k) is greater (respectively, smaller) than w_{k+1} , then split γ_k after each letter greater (respectively, smaller) than w_{k+1} . These splits divide γ_k into compartments. Cyclically shift each compartment of γ_k one unit to the right, and place w_{k+1} at the end. Let γ_{k+1} be the word thus obtained. Set $\varphi(w) = \gamma_n$.

1.4.7 Example. Before analyzing the map φ , let us first give an example. Let $w = 683941725 \in \mathfrak{S}_9$. Then $\gamma_1 = 6$. It is irrelevant at this point whether $6 < w_2$ or $6 > w_2$ since there can be only one compartment, and $\gamma_2 = 68$. Now $8 > w_3 = 3$, so we split 68 after numbers greater than 3, getting $6|8$. Cyclically shifting the two compartments of length one leaves them unchanged, so $\gamma_3 = 683$. Now $3 < w_4 = 9$, so we split 683 after numbers less than 9. We get $6|8|3$ and $\gamma_4 = 6839$. Now $9 > w_5 = 4$, so we split 6839 after numbers greater than 4, giving $6|8|39$. The cyclic shift of 39 is 93, so $\gamma_5 = 68934$. Continuing in this manner gives the following

sequence of γ_i 's and compartments:

$$\begin{array}{cccccccc}
 6 & & & & & & & \\
 6 & | & 8 & & & & & \\
 6 & | & 8 & | & 3 & & & \\
 6 & | & 8 & | & 3 & & 9 & \\
 6 & | & 8 & | & 9 & | & 3 & | & 4 & & & \\
 6 & | & 8 & & 9 & & 3 & | & 4 & | & 1 & \\
 6 & | & 3 & | & 8 & | & 9 & | & 4 & | & 1 & 7 & \\
 6 & & 3 & | & 8 & & 9 & & 4 & | & 7 & 1 & | & 2 & \\
 3 & & 6 & & 4 & & 8 & & 9 & & 1 & 7 & & 2 & 5
 \end{array}$$

Hence, $\varphi(w) = 364891725$. Note that $\text{maj}(w) = \text{inv}(\varphi(w)) = 18$.

Returning to the proof of Proposition 1.4.6, we claim that φ is a bijection transforming maj to inv , that is,

$$\text{maj}(w) = \text{inv}(\varphi(w)). \quad (1.43)$$

We have defined inv and maj for permutations $w \in \mathfrak{S}_n$, but precisely the same definition can be made for *any* sequence $w = w_1 \cdots w_n$ of integers. Namely,

$$\begin{aligned}
 \text{inv}(w) &= \#\{(i, j) : i < j, w_i > w_j\}, \\
 \text{maj}(w) &= \sum_{i: w_i > w_{i+1}} i.
 \end{aligned}$$

Let $\eta_k = w_1 w_2 \cdots w_k$. We then prove by induction on k that $\text{inv}(\gamma_k) = \text{maj}(\eta_k)$, from which the proof follows by letting $k = n$.

Clearly $\text{inv}(\gamma_1) = \text{maj}(\eta_1) = 0$. Assume that $\text{inv}(\gamma_k) = \text{maj}(\eta_k)$ for some $k < n$. First, suppose that the last letter w_k of γ_k is greater than w_{k+1} . Thus, $k \in D(w)$, so we need to show that $\text{inv}(\gamma_{k+1}) = k + \text{inv}(\gamma_k)$. The last letter of any compartment C of γ_k is the largest letter of the compartment. Hence, when we cyclically shift this compartment, we create $\#C - 1$ new inversions. Each compartment contains exactly one letter larger than w_{k+1} , so when we append w_{k+1} to the end of γ_k , the number of new inversions $(i, k+1)$ is equal to the number m of compartments. Thus, altogether we have created

$$\sum_C (\#C - 1) + m = k$$

new inversions, as desired. The proof for the case $w_k < w_{k+1}$ is similar and will be omitted.

It remains to show that φ is a bijection. To do so we define φ^{-1} . Let $v = v_1 v_2 \cdots v_n \in \mathfrak{S}_n$. We want to find a (unique) $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$ so that $\varphi(w) = v$. Let $\delta_{n-1} = v_1 v_2 \cdots v_{n-1}$ and $w_n = v_n$. Now suppose that δ_k and $w_{k+1}, w_{k+2}, \dots, w_n$ have been defined for some $1 \leq k < n$. If the *first* letter of δ_k

is greater (respectively, smaller) than w_{k+1} , then split δ_k before each letter greater (respectively, smaller) than w_{k+1} . Then in each compartment of δ_k thus formed, cyclically shift the letters one unit to the *left*. Let the last letter of the word thus formed be w_k , and remove this last letter to obtain δ_{k-1} . It is easily verified that this procedure simply reverses the procedure used to obtain $v = \varphi(w)$ from w , completing the proof. \square

Proposition 1.4.6 establishes the equidistribution of inv and maj on \mathfrak{S}_n . Whenever we have two equidistributed statistics $f, g : S \rightarrow \mathbb{N}$ on a set S , we can ask whether a stronger result holds, namely, whether f and g have a *symmetric joint distribution*. This means that for all j, k we have

$$\#\{x \in S : f(x) = j, g(x) = k\} = \#\{x \in S : f(x) = k, g(x) = j\}. \quad (1.44)$$

This condition can be restated in terms of generating functions as

$$\sum_{x \in S} q^{f(x)} t^{g(x)} = \sum_{x \in S} q^{g(x)} t^{f(x)}.$$

The best way to prove (1.44) is to find a bijection $\psi : S \rightarrow S$ such that for all $x \in S$, we have $f(x) = g(\psi(x))$ and $g(x) = f(\psi(x))$. In other words, ψ interchanges the two statistics f and g .

Our next goal is to show that inv and maj have a symmetric joint distribution on \mathfrak{S}_n . We will not give an explicit bijection $\psi : \mathfrak{S}_n \rightarrow \mathfrak{S}_n$ interchanging inv and maj , but rather we will deduce it from a surprising property of the bijection φ defined in the proof of Proposition 1.4.6. To explain this property, define the *inverse descent set* $\text{ID}(w)$ of $w \in \mathfrak{S}_n$ by $\text{ID}(w) = D(w^{-1})$. Alternatively, we may think of $\text{ID}(w)$ as the “reading set” of w as follows. We read the numbers $1, 2, \dots, n$ in w from left-to-right in their standard order, going back to the beginning of w when necessary. For instance, if $w = 683941725$, then we first read 12, then 345, then 67, and finally 89. The cumulative number of elements in these reading sequences, excluding the last, form the reading set of w . It is easy to see that this reading set is just $\text{ID}(w)$. For instance, $\text{ID}(683941725) = \{2, 5, 7\}$.

We can easily extend the definition of $\text{ID}(w)$ to arbitrary sequences $w_1 w_2 \cdots w_n$ of distinct integers. (We can even drop the condition that the w_i 's are distinct, but we have no need here for such generality.) Simply regard $w = w_1 w_2 \cdots w_n$ as a permutation of its elements written in increasing order, that is, if $S = \{w_1, \dots, w_n\} = \{u_1, \dots, u_n\}_<$, then identify w with the permutation of S defined by $w(u_i) = w_i$. We can then write w^{-1} as a word in the same way as w and hence can define $\text{ID}(w)$ as the descent set of w^{-1} written as a word. For instance, if $w = 74285$, then $w^{-1} = 54827$ and $\text{ID}(w) = \{1, 3\}$. We can obtain the same result by reading w in the increasing order of its elements as before, obtaining reading sequences $u_1 u_2 \cdots u_{i_1}$, $u_{i_1+1} \cdots u_{i_2}$, \dots , $u_{i_{j-1}+1} \cdots u_{i_n}$, and then obtaining $\text{ID}(w) = \{i_1, i_2, \dots, i_j\}$ (the cumulative numbers of elements in the reading sequences). For instance, with $w = 74285$ the reading sequences are 2, 45, 78, giving $\text{ID}(w) = \{1, 3\}$ as before.

1.4.8 Theorem. Let φ be the bijection defined in the proof of Proposition 1.4.6. Then for all $w \in \mathfrak{S}_n$, $\text{ID}(w) = \text{ID}(\varphi(w))$. In other words, φ preserves the inverse descent set.

Proof. Preserve the notation of the proof of Proposition 1.4.6. We prove by induction on k that $\text{ID}(\gamma_k) = \text{ID}(\eta_k)$, from which the proof follows by setting $k = n$. Clearly, $\text{ID}(\gamma_1) = \text{ID}(\eta_1) = \emptyset$. Assume that $\text{ID}(\gamma_k) = \text{ID}(\eta_k)$ for some $k < n$. First, suppose that the last letter w_k of γ_k is greater than w_{k+1} , so that the last letter of any compartment C of γ_k is the unique letter in the compartment larger than w_{k+1} . Consider the reading of η_{k+1} . It will proceed just as for η_k until we encounter the largest letter of η_k less than w_{k+1} , in which case we next read w_{k+1} and then return to the beginning. Exactly the same is true for reading γ_{k+1} , so by the induction hypothesis, the reading sets of η_{k+1} and γ_{k+1} are the same up to this point. Let L be the set of remaining letters to be read. The letters in L are those greater than w_{k+1} . The reading words of these letters are the same for η_k and γ_k by the induction hypothesis. But the letters of L appear in the same order in η_k and η_{k+1} by definition of η_j . Moreover, they also appear in the same order in γ_k and γ_{k+1} since each such letter appears in exactly one compartment, so cyclic shifts (or indeed any permutations) within each compartment of γ_k does not change their order in γ_{k+1} . Hence, the reading words of the letters in L are the same for η_{k+1} and γ_{k+1} , so the proof follows for the case $w_k > w_{k+1}$. The case $w_k < w_{k+1}$ is similar and will be omitted. \square

Let $\text{imaj}(w) = \text{maj}(w^{-1}) = \sum_{i \in \text{ID}(w)} i$. As an immediate corollary to Theorem 1.4.8, we get the symmetric joint distribution of three pairs of permutations statistics including (inv, maj) , thereby improving Proposition 1.4.6. For further information about the bidistribution of $(\text{maj}, \text{imaj})$, see Exercise 4.47 and Corollary 7.23.9.

1.4.9 Corollary. The three pairs of statistics (inv, maj) , $(\text{inv}, \text{imaj})$, and $(\text{maj}, \text{imaj})$ have symmetric joint distributions.

Proof. Let f be any statistic on \mathfrak{S}_n , and define g by $g(w) = f(w^{-1})$. Clearly (f, g) have a symmetric joint distribution, of which $(\text{maj}, \text{imaj})$ is a special case. By Theorem 1.4.8, φ transforms maj to inv while preserving imaj , so $(\text{inv}, \text{imaj})$ have a symmetric joint distribution. It then follows from Proposition 1.3.14 that (inv, maj) have a symmetric joint distribution. \square

We conclude this section by discussing a connection between permutations $w \in \mathfrak{S}_n$ and functions $f : [n] \rightarrow \mathbb{N}$ (the set \mathbb{N} could be replaced by any totally ordered set) in which the descent set plays a leading role.

1.4.10 Definition. Let $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$. We say that the function $f : [n] \rightarrow \mathbb{N}$ is w -compatible if the following two conditions hold.

- (a) $f(w_1) \geq f(w_2) \geq \cdots \geq f(w_n)$
 (b) $f(w_i) > f(w_{i+1})$ if $w_i > w_{i+1}$ (i.e., if $i \in D(w)$)

1.4.11 Lemma. Given $f : [n] \rightarrow \mathbb{N}$, there is a unique permutation $w \in \mathfrak{S}_n$ for which f is w -compatible.

Proof. An ordered partition or set composition of a (finite) set S is a vector (B_1, B_2, \dots, B_k) of subsets $B_i \subseteq S$ such that $B_i \neq \emptyset$, $B_i \cap B_j = \emptyset$ for $i \neq j$, and $B_1 \cup \cdots \cup B_k = S$. Clearly there is a unique ordered partition (B_1, \dots, B_k) of $[n]$ such that f is constant on each B_i and $f(B_1) > f(B_2) > \cdots > f(B_k)$ (where $f(B_i)$ means $f(m)$ for any $m \in B_i$). Then w is obtained by arranging the elements of B_1 in increasing order, then the elements of B_2 in increasing order, and so on. \square

The enumeration of certain natural classes of w -compatible functions is closely related to the statistics des and maj , as shown by the next lemma. Further enumerative results concerning w -compatible functions appear in Subsection 3.15.1. For $w \in \mathfrak{S}_n$, let $\mathcal{A}(w)$ denote the set of all w -compatible functions $f : [n] \rightarrow \mathbb{N}$; and for $w \in \mathfrak{S}_d$, let $\mathcal{A}_m(w)$ denote the set of w -compatible functions $f : [d] \rightarrow [m]$, i.e., $\mathcal{A}_m(w) = \mathcal{A}(w) \cap [m]^{[d]}$, where in general if X and Y are sets then Y^X denotes the set of all functions $f : X \rightarrow Y$. Note that $\mathcal{A}_0(w) = \emptyset$.

1.4.12 Lemma. (a) For $w \in \mathfrak{S}_d$ and $m \geq 0$, we have

$$\#\mathcal{A}_m(w) = \binom{m+d-1-\text{des}(w)}{d} = \left(\binom{m-\text{des}(w)}{d} \right) \quad (1.45)$$

and

$$\sum_{m \geq 1} \#\mathcal{A}_m(w) \cdot x^m = \frac{x^{1+\text{des}(w)}}{(1-x)^{d+1}}. \quad (1.46)$$

(If $0 \leq m < \text{des}(w)$, then we set $\left(\binom{m-\text{des}(w)}{d} \right) = 0$.)

(b) For $f : [n] \rightarrow \mathbb{N}$, write $|f| = \sum_{i=1}^n f(i)$. Then for $w \in \mathfrak{S}_n$, we have

$$\sum_{f \in \mathcal{A}(w)} q^{|f|} = \frac{q^{\text{maj}(w)}}{(1-q)(1-q^2) \cdots (1-q^n)}. \quad (1.47)$$

Proof. The basic idea of both proofs is to convert “partially strictly decreasing” sequences to weakly decreasing sequences similar to our first direct proof in Section 1.2 of the formula $\left(\binom{n}{k} \right) = \binom{n+k-1}{k}$. We will give “proofs by example” that should make the general case clear.

(a) Let $w = 4632715$. Then $f \in \mathcal{A}_m(w)$ if and only if

$$m \geq f(4) \geq f(6) > f(3) > f(2) \geq f(7) > f(1) \geq f(5) \geq 1. \quad (1.48)$$

Let $g(5) = f(5)$, $g(1) = f(1)$, $g(7) = f(7) - 1$, $g(2) = f(2) - 1$, $g(3) = f(3) - 2$, $g(6) = f(6) - 3$, $g(4) = f(4) - 3$. In general, $g(j) = f(j) - h_j$, where h_j is the

number of descents of w to the right of j . Equation (1.48) becomes

$$m - 3 \geq g(4) \geq g(6) \geq g(3) \geq g(2) \geq g(7) \geq g(1) \geq g(5) \geq 1.$$

Clearly the number of such g is $\left(\binom{m-3}{7}\right) = \left(\binom{m-\text{des}(w)}{d}\right)$, and (1.45) follows. There are numerous ways to obtain equation (1.46) from equation (1.45), for example, by observing that

$$\left(\binom{m-\text{des}(w)}{d}\right) = (-1)^{m-\text{des}(w)-1} \binom{-(d+1)}{m-\text{des}(w)-1}$$

and using (1.20).

(b) Let $w = 4632715$ as in (a). Then $f \in \mathcal{A}(w)$ if and only

$$f(4) \geq f(6) > f(3) > f(2) \geq f(7) > f(1) \geq f(5) \geq 0. \quad (1.49)$$

Defining g as in (a), equation (1.49) becomes

$$g(4) \geq g(6) \geq g(3) \geq g(2) \geq g(7) \geq g(1) \geq g(5) \geq 0.$$

Moreover, $\sum f(i) = \sum g(i) + 10 = \sum g(i) + \text{maj}(w)$. Hence,

$$\sum_{f \in \mathcal{A}(w)} q^{|f|} = q^{\text{maj}(w)} \sum_{g(4) \geq g(6) \geq g(3) \geq g(2) \geq g(7) \geq g(1) \geq g(5) \geq 0} q^{g(4) + \dots + g(5)}.$$

The latter sum is easy to evaluate in a number of ways, for example, as an iterated geometric progression (i.e., first sum on $g(4) \geq g(6)$, then on $g(6) \geq g(3)$, etc.). It also is equivalent to equation (1.76). The proof follows. \square

Let $\mathbb{N}^{[n]}$ denote the set of all functions $f : [n] \rightarrow \mathbb{N}$, and let $\mathcal{A}(w)$ denote those $f \in \mathbb{N}^{[n]}$ that are compatible with $w \in \mathfrak{S}_n$. Lemma 1.4.11 then says that we have a disjoint union

$$\mathbb{N}^{[n]} = \bigcup_{w \in \mathfrak{S}_n} \mathcal{A}(w). \quad (1.50)$$

It also follows that

$$[m]^{[d]} = \bigcup_{w \in \mathfrak{S}_d} \mathcal{A}_m(w). \quad (1.51)$$

We now are in a position to give more conceptual proofs of Propositions 1.4.4 and 1.4.6. Take the cardinality of both sides of (1.51), multiply by x^m , and sum on $m \geq 0$. We get

$$\sum_{m \geq 0} m^d x^m = \sum_{w \in \mathfrak{S}_d} \# \mathcal{A}_m(w) \cdot x^m.$$

The proof of Proposition 1.4.4 now follows from equation (1.46). Similarly, by (1.50) we have

$$\sum_{f \in \mathbb{N}^{[n]}} q^{|f|} = \sum_{w \in \mathfrak{S}_n} \sum_{f \in \mathcal{A}(w)} q^{|f|}.$$

The left-hand side is clearly $1/(1-q)^n$, whereas by equation (1.47) the right-hand side is

$$\sum_{w \in \mathfrak{S}_n} \frac{q^{\text{maj}(w)}}{(1-q)(1-q^2) \cdots (1-q^n)}.$$

Hence

$$\frac{1}{(1-q)^n} = \frac{\sum_{w \in \mathfrak{S}_n} q^{\text{maj}(w)}}{(1-q)(1-q^2) \cdots (1-q^n)}.$$

Multiplying by $(1-q)(1-q^2) \cdots (1-q^n)$ and simplifying gives Proposition 1.4.6.

1.5 Geometric Representations of Permutations

We have seen that a permutation can be regarded as either a function, a word, or a sequence (the inversion table). In this section, we will consider four additional ways of representing permutations and will illustrate the usefulness of each such representation.

The first representation is the most obvious, namely, a permutation matrix. Specifically, if $w \in \mathfrak{S}_n$, then define the $n \times n$ matrix P_w , with rows and columns indexed by $[n]$, as follows:

$$(P_w)_{ij} = \begin{cases} 1, & \text{if } w(i) = j \\ 0, & \text{otherwise.} \end{cases}$$

The matrix P_w is called the *permutation matrix* corresponding to w . Clearly, a square $(0,1)$ -matrix is a permutation matrix if and only if it has exactly one 1 in every row and column. Sometimes it is more convenient to replace the 0's and 1's with some other symbols. For instance, the matrix P_w could be replaced by a $n \times n$ grid, where each square indexed by $(i, w(i))$ is filled in. Figure 1.3 shows the matrix P_w corresponding to $w = 795418362$, together with the equivalent representation as a grid with certain squares filled in.

To illustrate the use of permutation matrices as geometric objects per se, define a *decreasing subsequence* of length k of a permutation $w = w_1 \cdots w_n \in \mathfrak{S}_n$ to be a subsequence $w_{i_1} > w_{i_2} > \cdots > w_{i_k}$ (so $i_1 < i_2 < \cdots < i_k$ by definition of subsequence). (*Increasing subsequence* is similarly defined, though we have no need for this concept in the present example.) Let $f(n)$ be the number of permutations $w \in \mathfrak{S}_n$ with no decreasing subsequence of length three. For instance,

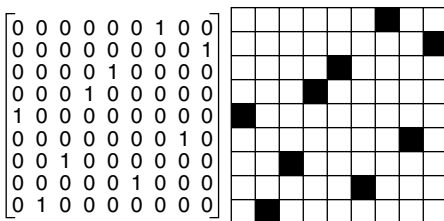


Figure 1.3 The permutation matrix of the permutation $w = 795418362$.

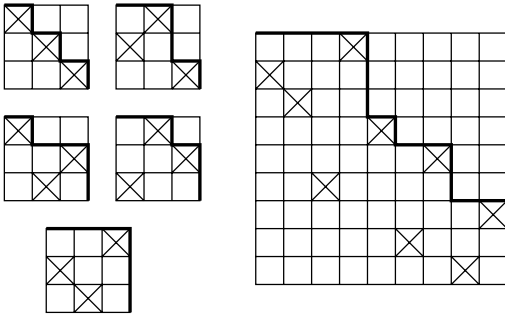


Figure 1.4 Lattice paths corresponding to 321-avoiding permutations.

$f(3) = 5$ since 321 is the only excluded permutation. Let w be a permutation with no decreasing subsequence of length three, and let P_w be its permutation matrix, where for better visualization we replace the 1's in P_w by X's. Draw a lattice path L_w from the upper-left corner of P_w to the lower-right corner, where each step is one unit to the right (east) or down (south), and where the “outside corners” (consisting of a right step followed by a down step) of L_w occur at the top and right of each square on or above the main diagonal containing an X. We trust that Figure 1.4 will make this definition clear; it shows the five paths for $w \in \mathfrak{S}_3$ as well as the path for $w = 412573968$. It is not hard to see that the lattice paths so obtained are exactly those that do not pass below the main diagonal. Conversely, it is also not hard to see that given a lattice path L not passing below the main diagonal, there is a unique permutation $w \in \mathfrak{S}_n$ with no decreasing subsequence of length three for which $L = L_w$.

We have converted our permutation enumeration problem to a much more tractable lattice path counting problem. It is shown in Corollary 6.2.3 that the number of such paths is the *Catalan number* $C_n = \frac{1}{n+1} \binom{2n}{n}$, so we have shown that

$$f(n) = C_n. \quad (1.52)$$

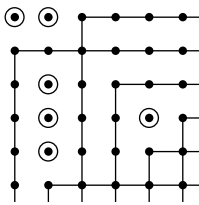
The growth diagrams discussed in Section 7.13 show a more sophisticated use of permutation matrices.

NOTE. The Catalan numbers form one of the most interesting and ubiquitous sequences in enumerative combinatorics; see Chapter 6, especially Corollary 6.2.3 and Exercise 6.19, for further information.

An object closely related to the permutation matrix P_w is the diagram of $w \in \mathfrak{S}_n$. Represent the set $[n] \times [n]$ as an $n \times n$ array of dots, using matrix coordinates, so the upper-left dot represents $(1, 1)$, the dot to its right is $(1, 2)$, and so on. If $w(i) = j$, then from the point (i, j) draw a horizontal line to the right and vertical line to the bottom. Figure 1.5 illustrates the case $w = 314652$. The set of dots that are not covered by lines is called the *diagram* D_w of w . For instance, Figure 1.5 shows that

$$D_{314652} = \{(1, 1), (1, 2), (3, 2), (4, 2), (4, 5), (5, 2)\}.$$

The dots of the diagram are circled for greater clarity.

Figure 1.5 The diagram of the permutation $w = 314652$.

It is easy to see that if a_j denotes the number of elements of D_w in column j , then the inversion table of w is given by $I(w) = (a_1, a_2, \dots, a_n)$. Similarly, if c_i is the number of elements in the i th row of D_w then $\text{code}(w) = (c_1, c_2, \dots, c_n)$. If D_w^t denotes the transpose (reflection about the main diagonal) of D_w , then $D_w^t = D_{w^{-1}}$.

As an illustration of the use of the diagram D_w , define a permutation $w = w_1 \cdots w_n \in \mathfrak{S}_n$ to be *132-avoiding* if there does not exist $i < j < k$ with $w_i < w_k < w_j$. In other words, no subsequence of w of length three has its terms in the same relative order as 132. Clearly, this definition can be generalized to define *u -avoiding* permutations, where $u \in \mathfrak{S}_k$. For instance, the previously considered permutations with no decreasing subsequence of length three are just 321-avoiding permutations.

It is not hard to see that w is 132-avoiding if and only if there exists integers $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ such that for all $i \geq 0$ the i th row of D_w consists of the first λ_i dots in that row. In symbols,

$$D_w = \{(i, j) : 1 \leq j \leq \lambda_i\}.$$

Equivalently, if $(i, j) \in D_w$ and $i' \leq i$, $j' \leq j$, then $(i', j') \in D_w$. In the terminology of Section 1.7, the sequence $\lambda = (\lambda_1, \lambda_2, \dots)$ is a *partition* of $\sum \lambda_i = \text{inv}(w)$, and D_w is the *Ferrers diagram* of λ . In this sense, diagrams of permutations are generalizations of diagrams of partitions. Note that in any $n \times n$ diagram D_w , where $w \in \mathfrak{S}_n$, there are at least i dots in the i th row that do not belong to D_w . Hence if w is 132-avoiding then the corresponding partition $\lambda = (\lambda_1, \dots, \lambda_n)$ satisfies $\lambda_i \leq n - i$. Conversely, it is easy to see that if λ satisfies $\lambda_i \leq n - i$, then the Ferrers diagram of λ is the diagram of a (necessarily 132-avoiding) permutation $w \in \mathfrak{S}_n$. Hence, the number of 132-avoiding permutations in \mathfrak{S}_n is equal to the number of integer sequences $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ such that $\lambda_i \leq n - i$. It follows from Exercise 6.19(s) that the number of such sequences is the Catalan number $C_n = \frac{1}{n+1} \binom{2n}{n}$. (There is also a simple bijection with the lattice paths that we put in one-to-one correspondence with 321-avoiding permutations. In fact, the lattice path construction we applied to 321-avoiding permutations works equally well for 132-avoiding permutations if our paths go from the upper right to lower left; see Figure 1.6.) Hence by equation (1.52) the number of 132-avoiding permutations in \mathfrak{S}_n is the same as the number of 321-avoiding permutations in \mathfrak{S}_n (i.e., permutations in \mathfrak{S}_n with no decreasing subsequence of length three). Simple symmetry

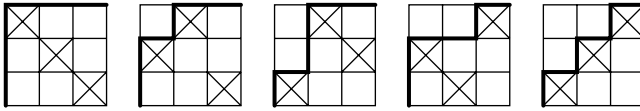


Figure 1.6 Lattice paths corresponding to 132-avoiding permutations in \mathfrak{S}_3 .

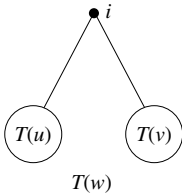


Figure 1.7 The definition of $T(w)$.

arguments (e.g., replacing $w_1 w_2 \cdots w_n$ with $w_n \cdots w_2 w_1$) then show that, for *any* $u \in \mathfrak{S}_3$, the number of u -avoiding permutations $w \in \mathfrak{S}_n$ is C_n .

Since $\#D_w = \text{inv}(w)$, the preceding characterization of diagrams of 132-avoiding permutations $w \in \mathfrak{S}_n$ yields the following refinement of the enumeration of such w .

1.5.1 Proposition. *Let $\mathcal{S}_{132}(n)$ denote the set of 132-avoiding $w \in \mathfrak{S}_n$. Then*

$$\sum_{w \in \mathcal{S}_{132}(n)} q^{\text{inv}(w)} = \sum_{\lambda} q^{|\lambda|},$$

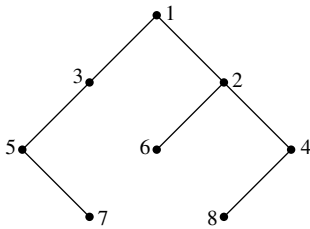
where λ ranges over all integer sequences $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ satisfying $\lambda_i \leq n - i$, and where $|\lambda| = \sum \lambda_i$.

For further information on the sums appearing in Proposition 1.5.1, see Exercise 6.34(a).

We now consider two ways to represent a permutation w as a tree T and discuss how the structure of T interacts with the combinatorial properties of w . Let $w = w_1 w_2 \cdots w_n$ be any word on the alphabet \mathbb{P} with no repeated letters. Define a binary tree $T(w)$ as follows. If $w = \emptyset$, then $T(w) = \emptyset$. If $w \neq \emptyset$, then let i be the least element (letter) of w . Thus, w can be factored uniquely in the form $w = uiv$. Now let i be the root of $T(w)$, and let $T(u)$ and $T(v)$ be the left and right subtrees of i ; see Figure 1.7. This procedure yields an inductive definition of $T(w)$. The left successor of a vertex j is the least element k to the left of j in w such that all elements of w between k and j (inclusive) are $\geq j$, and similarly for the right successor.

1.5.2 Example. Let $w = 57316284$. Then $T(w)$ is given by Figure 1.8.

The correspondence $w \mapsto T(w)$ is a bijection between \mathfrak{S}_n and *increasing binary trees* on n vertices; that is, binary trees with n vertices labeled $1, 2, \dots, n$ such that the labels along any path from the root are increasing. To obtain w from $T(w)$,

Figure 1.8 The increasing binary tree $T(57316284)$.

read the labels of w in *symmetric order*, that is, first the labels of the left subtree (in symmetric order, recursively), then the label of the root, and then the labels of the right subtree.

Let $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$. Define the element w_i of w to be

- a *double rise* or *double ascent*, if $w_{i-1} < w_i < w_{i+1}$
- a *double fall* or *double descent*, if $w_{i-1} > w_i > w_{i+1}$
- a *peak*, if $w_{i-1} < w_i > w_{i+1}$
- a *valley*, if $w_{i-1} > w_i < w_{i+1}$,

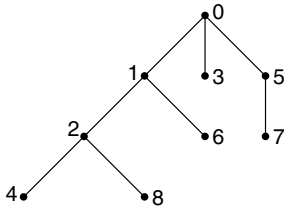
where we set $w_0 = w_{n+1} = 0$. It is easily seen that the property listed below of an element i of w corresponds to the given property of the vertex i of $T(w)$.

Element i of w	Vertex i of $T(w)$ has precisely the following successors
double rise	right
double fall	left
valley	left and right
peak	none

From this discussion of the bijection $w \mapsto T(w)$, a large number of otherwise mysterious properties of increasing binary trees can be trivially deduced. The following proposition gives a sample of such results. Exercise 1.61 provides a further application of $T(w)$.

- 1.5.3 Proposition.** (a) *The number of increasing binary trees with n vertices is $n!$.*
 (b) *The number of such trees for which exactly k vertices have left successors is the Eulerian number $A(n, k+1)$.*
 (c) *The number of complete (i.e., every vertex is either an endpoint or has two successors) increasing binary trees with $2n+1$ vertices is equal to the number E_{2n+1} of alternating permutations in \mathfrak{S}_{2n+1} .*

Let us now consider a second way to represent a permutation by a tree. Given $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$, construct an (unordered) tree $T'(w)$ with vertices $0, 1, \dots, n$

Figure 1.9 The unordered increasing tree $T'(57316284)$.

by defining vertex i to be the successor of the rightmost element j of w which precedes i and which is less than i . If there is no such element j , then let i be the successor of the root 0.

1.5.4 Example. Let $w = 57316284$. Then $T'(w)$ is given by Figure 1.9.

The correspondence $w \mapsto T'(w)$ is a bijection between \mathfrak{S}_n and increasing trees on $n + 1$ vertices. It is easily seen that the successors of 0 are just the *left-to-right minima* (or *retreating elements*) of w (i.e., elements w_i such that $w_i < w_j$ for all $j < i$, where $w = w_1 w_2 \cdots w_n$). Moreover, the endpoints of $T'(w)$ are just the elements w_i for which $i \in D(w)$ or $i = n$. Thus, in analogy to Proposition 1.5.3 (using Proposition 1.3.1 and the obvious symmetry between left-to-right maxima and left-to-right minima), we obtain the following result.

- 1.5.5 Proposition.** (a) *The number of unordered increasing trees on $n + 1$ vertices is $n!$.*
 (b) *The number of such trees for which the root has k successors is the signless Stirling number $c(n, k)$.]*
 (c) *The number of such trees with k endpoints is the Eulerian number $A(n, k)$.*

1.6 Alternating Permutations, Euler Numbers, and the cd -Index of \mathfrak{S}_n

In this section we consider enumerative properties of alternating permutations, as defined in Section 1.4. Recall that a permutation $w \in \mathfrak{S}_n$ is alternating if $D(w) = \{1, 3, 5, \dots\} \cap [n - 1]$, and reverse alternating if $D(w) = \{2, 4, 6, \dots\} \cap [n - 1]$.

1.6.1 Basic Properties

Recall that E_n denotes the number of alternating permutations (or reverse alternating permutations) $w \in \mathfrak{S}_n$ (with $E_0 = 1$) and is called an Euler number. The exponential generating function for Euler numbers is very elegant and surprising.

1.6.1 Proposition. *We have*

$$\begin{aligned} \sum_{n \geq 0} E_n \frac{x^n}{n!} &= \sec x + \tan x \\ &= 1 + x + \frac{x^2}{2!} + 2 \frac{x^3}{3!} + 5 \frac{x^4}{4!} + 16 \frac{x^5}{5!} + 61 \frac{x^6}{6!} + 272 \frac{x^7}{7!} + 1385 \frac{x^8}{8!} + \cdots \end{aligned}$$

Note that $\sec x$ is an even function (i.e., $\sec(-x) = \sec x$), while $\tan x$ is odd ($\tan(-x) = -\tan x$). It follows from Proposition 1.6.1 that

$$\sum_{n \geq 0} E_{2n} \frac{x^{2n}}{(2n)!} = \sec x, \quad (1.53)$$

$$\sum_{n \geq 0} E_{2n+1} \frac{x^{2n+1}}{(2n+1)!} = \tan x. \quad (1.54)$$

For this reason E_{2n} is sometimes called a *secant number* and E_{2n+1} a *tangent number*.

Proof of Proposition 1.6.1. Let $0 \leq k \leq n$. Choose a k -subset S of $[n]$ in $\binom{n}{k}$ ways, and set $\bar{S} = [n] - S$. Choose a reverse alternating permutation u of S in E_k ways, and choose a reverse alternating permutation v of \bar{S} in E_{n-k} ways. Let w be the concatenation $u^r, n+1, v$, where u^r denotes the reverse of u (i.e., if $u = u_1 \cdots u_k$, then $u^r = u_k \cdots u_1$). When $n \geq 2$, we obtain in this way every alternating and every reverse alternating permutation w exactly once. Since there is a bijection between alternating and reverse alternating permutations of any finite (ordered) set, the number of w obtained is $2E_{n+1}$. Hence,

$$2E_{n+1} = \sum_{k=0}^n \binom{n}{k} E_k E_{n-k}, \quad n \geq 1. \quad (1.55)$$

Set $y = \sum_{n \geq 0} E_n x^n / n!$. Taking into account the initial conditions $E_0 = E_1 = 1$, equation (1.55) becomes the differential equation

$$2y' = y^2 + 1, \quad y(0) = 1.$$

The unique solution is $y = \sec x + \tan x$. □

NOTE. The clever counting of both alternating and reverse alternating permutations in the proof of Proposition 1.6.1 can be avoided at the cost of a little elegance. Namely, by considering the position of 1 in an alternating permutation w , we obtain the recurrence

$$E_{n+1} = \sum_{\substack{1 \leq j \leq n \\ j \text{ odd}}} \binom{n}{j} E_j E_{n-j}, \quad n \geq 1.$$

This recurrence leads to a system of differential equations for the power series $\sum_{n \geq 0} E_{2n} x^{2n} / (2n)!$ and $\sum_{n \geq 0} E_{2n+1} x^{2n+1} / (2n+1)!$.

Note that equations (1.53) and (1.54) could in fact be used to *define* $\sec x$ and $\tan x$ in terms of alternating permutations. We can then try to develop as much trigonometry as possible (e.g., the identity $1 + \tan^2 x = \sec^2 x$) using this definition, thereby defining the subject of *combinatorial trigonometry*. For the first steps in this direction, see Exercise 5.7.

It is natural to ask whether Proposition 1.6.1 has a more conceptual proof. The proof preceding does not explain why we ended up with such a simple generating function. To be even more clear about this point, rewrite equation (1.53) as

$$\sum_{n \geq 0} E_{2n} \frac{x^{2n}}{(2n)!} = \frac{1}{\sum_{n \geq 0} (-1)^n \frac{x^{2n}}{(2n)!}}. \quad (1.56)$$

Compare this equation with the exponential generating function for the number of permutations in \mathfrak{S}_n with descent set $[n-1]$:

$$\sum_{n \geq 0} \frac{x^n}{n!} = \frac{1}{\sum_{n \geq 0} (-1)^n \frac{x^n}{n!}}. \quad (1.57)$$

Could there be a reason why having descents in every second position corresponds to taking every second term in the denominator of (1.57) and keeping the signs alternating? Possibly the similarity between (1.56) and (1.57) is just a coincidence. All doubts are dispelled, however, by the following generalization of equation (1.56). Let $f_k(n)$ denote the number of permutations $w \in \mathfrak{S}_n$ satisfying

$$D(w) = \{k, 2k, 3k, \dots\} \cap [n-1]. \quad (1.58)$$

Then

$$\sum_{n \geq 0} f_k(kn) \frac{x^{kn}}{(kn)!} = \frac{1}{\sum_{n \geq 0} (-1)^n \frac{x^{kn}}{(kn)!}}. \quad (1.59)$$

Such a formula cries out for a more conceptual proof. One such proof is given in Section 3.19. Exercise 2.22 gives a further proof for $k=2$ (easily extended to any k) based on Inclusion–Exclusion. Another enlightening proof, less elegant but more straightforward than the one in Section 3.19, is the following.

Proof of equation (1.59). We have

$$\begin{aligned} \frac{1}{\sum_{n \geq 0} (-1)^n \frac{x^{kn}}{(kn)!}} &= \frac{1}{1 - \sum_{n \geq 1} (-1)^{n-1} \frac{x^{kn}}{(kn)!}} \\ &= \sum_{j \geq 0} \left(\sum_{n \geq 1} (-1)^{n-1} \frac{x^{kn}}{(kn)!} \right)^j \\ &= \sum_{j \geq 0} \sum_{N \geq j} \sum_{\substack{a_1 + \dots + a_j = N \\ a_i \geq 1}} \binom{kN}{ka_1, \dots, ka_j} (-1)^{N-j} \frac{x^{kN}}{(kN)!}. \end{aligned}$$

Comparing (carefully) with equations (1.34) and (1.35) completes the proof. \square

A similar proof can be given of equation (1.54) and its extension to permutations in \mathfrak{S}_{kn+i} with descent set $\{k, 2k, 3k, \dots\} \cap [kn+i-1]$ for $1 \leq i \leq k-1$. Details are left as an exercise (Exercise 1.146).

1.6.2 Flip Equivalence of Increasing Binary Trees

Alternating permutations appear as the number of equivalence classes of certain naturally defined equivalence relations. (For an example unrelated to this section, see Exercise 3.127(b).) We will give an archetypal example in this subsection. In the next subsection, we will give a similar result, which has an application to the numbers $\beta_n(S)$ of permutations $w \in \mathfrak{S}_n$ with descent set S .

Recall that in Section 1.5 we associated an increasing binary tree $T(w)$ with a permutation $w \in \mathfrak{S}_n$. The *flip* of a binary tree at a vertex v is the binary tree obtained by interchanging the left and right subtrees of v . Define two increasing binary trees T and T' on the vertex set $[n]$ to be *equivalent* if T' can be obtained from T by a sequence of flips. Clearly, this definition of equivalence is an equivalence relation, and the number of increasing binary trees equivalent to T is $2^{n-\epsilon(T)}$, where $\epsilon(T)$ is the number of endpoints of T . The equivalence classes are in an obvious bijection with increasing $(1,2)$ -trees on the vertex set $[n]$, that is, increasing (rooted) trees so that every non-endpoint vertex has one or two children. (These are not plane trees, i.e., the order in which we write the children of a vertex is irrelevant.) Figure 1.10 shows the five increasing $(1,2)$ -trees on four vertices, so $f(4) = 5$. Let $f(n)$ denote the number of equivalence classes (i.e., the number of increasing $(1,2)$ -trees on the vertex set $[n]$).

1.6.2 Proposition. *We have $f(n) = E_n$ (an Euler number).*

Proof. Perhaps the most straightforward proof is by generating functions. Let

$$y = \sum_{n \geq 1} f(n) \frac{x^n}{n!} = x + \frac{x^2}{2} + 2\frac{x^3}{6} + \dots$$

Then $y' = \sum_{n \geq 0} f(n+1)x^n/n!$. Every increasing $(1,2)$ -tree with $n+1$ vertices either (a) is a single vertex ($n=0$), (b) has one subtree of the root, which is an increasing $(1,2)$ -tree with n vertices, or (c) has two subtrees of the root, each of which is an increasing $(1,2)$ -tree, with n vertices in all. The order of the two

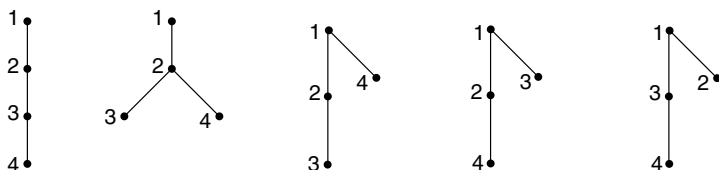


Figure 1.10 The five increasing $(1,2)$ -trees with four vertices.

subtrees is irrelevant. From this observation, we obtain the differential equation $y' = 1 + y + \frac{1}{2}y^2$, $y(0) = 0$. The unique solution is $y = \sec x + \tan x - 1$, and the proof follows from Proposition 1.6.1. \square

ALGEBRAIC NOTE. Let \mathcal{T}_n be the set of all increasing binary tree with vertex set $[n]$. For $T \in \mathcal{T}_n$ and $1 \leq i \leq n$, let $\omega_i T$ be the flip of T at vertex i . Then clearly the ω_i 's generate a group isomorphic to $(\mathbb{Z}/2\mathbb{Z})^n$ acting on \mathcal{T}_n , and the orbits of this action are the flip equivalence classes.

1.6.3 Min-Max Trees and the cd -Index

We now consider a variant of the bijection $w \mapsto T(w)$ between permutations and increasing binary trees defined in Section 1.5 that has an interesting application to descent sets of permutations. We will just sketch the basic facts and omit most details of proofs (all of which are fairly straightforward). We define the *min-max tree* $M(w)$ associated with a sequence $w = a_1 a_2 \cdots a_n$ of distinct integers as follows. First, $M(w)$ is a binary tree with vertices labeled a_1, a_2, \dots, a_n . Let j be the least integer for which *either* $a_j = \min\{a_1, \dots, a_n\}$ or $a_j = \max\{a_1, \dots, a_n\}$. Define a_j to be the root of $M(w)$. Then define (recursively) $M(a_1, \dots, a_{j-1})$ to be the left subtree of a_j , and $M(a_{j+1}, \dots, a_n)$ to be the right subtree. Figure 1.11(a) shows $M(5, 10, 4, 6, 7, 2, 12, 1, 8, 11, 9, 3)$. Note that no vertex of a min-max tree $M(w)$ has only a left successor. Note also that every vertex v is either the minimum or maximum element of the subtree with root v .

Given the min-max tree $M(w)$ where $w = a_1 \cdots a_n$, we will define operators ψ_i , $1 \leq i \leq n$, that permute the labels of $M(w)$, creating a new min-max tree $\psi_i M(w)$. The operator ψ_i only permutes the label of the vertex of $M(w)$ labeled a_i and the labels of the right subtree of this vertex. (Note that the vertex labeled a_i depends only on i and the tree $M(w)$, not on the permutation w .) All other vertices are fixed by ψ_i . In particular, if a_i is an endpoint, then $\psi_i M(w) = M(w)$. We denote by M_{a_i} the subtree of $M(w)$ consisting of a_i and the right subtree of a_i . Thus, a_i is either the minimum or maximum element of M_{a_i} . Suppose that a_i is the minimum element of M_{a_i} . Then replace a_i with the *largest* element of M_{a_i} , and permute the remaining elements of M_{a_i} so that they keep their same relative order. This defines $\psi_i M(w)$. Similarly, suppose that a_i is the maximum element of the subtree M_{a_i} with root a_i . Then replace a_i with the *smallest* element of M_{a_i} , and permute the

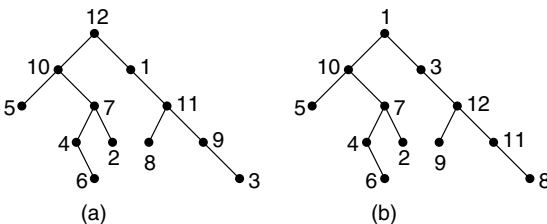


Figure 1.11 (a) The min-max tree $M = M(5, 10, 4, 6, 7, 2, 12, 1, 8, 11, 9, 3)$; (b) The transformed tree $\psi_7 M = M(5, 10, 4, 6, 7, 2, 1, 3, 9, 12, 11, 8)$.

remaining elements of M_{a_i} so that they keep their same relative order. Again this defines $\psi_i M(w)$. Figure 1.11(b) shows that $\psi_7 M(5, 10, 4, 6, 7, 2, 12, 1, 8, 11, 9, 3) = M(5, 10, 4, 6, 7, 2, 1, 3, 9, 12, 11, 8)$. We have $a_7 = 12$, so ψ_7 permutes vertex 12 and the vertices on the right subtree of 12. Vertex 12 is replaced by 1, the smallest vertex of the right subtree. The remaining elements 1, 3, 8, 9, 11 get replaced with 3, 8, 9, 11, 12 in that order.

Fact #1. *The operators ψ_i are commuting involutions and hence generate an (abelian) group \mathfrak{G}_w isomorphic to $(\mathbb{Z}/2\mathbb{Z})^{\iota(w)}$, where $\iota(w)$ is the number of internal vertices of $M(w)$. Those ψ_i for which a_i is not an endpoint are a minimal set G_w of generators for \mathfrak{G}_w . Hence, there are precisely $2^{\iota(w)}$ different trees $\psi M(w)$ for $\psi \in \mathfrak{G}_w$, given by $\psi_{i_1} \cdots \psi_{i_j} M(w)$, where $\{\psi_{i_1}, \dots, \psi_{i_j}\}$ ranges over all subsets of G_w .*

Given a permutation $w \in \mathfrak{S}_n$ and an operator $\psi \in \mathfrak{G}_w$, we define the permutation ψw by $\psi M(w) = M(\psi w)$. Define two permutations $v, w \in \mathfrak{S}_n$ to be M -equivalent, denoted $v \stackrel{M}{\sim} w$, if $v = \psi w$ for some $\psi \in \mathfrak{G}_w$. Clearly $\stackrel{M}{\sim}$ is an equivalence relation, and by Fact #1 the size of the equivalence class $[w]$ containing w is $2^{\iota(w)}$.

There is a simple connection between the descent sets of w and $\psi_i w$.

Fact #2. *Let a_i be an internal vertex of $M(w)$ with only a right child. Then*

$$D(\psi_i w) = \begin{cases} D(w) \cup \{i\}, & \text{if } i \notin D(w), \\ D(w) - \{i\}, & \text{if } i \in D(w). \end{cases}$$

Let a_i be an internal vertex of $M(w)$ with both a left and right child. Then exactly one of $i - 1, i$ belongs to $D(w)$, and we have

$$D(\psi_i w) = \begin{cases} (D(w) \cup \{i\}) - \{i - 1\}, & \text{if } i \notin D(w), \\ (D(w) \cup \{i - 1\}) - \{i\}, & \text{if } i \in D(w). \end{cases}$$

Note that if a_i is a vertex with two children, then a_{i-1} will always be an endpoint on the left subtree of a_i . It follows that the changes in the descent sets described by Fact #2 take place independently of each other. (In fact, this independence is equivalent to the commutativity of the ψ_i 's.) The different descent sets $D(w)$, where w ranges over an M -equivalence class, can be conveniently encoded by a noncommutative polynomial in the letters a and b . Given a set $S \subseteq [n - 1]$, define its *characteristic monomial* (or *variation*) to be the noncommutative monomial

$$u_S = e_1 e_2 \cdots e_{n-1}, \quad (1.60)$$

where

$$e_i = \begin{cases} a, & \text{if } i \notin S \\ b, & \text{if } i \in S. \end{cases}$$

For instance, $u_{D(37485216)} = ababbba$.

Now let $w = a_1 a_2 \cdots a_n \in \mathfrak{S}_n$, and let c, d, e be noncommutative indeterminates. For $1 \leq i \leq n$, define

$$f_i = f_i(w) = \begin{cases} c, & \text{if } a_i \text{ has only a right child in } M(w), \\ d, & \text{if } a_i \text{ has a left and right child,} \\ e, & \text{if } a_i \text{ is an endpoint.} \end{cases}$$

Let $\Phi'_w = \Phi'_w(c, d, e) = f_1 f_2 \cdots f_n$, and let $\Phi_w = \Phi_w(c, d) = \Phi'(c, d, 1)$, where 1 denotes the empty word. In other words, Φ_w is obtained from Φ'_w by deleting the e 's. For instance, consider the permutation $w = 5, 10, 4, 6, 7, 2, 12, 1, 8, 11, 9, 3$ of Figure 1.11. The degrees (number of children) of the vertices a_1, a_2, \dots, a_{12} are 0, 2, 1, 0, 2, 0, 2, 1, 0, 2, 1, 0, respectively. Hence,

$$\begin{aligned} \Phi'_w &= edcededcedce, \\ \Phi_w &= dcddcdc. \end{aligned} \tag{1.61}$$

It is clear that if $v \stackrel{M}{\sim} w$, then $\Phi'_v = \Phi'_w$ and $\Phi_v = \Phi_w$, since Φ'_w depends only on $M(w)$ regarded as an *unlabeled* tree.

From Fact #2, we obtain the following result.

Fact #3. Let $w \in \mathfrak{S}_n$, and let $[w]$ be the M -equivalence class containing w . Then

$$\Phi_w(a + b, ab + ba) = \sum_{v \in [w]} u_{D(v)}. \quad \square \tag{1.62}$$

For instance, from equation (1.61), we have

$$\sum_{v \in [w]} u_{D(v)} = (ab + ba)(a + b)(ab + ba)(ab + ba)(a + b)(ab + ba)(a + b).$$

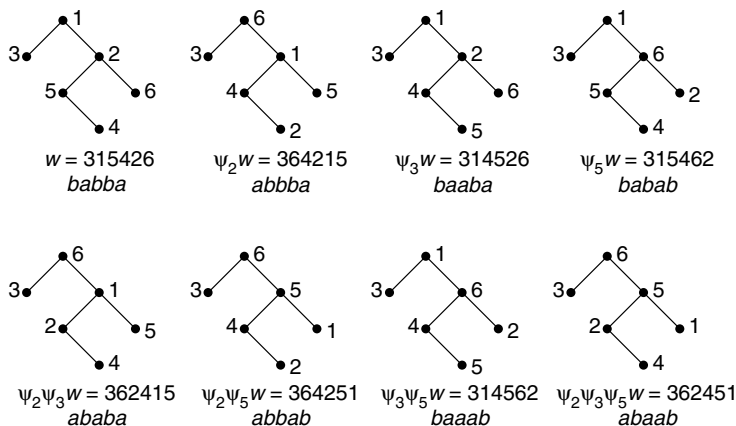
As a further example, Figure 1.12 shows the eight trees $M(v)$ in the M -equivalence class [315426], together with corresponding characteristic monomial $u_{D(v)}$. We see that

$$\begin{aligned} \sum_{v \in [315426]} u_{D(v)} &= babba + abbba + baaba + babab + ababa + abbab \\ &\quad + baaab + abaab = (ab + ba)(a + b)(ab + ba), \end{aligned}$$

whence $\Phi_w = dcd$.

Fact #4. Each equivalence class $[w]$ contains exactly one alternating permutation (as well as one reverse alternating permutation). Hence, the number of M -equivalence classes of permutations $w \in \mathfrak{S}_n$ is the Euler number E_n .

It is not difficult to prove Fact #4 directly from the definition of the tree $M(w)$ and the group \mathfrak{G}_w , but it is also immediate from Fact #3. For in the expansion of $\Phi_w(a + b, ab + ba)$, there will be exactly one alternating term $bababa \cdots$ and one term $ababab \cdots$.

Figure 1.12 The M -equivalence class $[315426]$.

Now consider the generating function

$$\begin{aligned}\Psi_n &= \Psi_n(a, b) = \sum_{w \in \mathfrak{S}_n} u_{D(w)} \\ &= \sum_{S \subseteq [n-1]} \beta(S) u_S.\end{aligned}\quad (1.63)$$

Thus, Ψ_n is a noncommutative generating function for the numbers $\beta(S)$. For instance, $\Psi_3 = aa + 2ab + 2ba + bb$. The polynomial Ψ_n is called the ab -index of the symmetric group \mathfrak{S}_n . (In the more general context of Section 3.17, Ψ_n is called the ab -index of the boolean algebra B_n .) We can group the terms of Ψ_n according to the M -equivalence classes $[w]$, that is,

$$\Psi_n = \sum_{[w]} \sum_{v \in [w]} u_{D(v)}, \quad (1.64)$$

where the outer sum ranges over all distinct M -equivalence classes $[w]$ of permutations in \mathfrak{S}_n . Now by equation (1.62) the inner sum is just $\Phi_w(a + b, ab + ba)$. Hence, we have established the following result.

1.6.3 Theorem. *The ab -index Ψ_n can be written as a polynomial Φ_n in the variables $c = a + b$ and $d = ab + ba$. This polynomial is a sum of E_n monomials.*

The polynomial Φ_n is called the cd -index of the symmetric group \mathfrak{S}_n (or boolean algebra B_n). It is a surprisingly compact way of codifying the numbers $\beta_n(S)$. The number of distinct terms in Φ_n is the Fibonacci number F_n (the number of cd -monomials of degree n , where $\deg c = 1$ and $\deg d = 2$; see Exercise 1.35(c)),

compared with the 2^{n-1} terms of the ab -index Ψ_n . For instance,

$$\Phi_1 = 1,$$

$$\Phi_2 = c,$$

$$\Phi_3 = c^2 + d,$$

$$\Phi_4 = c^3 + 2cd + 2dc,$$

$$\Phi_5 = c^4 + 3c^2d + 5cdc + 3dc^2 + 4d^2,$$

$$\Phi_6 = c^5 + 4c^3d + 9c^2dc + 9cdc^2 + 4dc^3 + 12cd^2 + 10dcd + 12d^2c.$$

One nice application of the cd -index concerns inequalities among the number $\beta_n(S)$. Given $S \subseteq [n-1]$, define $\omega(S) \subseteq [n-2]$ by the condition $i \in \omega(S)$ if and only if exactly one of i and $i+1$ belongs to S , for $1 \leq i \leq n-2$. For instance, if $n=9$ and $S = \{2, 4, 5, 8\}$, then $\omega(S) = \{1, 2, 3, 5, 7\}$. Note that

$$\omega(S) = [n-2] \iff S = \{1, 3, 5, \dots\} \cap [n-1] \text{ or } S = \{2, 4, 6, \dots\} \cap [n-1]. \quad (1.65)$$

1.6.4 Proposition. *Let $S, T \subseteq [n-1]$. If $\omega(S) \subset \omega(T)$, then $\beta_n(S) < \beta_n(T)$.*

Proof. Let $w \in \mathfrak{S}_n$ and $\Phi'_w = f_1 f_2 \cdots f_n$, so each $f_i = c, d$, or e . Define

$$S_w = \{i-1 : f_i = d\}.$$

It is easy to see that

$$\Phi_w = \sum_{\omega(X) \supseteq S_w} u_X.$$

Since Φ_n has nonnegative coefficients, it follows that if $\omega(S) \subseteq \omega(T)$, then $\beta_n(S) \leq \beta_n(T)$. Now assume that S and T are any subsets of $[n-1]$ for which $\omega(S) \subset \omega(T)$ (strict containment). We can easily find a cd -word Φ_w for which $\omega(T) \supseteq S_w$ but $\omega(S) \not\supseteq S_w$. For instance, if $i \in \omega(T) - \omega(S)$, then let $\Phi_w = c^{i-1} d c^{n-2-i}$, so $S_w = \{i\}$. It follows that $\beta_n(S) < \beta_n(T)$. \square

1.6.5 Corollary. *Let $S \subseteq [n-1]$. Then $\beta_n(S) \leq E_n$, with equality if and only if $S = \{1, 3, 5, \dots\} \cap [n-1]$ or $S = \{2, 4, 6, \dots\} \cap [n-1]$.*

Proof. Immediate from Proposition 1.6.4 and equation (1.65). \square

1.7 Permutations of Multisets

Much of what we have done concerning permutations of sets can be generalized to multisets. For instance, there are *two* beautiful theories of cycle decomposition for permutations of multisets (see Exercise 1.62 for one of them, and its solution for a reference to the other). In this section, however, we will only discuss some topics that will be of use later.

First, it is clear that we can define the descent set $D(w)$ of a permutation w of a (finite) multiset M on a totally ordered set (such as \mathbb{P}) exactly as we did for sets. Namely, if $w = w_1 w_2 \cdots w_n$, then

$$D(w) = \{i : w_i > w_{i+1}\}.$$

Thus we also have the concept of $\alpha(S) = \alpha_M(S)$ and $\beta(S) = \beta_M(S)$ for a multiset M , as well as the number $\text{des}(w)$ of descents, the major index $\text{maj}(w)$ and the multiset Eulerian polynomial

$$A_M(x) = \sum_{w \in \mathfrak{S}_M} x^{1+\text{des}(w)},$$

and so on. In Section 4.4.5 we will consider a vast generalization of these concepts. Note for now that there is no obvious analogue of Proposition 1.4.1—that is, an explicit formula for the number $\alpha_M(S)$ of permutations $w \in \mathfrak{S}_M$ with descent set contained in S .

We can also define an *inversion* of $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_M$ as a 4-tuple (i, j, w_i, w_j) for which $i < j$ and $w_i > w_j$, and as before we define $\text{inv}(w)$ to be the number of inversions of w . Note that unlike the case for permutations we shouldn't define an inversion to be just the pair (w_i, w_j) since we can have $(w_i, w_j) = (w_k, w_l)$ but $(i, j) \neq (k, l)$. We wish to generalize Corollary 1.3.13 to multisets. To do so we need a fundamental definition. If (a_1, \dots, a_m) is a sequence of nonnegative integers summing to n , then define the q -multinomial coefficient

$$\binom{n}{a_1, \dots, a_m} = \frac{(n)!}{(a_1)! \cdots (a_m)!}.$$

It is immediate from the definition that $\binom{n}{a_1, \dots, a_m}$ is a rational function of q which, when evaluated at $q = 1$, becomes the ordinary multinomial coefficient $\binom{n}{a_1, \dots, a_m}$. In fact, it is not difficult to see that $\binom{n}{a_1, \dots, a_m}$ is a polynomial in q whose coefficients are nonnegative integers. One way to do this is as follows. Write $\binom{n}{k}$ as short for $\binom{n}{k, n-k}$ (exactly in analogy with the notation $\binom{n}{k}$ for binomial coefficients). The expression $\binom{n}{k}$ is called a q -binomial coefficient (or *Gaussian polynomial*). It is straightforward to verify that

$$\binom{n}{a_1, \dots, a_m} = \binom{n}{a_1} \binom{n-a_1}{a_2} \binom{n-a_1-a_2}{a_3} \cdots \binom{a_m}{a_m} \quad (1.66)$$

and

$$\binom{n}{k} = \binom{n-1}{k} + q^{n-k} \binom{n-1}{k-1}. \quad (1.67)$$

From these equations and the “initial conditions” $\binom{n}{0} = 1$, it follows by induction that $\binom{n}{a_1, \dots, a_m}$ is a polynomial in q with nonnegative integer coefficients.

1.7.1 Proposition. Let $M = \{1^{a_1}, \dots, m^{a_m}\}$ be a multiset of cardinality $n = a_1 + \dots + a_m$. Then

$$\sum_{w \in \mathfrak{S}_M} q^{\text{inv}(w)} = \binom{n}{a_1, \dots, a_m}. \quad (1.68)$$

First Proof. Denote the left-hand side of (1.68) by $P(a_1, \dots, a_m)$ and write $Q(n, k) = P(k, n - k)$. Clearly $Q(n, 0) = 1$. Hence in view of (1.66) and (1.67) it suffices to show that

$$P(a_1, \dots, a_m) = Q(n, a_1)P(a_2, a_3, \dots, a_m), \quad (1.69)$$

$$Q(n, k) = Q(n - 1, k) + q^{n-k} Q(n - 1, k - 1). \quad (1.70)$$

If $w \in \mathfrak{S}_M$, then let w' be the permutation of $M' = \{2^{a_2}, \dots, m^{a_m}\}$ obtained by removing the 1's from w , and let w'' be the permutation of $M'' = \{1^{a_1}, 2^{n-a_1}\}$ obtained from w by changing every element greater than 2 to 2. Clearly w is uniquely determined by w' and w'' , and $\text{inv}(w) = \text{inv}(w') + \text{inv}(w'')$. Hence,

$$\begin{aligned} P(a_1, \dots, a_m) &= \sum_{w' \in \mathfrak{S}_{M'}} \sum_{w'' \in \mathfrak{S}_{M''}} q^{\text{inv}(w') + \text{inv}(w'')} \\ &= Q(n, a_1)P(a_2, a_3, \dots, a_m), \end{aligned}$$

which is (1.69).

Now let $M = \{1^k, 2^{n-k}\}$. Let $\mathfrak{S}_{M,i}$, $1 \leq i \leq 2$, consist of those $w \in \mathfrak{S}_M$ whose last element is i , and let $M_1 = \{1^{k-1}, 2^{n-k}\}$, $M_2 = \{1^k, 2^{n-k-1}\}$. If $w \in \mathfrak{S}_{M,1}$ and $w = u1$, then $u \in \mathfrak{S}_{M_1}$ and $\text{inv}(w) = n - k + \text{inv}(u)$. If $w \in \mathfrak{S}_{M,2}$ and $w = v2$, then $v \in \mathfrak{S}_{M_2}$ and $\text{inv}(w) = \text{inv}(v)$. Hence,

$$\begin{aligned} Q(n, k) &= \sum_{u \in \mathfrak{S}_{M_1}} q^{\text{inv}(u) + n - k} + \sum_{v \in \mathfrak{S}_{M_2}} q^{\text{inv}(v)} \\ &= q^{n-k} Q(n - 1, k - 1) + Q(n - 1, k), \end{aligned}$$

which is (1.70). □

Second Proof. Define a map

$$\begin{aligned} \phi : \mathfrak{S}_M \times \mathfrak{S}_{a_1} \times \dots \times \mathfrak{S}_{a_m} &\rightarrow \mathfrak{S}_n \\ (w_0, w_1, \dots, w_m) &\mapsto w \end{aligned}$$

by converting the a_i i 's in w_0 to the numbers $a_1 + \dots + a_{i-1} + 1, a_1 + \dots + a_{i-1} + 2, a_1 + \dots + a_{i-1} + a_i$ in the order specified by w_i . For instance $(21331223, 21, 231, 312) \mapsto 42861537$. We have converted 11 to 21 (preserving the relative order of the terms of $w_1 = 21$), 222 to 453 (preserving the order 231), and 333 to 867 (preserving 312). It is easily verified that ϕ is a bijection, and that

$$\text{inv}(w) = \text{inv}(w_0) + \text{inv}(w_1) + \dots + \text{inv}(w_m). \quad (1.71)$$

By Corollary 1.3.13, we conclude

$$\left(\sum_{w \in \mathfrak{S}_M} q^{\text{inv}(w)} \right) (a_1)! \cdots (a_m)! = (n)!,$$

and the proof follows. \square

NOTE. If w_1, \dots, w_m are all identity permutations, then we obtain a map $\psi : \mathfrak{S}_M \rightarrow \mathfrak{S}_n$ known as *standardization*. For instance, $\psi(14214331) = 17428563$. Standardization is a very useful technique for reducing problems about multisets to sets. For a significant example, see Lemma 7.11.6.

The first proof of Proposition 1.7.1 can be classified as “semicombinatorial.” We did not give a direct proof of (1.68) itself, but rather of the two recurrences (1.69) and (1.70). At this stage it would be difficult to give a direct combinatorial proof of (1.68) since there is no “obvious” combinatorial interpretation of the coefficients of $(a_1, \dots, a_m)^n$ nor of the value of this polynomial at $q \in \mathbb{N}$. Thus, we will now discuss the problem of giving a combinatorial interpretation of $\binom{n}{k}$ for certain $q \in \mathbb{N}$, which will lead to a combinatorial proof of (1.68) when $m = 2$. Combined with our proof of (1.69), this yields a combinatorial proof of (1.68) in general. The reader unfamiliar with finite fields may skip the rest of this section, except for the brief discussion of partitions.

Let q be a prime power, and denote by \mathbb{F}_q a finite field with q elements (all such fields are of course isomorphic) and by \mathbb{F}_q^n the n -dimensional vector space of all n -tuples $(\alpha_1, \dots, \alpha_n)$, where $\alpha_i \in \mathbb{F}_q$.

1.7.2 Proposition. *The number of k -dimensional subspaces of \mathbb{F}_q^n is $\binom{n}{k}$.*

Proof. Denote the number in question by $G(n, k)$, and let $N = N(n, k)$ equal the number of ordered k -tuples (v_1, \dots, v_k) of linearly independent vectors in \mathbb{F}_q^n . We may choose v_1 in $q^n - 1$ ways, then v_2 in $q^n - q$ ways, and so on, yielding

$$N = (q^n - 1)(q^n - q) \cdots (q^n - q^{k-1}). \quad (1.72)$$

On the other hand, we may choose (v_1, \dots, v_k) by first choosing a k -dimensional subspace W of \mathbb{F}_q^n in $G(n, k)$ ways, and then choosing $v_1 \in W$ in $q^k - 1$ ways, $v_2 \in W$ in $q^k - q$ ways, and so on. Hence,

$$N = G(n, k)(q^k - 1)(q^k - q) \cdots (q^k - q^{k-1}). \quad (1.73)$$

Comparing (1.72) and (1.73) yields

$$\begin{aligned} G(n, k) &= \frac{(q^n - 1)(q^n - q) \cdots (q^n - q^{k-1})}{(q^k - 1)(q^k - q) \cdots (q^k - q^{k-1})} \\ &= \frac{(n)!}{(k)!(n-k)!} = \binom{n}{k}. \end{aligned} \quad \square$$

Note that the above proof is completely analogous to the proof we gave in Section 1.2 that $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. We may consider our proof of Proposition 1.7.2 to be the “ q -analogue” of the proof that $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Now define a *partition* of $n \in \mathbb{N}$ to be a sequence $\lambda = (\lambda_1, \lambda_2, \dots)$ of integers λ_i satisfying $\sum \lambda_i = n$ and $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$. We also write $\lambda = (\lambda_1, \dots, \lambda_k)$ if $\lambda_{k+1} = \lambda_{k+2} = \dots = 0$. Thus, for example,

$$(3, 3, 2, 1, 0, 0, \dots) = (3, 3, 2, 1, 0, 0) = (3, 3, 2, 1),$$

as partitions of 9. We may also informally regard a partition $\lambda = (\lambda_1, \dots, \lambda_k)$ of n (say with $\lambda_k > 0$) as a way of writing n as a sum $\lambda_1 + \dots + \lambda_k$ of positive integers, *disregarding the order of the summands* (since there is a unique way of writing the summands in weakly decreasing order, where we don’t distinguish between equal summands). Compare with the definition of a composition of n , in which the order of the parts is essential. If λ is a partition of n , then we write either $\lambda \vdash n$ or $|\lambda| = n$. The nonzero terms λ_i are called the *parts* of λ , and we say that λ has k parts where $k = \#\{i : \lambda_i > 0\}$. The number of parts of λ is also called the *length* of λ and denoted $\ell(\lambda)$. If the partition λ has m_i parts equal to i , then we write $\lambda = \langle 1^{m_1}, 2^{m_2}, \dots \rangle$, where terms with $m_i = 0$ and the superscript $m_i = 1$ may be omitted. For instance,

$$(4, 4, 2, 2, 2, 1) = \langle 1^1, 2^3, 3^0, 4^2 \rangle = \langle 1, 2^3, 4^2 \rangle \vdash 15. \quad (1.74)$$

We also write $p(n)$ for the total number of partitions of n , $p_k(n)$ for the number of partitions of n with exactly k parts, and $p(j, k, n)$ for the number of partitions of n into at most k parts, with largest part at most j . For instance, there are seven partitions of 5, given by (omitting parentheses and commas from the notation) 5, 41, 32, 311, 221, 2111, 11111, so $p(5) = 7$, $p_1(5) = 1$, $p_2(5) = 2$, $p_3(5) = 2$, $p_4(5) = 1$, $p_5(5) = 1$, $p(3, 3, 5) = 3$, and so on. By convention we agree that $p_0(0) = p(0) = 1$. Note that $p_n(n) = 1$, $p_{n-1}(n) = 1$ if $n > 1$, $p_1(n) = 1$, $p_2(n) = \lfloor n/2 \rfloor$. It is easy to verify the recurrence

$$p_k(n) = p_{k-1}(n-1) + p_k(n-k),$$

which provides a convenient method for making a table of the numbers $p_k(n)$ for n, k small.

Let $(\lambda_1, \lambda_2, \dots) \vdash n$. The *Ferrers diagram* or *Ferrers graph* of λ is obtained by drawing a left-justified array of n dots with λ_i dots in the i th row. For instance, the Ferrers diagram of the partition 6655321 is given by Figure 1.13(a). If we replace the dots by juxtaposed squares, then we call the resulting diagram the *Young diagram* of λ . For instance, the Young diagram of 6655321 is given by Figure 1.13(b). We will have more to say about partitions in various places throughout this book and especially in the next two sections. However, we will not attempt a systematic investigation of this enormous and fascinating subject.

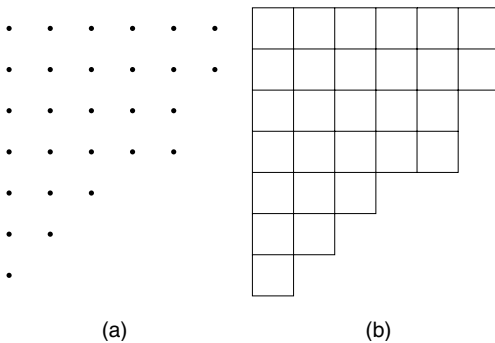


Figure 1.13 The Ferrers diagram and Young diagram of the partition 6655321.

The next result shows the relevance of partitions to the q -binomial coefficients.

1.7.3 Proposition. Fix $j, k \in \mathbb{N}$. Then

$$\sum_{n \geq 0} p(j, k, n) q^n = \binom{j+k}{j}_q.$$

Proof. While it is not difficult to give a proof by induction using (1.67), we prefer a direct combinatorial proof based on Proposition 1.7.2. To this end, let $m = j + k$ and recall from linear algebra that any k -dimensional subspace of \mathbb{F}_q^m (or of the m -dimensional vector space F^m over any field F) has a unique ordered basis (v_1, \dots, v_k) for which the matrix

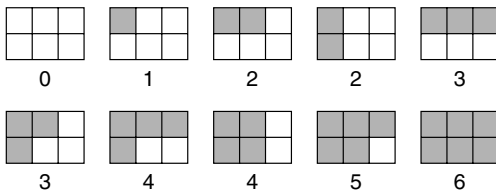
$$M = \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix} \quad (1.75)$$

is in *row-reduced echelon form*. This means: (a) the first nonzero entry of each v_i is a 1; (b) the first nonzero entry of v_{i+1} appears in a column to the right of the first nonzero entry of v_i ; and (c) in the column containing the first nonzero entry of v_i , all other entries are 0.

Now suppose that we are given an integer sequence $1 \leq a_1 < a_2 < \dots < a_k \leq m$, and consider all row-reduced echelon matrices (1.75) over \mathbb{F}_q for which the first nonzero entry of v_i occurs in the a_i th column. For instance, if $m = 7$, $k = 4$, and $(a_1, \dots, a_4) = (1, 3, 4, 6)$, then M has the form

$$\begin{bmatrix} 1 & * & 0 & 0 & * & 0 & * \\ 0 & 0 & 1 & 0 & * & 0 & * \\ 0 & 0 & 0 & 1 & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * \end{bmatrix}$$

where the symbol $*$ denotes an arbitrary entry of \mathbb{F}_q . The number λ_i of $*$'s in row i is $j - a_i + i$, and the sequence $(\lambda_1, \lambda_2, \dots, \lambda_k)$ defines a partition of some integer $n = \sum \lambda_i$ into at most k parts, with largest part at most j . The total number

Figure 1.14 Partitions in a 2×3 rectangle.

of matrices (1.75) with a_1, \dots, a_k specified as earlier is $q^{|\lambda|}$. Conversely, given any partition λ into at most k parts with largest part at most j , we can define $a_i = j - \lambda_i + i$, and there exists exactly $q^{|\lambda|}$ row-reduced matrices (1.75) with a_1, \dots, a_k having their meaning as earlier.

Since the number of row-reduced echelon matrices (1.75) is equal to the number $\binom{j+k}{k}$ of k -dimensional subspaces of \mathbb{F}_q^m , we get

$$\binom{j+k}{k} = \sum_{\substack{\lambda \\ \leq k \text{ parts} \\ \text{largest part} \leq j}} q^{|\lambda|} = \sum_{n \geq 0} p(j, k, n) q^n.$$

□

For readers familiar with this area, let us remark that the proof of Proposition 1.7.3 essentially constructs the well-known cellular decomposition of the Grassmann variety G_{km} .

The partitions λ enumerated by $p(j, k, n)$ may be described as those partitions of n whose Young diagram fits in a $k \times j$ rectangle. For instance, if $k = 2$ and $j = 3$, then Figure 1.14 shows the $\binom{5}{2} = 10$ partitions that fit in a 2×3 rectangle. The value of $|\lambda|$ is written beneath the diagram. It follows that

$$\binom{5}{2} = 1 + q + 2q^2 + 2q^3 + 2q^4 + q^5 + q^6.$$

It remains to relate Propositions 1.7.1 and 1.7.3 by showing that $p(j, k, n)$ is the number of permutations w of the multiset $M = \{1^j, 2^k\}$ with n inversions. Given a partition λ of n with at most k parts and largest part at most j , we will describe a permutation $w = w(\lambda) \in \mathfrak{S}_M$ with n inversions, leaving to the reader the easy proof that this correspondence is a bijection. Regard the Young diagram Y of λ as being contained in a $k \times j$ rectangle, and consider the lattice path L from the upper-right-hand corner to the lower-left-hand corner of the rectangle that travels along the boundary of Y . Walk along L , and write down a 1 whenever one takes a horizontal step and a 2 whenever one takes a vertical step. This process yields the desired permutation w . For instance, if $k = 3$, $j = 5$, $\lambda = 431$, then Figure 1.15 shows that path L and its labeling by 1's and 2's. We can also describe w by the condition that the 2's appear in positions $j - \lambda_i + i$, where $\lambda = (\lambda_1, \dots, \lambda_k)$.

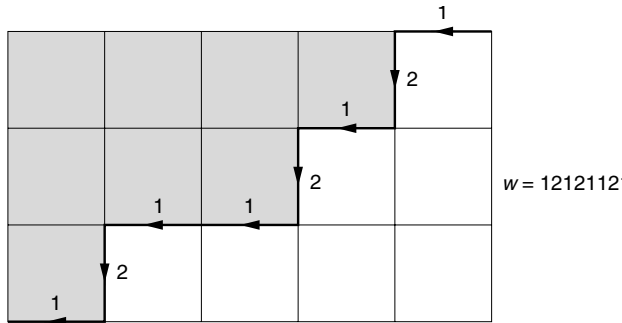


Figure 1.15 The lattice path associated with the partition 431.

1.8 Partition Identities

In the previous section, we defined a partition λ of $n \in \mathbb{N}$ and described its Ferrers diagram and Young diagram. In this section, we develop further the theory of partitions, in particular, the fascinating interaction between generating function identities and bijective proofs.

Let us begin by describing a fundamental involution on the set of partitions of n . Namely, if $\lambda \vdash n$, then define the *conjugate* partition λ' to be the partition whose Ferrers (or Young) diagram is obtained from that of λ by interchanging rows and columns. Equivalently, the diagram (Ferrers or Young) of λ' is the reflection of that of λ about the main diagonal. If $\lambda = (\lambda_1, \lambda_2, \dots)$, then the number of parts of λ' that equal i is $\lambda_i - \lambda_{i+1}$. This description of λ' provides a convenient method of computing λ' from λ without drawing a diagram. For instance, if $\lambda = (4, 3, 1, 1, 1)$, then $\lambda' = (5, 2, 2, 1)$.

Recall that $p_k(n)$ denotes the number of partitions of n into k parts. Similarly, let $p_{\leq k}(n)$ denote the number of partitions of n into at most k parts, that is, $p_{\leq k}(n) = p_0(n) + p_1(n) + \dots + p_k(n)$. Now λ has at most k parts if and only if λ' has largest part at most k . This observation enables us to compute the generating function $\sum_{n \geq 0} p_{\leq k}(n)q^n$. A partition of n with largest part at most k may be regarded as a solution in nonnegative integers to $m_1 + 2m_2 + \dots + km_k = n$. Here m_i is the number of times that the part i appears in the partition λ , that is, $\lambda = \langle 1^{m_1} 2^{m_2} \dots k^{m_k} \rangle$. Hence,

$$\begin{aligned}
 \sum_{n \geq 0} p_{\leq k}(n)q^n &= \sum_{n \geq 0} \sum_{m_1 + \dots + km_k = n} q^n \\
 &= \sum_{m_1 \geq 0} \sum_{m_2 \geq 0} \dots \sum_{m_k \geq 0} q^{m_1 + 2m_2 + \dots + km_k} \\
 &= \left(\sum_{m_1 \geq 0} q^{m_1} \right) \left(\sum_{m_2 \geq 0} q^{2m_2} \right) \dots \left(\sum_{m_k \geq 0} q^{km_k} \right) \\
 &= \frac{1}{(1-q)(1-q^2) \dots (1-q^k)}. \tag{1.76}
 \end{aligned}$$

This computation is just a precise way of writing the intuitive fact that the most natural way of computing the coefficient of q^n in $1/(1-q)(1-q^2)\cdots(1-q^k)$ entails computing all the partitions of n with largest part at most k . If we let $k \rightarrow \infty$, then we obtain the famous generating function

$$\sum_{n \geq 0} p(n)q^n = \prod_{i \geq 1} \frac{1}{1-q^i}. \quad (1.77)$$

Equations (1.76) and (1.77) can be considerably generalized. The following result, although by no means the most general possible, will suffice for our purposes.

1.8.1 Proposition. *For each $i \in \mathbb{P}$, fix a set $S_i \subseteq \mathbb{N}$. Let $\mathcal{S} = (S_1, S_2, \dots)$, and define $P(\mathcal{S})$ to be the set of all partitions λ such that if the part i occurs $m_i = m_i(\lambda)$ times, then $m_i \in S_i$. Define the generating function in the variables $\mathbf{q} = (q_1, q_2, \dots)$,*

$$F(\mathcal{S}, \mathbf{q}) = \sum_{\lambda \in P(\mathcal{S})} q_1^{m_1(\lambda)} q_2^{m_2(\lambda)} \cdots.$$

Then

$$F(\mathcal{S}, \mathbf{q}) = \prod_{i \geq 1} \left(\sum_{j \in S_i} q_i^j \right). \quad (1.78)$$

Proof. The reader should be able to see the validity of this result by “inspection.” The coefficient of $q_1^{m_1} q_2^{m_2} \cdots$ in the right-hand side of (1.78) is 1 if each $m_i \in S_i$, and 0 otherwise, which yields the desired result. \square

1.8.2 Corollary. *Preserve the notation of the previous proposition, and let $p(\mathcal{S}, n)$ denote the number of partitions of n belonging to $P(\mathcal{S})$, that is,*

$$p(\mathcal{S}, n) = \#\{\lambda \vdash n : \lambda \in P(\mathcal{S})\}.$$

Then

$$\sum_{n \geq 0} p(\mathcal{S}, n)q^n = \prod_{i \geq 1} \left(\sum_{j \in S_i} q^{ij} \right).$$

Proof. Put each $q_i = q^i$ in Proposition 1.8.1. \square

Let us now give a sample of some of the techniques and results from the theory of partitions. First, we give an idea of the usefulness of Young diagrams and Ferrers diagrams.

1.8.3 Proposition. *For any partition $\lambda = (\lambda_1, \lambda_2, \dots)$ we have*

$$\sum_{i \geq 1} (i-1)\lambda_i = \sum_{i \geq 1} \binom{\lambda'_i}{2}. \quad (1.79)$$

Proof. Place an $i - 1$ in each square of row i of the Young diagram of λ . For instance, if $\lambda = 5322$ we get

0	0	0	0	0
1	1	1		
2	2			
3	3			

If we add up all the numbers in the diagram by rows, then we obtain the left-hand side of (1.79). If we add up by columns, then we obtain the right-hand side. \square

1.8.4 Proposition. Let $c(n)$ denote the number of self-conjugate partitions λ of n , that is, $\lambda = \lambda'$. Then

$$\sum_{n \geq 0} c(n)q^n = (1+q)(1+q^3)(1+q^5)\cdots. \quad (1.80)$$

Proof. Let λ be a self-conjugate partition. Consider the “diagonal hooks” of the Ferrers diagram of $\lambda \vdash n$, as illustrated in Figure 1.16 for the partition $\lambda = 54431$. The number of dots in each hook form a partition μ of n into distinct odd parts. For Figure 1.16 we have $\mu = 953$. The map $\lambda \mapsto \mu$ is easily seen to be a bijection from self-conjugate partitions of n to partitions of n into distinct odd parts. The proof now follows from the special case $S_i = \{0, 1\}$ if i is odd, and $S_i = \{0\}$ if i is even, of Corollary 1.8.2 (though it should be obvious by inspection that the right-hand side of (1.80) is the generating function for the number of partitions of n into distinct odd parts). \square

There are many results in the theory of partitions that assert the equicardinality of two classes of partitions. The quintessential example is given by the following result.

1.8.5 Proposition. Let $q(n)$ denote the number of partitions of n into distinct parts and $p_{\text{odd}}(n)$ the number of partitions of n into odd parts. Then $q(n) = p_{\text{odd}}(n)$ for all $n \geq 0$.

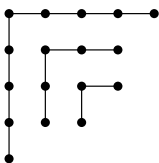


Figure 1.16 The diagonal hooks of the self-conjugate partition 54431.

First Proof (generating functions). Setting each $S_i = \{0, 1\}$ in Corollary 1.8.2 (or by direct inspection), we have

$$\begin{aligned} \sum_{n \geq 0} q(n)q^n &= (1+q)(1+q^2)(1+q^3)\cdots \\ &= \frac{1-q^2}{1-q} \cdot \frac{1-q^4}{1-q^2} \cdot \frac{1-q^6}{1-q^3} \cdots \\ &= \frac{\prod_{n \geq 1} (1-q^{2n})}{\prod_{n \geq 1} (1-q^n)} \\ &= \frac{1}{(1-q)(1-q^3)(1-q^5)\cdots}. \end{aligned} \quad (1.81)$$

Again by Corollary 1.8.2 or by inspection, we have

$$\frac{1}{(1-q)(1-q^3)(1-q^5)\cdots} = \sum_{n \geq 0} p_{\text{odd}}(n)q^n,$$

and the proof follows. \square

Second Proof (bijective). Naturally a combinatorial proof of such a simple and elegant result is desired. Perhaps the simplest is the following. Let λ be a partition of n into odd parts, with the part $2j-1$ occurring r_j times. Define a partition μ of n into distinct parts by requiring that the part $(2j-1)2^k$, $k \geq 0$, appears in μ if and only the binary expansion of r_j contains the term 2^k . We leave the reader to check the validity of this bijection, which rests on the fact that every positive integer can be expressed uniquely as a product of an odd positive integer and a power of 2. For instance, if $\lambda = \langle 9^5, 5^{12}, 3^2, 1^3 \rangle \vdash 114$, then

$$\begin{aligned} 114 &= 9(1+4) + 5(4+8) + 3(2) + 1(1+2) \\ &= 9 + 36 + 20 + 40 + 6 + 1 + 2, \end{aligned}$$

so $\mu = (40, 36, 20, 9, 6, 2, 1)$. \square

Third Proof (bijective). There is a completely different bijective proof which is a good example of “diagram cutting.” Identify a partition λ into odd parts with its Ferrers diagram. Take each row of λ , convert it into a self-conjugate hook, and arrange these hooks diagonally in decreasing order. Now connect the upper-left-hand corner u with all dots in the “shifted hook” of u , consisting of all dots directly to the right of u and directly to the southeast of u . For the dot v directly below u (when $|\lambda| > 1$), connect it to all the dots in the conjugate shifted hook of u . Now take the northwest-most remaining dot above the main diagonal and connect it to its shifted hook, and similarly connect the northwest-most dot below the main diagonal with its conjugate shifted hook. Continue until all the entire diagram has been partitioned into shifted hooks and conjugate shifted hooks. The number of dots in these hooks form the parts of a partition μ of n into distinct parts.

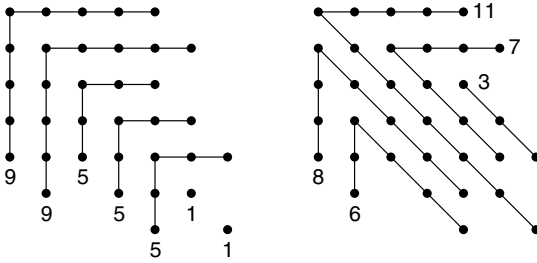


Figure 1.17 A second bijective proof that $q(n) = p_{\text{odd}}(n)$.

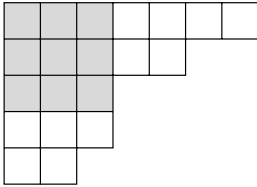


Figure 1.18 The Durfee square of the partition 75332.

Figure 1.17 shows the case $\lambda = 9955511$ and $\mu = (11, 8, 7, 6, 3)$. We trust that this figure will make the preceding rather vague description of the map $\lambda \mapsto \mu$ clear. It is easy to check that this map is indeed a bijection from partitions of n into odd parts to partitions of n into distinct parts. \square

There are many combinatorial identities asserting that a product is equal to a sum that can be interpreted in terms of partitions. We give three of the simplest in proposition 1.8.6, relegating some more interesting and subtle identities to the exercises. The second identity is related to the concept of the *rank* $\text{rank}(\lambda)$ of a partition $\lambda = (\lambda_1, \lambda_2, \dots)$, defined to be the largest i for which $\lambda_i \geq i$. Equivalently, $\text{rank}(\lambda)$ is the length of the main diagonal in the (Ferrers or Young) diagram of λ . It is also the side length of the largest square in the diagram of λ . We can place this square to include the first dot or box in the first row of the diagram, in which case it is called the *Durfee square* of λ . Figure 1.18 shows the Young diagram of the partition $\lambda = 75332$ of rank 3, with the Durfee square shaded.

1.8.6 Proposition. (a) *We have*

$$\frac{1}{\prod_{i \geq 1} (1 - xq^i)} = \sum_{k \geq 0} \frac{x^k q^{k^2}}{(1 - q)(1 - q^2) \cdots (1 - q^k)}. \quad (1.82)$$

(b) *We have*

$$\frac{1}{\prod_{i \geq 1} (1 - xq^i)} = \sum_{k \geq 0} \frac{x^k q^{k^2}}{(1 - q) \cdots (1 - q^k)(1 - xq) \cdots (1 - xq^k)}.$$

(c) We have

$$\prod_{i \geq 1} (1 + xq^i) = \sum_{k \geq 0} \frac{x^k q^{\binom{k+1}{2}}}{(1-q)(1-q^2) \cdots (1-q^k)}. \quad (1.83)$$

Proof. (a) It should be clear by inspection that

$$\frac{1}{\prod_{i \geq 1} (1 - xq^i)} = \sum_{\lambda} x^{\ell(\lambda)} q^{|\lambda|}, \quad (1.84)$$

where λ ranges over all partitions of all $n \geq 0$. We can obtain λ by first choosing $\ell(\lambda) = k$. It follows from equation (1.76) that

$$\sum_{\substack{\lambda \\ \ell(\lambda)=k}} q^{|\lambda|} = \frac{q^k}{(1-q)(1-q^2) \cdots (1-q^k)},$$

and the proof follows.

We should also indicate how (1.82) can be proved nonbijectively, since the technique is useful in other situations. Let

$$F(x, q) = \frac{1}{\prod_{i \geq 1} (1 - xq^i)}.$$

Clearly, $F(x, q) = F(xq, q)/(1 - xq)$, and $F(x, q)$ is uniquely determined by this functional equation and the initial condition $F(0, q) = 1$. Now let

$$G(x, q) = \sum_{k \geq 0} \frac{x^k q^k}{(1-q)(1-q^2) \cdots (1-q^k)}.$$

Then

$$\begin{aligned} G(xq, q) &= \sum_{k \geq 0} \frac{x^k q^{2k}}{(1-q)(1-q^2) \cdots (1-q^k)} \\ &= \sum_{k \geq 0} \frac{x^k q^k}{(1-q)(1-q^2) \cdots (1-q^{k-1})} \left(\frac{1}{1-q^k} - 1 \right) \\ &= G(x, q) - xqG(x, q) \\ &= (1 - xq)G(x, q). \end{aligned}$$

Since $G(x, 0) = 1$, the proof follows.

(b) Again we use (1.84), but now the terms on the right-hand side will correspond to $\text{rank}(\lambda)$ rather than $\ell(\lambda)$. If $\text{rank}(\lambda) = k$, then when we remove the Durfee square from the diagram of λ , we obtain disjoint diagrams of partitions μ and ν such that $\ell(\mu) \leq k$ and $\nu_1 = \ell(\nu') \leq k$. (For the partition $\lambda = 75332$ of Figure 1.18 we have

$\mu = 42$ and $\nu = 32$.) Every λ of rank k is obtained uniquely from such μ and ν . Moreover, $|\lambda| = k^2 + |\mu| + |\nu|$ and $\ell(\lambda) = k + \ell(\nu)$. It follows that

$$\sum_{\substack{\lambda \\ \text{rank}(\lambda)=k}} x^{\ell(\lambda)} q^{|\lambda|} = x^k q^{k^2} \frac{1}{(1-q) \cdots (1-q^k)} \cdot \frac{1}{(1-xq) \cdots (1-xq^k)}.$$

Summing over all $k \geq 0$ completes the proof.

(c) Now the coefficient of $x^k q^n$ in the left-hand side is the number of partitions of n into k distinct parts $\lambda_1 > \cdots > \lambda_k > 0$. Then $(\lambda_1 - k, \lambda_2 - k + 1, \dots, \lambda_k - 1)$ is a partition of $n - \binom{k+1}{2}$ into at most k parts, from which the proof follows easily. \square

The generating function (obtained, e.g., from (1.82) by substituting x/q for x , or by a simple modification of either of our two proofs of Proposition 1.8.6(a))

$$\frac{1}{\prod_{i \geq 0} (1 - xq^i)} = \sum_{k \geq 0} \frac{x^k}{(1-q)(1-q^2) \cdots (1-q^k)}$$

is known as the q -exponential function, since $(1-q)(1-q^2) \cdots (1-q^n) = (1-q)^n (n)!$. We could even replace x with $(1-q)x$, getting

$$\frac{1}{\prod_{i \geq 0} (1 - x(1-q)q^i)} = \sum_{k \geq 0} \frac{x^k}{(k)!}. \quad (1.85)$$

The right-hand side reduces to e^x upon setting $q = 1$, though we cannot also substitute $q = 1$ on the left-hand side to obtain e^x . It is an instructive exercise (Exercise 1.101) to work out why this is the case. In other words, why does substituting $(1-q)x$ for x and then setting $q = 1$ in two expressions for the same power series not maintain the equality of the two series?

A generating function of the form

$$F(x) = \sum_{n \geq 0} a_n \frac{x^n}{(1-q)(1-q^2) \cdots (1-q^n)}$$

is called an *Eulerian* or q -exponential generating function. It is the natural q -analogue of an exponential generating function. We could just as well use

$$F(x(1-q)) = \sum_{n \geq 0} a_n \frac{x^n}{(n)!} \quad (1.86)$$

in place of $F(x)$. The use of $F(x)$ is traditional, though $F(x(1-q))$ is more natural combinatorially and has the virtue that setting $q = 1$ in the right-hand side of (1.86) gives an exponential generating function. We will see especially in the general theory of generating functions developed in Section 3.18 why the right-hand side of (1.86) is combinatorially “natural.”

Propositions 1.8.6(a) and (c) have interesting “finite versions,” where in addition to the number of parts we also restrict the largest part. Recall that $p(j, k, n)$ denotes the number of partitions $\lambda \vdash n$ for which $\lambda_1 \leq j$ and $\ell(\lambda) \leq k$. The proof of Proposition 1.8.6(a) then generalizes *mutatis mutandis* to give the following formula:

$$\frac{1}{\prod_{i=0}^j (1 - xq^i)} = \sum_{k \geq 0} x^k \sum_{n \geq 0} p(j, k, n) q^n$$

$$= \sum_{k \geq 0} x^k \binom{j+k}{j}.$$

By exactly the same reasoning, using the proof of Proposition 1.8.6(c), we obtain

$$\prod_{i=0}^{j-1} (1 + xq^i) = \sum_{k=0}^j x^k q^{\binom{k}{2}} \binom{j}{k}. \quad (1.87)$$

Equation (1.87) is known as the *q-binomial theorem*, since setting $q = 1$ gives the usual binomial theorem. It is a good illustration of the difficulty of writing down a *q*-analogue of an identity by inspection; it is difficult to predict without any prior insight why the factor $q^{\binom{k}{2}}$ appears in the terms on the right.

Of course, there are many other ways to prove the *q*-binomial theorem, including a straightforward induction on j . We can also give a finite field proof, where we regard q as a prime power. For each factor $1 + xq^i$ of the left-hand side of (1.87), choose either the term 1 or xq^i . If the latter, then choose a row vector γ_i of length j whose first nonzero coordinate is a 1, which occurs in the $(j - i)$ th position. Thus, there are q^i choices for γ_i . After making this choice for all i , let V be the span in \mathbb{F}_q^j of the chosen γ_i 's. If we chose k of the γ_i 's, then $\dim V = k$. Let M be the $k \times j$ matrix whose rows are the γ_i 's in decreasing order of the index i . There is a unique $k \times k$ upper unitriangular matrix T (i.e., T is upper triangular with 1's on the main diagonal) for which TM is in row-reduced echelon form. Reversing these steps, for each k -dimensional subspace V of \mathbb{F}_q^j , let A be the unique $k \times j$ matrix in row-reduced echelon form whose row space is V . There are $q^{\binom{k}{2}}$ $k \times k$ upper unitriangular matrices T^{-1} , and for each of them the rows of $M = T^{-1}A$ define a choice of γ_i 's. It follows that we obtain every k -dimensional subspace of \mathbb{F}_q^j as a span of γ_i 's exactly $q^{\binom{k}{2}}$ times, and the proof follows.

Variant. There is a slight variant of the previous finite field proof of (1.87), which has less algebraic significance but is more transparent combinatorially. Namely, once we have chosen the $k \times j$ matrix M , change every entry above the first 1 in any row to 0. We then obtain a matrix in row-reduced echelon form. There are $\binom{k}{2}$ entries of M that are changed to 0, so we get every row-reduced echelon matrix with k rows exactly $q^{\binom{k}{2}}$ times. The proof follows as before.

For yet another proof of equation (1.87) based on finite fields, see Exercise 3.119.

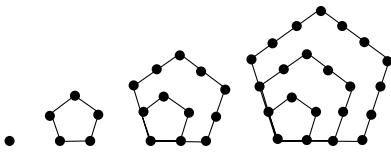


Figure 1.19 The pentagonal numbers 1, 5, 12, 22.

We next turn to a remarkable product expansion related to partitions. It is the archetype for a vast menagerie of similar results. We will give only a bijective proof; it is also an interesting challenge to find an algebraic proof. The result is called the *pentagonal number formula* or *pentagonal number theorem* because of the appearance of the numbers $k(3k-1)/2$, which are known as *pentagonal numbers*. See Figure 1.19 for an explanation of this terminology.

1.8.7 Proposition. *We have*

$$\prod_{k \geq 1} (1 - x^k) = \sum_{n \in \mathbb{Z}} (-1)^n x^{n(3n-1)/2} \quad (1.88)$$

$$= 1 + \sum_{n \geq 1} (-1)^n \left(x^{n(3n-1)/2} + x^{n(3n+1)/2} \right) \quad (1.89)$$

$$= 1 - x - x^2 + x^5 + x^7 - x^{12} - x^{15} + x^{22} + x^{26} - \dots$$

Proof. Let $f(n) = q_e(n) - q_o(n)$, where $q_e(n)$ (respectively, $q_o(n)$) is the number of partitions of n into an even (respectively, odd) number of distinct parts. It should be clear that

$$\prod_{k \geq 1} (1 - x^k) = \sum_{n \geq 0} f(n) x^n.$$

Hence we need to show that

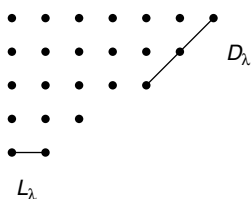
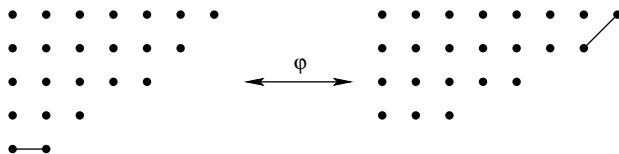
$$f(n) = \begin{cases} (-1)^k, & \text{if } n = k(3k \pm 1)/2, \\ 0, & \text{otherwise.} \end{cases} \quad (1.90)$$

Let $Q(n)$ denote the set of all partitions of n into distinct parts. We will prove (1.90) when $n \neq k(3k \pm 1)/2$ by defining an involution $\varphi : Q(n) \rightarrow Q(n)$ such that $\ell(\lambda) \not\equiv \ell(\varphi(\lambda)) \pmod{2}$ for all $\lambda \in Q(n)$. When $n = k(3k \pm 1)/2$, we will define a partition $\mu \in Q(n)$ and an involution $\varphi : Q(n) - \{\mu\} \rightarrow Q(n) - \{\mu\}$ such that $\ell(\lambda) \not\equiv \ell(\varphi(\lambda)) \pmod{2}$ for all $\lambda \in Q(n) - \{\mu\}$, and moreover $\ell(\mu) = k$. Such a method of proof is called a *sign-reversing involution* argument. The involution φ changes the sign of $(-1)^{\ell(\lambda)}$ and hence cancels out all terms in the expansion

$$\sum_{\lambda \in Q(n)} (-1)^{\ell(\lambda)}$$

except those terms indexed by partitions λ not in the domain of φ . These partitions form a much smaller set that can be analyzed separately.

The definition of φ is quite simple. Let L_λ denote the last row of the Ferrers diagram of λ , and let D_λ denote the set of last elements of all rows i for which

Figure 1.20 The sets L_λ and D_λ for $\lambda = 76532$.Figure 1.21 The involution φ from the proof of the Pentagonal Number Formula.

$\lambda_i = \lambda_1 - i + 1$. Figure 1.20 shows L_λ and D_λ for $\lambda = 76532$. If $\#D_\lambda < \#L_\lambda$, define $\varphi(\lambda)$ to be the partition obtained from (the Ferrers diagram of) λ by removing D_λ and replacing it under L_λ to form a new row. Similarly, if $\#L_\lambda \leq \#D_\lambda$, define $\varphi(\lambda)$ to be the partition obtained from (the Ferrers diagram of) λ by removing L_λ and replacing it parallel and to the right of D_λ , beginning at the top row. Clearly, $\varphi(\lambda) = \mu$ if and only if $\varphi(\mu) = \lambda$. See Figure 1.21 for the case $\lambda = 76532$ and $\mu = 8753$. It is evident that φ is an involution where it is defined; the problem is that the diagram defined by $\varphi(\lambda)$ may not be a valid Ferrers diagram. A little thought shows that there are exactly two situations when this is the case. The first case occurs when λ has the form $(2k-1, 2k-2, \dots, k)$. In this case $|\lambda| = k(3k-1)/2$ and $\ell(\lambda) = k$. The second bad case is $\lambda = (2k, 2k-1, \dots, k+1)$. Now $|\lambda| = k(3k+1)/2$ and $\ell(\lambda) = k$. Hence, φ is a sign-reversing involution on all partitions λ , with the exception of a single partition of length k of numbers of the form $k(3k \pm 1)/2$, and the proof follows. \square

We can rewrite the Pentagonal Number Formula (1.88) in the form

$$\left(\sum_{n \geq 0} p(n)x^n \right) \left(\sum_{n \in \mathbb{Z}} (-1)^n x^{n(3n-1)/2} \right) = 1.$$

If we equate coefficients of x^n on both sides, then we obtain a recurrence for $p(n)$:

$$p(n) = p(n-1) + p(n-2) - p(n-5) - p(n-7) + p(n-12) + p(n-15) - \dots \quad (1.91)$$

It is understood that $p(n) = 0$ for $n < 0$, so the number of terms on the right-hand side is roughly $2\sqrt{2n/3}$. For instance,

$$\begin{aligned} p(20) &= p(19) + p(18) - p(15) - p(13) + p(8) + p(5) \\ &= 490 + 385 - 176 - 101 + 22 + 7 \\ &= 627. \end{aligned}$$

Equation (1.91) affords the most efficient known method to compute all the numbers $p(1), p(2), \dots, p(n)$ for given n . Much more sophisticated methods (discussed briefly later) are known for computing $p(n)$ that don't involve computing smaller values. It is known, for instance, that

$$p(10^4) = 36167251325636293988820471890953695495016030339315650422081868605887952568754066420592310556052906916435144.$$

In fact, $p(10^{15})$ can be computed exactly, a number with exactly 35,228,031 decimal digits.

It is natural to ask for the rate of growth of $p(n)$. To this end we mention without proof the famous asymptotic formula

$$p(n) \sim \frac{e^{\pi\sqrt{2n/3}}}{4\sqrt{3}n}. \quad (1.92)$$

For instance, when $n = 100$ the ratio of the right-hand side to the left is $1.0457\dots$, whereas when $n = 1000$, it is $1.0141\dots$. When $n = 10000$ the ratio is $1.00444\dots$. There is in fact an asymptotic series for $p(n)$ that actually converges rapidly to $p(n)$. (Typically, an asymptotic series is divergent.) This asymptotic series is the best-known method for evaluating $p(n)$ for large n .

1.9 The Twelfold Way

In this section we will be concerned with counting functions between two sets. Let N and X be finite sets with $\#N = n$ and $\#X = x$. We wish to count the number of functions $f : N \rightarrow X$ subject to certain restrictions. There will be three restrictions on the functions themselves and four restrictions on when we consider two functions to be the same. This gives a total of twelve counting problems, whose solution is called the *Twelfold Way*.

The three restrictions on the functions $f : N \rightarrow X$ are the following.

- (a) f is arbitrary (no restriction).
- (b) f is injective (one-to-one).
- (c) f is surjective (onto).

The four interpretations as to when two functions are the same (or *equivalent*) come about from regarding the elements of N and X as “distinguishable” or “indistinguishable.” Think of N as a set of balls and X as a set of boxes. A function $f : N \rightarrow X$ consists of placing each ball into some box. If we can tell the balls apart, then the elements of N are called *distinguishable*, otherwise *indistinguishable*. Similarly if we can tell the boxes apart, then elements of X are called *distinguishable*, otherwise *indistinguishable*. For example, suppose $N = \{1, 2, 3\}$,

$X = \{a, b, c, d\}$, and define functions $f, g, h, i : N \rightarrow X$ by

$$\begin{aligned} f(1) &= f(2) = a, & f(3) &= b, \\ g(1) &= g(3) = a, & g(2) &= b, \\ h(1) &= h(2) = b, & h(3) &= d, \\ i(2) &= i(3) = b, & i(1) &= c. \end{aligned}$$

If the elements of both N and X are distinguishable, then the functions have the “pictures” shown by Figure 1.22. All four pictures are different, and the four functions are inequivalent. Now suppose that the elements of N (but not X) are indistinguishable. This assumption corresponds to erasing the labels on the balls. The pictures for f and g both become as shown in Figure 1.23, so f and g are equivalent. However, f , h , and i remain inequivalent. If the elements of X (but not N) are indistinguishable, then we erase the labels on the boxes. Thus, f and h both have the picture shown in Figure 1.24. (The order of the boxes is irrelevant if we can’t tell them apart.) Hence, f and h are equivalent, but f , g , and i are inequivalent. Finally, if the elements of both N and X are indistinguishable, then all four functions have the picture shown in Figure 1.25, so all four are equivalent.

A rigorous definition of the above notions of equivalence is desirable. Two functions $f, g : N \rightarrow X$ are said to be *equivalent with N indistinguishable* if there is a bijection $u : N \rightarrow N$ such that $f(u(a)) = g(a)$ for all $a \in N$. Similarly, f and g are *equivalent with X indistinguishable* if there is a bijection $v : X \rightarrow X$ such that $v(f(a)) = g(a)$ for all $a \in N$. Finally, f and g are *equivalent with N and X indistinguishable* if there are bijections $u : N \rightarrow N$ and $v : X \rightarrow X$ such



Figure 1.22 Four functions with distinguishable balls and boxes.

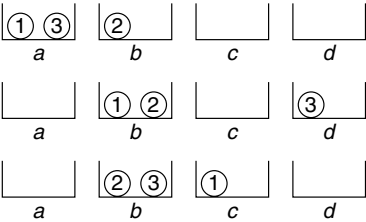


Figure 1.23 Balls indistinguishable.



Figure 1.24 Boxes indistinguishable.



Figure 1.25 Balls and boxes indistinguishable.

that $v(f(u(a))) = g(a)$ for all $a \in N$. These three notions of equivalence are all equivalence relations, and the number of “different” functions with respect to one of these equivalences simply means the number of equivalence classes. If f and g are equivalent (in any of the foregoing ways), then f is injective (respectively, surjective) if and only if g is injective (respectively, surjective). We, therefore, say that the notions of injectivity and surjectivity are *compatible* with the equivalence relation. By the “number of inequivalent injective functions $f : N \rightarrow X$,” we mean the number of equivalence classes all of whose elements are injective.

We are now ready to present the Twelfold Way. The twelve entries are numbered and will be discussed individually. The table gives the number of inequivalent functions $f : N \rightarrow X$ of the appropriate type, where $\#N = n$ and $\#X = x$.

The Twelfold Way

Elements of N	Elements of X	Any f	Injective f	Surjective f
dist.	dist.	1. x^n	2. $(x)_n$	3. $x!S(n, x)$
indist.	dist.	4. $\left(\binom{x}{n}\right)$	5. $\binom{x}{n}$	6. $\left(\binom{x}{n-x}\right)$
dist.	indist.	7. $S(n, 0) + S(n, 1) + \cdots + S(n, x)$	8. 1 if $n \leq x$ 0 if $n > x$	9. $S(n, x)$
indist.	indist.	10. $p_0(n) + p_1(n) + \cdots + p_x(n)$	11. 1 if $n \leq x$ 0 if $n > x$	12. $p_x(n)$

Discussion of Twelfold Way Entries

- For each $a \in N$, $f(a)$ can be any of the x elements of X . Hence, there are x^n functions.
- Say $N = \{a_1, \dots, a_n\}$. Choose $f(a_1)$ in x ways, then $f(a_2)$ in $x - 1$ ways, and so on, giving $x(x - 1) \cdots (x - n + 1) = (x)_n$ choices in all.
- * A *partition* of a finite set N is a collection $\pi = \{B_1, B_2, \dots, B_k\}$ of subsets of N such that
 - $B_i \neq \emptyset$ for each i ,
 - $B_i \cap B_j = \emptyset$ if $i \neq j$,
 - $B_1 \cup B_2 \cup \cdots \cup B_k = N$.

(Contrast this definition with that of an *ordered partition* in the proof of Lemma 1.4.11, for which the subsets B_1, \dots, B_k are linearly ordered.) We call B_i a *block* of π , and we say that π has k blocks, denoted $|\pi| = k$. Define $S(n, k)$ to be the number of partitions of an n -set into k -blocks. The number $S(n, k)$ is called a *Stirling number of the second kind*. (Stirling numbers of the first kind were defined preceding Lemma 1.3.6.) By convention, we put $S(0, 0) = 1$. We use

* Discussion of entry 4 begins on page 79.

notation such as 135-26-4 to denote the partition of [6] with blocks $\{1, 3, 5\}$, $\{2, 6\}$, $\{4\}$. For instance, $S(4, 2) = 7$, corresponding to the partitions 123-4, 124-3, 134-2, 234-1, 12-34, 13-24, 14-23. The reader should check that for $n \geq 1$, $S(n, k) = 0$ if $k > n$, $S(n, 0) = 0$, $S(n, 1) = 1$, $S(n, 2) = 2^{n-1} - 1$, $S(n, n) = 1$, $S(n, n-1) = \binom{n}{2}$, and $S(n, n-2) = \binom{n}{3} + 3\binom{n}{4}$. (See Exercise 43.)

NOTE. There is a simple bijection between the equivalence relations \sim on a set X (which may be infinite) and the partitions of X , namely, the equivalence classes of \sim form a partition of X .

The Stirling numbers of the second kind satisfy the following basic recurrence:

$$S(n, k) = kS(n-1, k) + S(n-1, k-1). \quad (1.93)$$

Equation (1.93) is proved as follows. To obtain a partition of $[n]$ into k blocks, we can partition $[n-1]$ into k blocks and place n into any of these blocks in $kS(n-1, k)$ ways, or we can put n in a block by itself and partition $[n-1]$ into $k-1$ blocks in $S(n-1, k-1)$ ways. Hence, (1.93) follows. The recurrence (1.93) allows one to prove by induction many results about the numbers $S(n, k)$, though frequently there will be preferable combinatorial proofs. The *total* number of partitions of an n -set is called a *Bell number* and is denoted $B(n)$. Thus, $B(n) = \sum_{k=1}^n S(n, k)$, $n \geq 1$. The values of $B(n)$ for $1 \leq n \leq 10$ are given by the following table.

n	1	2	3	4	5	6	7	8	9	10
$B(n)$	1	2	5	15	52	203	877	4140	21147	115975

The following is a list of some basic formulas concerning $S(n, k)$ and $B(n)$:

$$S(n, k) = \frac{1}{k!} \sum_{i=0}^k (-1)^{k-i} \binom{k}{i} i^n, \quad (1.94a)$$

$$\sum_{n \geq k} S(n, k) \frac{x^n}{n!} = \frac{1}{k!} (e^x - 1)^k, \quad k \geq 0, \quad (1.94b)$$

$$\sum_{n \geq k} S(n, k) x^n = \frac{x^k}{(1-x)(1-2x) \cdots (1-kx)}, \quad (1.94c)$$

$$x^n = \sum_{k=0}^n S(n, k) (x)_k, \quad (1.94d)$$

$$B(n+1) = \sum_{i=0}^n \binom{n}{i} B(i), \quad n \geq 0, \quad (1.94e)$$

$$\sum_{n \geq 0} B(n) \frac{x^n}{n!} = \exp(e^x - 1). \quad (1.94f)$$

We now indicate the proofs of (1.94a)–(1.94f). For all except (1.94d), we describe noncombinatorial proofs, though with a bit more work combinatorial proofs

can be given (see e.g. Example 5.2.4). Let $F_k(x) = \sum_{n \geq k} S(n, k)x^n/n!$. Clearly, $F_0(x) = 1$. From (1.93) we have

$$F_k(x) = k \sum_{n \geq k} S(n-1, k) \frac{x^n}{n!} + \sum_{n \geq k} S(n-1, k-1) \frac{x^n}{n!}.$$

Differentiate both sides to obtain

$$F'_k(x) = kF_k(x) + F_{k-1}(x). \quad (1.95)$$

Assume by induction that $F_{k-1}(x) = \frac{1}{(k-1)!}(e^x - 1)^{k-1}$. Then the unique solution to (1.95) whose coefficient of x^k is $1/k!$ is given by $F_k(x) = \frac{1}{k!}(e^x - 1)^k$. Hence (1.94b) is true by induction. To prove (1.94a), write

$$\frac{1}{k!}(e^x - 1)^k = \frac{1}{k!} \sum_{j=0}^k (-1)^{k-j} \binom{k}{j} e^{jx}$$

and extract the coefficient of x^n . To prove (1.94f), sum (1.94b) on k to obtain

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

Equation (1.94e) may be proved by differentiating (1.94f) and comparing coefficients, and it is also quite easy to give a direct combinatorial proof (Exercise 107). Equation (1.94c) is proved analogously to our proof of (1.94b) and can also be given a proof analogous to that of Proposition 1.3.7 (Exercise 45). It remains to prove (1.94d), and this will be done following the next paragraph.

We now verify entry 3 of the Twelfold Way. We have to show that the number of surjective functions $f: N \rightarrow X$ is $x!S(n, x)$. Now $x!S(n, x)$ counts the number of ways of partitioning N into x blocks and then linearly ordering the blocks, say (B_1, B_2, \dots, B_x) . Let $X = \{b_1, b_2, \dots, b_x\}$. We associate the ordered partition (B_1, B_2, \dots, B_x) with the surjective function $f: N \rightarrow X$ defined by $f(i) = b_j$ if $i \in B_j$. (More succinctly, we can write $f(B_j) = b_j$.) This establishes the desired correspondence.

We can now give a simple combinatorial proof of (1.94d). The left-hand side is the total number of functions $f: N \rightarrow X$. Each such function is surjective onto a unique subset $Y = f(N)$ of X satisfying $\#Y \leq n$. If $\#Y = k$, then there are $k!S(n, k)$ such functions, and there are $\binom{x}{k}$ choices of subsets Y of X with $\#Y = k$. Hence,

$$x^n = \sum_{k=0}^n k!S(n, k) \binom{x}{k} = \sum_{k=0}^n S(n, k)(x)_k. \quad (1.96)$$

Equation (1.94d) has the following additional interpretation. The set $\mathcal{P} = K[x]$ of all polynomials in the indeterminate x with coefficients in the field K forms a vector space over K . The sets $B_1 = \{1, x, x^2, \dots\}$ and $B_2 = \{1, (x)_1, (x)_2, \dots\}$ are both bases for \mathcal{P} . Then (1.94d) asserts that the (infinite) matrix $S = [S(n, k)]_{k, n \in \mathbb{N}}$

is the transition matrix between the basis B_2 and the basis B_1 . Now consider again equation (1.28) from Section 1.3. If we change x to $-x$ and multiply by $(-1)^n$, we obtain

$$\sum_{k=0}^n s(n, k)x^k = (x)_n.$$

Thus, the matrix $s = [s(n, k)]_{k, n \in \mathbb{N}}$ is the transition matrix from B_1 to B_2 and is, therefore, the *inverse* to the matrix S .

The assertion that the matrices S and s are inverses leads to the following result.

1.9.1 Proposition. *a. For all $m, n \in \mathbb{N}$, we have*

$$\sum_{k \geq 0} S(m, k)s(k, n) = \delta_{mn}.$$

b. Let a_0, a_1, \dots and b_0, b_1, \dots be two sequences of elements of a field K . The following two conditions are equivalent:

i. For all $n \in \mathbb{N}$,

$$b_n = \sum_{k=0}^n S(n, k)a_k.$$

ii. For all $n \in \mathbb{N}$,

$$a_n = \sum_{k=0}^n s(n, k)b_k.$$

Proof.

- a. This is just the assertion that the product of the two matrices S and s is the identity matrix $[\delta_{mn}]$.
- b. Let \mathbf{a} and \mathbf{b} denote the (infinite) column vectors $(a_0, a_1, \dots)^t$ and $(b_0, b_1, \dots)^t$, respectively (where t denotes transpose). Then (i) asserts that $S\mathbf{a} = \mathbf{b}$. Multiply on the left by s to obtain $\mathbf{a} = s\mathbf{b}$, which is (ii). Similarly (ii) implies (i).

□

The matrices S and s look as follows:

$$S = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 7 & 6 & 1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 15 & 25 & 10 & 1 & 0 & 0 \\ 0 & 1 & 31 & 90 & 65 & 15 & 1 & 0 \\ 0 & 1 & 63 & 301 & 350 & 140 & 21 & 1 \\ & & & & \vdots & & & \end{bmatrix}$$

$$s = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & -3 & 1 & 0 & 0 & 0 & 0 \\ 0 & -6 & 11 & -6 & 1 & 0 & 0 & 0 \dots \\ 0 & 24 & -50 & 35 & -10 & 1 & 0 & 0 \\ 0 & -120 & 274 & -225 & 85 & -15 & 1 & 0 \\ 0 & 720 & -1764 & 1624 & -735 & 175 & -21 & 1 \\ & & & & \vdots & & & \end{bmatrix}$$

Equations (1.28) and (1.94d) also have close connections with the *calculus of finite differences*, about which we will say a very brief word here. Given a function $f: \mathbb{Z} \rightarrow K$ (or possibly $f: \mathbb{N} \rightarrow K$), where K is a field of characteristic 0, define a new function Δf , called the *first difference* of f , by

$$\Delta f(n) = f(n+1) - f(n).$$

We call Δ the *first difference operator*, and a succinct but greatly oversimplified definition of the calculus of finite differences would be that it is the study of the operator Δ . We may iterate Δ k times to obtain the *k -th difference operator*,

$$\Delta^k f = \Delta(\Delta^{k-1} f).$$

The field element $\Delta^k f(0)$ is called the *k -th difference of f at 0*. Define another operator E , called the *shift operator*, by $Ef(n) = f(n+1)$. Thus, $\Delta = E - 1$, where 1 denotes the identity operator. We now have

$$\begin{aligned} \Delta^k f(n) &= (E - 1)^k f(n) \\ &= \sum_{i=0}^k (-1)^{k-i} \binom{k}{i} E^i f(n) \\ &= \sum_{i=0}^k (-1)^{k-i} \binom{k}{i} f(n+i). \end{aligned} \tag{1.97}$$

In particular,

$$\Delta^k f(0) = \sum_{i=0}^k (-1)^{k-i} \binom{k}{i} f(i), \tag{1.98}$$

which gives an explicit formula for $\Delta^k f(0)$ in terms of the values $f(0), f(1), \dots, f(k)$. We can easily invert (1.97) and express $f(n)$ in terms of the numbers $\Delta^i f(0)$. Namely,

$$\begin{aligned} f(n) &= E^n f(0) \\ &= (1 + \Delta)^n f(0) \\ &= \sum_{k=0}^n \binom{n}{k} \Delta^k f(0). \end{aligned} \tag{1.99}$$

NOTE. The operator Δ is a “discrete analogue” of the derivative operator $D = \frac{d}{dx}$. It is an instructive exercise to find finite difference analogues of concepts and results from calculus. For instance, the finite difference analogue of e^x is 2^n , since $De^x = e^x$ and $\Delta 2^n = 2^n$. Similarly, the finite difference analogue of x^n is $(x)_n$, since $Dx^n = nx^{n-1}$ and $\Delta(x)_n = n(x)_{n-1}$. The finite difference analogue of the Taylor series expansion

$$f(x) = \sum_{k \geq 0} \frac{1}{k!} (D^k f(0)) x^k$$

is just equation (1.99), where we should write $\binom{n}{k} = \frac{1}{k!} (n)_k$ to make the analogy even more clear. A unified framework for working with operators such as D and Δ is provided by Exercise 5.37.

Now given the function $f : \mathbb{Z} \rightarrow K$, write on a line the values

$$\cdots f(-2) f(-1) f(0) f(1) f(2) \cdots$$

If we write below the space between any two consecutive terms $f(i), f(i+1)$ their difference $f(i+1) - f(i) = \Delta f(i)$, then we obtain the sequence

$$\cdots \Delta f(-2) \Delta f(-1) \Delta f(0) \Delta f(1) \Delta f(2) \cdots$$

Iterating this procedure yields the *difference table* of the function f . The k th row (regarding the top row as row 0) consists of the values $\Delta^k f(n)$. The diagonal beginning with $f(0)$ and extending down and to the right consists of the differences at 0 (i.e., $\Delta^k f(0)$). For instance, let $f(n) = n^4$ (where $K = \mathbb{Q}$, say). The difference table (beginning with $f(0)$) looks like

$$\begin{array}{cccccccc} 0 & 1 & 16 & 81 & 256 & 625 & \cdots \\ & 1 & 15 & 65 & 175 & 369 & \\ & & 14 & 50 & 110 & 194 & \\ & & & 36 & 60 & 84 & \\ & & & & 24 & 24 & \\ & & & & & 0 & \\ & & & & & & \ddots \end{array}$$

Hence by (1.99),

$$n^4 = \binom{n}{1} + 14 \binom{n}{2} + 36 \binom{n}{3} + 24 \binom{n}{4} + 0 \binom{n}{5} + \cdots$$

In this case, since n^4 is a polynomial of degree 4 and $\binom{n}{k}$, for fixed k , is a polynomial of degree k , the preceding expansion stops after the term $24 \binom{n}{4}$, that is, $\Delta^k 0^4 = 0$ if $k > 4$ (or more generally, $\Delta^k n^4 = 0$ if $k > 4$). Note that by (1.94d) we have

$$n^4 = \sum_{k=0}^4 k! S(4, k) \binom{n}{k},$$

so we conclude $1!S(4, 1) = 1$, $2!S(4, 2) = 14$, $3!S(4, 3) = 36$, $4!S(4, 4) = 24$.

There was of course nothing special about the function n^4 in the preceding discussion. The same reasoning establishes the following result.

1.9.2 Proposition. *Let K be a field of characteristic 0.*

- (a) *A function $f : \mathbb{Z} \rightarrow K$ is a polynomial of degree at most d if and only if $\Delta^{d+1} f(n) = 0$ (or $\Delta^d f(n)$ is constant).*
- (b) *If the polynomial $f(n)$ of degree at most d is expanded in terms of the basis $\binom{n}{k}$, $0 \leq k \leq d$, then the coefficients are $\Delta^k f(0)$; that is,*

$$f(n) = \sum_{i=0}^d \Delta^i f(0) \cdot \binom{n}{i}.$$

- (c) *In the special case $f(n) = n^d$, we have*

$$\Delta^k 0^d = k! S(d, k).$$

1.9.3 Corollary. *Let $f : \mathbb{Z} \rightarrow K$ be a polynomial of degree d , where $\text{char}(K) = 0$. A necessary and sufficient condition that $f(n) \in \mathbb{Z}$ for all $n \in \mathbb{Z}$ is that $\Delta^k f(0) \in \mathbb{Z}$, $0 \leq k \leq d$. (In algebraic terms, the abelian group of all polynomials $f : \mathbb{Z} \rightarrow \mathbb{Z}$ of degree at most d is free with basis $\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{d}$.)*

Let us now proceed to the next entry of the Twelfold Way.

4. The “balls” are indistinguishable, so we are only interested in *how many* balls go into each box b_1, b_2, \dots, b_x . If $v(b_i)$ balls go into box b_i , then v defines an n -element multiset on X . The number of such multisets is $\binom{x}{n}$.
5. This is similar to 4, except that each box contains at most one ball. Thus our multiset becomes a set, and there are $\binom{x}{n}$ n -element subsets of X .
6. Each box b_i must contain at least one ball. If we remove one ball from each box, then we obtain an $(n-x)$ -element multiset on X . The number of such multisets is $\binom{x}{n-x}$. Alternatively, we can clearly regard a ball placement as a composition of n into x parts, whose number is $\binom{n-1}{x-1} = \binom{x}{n-x}$.
7. Since the boxes are indistinguishable, a function $f : N \rightarrow X$ is determined by the nonempty sets $f^{-1}(b)$, $b \in X$, where $f^{-1}(b) = \{a \in N : f(a) = b\}$. These sets form a partition π of N , called the *kernel* or *coimage* of f . The only restriction on π is that it can contain no more than x blocks. The number of partitions of N into at most x blocks is $S(n, 0) + S(n, 1) + \dots + S(n, x)$.
8. Each block of the coimage π of f must have one element. There is one such π if $x \geq n$; otherwise, there is no such π .
9. If f is surjective, then none of the sets $f^{-1}(b)$ is empty. Hence, the coimage π contains exactly x blocks. The number of such π is $S(n, x)$.
10. Let $p_k(n)$ denote the number of partitions of n into k parts, as defined in Section 1.7. A function $f : N \rightarrow X$ with N and X both indistinguishable is

determined only by the *number of elements* in each block of its coimage π . The actual elements themselves are irrelevant. The only restriction on these numbers is that they be positive integers summing to n , and that there can be no more than x of them. In other words, the numbers form a partition of n into at most x parts. The number of such partitions is $p_0(n) + p_1(n) + \cdots + p_x(n)$. Note that this number is equal to $p_x(n+x)$ (Exercise 66).

11. Same argument as 8.
12. Analogous argument to 9. If $f : N \rightarrow X$ is surjective, then the coimage π of f has exactly x blocks, so their cardinalities form a partition of n into exactly x parts.

There are many possible generalizations of the Twelvelfold Way and its individual entries. See the Notes for an extension of the Twelvelfold Way to a “Thirtyfold Way.” Another very natural generalization of some of the Twelvelfold Way entries is the following. Let $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m$ and $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{N}^n$. Suppose that we have α_i balls of color i , $1 \leq i \leq m$. Balls of the same color are indistinguishable. We also have n distinguishable boxes B_1, \dots, B_n . In how many ways can we place the balls into the boxes so that box B_j has exactly β_j balls? Call this number $N_{\alpha\beta}$. Similarly define $M_{\alpha\beta}$ to be the number of such placements with the further condition that each box can contain at most one ball of each color. Clearly, $N_{\alpha\beta} = M_{\alpha\beta} = 0$ unless $\sum \alpha_i = \sum \beta_j$ (the total number of balls). Given a placement of the balls into the boxes, let A be the $m \times n$ matrix such that A_{ij} is the number of balls colored i that are placed in box B_j . It is easy to see that this placement is enumerated by $N_{\alpha\beta}$ if and only if the i th row sum of A is α_i and the j th column sum is β_j . In other words, A has *row sum vector* $\text{row}(A) = \alpha$ and *column sum vector* $\text{col}(A) = \beta$. Thus, $N_{\alpha\beta}$ is the number of $m \times n$ \mathbb{N} -matrices with $\text{row}(A) = \alpha$ and $\text{col}(A) = \beta$. Similarly, $M_{\alpha\beta}$ is the number of $m \times n$ $(0, 1)$ -matrices with $\text{row}(A) = \alpha$ and $\text{col}(A) = \beta$. In general, there is no simple formula for $N_{\alpha\beta}$ or $M_{\alpha\beta}$, but there are many interesting special cases, generating functions, algebraic connections, and the like. See for instance Proposition 4.6.2, Proposition 5.5.8–Corollary 5.5.11, and the many appearances of $N_{\alpha\beta}$ and $M_{\alpha\beta}$ in Chapter 7.

1.10 Two q -Analogues of Permutations

We have seen that the vector space \mathbb{F}_q^n is a good q -analogue of the n -element set $[n]$, and a k -dimensional subspace of \mathbb{F}_q^n is a good q -analogue of a k -element subset of $[n]$. See in particular the finite field proofs of Proposition 1.7.3 and the q -binomial theorem (equation (1.87)). In this section, we pursue this line of thought further by considering the q -analogue of a permutation of the set $[n]$. It turns out that there are *two* good q -analogues that are closely related. This section involves some linear algebra over finite fields and is unrelated to the rest of the text; it may be omitted without significant loss of continuity.

1.10.1 A q -Analogue of Permutations as Bijections

A permutation w of the set $[n]$ may be regarded as an automorphism of $[n]$ (i.e., a bijection $w : [n] \rightarrow [n]$ preserving the “structure” of $[n]$). Since $[n]$ is being regarded simply as a set, any bijection $w : [n] \rightarrow [n]$ preserves the structure. Hence, one q -analogue of a permutation w is a bijection $A : \mathbb{F}_q^n \rightarrow \mathbb{F}_q^n$ preserving the structure of \mathbb{F}_q^n . The structure under consideration is that of a vector space, so A is simply an invertible linear transformation on \mathbb{F}_q^n . The set of all such linear transformations is denoted $\text{GL}(n, q)$, the *general linear group* of degree n over \mathbb{F}_q . Thus, $\text{GL}(n, q)$ is a q -analogue of the symmetric group \mathfrak{S}_n . We will sometimes identify a linear transformation $A \in \text{GL}(n, q)$ with its matrix with respect to the standard basis e_1, \dots, e_n of \mathbb{F}_q^n , that is, e_i is the i th unit coordinate vector $(0, 0, \dots, 0, 1, 0, \dots, 0)$ (with 1 in the i th coordinate). Hence, $\text{GL}(n, q)$ may be identified with the group of all $n \times n$ invertible matrices over \mathbb{F}_q .

For any of the myriad properties of permutations, we can try to find a corresponding property of linear transformations over \mathbb{F}_q . Here we will consider the following two properties: the total number of permutations in \mathfrak{S}_n , and the distribution of permutations by cycle type. The total number of elements (i.e., the order) of $\text{GL}(n, q)$ is straightforward to compute.

1.10.1 Proposition. *We have*

$$\begin{aligned} \#\text{GL}(n, q) &= (q^n - 1)(q^n - q)(q^n - q^2) \cdots (q^n - q^{n-1}) \\ &= q^{\binom{n}{2}} (q - 1)^n (n)!. \end{aligned} \quad (1.100)$$

Proof. Regard $A \in \text{GL}(n, q)$ as an $n \times n$ matrix. An arbitrary $n \times n$ matrix over \mathbb{F}_q is invertible if and only if its rows are linearly independent. There are, therefore, $q^n - 1$ choices for the first row; it can be any nonzero element of \mathbb{F}_q^n . There are q vectors in \mathbb{F}_q^n linearly dependent on the first row, so there are $q^n - q$ choices for the second row. Since the first two rows are linearly independent, they span a subspace V of \mathbb{F}_q^n of dimension 2. The third row can be any vector in \mathbb{F}_q^n not in V , so there are $q^n - q^2$ choices for the third row. Continuing this line of reasoning, there will be $q^n - q^{i-1}$ choices for the i th row, so we obtain (1.100). \square

The q -analogue of the cycle type of a permutation is more complicated. Two elements $u, v \in \mathfrak{S}_n$ have the same cycle type if and only if they are *conjugate* in \mathfrak{S}_n (i.e., if and only if there exists a permutation $w \in \mathfrak{S}_n$ such that $v = wuw^{-1}$). Hence, a reasonable analogue of cycle type for $\text{GL}(n, q)$ is the conjugacy class of an element of $\text{GL}(n, q)$. In this context, it is better to work with *all* $n \times n$ matrices over \mathbb{F}_q and then specialize to invertible matrices. Let $\text{Mat}(n, q)$ denote the set (in fact, an \mathbb{F}_q -algebra of dimension n^2) of all $n \times n$ matrices over \mathbb{F}_q . We briefly review the theory of the adjoint action of $\text{GL}(n, q)$ on $\text{Mat}(n, q)$. The proper context for understanding this theory is commutative algebra, so we first review the relevant background. There is nothing special about finite fields in this theory,

so we work over any field K , letting $\mathrm{GL}(n, K)$ (respectively, $\mathrm{Mat}(n, K)$) denote the set of invertible (respectively, arbitrary) $n \times n$ matrices over K .

Let R be a principal ideal domain (PID) that is not a field, and let M be a finitely generated R -module. Two irreducible (= prime, for PIDs) elements $x, y \in R$ are *equivalent* if $xR = yR$ (i.e., if $y = ex$ for some unit e). Let \mathcal{P} be a maximal set of inequivalent irreducible elements of R . The structure theorem for finitely generated modules over PIDs asserts that M is isomorphic to a (finite) direct sum of copies of R and $R/x^i R$ for $x \in \mathcal{P}$ and $i \geq 1$. Moreover, the terms in this direct sum are unique up to the order of summands. Thus, there is a unique $k \geq 0$, and for each $x \in \mathcal{P}$ there is a unique partition $\lambda(x) = (\lambda_1(x), \lambda_2(x), \dots)$ (which may be the empty partition) such that

$$M \cong R^k \oplus \bigoplus_{x \in \mathcal{P}} \bigoplus_{i \geq 1} R/x^i R.$$

If moreover M has finite length d (i.e., d is the largest integer j for which there is a proper chain $M_0 \subset M_1 \subset \dots \subset M_j$ of submodules of M), then $k = 0$.

Now consider the case where $R = K[u]$, well-known to be a PID. Let $\mathcal{I} = \mathcal{I}(K)$ (abbreviated to $\mathcal{I}(q)$ when $K = \mathbb{F}_q$) denote the set of all nonconstant monic irreducible polynomials $f(u)$ over K , and let Par denote the set of all partitions of all nonnegative integers. Given $M \in \mathrm{Mat}(n, K)$, then M defines a $K[u]$ -module structure on K^n , where the action of u is that of M . Let us denote this $K[u]$ -module by $K[M]$. Since $K[M]$ has finite length as a $K[u]$ -module (or even as a vector space over K), we have an isomorphism

$$K[M] \cong \bigoplus_{f \in \mathcal{I}(K)} \bigoplus_{i \geq 1} K[u]/\left(f(u)^{\lambda_i(f)}\right). \quad (1.101)$$

Moreover, the characteristic polynomial $\det(zI - M)$ of M is given by

$$\det(zI - M) = \prod_{f \in \mathcal{I}(K)} f(z)^{|\lambda(f)|}.$$

Now $\mathrm{GL}(n, K)$ acts on $\mathrm{Mat}(n, K)$ by conjugation, that is, if $A \in \mathrm{GL}(n, K)$ and $M \in \mathrm{Mat}(n, K)$, then $A \cdot M = AMA^{-1}$. (This action is called the *adjoint representation* or *adjoint action* of $\mathrm{GL}(n, K)$.) Moreover, two matrices M and N in $\mathrm{Mat}(n, K)$ are in the same orbit of this action if and only if $K[M]$ and $K[N]$ are isomorphic as $K[u]$ -modules. Hence, by equation (1.101), we can index the orbit of M by a function

$$\Phi_M : \mathcal{I}(K) \rightarrow \mathrm{Par},$$

where

$$\sum_{f \in \mathcal{I}(K)} |\Phi_M(f)| \cdot \deg(f) = n, \quad (1.102)$$

namely, $\Phi_M(f) = \lambda(f)$. We call the function $\Phi = \Phi_M$ the *orbit type* of M . It is the analogue for $\mathrm{Mat}(n, K)$ of the cycle type of a permutation $w \in \mathfrak{S}_n$.

We now restrict ourselves to the case $K = \mathbb{F}_q$. As a first application of the description of the orbits of $\mathrm{GL}(n, q)$ acting adjointly on $\mathrm{Mat}(n, q)$, we can find the number of orbits. To do so, define $\beta(n, q) = \beta(n)$ to be the number of monic irreducible polynomials $f(z)$ of degree n over \mathbb{F}_q . It is well known (see Exercise 2.7) that

$$\beta(n, q) = \frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}. \quad (1.103)$$

1.10.2 Proposition. *Let $\omega(n, q)$ denote the number of orbits of the adjoint action of $\mathrm{GL}(n, q)$ on $\mathrm{Mat}(n, q)$, or equivalently, the number of different functions $\Phi : \mathcal{I}(q) \rightarrow \mathrm{Par}$ satisfying (1.102). Then*

$$\omega(n, q) = \sum_j p_j(n) q^j,$$

where $p_j(n)$ denotes the number of partitions of n into j parts. Equivalently,

$$\sum_{n \geq 0} \omega(n, q) x^n = \prod_{j \geq 1} (1 - qx^j)^{-1}.$$

Proof. We have

$$\begin{aligned} \sum_{n \geq 0} \omega(n, q) x^n &= \sum_{\Phi: \mathcal{I} \rightarrow \mathrm{Par}} x^{\sum_{f \in \mathcal{I}} |\Phi(f)| \cdot \deg(f)} \\ &= \prod_{f \in \mathcal{I}} \left(\sum_{\lambda \in \mathrm{Par}} x^{|\lambda| \cdot \deg(f)} \right) \\ &= \prod_{f \in \mathcal{I}} \prod_{j \geq 1} (1 - x^{j \cdot \deg(f)})^{-1} \quad (\text{by (1.77)}) \\ &= \prod_{n \geq 1} \prod_{j \geq 1} (1 - x^{jn})^{-\beta(n)}. \end{aligned}$$

Now using the formula (1.103) for $\beta(n)$, we get

$$\begin{aligned} \log \sum_{n \geq 0} \omega(n, q) x^n &= \sum_{n \geq 1} \sum_{j \geq 1} \beta(n) \log(1 - x^{jn})^{-1} \\ &= \sum_{n \geq 1} \sum_{j \geq 1} \frac{1}{n} \sum_{d|n} \mu(n/d) q^d \sum_{i \geq 1} \frac{x^{ijn}}{i}. \end{aligned}$$

Extract the coefficient $c(d, N)$ of $q^d x^N$. Clearly, $c(d, N) = 0$, when $d \nmid N$, so assume $d|N$. We get

$$\begin{aligned} c(d, N) &= \sum_{i|N} \frac{1}{i} \sum_{n: d|n} \frac{1}{n} \mu(n/d) \\ &= \sum_{i|\frac{N}{d}} \frac{1}{i} \sum_{m|\frac{N}{id}} \frac{1}{dm} \mu(m). \end{aligned}$$

An elementary and basic result of number theory asserts that

$$\sum_{k|r} \frac{\mu(k)}{k} = \frac{\phi(r)}{r},$$

where ϕ denotes the Euler phi-function. Hence,

$$c(d, N) = \frac{1}{d} \sum_{i|N/d} \frac{\phi(N/id)}{N/d}.$$

Another standard result of elementary number theory states that

$$\sum_{k|r} \phi(r/k) = \sum_{k|r} \phi(k) = r,$$

so we finally obtain

$$c(d, N) = \frac{1}{d} \frac{N/d}{N/d} = \frac{1}{d}.$$

On the other hand, we have

$$\log \prod_{n \geq 1} (1 - qx^n)^{-1} = \sum_{n \geq 1} \sum_{d \geq 1} \frac{q^d x^{nd}}{d}.$$

The coefficient $c'(d, N)$ of $q^d x^N$ is 0 unless $d|N$; otherwise, it is $1/d$. Hence, $c(d, n) = c'(d, n)$, and the proof follows. \square

NOTE. Proposition 1.10.2 shows that, insofar as the number of conjugacy classes is concerned, the “correct” q -analogue of \mathfrak{S}_n is not the group $\mathrm{GL}(n, q)$ itself, but rather its adjoint action on $\mathrm{Mat}(n, q)$. The number of orbits $\omega(n, q)$ is a completely satisfactory q -analogue of $p(n)$, the number of conjugacy classes in \mathfrak{S}_n , since $\omega(n, q)$ is a polynomial in q with nonnegative integer coefficients satisfying $\omega(n, 1) = p(n)$. Note that if $\omega^*(n, q)$ denotes the number of conjugacy classes in $\mathrm{GL}(n, q)$, then $\omega^*(n, q)$ is a polynomial in q satisfying $\omega^*(n, 1) = 0$ (Exercise 1.190). For more conceptual proofs of Proposition 1.10.2, see Exercise 1.191.

We next define a “cycle indicator” of $M \in \mathrm{Mat}(n, q)$ that encodes the orbit of M . For every $f \in \mathcal{I}$ and every partition $\lambda \neq \emptyset$, let $t_{f, \lambda}$ be an indeterminate. If $\lambda = \emptyset$, then set $t_{f, \lambda} = 1$. Let $\Phi_M : \mathcal{I} \rightarrow \mathrm{Par}$ be the orbit type of M . Define

$$t^{\Phi M} = \prod_{f \in \mathcal{I}} t_{f, \Phi_M(f)}.$$

Set

$$\gamma(n) = \gamma(n, q) = \#\mathrm{GL}(n, q).$$

We now define the *cycle indicator* (or *cycle index*) of $\text{Mat}(n, q)$ to be the polynomial

$$Z_n(t; q) = Z_n(\{t_{f, \lambda}\}; q) = \frac{1}{\gamma(n)} \sum_{M \in \text{Mat}(n, q)} t^{\Phi M}.$$

(Set $Z_0(t; q) = 1$.)

1.10.3 Example. (a) Let M be the diagonal matrix $\text{diag}(1, 1, 3)$. Then $t^{\Phi M} = t_{z-1, (1,1)} t_{z-3, (1)}$ if $q \neq 2^m$; otherwise, $t^{\Phi M} = t_{z-1, (1,1,1)}$.

(b) Let $n = q = 2$. Then

$$\begin{aligned} Z_2(t; 2) = \frac{1}{6} & \left(t_{z, (1,1)} + 3t_{z, (2)} + 6t_{z, (1)} t_{z+1, (1)} + t_{z+1, (1,1)} \right. \\ & \left. + 3t_{z+1, (2)} + 2t_{z+2, (1,1)} \right). \end{aligned} \quad (1.104)$$

We now give a q -analogue of Theorem 1.3.3, in other words, a generating function for the polynomials $Z_n(t; q)$. To see the analogy more clearly, recall from equation (1.27) that

$$\sum_{n \geq 0} Z_n(t; q) x^n = \prod_{i \geq 1} \sum_{j \geq 0} t_i^j \frac{x^{ij}}{i^j j!}.$$

The denominator $i^j j!$ is the number of permutations $w \in \mathfrak{S}_{ij}$ that commute with a fixed permutation with j i -cycles.

1.10.4 Theorem. *We have*

$$\sum_{n \geq 0} Z_n(t; q) x^n = \prod_{f \in \mathcal{I}} \sum_{\lambda \in \text{Par}} \frac{t_{f, \lambda} x^{|\lambda| \cdot \deg(f)}}{c_f(\lambda)}, \quad (1.105)$$

where $c_f(\lambda)$ is the number of matrices in $\text{GL}(n, q)$ commuting with a fixed matrix M of size $|\lambda(f)| \cdot \deg(f)$ satisfying

$$\Phi_M(g) = \begin{cases} \lambda, & g = f, \\ \emptyset, & g \neq f. \end{cases}$$

Equivalently, $c_f(\lambda)$ is the number of \mathbb{F}_q -linear automorphisms of the ring

$$\mathbb{F}_q[M] \cong \bigoplus_{i \geq 1} \mathbb{F}_q[u] / \left(f(u)^{\lambda_i(f)} \right)$$

appearing in equation (1.101).

Proof (sketch). Let G be a finite group acting on a finite set X . For $a \in X$, let $Ga = \{g \cdot a : g \in G\}$, the *orbit* of G containing a . Also let $G_a = \{g \in G : g \cdot a = a\}$, the *stabilizer* of a . A basic and elementary result in group theory asserts that $\#Ga \cdot \#G_a = \#G$. Consider the present situation, where $G = \text{GL}(n, q)$ is acting on

$\text{Mat}(n, q)$. Let $M \in \text{Mat}(n, q)$. Then $A \in G_M$ if and only if $AMA^{-1} = M$ (i.e., if and only if A and M commute). Hence,

$$\#GM = \frac{\#G}{c_G(M)}, \quad (1.106)$$

where $c_G(M)$ is the number of elements of G commuting with M .

We have a unique direct sum decomposition

$$\mathbb{F}_q^n = \bigoplus_{f \in \mathcal{I}} V_f,$$

where

$$V_f = \{v \in \mathbb{F}_q^n : f(M)^r(v) = 0 \text{ for some } r \geq 1\}.$$

Thus, $M = \bigoplus_{f \in \mathcal{I}} M_f$, where $M_f V_f \subseteq V_f$ and $M_f V_g = \{0\}$ if $g \neq f$. A matrix A commuting with M respects this decomposition (i.e., $AV_f \subseteq V_f$ for all $f \in \mathcal{I}$). Thus, $A = \bigoplus_{f \in \mathcal{I}} A_f$, where $A_f V_f \subseteq V_f$ and $A_f V_g = \{0\}$ if $g \neq f$. Then A commutes with M if and only if A_f commutes with M_f for all f . In particular,

$$c_G(M) = \prod_{f \in \mathcal{I}} c_f(\Phi_M(f)).$$

It follows from equation (1.106) that the number of conjugates of M (i.e., the size of the orbit GM) is given by

$$\#GM = \frac{\gamma(n)}{\prod_f c_f(\Phi_M(f))}. \quad (1.107)$$

This number is precisely the coefficient of $t^{\Phi_M} / \gamma(n)$ in equation (1.105), and the proof follows. \square

In order for Theorem 1.10.4 to be of any use, it is necessary to find a formula for the numbers $c_f(\lambda)$. There is one special case that is quite simple.

1.10.5 Lemma. *Let $f(z) = z - a$ for some $a \in \mathbb{F}_q$, and let $\langle 1^k \rangle$ denote the partition with k parts equal to 1. Then $c_f(\langle 1^k \rangle) = \gamma(k)$.*

Proof. We are counting matrices $A \in \text{GL}(k, q)$ that commute with a $k \times k$ diagonal matrix with a 's on the diagonal, so A can be any matrix in $\text{GL}(k, q)$. \square

1.10.6 Corollary. *Let d_n denote the number of diagonalizable (over \mathbb{F}_q) matrices $M \in \text{Mat}(n, q)$. Then*

$$\sum_{n \geq 0} d_n \frac{x^n}{\gamma(n)} = \left(\sum_{k \geq 0} \frac{x^k}{\gamma(k)} \right)^q.$$

Proof. A matrix M is diagonalizable over \mathbb{F}_q if and only if its corresponding orbit type $\Phi_M : \mathcal{I} \rightarrow \text{Par}$ satisfies $\Phi_M(f) = \emptyset$ unless $f = z - a$ for $a \in \mathbb{F}_q$, and

$\Phi_M(z - a) = \langle 1^k \rangle$ in the notation of equation (1.74) (where we may have $k = 0$, i.e., a is not an eigenvalue of M). Hence,

$$d_n = \gamma(n) Z_n(t; q) \Big|_{t_{z-a, \langle 1^k \rangle} = 1, t_{f, \lambda} = 0 \text{ otherwise}}.$$

Making the substitution $t_{z-a, \langle 1^k \rangle} = 1, t_{f, \lambda} = 0$ otherwise into Theorem 1.10.4 yields

$$\sum_{n \geq 0} d_n \frac{x^n}{\gamma(n)} = \prod_{a \in \mathbb{F}_q} \sum_{k \geq 0} \frac{x^k}{c_{z-a}(\langle 1^k \rangle)}.$$

The proof follows from Lemma 1.10.5. \square

The evaluation of $c_f(\lambda)$ for arbitrary f and λ is more complicated. It may be regarded as the q -analogue of Proposition 1.3.2, since equation (1.107) shows that the number of conjugates of a matrix M is determined by the numbers $c_f(\Phi_M(f))$. This formula for $c_f(\lambda)$ is a fundamental enumerative result on enumerating classes of matrices in $\text{Mat}(n, q)$, from which a host of other enumerative results can be derived. Let $\lambda' = (\lambda'_1, \lambda'_2, \dots)$ denote the conjugate partition to λ , and let $m_i = m_i(\lambda) = \lambda'_i - \lambda'_{i+1}$ be the number of parts of λ of size i . Set

$$h_i = \lambda'_1 + \lambda'_2 + \dots + \lambda'_i,$$

and let $d = \deg(f)$.

1.10.7 Theorem. *We have*

$$c_f(\lambda) = \prod_{i \geq 1} \prod_{j=1}^{m_i} \left(q^{h_i d} - q^{(h_i - j)d} \right). \quad (1.108)$$

1.10.8 Example. (a) Let $\lambda = (4, 2, 2, 2, 1)$, so $\lambda' = (5, 4, 1, 1)$, $h_1 = 5$, $h_2 = 9$, $h_3 = 10$, $h_4 = 11$, $m_1 = 1$, $m_2 = 3$, $m_4 = 1$. Thus, for $\deg(f) = 1$, we have

$$c_f(4, 2, 2, 2, 1) = (q^5 - q^4)(q^9 - q^8)(q^9 - q^7)(q^9 - q^6)(q^{11} - q^{10}).$$

(b) Let $\lambda = (k)$, so $\lambda' = \langle 1^k \rangle$, $h_i = i$ for $1 \leq i \leq k$, and $m_k = 1$. For $\deg(f) = 1$, we get $c_f(k) = q^k - q^{k-1}$. Indeed, we are asking for the number of matrices $A \in \text{GL}(k, q)$ commuting with a $k \times k$ Jordan block. Such matrices are easily seen to be upper triangular with constant diagonals (parallel to the main diagonal). There are $q - 1$ choices for the main diagonal and q choices for each of the $k - 1$ diagonals above the main diagonal, giving $(q - 1)q^{k-1} = q^k - q^{k-1}$ choices in all.

Proof of Theorem 1.10.7. The proof is analogous to that of Proposition 1.3.2. We write down some data that determine a linear transformation $M \in \text{Mat}(nd, q)$ for which $\Phi_M(f) = \lambda \vdash n$, and then we count in how many ways we obtain the same linear transformation M . Let $\ell = \ell(\lambda)$, the number of parts of λ , and similarly $k = \ell(\lambda') = \lambda_1$.

Now let

$$\mathbf{v} = \{v_{ij} : 1 \leq i \leq \ell, 1 \leq j \leq d\lambda_i\}$$

be a basis B for \mathbb{F}_q^{nd} , together with the indexing v_{ij} of the basis elements. Thus, the number $N(n, d, q)$ of possible \mathbf{v} is the number of *ordered* bases of \mathbb{F}_q^{nd} , namely,

$$N(n, d, q) = (q^{nd} - 1)(q^{nd} - q) \cdots (q^{nd} - q^{nd-1}) = \#\mathrm{GL}(nd, q). \quad (1.109)$$

Let $M = M_{\mathbf{v}}$ be the unique linear transformation satisfying the following three properties:

- The characteristic polynomial $\det(zI - M)$ of M is $f(z)^n$.
- For all $1 \leq i \leq \ell$ and $1 \leq j < \lambda_i d$, we have $M(v_{ij}) = v_{i, j+1}$.
- For all $1 \leq i \leq \ell$, we have that $M(v_{i, \lambda_i d})$ is a linear combination of the v_{ij} 's for $1 \leq j \leq \lambda_i d$.

It is not hard to see that M is indeed unique and that $\Phi_M(f) = \lambda$.

We now consider how many indexed bases $\mathbf{v} = (v_{ij})$ determine the same linear transformation M . Given M , define

$$V_i = \{v \in \mathbb{F}_q^{nd} : f(M)^i(v) = 0\}, \quad 1 \leq i \leq k.$$

It is clear that

$$\begin{aligned} V_1 &\subset V_2 \subset \cdots \subset V_k, \\ \dim V_i &= (\lambda'_1 + \lambda'_2 + \cdots + \lambda'_i)d = h_i d, \\ \dim(V_i/V_{i-1}) &= \lambda'_i d. \end{aligned}$$

If B is a subset of \mathbb{F}_q^n , then set

$$f(M)B = \{f(M)v : v \in B\}.$$

There are $q^{\dim(V_k)d} - q^{\dim(V_{k-1})d} = q^{h_k d} - q^{h_{k-1} d}$ choices for v_{11} (since v_{11} can be any vector in V_k not in V_{k-1}), after which all other v_{ij} are determined. There are then $q^{h_k d} - q^{(h_{k-1}+1)d}$ choices for v_{21} (since v_{21} can be any vector in V_k not in the span of V_{k-1} and $\{v_{11}, v_{12}, \dots, v_{1d}\}$), and so on, down to $q^{h_k d} - q^{(h_{k-1}+m_k)d}$ choices for $v_{m_k, 1}$.

Let

$$B_1 := \{v_{i1}, v_{i2}, \dots, v_{id} : 1 \leq i \leq \lambda'_k\}.$$

Thus, B_1 is a subset of V_k whose image in V_k/V_{k-1} is a basis for V_k/V_{k-1} . Now $v_{m_k+1, 1} (= v_{\lambda'_k+1, 1})$ can be any vector in V_{k-1} not in the span of $f(M)B_1 \cup V_{k-2}$, so there are

$$\begin{aligned} q^{\dim(V_{k-1})d} - q^{\#B_1 + \dim(V_{k-2})d} &= q^{h_{k-1}d} - q^{m_k d + h_{k-2}d} = q^{h_{k-1}d} - q^{(h_{k-1} - m_{k-1})d} \\ &\text{choices for } v_{\lambda'_k+1, 1}. \text{ There are then } q^{h_{k-1}d} - q^{(h_{k-1} - m_{k-1} + 1)d} \text{ choices for } v_{\lambda'_k+2, 1}, \\ &\text{then } q^{h_{k-1}d} - q^{(h_{k-1} - m_{k-1} + 2)d} \text{ choices for } v_{\lambda'_k+3, 1}, \text{ etc., down to } q^{h_{k-1}d} - \\ &q^{(h_{k-1} - 1)d} \text{ choices for } v_{\lambda'_k-1, 1}. \end{aligned}$$

Let

$$B_2 = \{v_{i1}, v_{i2}, \dots, v_{id} : \lambda'_k + 1 \leq i \leq \lambda'_{k-1}\},$$

so $B_2 = \emptyset$ if $\lambda'_k = \lambda'_{k-1}$. Then $f(M)(B_1 \cup B_2)$ is a subset of V_{k-1} whose image in V_{k-1}/V_{k-2} is a basis for V_{k-1}/V_{k-2} . Now $v_{\lambda'_{k-1}+1,1}$ can be any vector in V_{k-2} not in the span of $f(M)(B_1 \cup B_2) \cup V_{k-3}$, so there are

$$\begin{aligned} q^{\dim(V_{k-2}) - \#B_1 + \#B_2 + \dim(V_{k-3})} &= q^{h_{k-2}d - q^{m_{k-1}d + h_{k-3}d}} \\ &= q^{h_{k-2}d} - q^{(h_{k-2} - m_{k-2})d} \end{aligned}$$

choices for $v_{\lambda'_{k-1}+1,1}$. There are then $q^{h_{k-2}d} - q^{(h_{k-2} - m_{k-2} + 1)d}$ choices for $v_{\lambda'_{k-1}+2,1}$, then $q^{h_{k-2}d} - q^{(h_{k-2} - m_{k-2} + 2)d}$ choices for $v_{\lambda'_{k-1}+3,1}$, and so on, down to $q^{h_{k-2}d} - q^{(h_{k-2} - 1)d}$ choices for $v_{\lambda'_{k-2},1}$.

Continuing in this manner shows that the total number of choices for \mathbf{v} is given by the right-hand side of equation (1.108).

We have shown that each indexed basis \mathbf{v} of \mathbb{F}_q^{nd} defines a matrix $M \in \text{Mat}(nd, q)$ with $\Phi_M(f) = \lambda$. Moreover, every matrix satisfying $\Phi_M(f) = \lambda$ occurs the same number $L(n, d, q)$ of times, given by the right-hand side of (1.108). Since by (1.109) the number of indexed bases is $\#\text{GL}(nd, q)$, we get that the number of matrices M satisfying $\Phi_M(f) = \lambda$ is equal to $\text{GL}(nd, q)/L(nd, q)$. It follows from equation (1.106) that $L(nd, q) = c_f(\lambda)$, completing the proof.

As a slight variation, we can see directly that $L(nd, q) = c_f(\lambda)$ as follows. Let $\mathbf{v} = (v_{ij})$ be a fixed indexed basis for \mathbb{F}_q^{nd} with $M = M(\mathbf{v})$. Let $\mathbf{v}' = (v'_{ij})$ be another indexed basis satisfying $M = M(\mathbf{v}')$. Then the linear transformation $A \in \text{GL}(nd, q)$ satisfying $A(v_{ij}) = v'_{ij}$ for all i, j commutes with M , and all matrices commuting with M arise in this way. Hence once again $L(nd, q) = c_f(\lambda)$. \square

Even with the preceding formula for $c_f(\lambda)$, equation (1.105) is difficult to work with in its full generality. However, if we specialize each variable $t_{f,\lambda}$ to $t_f^{|\lambda|}$, then the following lemma allows a simplification of (1.105).

1.10.9 Lemma. *For any $f \in \mathcal{I}$ of degree d we have*

$$\sum_{\lambda \in \text{Par}} \frac{x^{|\lambda|}}{c_f(\lambda)} = \prod_{r \geq 1} \left(1 - \frac{x}{q^r d}\right)^{-1}.$$

Proof. By Theorem 1.10.7, it suffices to assume $d = 1$. Our computations take place in the ring $\mathbb{C}(q)[[x]]$ (i.e., power series in x whose coefficients are rational functions in q with complex coefficients). It follows from Proposition 1.8.6(c) that

$$\begin{aligned} \prod_{r \geq 1} \left(1 - \frac{x}{q^r}\right)^{-1} &= \sum_{n \geq 0} \frac{x^n q^{-n}}{(1 - q^{-1}) \cdots (1 - q^{-n})} \\ &= \sum_{n \geq 0} \frac{(-1)^n x^n q^{\binom{n}{2}}}{(1 - q)(1 - q^2) \cdots (1 - q^n)}. \end{aligned}$$

Hence, by Theorem 1.10.7, we need to prove that

$$\sum_{\lambda \vdash n} \prod_{i \geq 1} \prod_{j=1}^{m_i(\lambda)} \frac{1}{q^{h_i(\lambda)} - q^{h_i(\lambda)-j}} = \frac{(-1)^n q^{\binom{n}{2}}}{(1-q)(1-q^2) \cdots (1-q^n)}. \quad (1.110)$$

Substitute $1/q$ for q in equation (1.110). We will simply write $h_i = h_i(\lambda)$ and $m_i = m_i(\lambda)$. Since

$$\frac{1}{q^{-h_i} - q^{-(h_i-j)}} = \frac{q^{h_i}}{1 - q^j},$$

the left-hand side of (1.110) becomes

$$\sum_{\lambda \vdash n} \prod_{i \geq 1} \frac{q^{m_i h_i}}{(1-q) \cdots (1-q^{m_i})}.$$

It is easy to see that

$$\sum_{i \geq 1} m_i h_i = \sum_{i \geq 1} (\lambda'_i)^2,$$

which we denote by $\langle \lambda', \lambda' \rangle$.

Under the substitution $q \rightarrow 1/q$, the right-hand side of (1.110) becomes $q^n / (1-q) \cdots (1-q^n)$. Thus, we are reduced to proving that

$$\sum_{\lambda \vdash n} q^{\langle \lambda', \lambda' \rangle} \prod_{i \geq 1} \frac{1}{(1-q) \cdots (1-q^{m_i})} = \frac{q^n}{(1-q) \cdots (1-q^n)}. \quad (1.111)$$

We can replace $\langle \lambda', \lambda' \rangle$ by $\langle \lambda, \lambda \rangle$ since this substitution merely permutes the terms in the sum. Set $m'_i = m_i(\lambda') = \lambda_i - \lambda_{i+1}$. Then

$$\sum_{\lambda \vdash n} q^{\langle \lambda, \lambda \rangle} \prod_{i \geq 1} \frac{1}{(1-q) \cdots (1-q^{m'_i})} = \sum_{\lambda \vdash n} \frac{q^{\langle \lambda, \lambda \rangle}}{(1-q) \cdots (1-q^{\lambda_1})} \binom{\lambda_1}{\lambda_2} \binom{\lambda_2}{\lambda_3} \cdots.$$

The coefficient of q^k in the right-hand side of (1.111) is equal to $p_n(k)$, the number of partitions of k with largest part n . Given such a partition $\mu = (\mu_1, \mu_2, \dots)$, associate a partition $\lambda \vdash \mu_1$ by taking the rank (= length of the Durfee square) of μ , then the rank of the partition whose diagram is to the right of the Durfee square of μ , and so on. For instance, if $\mu = (7, 7, 5, 4, 3, 2)$, then $\lambda = (4, 2, 1)$ as indicated by Figure 1.26. Given λ , the generating function $\sum_{\mu} q^{|\mu|}$ for all corresponding μ is

$$\frac{q^{\langle \lambda, \lambda \rangle}}{(1-q) \cdots (1-q^{\lambda_1})} \binom{\lambda_1}{\lambda_2} \binom{\lambda_2}{\lambda_3} \cdots,$$

as indicated by Figure 1.27 (using Proposition 1.7.3), and the proof follows. \square

Now let

$$\widehat{Z}_n(t; q) = Z_n(t; q)|_{t_{f, \lambda} = | \lambda |}.$$

For instance, from equation (1.104) we have

$$\widehat{Z}_2(t; 2) = \frac{1}{6} \left(4t_z^2 + 4t_{z+1}^2 + 6t_z t_{z+1} + 2t_{z^2+z+1} \right).$$

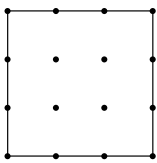


Figure 1.26 The “successive Durfee squares” of $\mu = (7, 7, 5, 4, 3, 2)$.

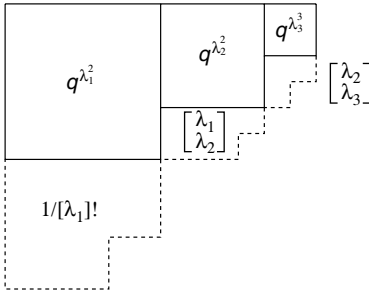


Figure 1.27 The “successive Durfee square decomposition” of λ .

Let $f = \prod_{f_i \in \mathcal{I}} f_i^{a_i}$, with $\deg f = n$. Then the coefficient of $t_{f_1}^{a_1} t_{f_2}^{a_2} \cdots$ in $\gamma(n, q) \widehat{Z}(t; q)$ is just the number of matrices $M \in \text{Mat}(n, q)$ with characteristic polynomial f . Note that in general if we define $\deg(t_f) = \deg(f)$, then $\widehat{Z}_n(t; q)$ is homogeneous of degree n .

The following corollary is an immediate consequence of Theorem 1.10.4 and Lemma 1.10.9.

1.10.10 Corollary. *We have*

$$\sum_{n \geq 0} \widehat{Z}_n x^n = \prod_{f \in \mathcal{I}} \prod_{r \geq 1} \left(1 - \frac{t_f x^{\deg(f)}}{q^r \deg(f)} \right)^{-1}.$$

Many interesting enumerative results can be obtained from Theorem 1.10.4 and Corollary 1.10.10. We give a couple here and some more in the Exercises (193–195). Let $\beta^*(n, q)$ denote the number of monic irreducible polynomials $f(z) \neq z$ of degree n over \mathbb{F}_q . It follows from (1.103) that

$$\beta^*(n, q) = \begin{cases} q - 1, & n = 1, \\ \frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}, & n > 1. \end{cases} \quad (1.112)$$

1.10.11 Corollary. (a) *We have*

$$\frac{1}{1-x} = \prod_{n \geq 1} \prod_{r \geq 1} (1 - q^{rn} x^n)^{-\beta^*(n, q)}$$

(b) Let $g(n)$ denote the number of nilpotent matrices $M \in \text{Mat}(n, q)$. (Recall that A is nilpotent if $A^m = 0$ for some $m \geq 1$.) Then $g(n) = q^{n(n-1)}$.

Proof. (a) Let $\mathcal{I}^* = \mathcal{I} - \{z\}$. Set $t_z = 0$ and $t_f = 1$ for $f \neq z$ in Corollary 1.10.10. Now

$$\widehat{Z}_n(t_z = 0, t_f = 1 \text{ if } f \neq z) = \frac{\gamma(n)}{\gamma'(n)} = 1,$$

so we get

$$\begin{aligned} \frac{1}{1-x} &= \prod_{f \in \mathcal{I}^*} \prod_{r \geq 1} \left(1 - \frac{x^{\deg(f)}}{q^{r \deg(f)}} \right) \\ &= \prod_{n \geq 1} \prod_{r \geq 1} (1 - q^{-rn} x^n)^{-\beta^*(n, q)}. \end{aligned}$$

Since the left-hand side is independent of q , we can substitute $1/q$ for q in the right-hand side without changing its value, and the proof follows. This result can also be proved by taking the logarithm of both sides and using the explicit formula for $\beta^*(n, q)$ given by equation (1.112).

(b) A matrix is nilpotent if and only if all its eigenvalues are 0. Hence,

$$g(n) = \gamma(n) \widehat{Z}(t; q) \Big|_{t_z=1, t_f=0 \text{ if } f \neq z}.$$

By Corollary 1.10.10 and Proposition 1.8.6(a) there follows

$$\begin{aligned} \sum_{n \geq 0} g(n) \frac{x^n}{\gamma(n)} &= \prod_{r \geq 1} \left(1 - \frac{x}{q^r} \right)^{-1} \\ &= \sum_{k \geq 0} \frac{q^{-k} x^k}{(1 - q^{-1}) \cdots (1 - q^{-k})} \\ &= \sum_{k \geq 0} q^{k(k-1)} \frac{x^k}{\gamma(k)}, \end{aligned}$$

and the proof follows. (For a more direct proof, see Exercise 1.188.)

□

1.10.2 A q -Analogue of Permutations as Words

We now discuss a second q -analogue of permutations (already discussed briefly after Corollary 1.3.13) and then connect it with one discussed earlier (matrices in $\text{GL}(n, q)$). Rather than regarding permutations of $1, 2, \dots, n$ as bijections

$w : [n] \rightarrow [n]$, we may regard them as words $a_1 a_2 \cdots a_n$. Equivalently, we can identify w with the maximal chain (or (*complete*) *flag*)

$$\emptyset = S_0 \subset S_1 \subset \cdots \subset S_n = [n] \quad (1.113)$$

of subsets of $[n]$, by the rule $\{a_i\} = S_i - S_{i-1}$. For instance, the flag $\emptyset \subset \{2\} \subset \{2, 4\} \subset \{1, 2, 4\} \subset \{1, 2, 3, 4\}$ corresponds to the permutation $w = 2413$. The natural q -analogue of a flag (1.113) is a maximal chain or (*complete*) *flag* of subspaces

$$\{0\} = V_0 \subset V_1 \subset \cdots \subset V_n = \mathbb{F}_q^n \quad (1.114)$$

of subspaces of \mathbb{F}_q^n , so $\dim V_i = i$. It is easy to count the number of such flags (as mentioned after Corollary 1.3.13).

1.10.12 Proposition. *The number $f(n, q)$ of complete flags (1.114) is given by*

$$f(n, q) = (n)! = (1+q)(1+q+q^2) \cdots (1+q+\cdots+q^{n-1}).$$

Proof. There are $\binom{n}{1} = (n)$ choices for V_1 , then $\binom{n-1}{1}$ choices for V_2 (since the quotient space \mathbb{F}_q^n/V_1 is an $(n-1)$ -dimensional vector space), etc. \square

Comparing Corollary 1.3.13 with Proposition 1.10.12, we see that

$$f(n, q) = \sum_{w \in \mathfrak{S}_n} q^{\text{inv}(w)}.$$

We can ask whether there is a bijective proof of this fact analogous to our proof of Proposition 1.7.3. In other words, letting $\mathcal{F}(n, q)$ denote the set of all flags (1.114), we want to find a map $\varphi : \mathcal{F}(n, q) \rightarrow \mathfrak{S}_n$ such that $\#\varphi^{-1}(w) = q^{\text{inv}(w)}$ for all $w \in \mathfrak{S}_n$. Such a map can be defined as follows. Let $F \in \mathcal{F}(n, q)$ be the flag (1.114). It is not hard to see that there is a unique ordered basis $\mathbf{v} = \mathbf{v}(F) = (v_1, v_2, \dots, v_n)$ for \mathbb{F}_q^n (where we regard each v_i as a column vector) satisfying the two conditions:

- $V_i = \text{span}\{v_1, \dots, v_i\}$, $1 \leq i \leq n$
- There is a unique permutation $\varphi(F) = w \in \mathfrak{S}_n$ for which the matrix $A = [v_1, \dots, v_n]^t$ satisfies (a) $A_{i, w(i)} = 1$ for $1 \leq i \leq n$, (b) $A_{i, j} = 0$ if $j > w(i)$, and (c) $A_{j, w(i)} = 0$ if $j > i$. In other words, A can be obtained from the permutation matrix P_w (as defined in Section 1.5) by replacing the entries A_{ij} for $(i, j) \in D_w$ (as defined in Section 1.5) by any elements of \mathbb{F}_q . We call A a w -reduced matrix.

For instance, suppose that $w = 314652$. Figure 1.5 shows that the possible matrices A have the form

$$A = \begin{bmatrix} * & * & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 1 & 0 & 0 \\ 0 & * & 0 & 0 & * & 1 \\ 0 & * & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Let Ω_w be the set of flags $F \in \mathcal{F}(n, q)$ for which $\varphi(F) = w$. Thus,

$$\mathcal{F}(n, q) = \bigcup_{w \in \mathfrak{S}_n} \Omega_w. \quad (1.115)$$

Since $\#D_w = \text{inv}(w)$, we have $\#\Omega_w = q^{\text{inv}(w)}$, so we have found the desired combinatorial interpretation of Proposition 1.10.12. The sets Ω_w are known as *Schubert cells*, and equation (1.115) gives the cellular decomposition of the flag variety $\mathcal{F}(n, q)$, completely analogous to the cellular decomposition of the Grassmann variety G_{km} given in the proof of Proposition 1.7.3. The canonical ordered basis $\nu(F)$ is the “flag analogue” of row-reduced echelon form, which gives a canonical ordered basis for a *subspace* (rather than a flag) of \mathbb{F}_q^n .

1.10.3 The Connection Between the Two q -Analogues

The order $\gamma(n, q)$ of $\text{GL}(n, q)$ and the number $f(n, q)$ of complete flags is related by

$$\gamma(n, q) = q^{\binom{n}{2}} (q-1)^n f(n, q).$$

Can we find a simple combinatorial explanation? We would like to find a map $\psi : \text{GL}(n, q) \rightarrow \mathcal{F}(n, q)$ satisfying $\#\psi^{-1}(F) = q^{\binom{n}{2}} (q-1)^n$ for all $F \in \mathcal{F}(n, q)$. The definition of ψ is quite simple: if $A = [v_1, \dots, v_n]^t$, then let $\psi(F)$ be the flag $\{0\} = V_0 \subset V_1 \subset \dots \subset V_n = \mathbb{F}_q^n$ given by $V_i = \text{span}\{v_1, \dots, v_i\}$. Given F , there are $q-1$ choices for v_1 , then $q^2 - q$ choices for v_2 , then $q^3 - q^2$ choices for v_3 , and so on, showing that $\#\psi^{-1}(F) = q^{\binom{n}{2}} (q-1)^n$ as desired.

We have constructed maps $\text{GL}(n, q) \xrightarrow{\psi} \mathcal{F}(n, q) \xrightarrow{\varphi} \mathfrak{S}_n$. Given $w \in \mathfrak{S}_n$, let $\Gamma_w = \psi^{-1}\varphi^{-1}(w)$. Thus,

$$\text{GL}(n, q) = \bigcup_{w \in \mathfrak{S}_n} \Gamma_w, \quad (1.116)$$

the *Bruhat decomposition* of $\text{GL}(n, q)$. (The Bruhat decomposition is usually defined more abstractly and in greater generality than we have done.) It is immediate from the formulas $\#\Omega_w = q^{\text{inv}(w)}$ and $\#\psi^{-1}(F) = q^{\binom{n}{2}} (q-1)^n$ that $\#\Gamma_w = q^{\binom{n}{2}} (q-1)^n q^{\text{inv}(w)}$ and

$$\gamma(n, q) = q^{\binom{n}{2}} (q-1)^n \sum_{w \in \mathfrak{S}_n} q^{\text{inv}(w)}. \quad (1.117)$$

Together with Corollary 1.3.13, equation (1.117) gives a second combinatorial proof of Proposition 1.10.1.

It is not difficult to give a concrete description of the “Bruhat cells” Γ_w . Namely, every element A of Γ_w can be uniquely written in the form $A = LM$, where L is a lower-triangular matrix in $\text{GL}(n, q)$ and M is a w -reduced matrix. We omit the straightforward proof.

1.10.13 Example. (a) Every matrix $A \in \text{GL}(2, q)$ can be uniquely written in one of the two forms

$$\begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & 0 \\ b & c \end{bmatrix}$$

$$\begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \begin{bmatrix} \alpha & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \alpha a & a \\ \alpha b + c & b \end{bmatrix},$$

where $b, \alpha \in \mathbb{F}_q$, $a, c \in \mathbb{F}_q^* = \mathbb{F}_q - \{0\}$.

(b) The cell Γ_{3142} consists of all matrices of the form

$$\begin{bmatrix} a & 0 & 0 & 0 \\ b & c & 0 & 0 \\ d & e & f & 0 \\ g & h & i & j \end{bmatrix} \begin{bmatrix} \alpha & \beta & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & \gamma & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \alpha a & \beta a & a & 0 \\ \alpha b + c & \beta b & b & 0 \\ \alpha d + e & \beta d + \gamma f & d & f \\ \alpha g + h & \beta g + \gamma i + j & g & i \end{bmatrix},$$

where $b, d, e, g, h, i, \alpha, \beta, \gamma \in \mathbb{F}_q$ and $a, c, f, j \in \mathbb{F}_q^*$.

The Bruhat decomposition (1.116) can be a useful tool for counting certain subsets S of $\text{GL}(n, q)$, by computing each $\#(S \cap \Gamma_w)$ and summing over all $w \in \mathfrak{S}_n$. Proposition 1.10.15 illustrates this technique. First we need a simple enumerative lemma.

1.10.14 Lemma. Fix q , and for any integer $n \geq 0$, let

$$a_n = \#\{(\alpha_1, \dots, \alpha_n) \in (\mathbb{F}_q^*)^n : \sum \alpha_i = 0\}.$$

Then $a_0 = 1$, and $a_n = \frac{1}{q}((q-1)^n + (q-1)(-1)^n)$ for $n > 0$.

Proof. Define $b_n = \sum_{k=0}^n \binom{n}{k} a_k$. Since every sequence $(\alpha_1, \dots, \alpha_n) \in \mathbb{F}_q^n$ satisfying $\sum \alpha_i = 0$ can be obtained by first specifying $n-k$ terms to be 0 in $\binom{n}{k}$ ways and then specifying the remaining k terms in a_k ways, we have

$$b_n = \begin{cases} 1, & n = 0 \\ q^{n-1}, & n \geq 1. \end{cases}$$

There are many ways to see (e.g., equations (2.9) and (2.10)) that we can invert this relationship between the a_n 's and b_n 's to obtain

$$\begin{aligned} a_n &= \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} b_k \\ &= \frac{1}{q} \left[\sum_{k=0}^n (-1)^{n-k} \binom{n}{k} q^k + (q-1)(-1)^n \right] \\ &= \frac{1}{q} ((q-1)^n + (q-1)(-1)^n). \end{aligned}$$

□

1.10.15 Proposition. Let $\text{GL}_0(n, q) = \{A \in \text{GL}(n, q) : \text{tr}(A) = 0\}$, where $\text{tr}(A)$ denotes the trace of A , and set $\gamma_0(n, q) = \#\text{GL}_0(n, q)$. Then

$$\gamma_0(n, q) = \frac{1}{q} \left(\gamma(n, q) + (-1)^n (q-1) q^{\binom{n}{2}} \right).$$

Proof. Let id denote the identity permutation $1, 2, \dots, n$, so $\text{inv}(\text{id}) = 0$. We will show that

$$\begin{aligned} \#(\text{GL}_0(n, q) \cap \Gamma_w) &= \frac{1}{q} \# \Gamma_w, \quad w \neq \text{id}, \\ \#(\text{GL}_0(n, q) \cap \Gamma_{\text{id}}) &= \frac{1}{q} \left(\# \Gamma_{\text{id}} + (-1)^n (q-1) q^{\binom{n}{2}} \right), \end{aligned}$$

from which the proof follows since $\sum_w \# \Gamma_w = \gamma(n, q)$.

Suppose that $w \neq \text{id}$. Let r be the least integer for which there exists an element $(r, s) \in D_w$, where D_w denotes the diagram of w . It is easy to see that then $(r, r) \in D_w$. Consider a general element $A = LM$ of Γ_w , so the entries L_{ij} satisfy $L_{ii} \in \mathbb{F}_q^*$, $L_{ij} \in \mathbb{F}_q$ if $i > j$, and $L_{ij} = 0$ if $i < j$. Similarly, $M_{i, w(i)} = 1$, $M_{ij} \in \mathbb{F}_q$ if $(i, j) \in D_w$, and $M_{ij} = 0$ otherwise. Thus, A_{rr} will be a polynomial in the L_{ij} 's and M_{ij} 's with a term $L_{rr} M_{rr}$. (In fact, it is not hard to see that $A_{rr} = L_{rr} M_{rr}$, though we don't need this stronger fact here.) There is no other occurrence of M_{rr} in a main diagonal term of A . If we choose all the free entries of L and M except M_{rr} (subject to the preceding conditions), then we can solve uniquely for M_{rr} (since its coefficient is $L_{rr} \neq 0$) so that $\text{tr}(A) = 0$. Thus, rather than q choices for M_{rr} for any $A \in \Gamma_w$, there is only one choice, so $\#(\text{GL}_0(n, q) \cap \Gamma_w) = \frac{1}{q} \# \Gamma_w$ as claimed.

Example. Consider the cell Γ_{3142} of Example 1.10.13(b). We have that $\#(\text{GL}_0(4, q) \cap \Gamma_{3142})$ is the number of 13-tuples $(a, \dots, j, \alpha, \beta, \gamma)$ such that $b, d, e, g, h, i, \alpha, \beta, \gamma \in \mathbb{F}_q$ and $a, c, f, j \in \mathbb{F}_q^*$, satisfying

$$\alpha a + \beta b + d + i = 0. \quad (1.118)$$

We have $r = 1$, so we can specify all 13 variables except α in $q^8(q-1)^4$ ways, and then solve equation (1.118) uniquely for α . Hence $\#(\text{GL}_0(4, q) \cap \Gamma_{3142}) = q^8(q-1)^4 = \frac{1}{q} \# \Gamma_{3142}$.

Now let $w = \text{id}$, so $A = L$. Hence, we can choose the elements of A below the diagonal in $q^{\binom{n}{2}}$ ways, while the number of choices for the diagonal elements is the number a_n of Lemma 1.10.14. Hence from Lemma 1.10.14 we get

$$\begin{aligned} \# \Gamma_{\text{id}} &= q^{\binom{n}{2}} a_n \\ &= q^{\binom{n}{2}} \frac{1}{q} ((q-1)^n + (q-1)(-1)^n), \end{aligned}$$

and the proof follows. □

Notes

It is not our intention here to trace the development of the basic ideas and results of enumerative combinatorics. It is interesting to note, however, that according to Heath [1.40, p. 319], a result of Xenocrates of Chalcedon (396–314 BCE) possibly “represents the first attempt on record to solve a difficult problem in permutations and combinations.” (See also Biggs [1.8, p. 113].) Moreover, Exercise 1.203 shows that Hipparchus (c. 190–after 127 BCE) certainly was successful in solving such a problem. We should also point out that the identity of Example 1.1.17 is perhaps the oldest of all binomial coefficient identities. It is called by such names as the *Chu-Vandermonde identity* or *Vandermonde’s theorem*, after Chu Shih-Chieh (Zhū Shìjié in Pinyin and 朱世杰 in simplified Chinese characters) (c. 1260–c. 1320) and Alexandre-Théophile Vandermonde (1735–1796).

Two valuable sources for the history of enumeration are Biggs [1.8] and Stein [1.71]. Knuth [1.49, §7.2.1.7] has written a fascinating history of the generation of combinatorial objects (such as all permutations of a finite set). Later we will give mostly references and comments not readily available in [1.8] and [1.71].

For further information on formal power series from a combinatorial viewpoint, see, for example Niven [1.60] and Tutte [1.73]. A rigorous algebraic approach appears in Bourbaki [1.12, Ch. IV, §5], and a further paper of interest is Bender [1.5]. Wilf [1.76] is a nice introduction to generating functions at the undergraduate level.

To illustrate the misconceptions (or at least infelicitous language) that can arise in dealing with formal power series, we offer the following quotations (anonymously) from the literature.

Since the sum of an infinite series is really not used, our viewpoint can be either rigorous or formal.

(1.3) demonstrates the futility of seeking a generating function, even an exponential one, for $IU(n)$; for it is so big that

$$F(z) = \sum_n IU(n)z^n/n!$$

fails to converge if $z \neq 0$. Any closed equation for F therefore has no solutions, and when manipulated by Taylor expansion, binomial theorem, etc., is bound to produce a heap of eggs (single -0- or double -∞-yolked). Try finding a generating function for 2^{2^n} .

Sometimes we have difficulties with convergence for some functions whose coefficients a_n grow too rapidly; then instead of the regular generating function we study the *exponential* generating function.

An analyst might at least raise the point that the only general techniques available for estimating the rate of growth of the coefficients of a power series require convergence (so that e.g. the apparatus of complex variable theory is available). There are, however, methods for estimating the coefficients of a divergent power series;

see Bender [1.6, §5] and Odlyzko [1.61, §7]. For further information on estimating coefficients of power series, see for instance Flajolet and Sedgewick [1.22], Odlyzko [1.61] and Pemantle and Wilson [1.63]. In particular, the asymptotic formula (1.12), due to Moser and Wyman [1.57], appears in [1.61, (8.49)].

The technique of representing combinatorial objects such as permutations by “models” such as words and trees has been extensively developed. A pioneering work in this area in the monograph [1.26] of Foata and Schützenberger. In particular, the “transformation fondamentale” on pp. 13–15 of this reference is essentially our map $w \mapsto \hat{w}$ of Proposition 1.3.1. Note, however, that this bijection was earlier used by Alfréd Rényi [1.67, §4] to prove Proposition 1.3.1. The history of the generating function for the cycle indicator of \mathfrak{S}_n (Theorem 1.3.3) is discussed in the first paragraph of the Notes to Chapter 5. The generating function for permutations by number of inversions (Corollary 1.3.13) appears in Rodrigues [1.68] and Netto [1.58, p. 73]. The generalization to multisets (Proposition 1.7.1) is due to MacMahon [1.54, §1]. It was rediscovered by Carlitz [1.16]. The second proof given here was suggested by A. Björner and M. L. Wachs [1.9, §3]. The cellular decomposition of the Grassmann variety (the basis for our second proof of Proposition 1.7.3) is discussed by S. L. Kleiman and D. Laksov [1.46]. For some further historical information on the results of Rodrigues and MacMahon, see the book review by Johnson [1.44]. The major index of a permutation was first considered by MacMahon [1.53], who used the term “greater index.” The terminology “major index” was introduced by Foata [1.25] in honor of MacMahon, who was a major in the British army. MacMahon’s main result on the major index is the equidistribution of $\text{inv}(w)$ and $\text{maj}(w)$ for $w \in \mathfrak{S}_n$. He gives the generating function (1.42) for $\text{maj}(w)$ in [1.53, §6] (where in fact w is a permutation of a multiset), and in [1.54] he shows the equidistribution with $\text{inv}(w)$. The bijective proof we have given here (proof of Proposition 1.4.6) appears in seminal papers [1.23][1.24] of Foata, which helped lay the groundwork for the modern theory of bijective proofs. The strengthening of Foata’s result given by Corollary 1.4.9 is due to Foata and Schützenberger [1.28].

The investigation of the descent set and number of descents of a permutation (of a set or multiset) was begun by MacMahon [1.52]. MacMahon apparently did not realize that the number of permutations of $[n]$ with k descents is an Eulerian number. The first written statement connecting Eulerian numbers with descents seems to have been by Carlitz and Riordan [1.17] in 1953. The fundamental Lemma 1.4.11 is due to MacMahon [1.53, p. 287]. Eulerian numbers occur in some unexpected contexts, such as cube slicing (Exercise 1.51), juggling sequences [1.15], and the statistics of carrying in the standard algorithm for adding integers (Exercise 1.52). MacMahon [1.55, vol. 1, p. 186] was also the first person to consider the excedance of a permutation (though he did not give it a name) and showed the equidistribution of the number of descents with the number of excedances (Proposition 1.4.3).

We will not attempt to survey the vast subject of representing permutations by other combinatorial objects, but let us mention that an important generalization of

the representation of permutations by plane trees is the paper of Cori [1.19]. The first result on pattern avoidance seems to be the proof of MacMahon [1.55, §97] that the number of 321-avoiding permutations $w \in \mathfrak{S}_n$ is the Catalan number C_n . MacMahon states his result not in terms of pattern avoidance, but rather in terms of permutations that are a union of two decreasing sequences. MacMahon's result was rediscovered by J. M. Hammersley [1.38], who stated it without proof. Proofs were later given by D. E. Knuth [1.48, §5.1.4] and D. Rotem [1.69]. For further information on 321-avoiding and 132-avoiding permutations, see Exercise 6.19(ee,ff) and the survey of Claesson and Kitaev [1.18]. For further information on pattern avoidance in general, see Exercises 57–59, as well as books by M. Bóna [1.11, Chs. 4–5] and by S. Heubach and T. Mansour [1.42].

Alternating permutations were first considered by D. André [1.1], who obtained the basic and elegant Proposition 1.6.1. (Note however that Ginsburg [1.34] asserts without giving a reference that Binet was aware before André that the coefficients of $\sec x$ count alternating permutations.) A combinatorial proof of Proposition 1.6.2 on flip equivalence is due to R. Donaghey [1.20]. Further information on the connection between alternating permutations and increasing trees appears in a paper of Kuznetsov, Pak, and Postnikov [1.51].

The cd -index $\Phi_n(c, d)$ was first considered by Foata and Schützenberger [1.27], who defined it in terms of certain permutations they called *André permutations*. Their term for the cd -index was “non-commutative André polynomial.” Foata and Strehl [1.29][1.30] further developed the theory of André polynomials, André permutations, and their connection with permutation statistics. Meanwhile Jonathan Fine [1.21] defined a noncommutative polynomial $\Phi_P(c, d)$ associated with certain partially ordered sets (posets) P . This polynomial was first systematically investigated by Bayer and Klapper [1.4] and later by Stanley [1.70], who extended the class of posets P which possessed a cd -index $\Phi_P(c, d)$ to Eulerian posets. The basic theory of the cd -index of an Eulerian poset is covered in Section 3.17. M. Purtill [1.65, Thm. 6.1] showed that the cd -index Φ_n that we have defined is just the cd -index Φ_{B_n} (in the sense of Fine and Bayer-Klapper) of the boolean algebra B_n (the poset of all subsets of $[n]$, ordered by inclusion). The approach to the cd -index Φ_n given here, based on min-max trees, is due to G. Hetyei and E. Reiner [1.41]. For some additional properties of min-max trees, see Bóna [1.10]. Corollary 1.6.5 was first proved by Niven [1.59] by a complicated induction. De Bruijn [1.13] gave a simpler proof and extended it to Proposition 1.6.4. A further proof is due to Viennot [1.75]. The proof we have given based on the cd -index appears in Stanley [1.70, pp. 495–496]. For a generalization see Exercise 3.55.

The theory of partitions of an integer originated in the work of Euler, if we ignore some unpublished work of Leibniz that was either trivial or wrong (see Knobloch [1.47]). An excellent introduction to this subject is the text by Andrews [1.2]. For a masterful survey of bijective proofs of partition identities, see Pak [1.62]. The latter two references provide historical information on the results appearing in

Section 1.8. The asymptotic formula (1.92) is due to Hardy and Ramanujan [1.39], and the asymptotic series mentioned after equation (1.92) is due to Rademacher [1.66]. More recently J. H. Bruinier and K. Ono,

(<http://www.aimath.org/news/partition/brunier-ono>),

have given an explicit finite formula for $p(n)$. For an exposition of partition asymptotics, see Andrews [1.2, Ch. 5].

The idea of the Twelvelfold Way (Section 1.9) is due to G.-C. Rota (in a series of lectures), while the terminology “Twelvelfold Way” was suggested by Joel Spencer. An extension of the Twelvelfold Way to a “Thirtyfold Way” (and suggestion of even more entries) is due to R. Proctor [1.64]. An interesting popular account of Bell numbers appears in an article by M. Gardner [1.33]. In particular, pictorial representations of the 52 partitions of a 5-element set are used as “chapter headings” for all but the first and last chapters of certain editions of *The Tale of Genji* by Lady Murasaki (c. 978–c. 1031 CE). A standard reference for the calculus of finite difference is the text by C. Jordan [1.45].

The cycle indicator $Z_n(t; q)$ of $\mathrm{GL}(n, q)$ was first explicitly defined by Kung [1.50]. The underlying algebra was known much earlier; for instance, according to Green [1.35, p. 407] the basic Theorem 1.10.7 is due to P. Hall [1.36] and is a simple consequence of earlier work of Frobenius (see Jacobson [1.43, Thm. 19, p. 111]). Green himself sketches a proof on page 409, op. cit. Further work on the cycle indicator of $\mathrm{GL}(n, q)$ was done by Stong [1.72] and Fulman [1.31]. A nice survey of enumeration of matrices over \mathbb{F}_q was given by Morrison [1.56], whom we have followed for Exercises 1.193–1.195. Our proof of Lemma 1.10.9 is equivalent to one given by P. Hall [1.37].

The cellular decomposition (1.115) of the flag variety $\mathcal{F}(n, q)$ and the Bruhat decomposition (1.116) of $\mathrm{GL}(n, K)$ (for any field K) are standard topics in Lie theory. See for instance Fulton and Harris [1.32, §23.4]. A complicated recursive description of the number of matrices in $\mathrm{GL}(n, q)$ with trace 0 and a given rank r was given by Buckheister [1.14]. Bender [1.7] used this recurrence to give a closed-form formula. The proof we have given of the case $k = 0$ (Proposition 1.10.15) based on Bruhat decomposition is new. For a generalization see Exercise 1.196.

Bibliography

- [1] D. André, Développement de $\sec x$ and $\mathrm{tg} x$, *C. R. Math. Acad. Sci. Paris* **88** (1879), 965–979.
- [2] G. E. Andrews, *The Theory of Partitions*, Addison-Wesley, Reading, Mass., 1976.
- [3] G. E. Andrews, ed., *Percy Alexander MacMahon, Collected Papers*, vol. 1, M.I.T. Press, Cambridge, Mass. 1978.
- [4] M. M. Bayer and A. Klapper, A new index for polytopes, *Discrete Comput. Geom.* **6** (1991), 33–47.
- [5] E. A. Bender, A lifting theorem for formal power series, *Proc. Amer. Math. Soc.* **42** (1974), 16–22.
- [6] E. A. Bender, Asymptotic methods in enumeration, *SIAM Review* **16** (1974), 485–515. Errata, *SIAM Review* **18** (1976), 292.

- [7] E. A. Bender, On Buckheister's enumeration of $n \times n$ matrices, *J. Combinatorial Theory, Ser. A* **17** (1974), 273–274.
- [8] N. L. Biggs, The roots of combinatorics, *Historia Math.* **6** (1979), 109–136.
- [9] A. Björner and M. L. Wachs, q -Hook length formulas for forests, *J. Combinatorial Theory Ser. A* **52** (1989), 165–187.
- [10] M. Bóna, A combinatorial proof of a result of Heteyi and Reiner on Foata-Strehl type permutation trees, *Ann. Combinatorics* **1** (1997), 119–122.
- [11] M. Bóna, *Combinatorics of Permutations*, Chapman and Hall/CRC, Boca Raton, 2004.
- [12] N. Bourbaki, *Éléments de Mathématique, Livre II, Algèbre*, Ch. 4–5, 2^e ed., Hermann, Paris, 1959.
- [13] N. G. de Bruijn, Permutations with given ups and downs, *Nieuw Arch. Wisk.* **18** (1970), 61–65.
- [14] P. G. Buckheister, The number of $n \times n$ matrices of rank r and trace α over a finite field, *Duke Math. J.* **39** (1972), 695–699.
- [15] J. Buhler, D. Eisenbud, R. Graham, and C. Wright, Juggling drops and descents, *Amer. Math. Monthly* **101** (1994), 507–519.
- [16] L. Carlitz, Sequences and inversions, *Duke Math. J.* **37** (1970), 193–198.
- [17] L. Carlitz and J. Riordan, Congruences for Eulerian numbers, *Duke Math. J.* **20** (1953), 339–344.
- [18] A. Claesson and S. Kitaev, Classification of bijections between 321- and 132-avoiding permutations, *Sém. Lothar. Combinatoire* **60** (2008), B60d.
- [19] R. Cori, Une code pour les graphes planaires et ses applications, *Astérisque*, no. 27, Société Mathématique de France, Paris, 1975.
- [20] R. Donaghey, Alternating permutations and binary increasing trees, *J. Combinatorial Theory Ser. A* **18** (1975), 141–148.
- [21] J. Fine, Morse theory for convex polytopes, unpublished manuscript, 1985.
- [22] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2009.
- [23] D. Foata, Sur un énoncé de MacMahon, *C. R. Acad. Sci. Paris* **258** (1964), 1672–1675.
- [24] D. Foata, On the Netto inversion number of a sequence, *Proc. Amer. Math. Soc.* **19** (1968), 236–240.
- [25] D. Foata, Distributions eulériennes et mahoniennes sur le groupe des permutations, in *Higher Combinatorics (Proc. NATO Advanced Study Inst., Berlin, 1976)*, Reidel, Dordrecht-Boston, Mass., 1977, pp. 27–49.
- [26] D. Foata and M.-P. Schützenberger, Théorie géométrique des polynômes Eulériens, *Lecture Notes in Math.*, no. 138, Springer, Berlin, 1970.
- [27] D. Foata and M.-P. Schützenberger, Nombres d'Euler et permutations alternantes, in *A Survey of Combinatorial Theory* (J. N. Srivastava et al., eds.), North-Holland, Amsterdam, 1973, pp. 173–187.
- [28] D. Foata and M.-P. Schützenberger, Major index and inversion number of permutations, *Math. Machr.* **83** (1978), 143–159.
- [29] D. Foata and V. Strehl, Euler numbers and variations of permutations, in *Colloquio Internazionale sulle Teorie Combinatorie (Roma, 1973)*, Tomo I, Atti dei Convegni Lincei, No. 17, Accad. Naz. Lincei, Rome, 1976, pp. 119–131.
- [30] D. Foata and V. Strehl, Rearrangements of the symmetric group and enumerative properties of the tangent and secant numbers, *Math. Z.* **137** (1974), 257–264.
- [31] J. Fulman, Random matrix theory over finite fields, *Bull. Amer. Math. Soc. (N.S.)* **39** (2002), 51–85.
- [32] W. Fulton and J. Harris, *Representation Theory*, Springer-Verlag, New York, 1991.
- [33] M. Gardner, Mathematical games, *Scientific American* **238** (May 1978), 24–30; reprinted (with an addendum) in *Fractal Music, Hypercards, and More*, Freeman, New York, 1992, pp. 24–38.
- [34] J. Ginsburg, Stirling numbers, *Encyclopedia Britannica*, 1965.
- [35] J. A. Green, The characters of the finite general linear groups, *Trans. Amer. Math. Soc.* **80** (1955), 402–447.

- [36] P. Hall, Abelian p -groups and related modules, unpublished manuscript.
- [37] P. Hall, A partition formula connected with Abelian groups, *Comm. Math. Helv.* **11** (1938/39), 126–129.
- [38] J. M. Hammersley, A few seedlings of research, in *Proc. Sixth Berkeley Symposium on Mathematical Statistics and Probability (Berkeley, 1970/1971)*, Vol. 1: *Theory of statistics*, Univ. California Press, Berkeley, Calif., 1972, pp. 345–394.
- [39] G. H. Hardy and S. Ramanujan, Asymptotic formulae in combinatory analysis, *Proc. London Math. Soc.* (2) **17** (1918), 75–117. Also in *Collected Papers of S. Ramanujan*, Cambridge University Press, London and New York, 1927; reprinted by Chelsea, New York, 1962.
- [40] T. L. Heath, *A History of Greek Mathematics*, vol. 1, Dover, New York, 1981.
- [41] G. Hetyei and E. Reiner, Permutations trees and variation statistics, *European J. Combinatorics* **19** (1998), 847–866.
- [42] S. Heubach and T. Mansour, *Combinatorics and Compositions on Words*, Chapman and Hall/CRC, Boca Raton, 2010.
- [43] N. Jacobson, *Lectures in Abstract Algebra*, vol. II—Linear Algebra, D. van Nostrand Co., Princeton, N.J., 1953.
- [44] W. P. Johnson, Review of *Mathematics and Social Utopias in France: Olinde Rodrigues and His Times*, *Amer. Math. Monthly* **114** (2007), 752–758.
- [45] C. Jordan, *Calculus of Finite Differences*, Chelsea, New York, 1965.
- [46] S. L. Kleiman and D. Laksov, Schubert calculus, *Amer. Math. Monthly* **79** (1972), 1061–1082.
- [47] E. Knobloch, Leibniz on combinatorics, *Historia Math.* **1** (1974), 409–430.
- [48] D. E. Knuth, *The Art of Computer Programming*, vol. 3, *Sorting and Searching*, Addison-Wesley, Reading, Mass., 1973; 2nd ed., 1998.
- [49] D. E. Knuth, *The Art of Computer Programming*, vol. 4, *Fascicle 4*, Addison-Wesley, Upper Saddle River, N.J., 2006.
- [50] J. P. S. Kung, The cycle structure of a linear transformation over a finite field, *Linear Algebra Appl.* **36** (1981), 141–155.
- [51] A. G. Kuznetsov, I. M. Pak, and A. E. Postnikov, Increasing trees and alternating permutations, *Russian Math. Surveys* **49:6** (1994), 79–114; translated from *Uspekhi Mat. Nauk* **49:6** (1994), 79–110.
- [52] P. A. MacMahon, Second memoir on the compositions of integers, *Phil. Trans.* **207** (1908), 65–134; reprinted in [1.3], pp. 687–756.
- [53] P. A. MacMahon, The indices of permutations and the derivation therefrom of functions of a single variable associated with the permutations of any assemblage of objects, *Amer. J. Math.* **35** (1913), 281–322; reprinted in [1.3], pp. 508–549.
- [54] P. A. MacMahon, Two applications of general theorems in combinatory analysis: (1) to the theory of inversions of permutations; (2) to the ascertainment of the numbers of terms in the development of a determinant which has amongst its elements an arbitrary number of zeros, *Proc. London Math. Soc.* (2) **15** (1916), 314–321; reprinted in [1.3], pp. 556–563.
- [55] P. A. MacMahon, *Combinatory Analysis*, vols. 1 and 2, Cambridge University Press, 1916; reprinted by Chelsea, New York, 1960, and by Dover, New York, 2004.
- [56] K. E. Morrison, Integer sequences and matrices over finite fields, *J. Integer Sequences* (electronic) **9** (2006), Article 06.2.1.
- [57] L. Moser and M. Wyman, On the solutions of $x^d = 1$ in symmetric groups, *Canad. J. Math.* **7** (1955), 159–168.
- [58] E. Netto, *Lehrbuch der Combinatorik*, Teubner, Leipzig, 1900.
- [59] I. Niven, A combinatorial problem of finite sequences, *Nieuw Arch. Wisk.* **16** (1968), 116–123.
- [60] I. Niven, Formal power series, *Amer. Math. Monthly* **76** (1969), 871–889.
- [61] A. Odlyzko, Asymptotic enumeration methods, in *Handbook of Combinatorics*, vol. 2 (R. L. Graham, M. Groetschel, and L. Lovász, eds.), Elsevier, Amsterdam, 1995, pp. 1063–1229.
- [62] I. Pak, Partition bijections. A survey, *Ramanujan J.* **12** (2006), 5–76.

- [63] R. Pemantle and M. Wilson, Twenty combinatorial examples of asymptotics derived from multivariate generating functions, *SIAM Review* **50** (2008), 199–272.
- [64] R. Proctor, Let's expand Rota's Twelvifold Way for counting partitions!, preprint; [arXiv:math/0606404](https://arxiv.org/abs/math/0606404).
- [65] M. Purtill, André permutations, lexicographic shellability and the cd -index of a convex polytope, *Trans. Amer. Math. Soc.* **338** (1993), 77–104.
- [66] H. Rademacher, On the partition function $p(n)$, *Proc. London Math. Soc.* (2) **43** (1937), 241–254.
- [67] A. Rényi, Théorie des éléments saillants d'une suite d'observations, in *Colloq. Combinatorial Methods in Probability Theory, August 1–10, 1962*, Matematisk Institut, Aarhus Universitet, 1962, pp. 104–115.
- [68] O. Rodrigues, Note sur les inversions, ou dérangements produits dans les permutations, *J. Math. Pures Appl.* **4** (1839), 236–240.
- [69] D. Rotem, On a correspondence between binary trees and a certain type of permutation, *Inf. Proc. Letters* **4** (1975/76), 58–61.
- [70] R. P. Stanley, Flag f -vectors and the cd -index, *Math. Z.* **216** (1994), 483–499.
- [71] P. R. Stein, A brief history of enumeration, in *Science and Computers, a volume dedicated to Nicolas Metropolis* (G.-C. Rota, ed.), Academic Press, New York, 1986, pp. 169–206.
- [72] R. A. Stong, Some asymptotic results on finite vector spaces, *Advances in Applied Math.* **9** (1988), 167–199.
- [73] W. T. Tutte, On elementary calculus and the Good formula, *J. Combinatorial Theory* **18** (1975), 97–137.
- [74] L. G. Valiant, The complexity of enumeration and reliability problems, *SIAM J. Comput.* **8** (1979), 410–421.
- [75] G. Viennot, Permutations ayant une forme donnée, *Discrete Math.* **26** (1979), 279–284.
- [76] H. S. Wilf, *Generatingfunctionology*, 3rd ed., A K Peters, Wellesley, Mass., 2006.

A Note About the Exercises

Each exercise is given a difficulty rating, as follows.

1. Routine, straightforward
2. Somewhat difficult or tricky
3. Difficult
4. Horrendously difficult
5. Unsolved

Further gradations are indicated by + and –. Thus, [1–] denotes an utterly trivial problem, and [5–] denotes an unsolved problem that has received little attention and may not be too difficult. A rating of [2+] denotes about the hardest problem that could be reasonably assigned to a class of graduate students. A few students may be capable of solving a [3–] problem, whereas almost none could solve a [3] in a reasonable period of time. Of course the ratings are subjective, and there is always the possibility of an overlooked simple proof that would lower the rating. Some problems (seemingly) require results or techniques from other branches of mathematics that are not usually associated with combinatorics. Here the rating is less meaningful – it is based on an assessment of how likely the reader is to discover for herself or himself the relevance of these outside techniques and results. An asterisk after the difficulty rating indicates that no solution is provided.

Exercises for Chapter 1

1. [1–] Let S and T be disjoint one-element sets. Find the number of elements of their union $S \cup T$.
2. [1+] We continue with a dozen simple numerical problems. Find as simple a solution as possible.

- a. How many subsets of the set $[10] = \{1, 2, \dots, 10\}$ contain at least one odd integer?
 - b. In how many ways can seven people be seated in a circle if two arrangements are considered the same whenever each person has the same neighbors (not necessarily on the same side)?
 - c. How many permutations $w : [6] \rightarrow [6]$ satisfy $w(1) \neq 2$?
 - d. How many permutations of $[6]$ have exactly two cycles (i.e., find $c(6, 2)$)?
 - e. How many partitions of $[6]$ have exactly three blocks (i.e., find $S(6, 3)$)?
 - f. There are four men and six women. Each man marries one of the women. In how many ways can this be done?
 - g. Ten people split up into five groups of two each. In how many ways can this be done?
 - h. How many compositions of 19 use only the parts 2 and 3?
 - i. In how many different ways can the letters of the word MISSISSIPPI be arranged if the four S's cannot appear consecutively?
 - j. How many sequences $(a_1, a_2, \dots, a_{12})$ are there consisting of four 0's and eight 1's, if no two consecutive terms are both 0's?
 - k. A box is filled with three blue socks, three red socks, and four chartreuse socks. Eight socks are pulled out, one at a time. In how many ways can this be done? (Socks of the same color are indistinguishable.)
 - l. How many functions $f : [5] \rightarrow [5]$ are at most two-to-one, that is, $\#f^{-1}(n) \leq 2$ for all $n \in [5]$?
3. Give *combinatorial* proofs of the following identities, where x, y, n, a, b are nonnegative integers.
- a. $[2-] \sum_{k=0}^n \binom{x+k}{k} = \binom{x+n+1}{n}$
 - b. $[1+] \sum_{k=0}^n k \binom{n}{k} = n2^{n-1}$
 - c. $[3] \sum_{k=0}^n \binom{2k}{k} \binom{2(n-k)}{n-k} = 4^n$
 - d. $[3-] \sum_{k=0}^m \binom{x+y+k}{k} \binom{y}{a-k} \binom{x}{b-k} = \binom{x+a}{b} \binom{y+b}{a}$, where $m = \min(a, b)$
 - e. $[1] 2 \binom{2n-1}{n} = \binom{2n}{n}$
 - f. $[2-] \sum_{k=0}^n (-1)^k \binom{n}{k} = 0, n \geq 1$
 - g. $[2+] \sum_{k=0}^n \binom{n}{k}^2 x^k = \sum_{j=0}^n \binom{n}{j} \binom{2n-j}{n} (x-1)^j$
 - h. $[3-] \sum_{i+j+k=n} \binom{i+j}{i} \binom{j+k}{j} \binom{k+i}{k} = \sum_{r=0}^n \binom{2r}{r}$, where $i, j, k \in \mathbb{N}$
4. $[2]^*$ Fix $j, k \in \mathbb{Z}$. Show that

$$\sum_{n \geq 0} \frac{(2n-j-k)!x^n}{(n-j)!(n-k)!(n-j-k)!n!} = \left[\sum_{n \geq 0} \frac{x^n}{n!(n-j)!} \right] \left[\sum_{n \geq 0} \frac{x^n}{n!(n-k)!} \right].$$

Any term with $(-r)!$ in the denominator, where $r > 0$, is set equal to 0.

5. [2]* Show that

$$\sum_{n_1, \dots, n_k \geq 0} \min(n_1, \dots, n_k) x_1^{n_1} \cdots x_k^{n_k} = \frac{x_1 \cdots x_k}{(1-x_1) \cdots (1-x_k)(1-x_1 x_2 \cdots x_k)}.$$

6. [3-]* For $n \in \mathbb{Z}$ let

$$J_n(2x) = \sum_{k \in \mathbb{Z}} \frac{(-1)^k x^{n+2k}}{k!(n+k)!},$$

where we set $1/j! = 0$ for $j < 0$. Show that

$$e^x = \sum_{n \geq 0} L_n J_n(2x),$$

where $L_0 = 1, L_1 = 1, L_2 = 3, L_{n+1} = L_n + L_{n-1}$ for $n \geq 2$. (The numbers L_n for $n \geq 1$ are *Lucas numbers*.)

7. [2]* Let

$$e^{x+\frac{x^2}{2}} = \sum_{n \geq 0} f(n) \frac{x^n}{n!}.$$

Find a simple expression for $\sum_{i=0}^n (-1)^{n-i} \binom{n}{i} f(i)$. (See equation (1.13).)

8. a. [2-] Show that

$$\frac{1}{\sqrt{1-4x}} = \sum_{n \geq 0} \binom{2n}{n} x^n.$$

b. [2-] Find $\sum_{n \geq 0} \binom{2n-1}{n} x^n$.

9. Let $f(m, n)$ be the number of paths from $(0, 0)$ to $(m, n) \in \mathbb{N} \times \mathbb{N}$, where each step is of the form $(1, 0)$, $(0, 1)$, or $(1, 1)$.

a. [1+]* Show that $\sum_{m \geq 0} \sum_{n \geq 0} f(m, n) x^m y^n = (1 - x - y - xy)^{-1}$.

b. [3-] Find a simple explicit expression for $\sum_{n \geq 0} f(n, n) x^n$.

10. [2+] Let $f(n, r, s)$ denote the number of subsets S of $[2n]$ consisting of r odd and s even integers, with no two elements of S differing by 1. Give a bijective proof that $f(n, r, s) = \binom{n-r}{s} \binom{n-s}{r}$.

11. a. [2+] Let $m, n \in \mathbb{N}$. Interpret the integral

$$B(m+1, n+1) = \int_0^1 u^m (1-u)^n du,$$

as a probability and evaluate it by combinatorial reasoning.

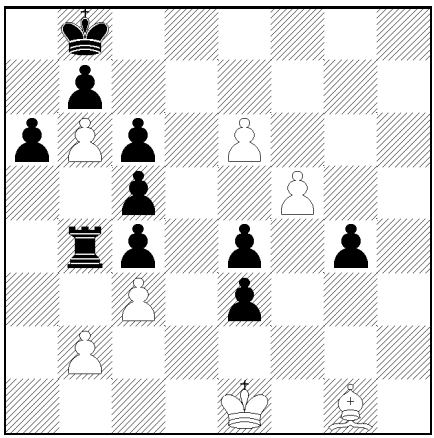
- b. [3+] Let $n \in \mathbb{P}$ and $r, s, t \in \mathbb{N}$. Let x, y_k, z_k and a_{ij} be indeterminates, with $1 \leq k \leq n$ and $1 \leq i < j \leq n$. Let M be the multiset with n occurrences of x , r occurrences of each y_k , s occurrences of each z_k , and $2t$ occurrences of each a_{ij} . Let $f(n, r, s, t)$ be the number of permutations w of M such that (i) all y_k 's appear before the k th x (reading the x 's from left-to-right in w), (ii) all z_k 's appear after the k th x , and (iii) all a_{ij} 's appear between the i th x and j th x . Show that

$$f(n, r, s, t) = \frac{[(r+s+1)n + tn(n-1)!]}{n! r!^n s!^n t!^n (2t)! \binom{n}{2}} \cdot \prod_{j=1}^n \frac{(r+(j-1)t)!(s+(j-1)t)!(jt)!}{(r+s+1+(n+j-2)t)!}. \quad (1.119)$$

- c. [3-] Consider the following chess position.

R. Stanley

Suomen Tehtävänikat, 2005



Black is to make 14 consecutive moves, after which White checkmates Black in one move. Black may not move into check, and may not check White (except possibly on his last move). Black and White are *cooperating* to achieve the aim of checkmate. (In chess problem parlance, this problem is called a *serieshelpmate in 14*.) How many different solutions are there?

12. [2+]* Choose n points on the circumference of a circle in “general position.” Draw all $\binom{n}{2}$ chords connecting two of the points. (“General position” means that no three of these chords intersect in a point.) Into how many regions will the interior of the circle be divided? Try to give an elegant proof avoiding induction, finite differences, generating functions, summations, and the like.
13. [2] Let p be prime and $a \in \mathbb{P}$. Show *combinatorially* that $a^p - a$ is divisible by p . (A combinatorial proof would consist of exhibiting a set S with $a^p - a$ elements and a partition of S into pairwise disjoint subsets, each with p elements.)
14. a. [2+] Let p be a prime, and let $n = \sum a_i p^i$ and $m = \sum b_i p^i$ be the p -ary expansions of the positive integers m and n . Show that

$$\binom{n}{m} \equiv \binom{a_0}{b_0} \binom{a_1}{b_1} \cdots \pmod{p}.$$

- b. [3–] Use (a) to determine when $\binom{n}{m}$ is odd. For what n is $\binom{n}{m}$ odd for all $0 \leq m \leq n$? In general, how many coefficients of the polynomial $(1+x)^n$ are odd?
- c. [2+] It follows from (a), and is easy to show directly, that $\binom{pa}{pb} \equiv \binom{a}{b} \pmod{p}$. Give a *combinatorial proof* that in fact $\binom{pa}{pb} \equiv \binom{a}{b} \pmod{p^2}$.
- d. [3–] If $p \geq 5$, then show in fact

$$\binom{pa}{pb} \equiv \binom{a}{b} \pmod{p^3}.$$

Is there a combinatorial proof?

- e. [3–] Give a simple description of the largest power of p dividing $\binom{n}{m}$.

15. a. [2] How many coefficients of the polynomial $(1+x+x^2)^n$ are not divisible by 3?
 b. [3–] How many coefficients of the polynomial $(1+x+x^2)^n$ are odd?
 c. [2+] How many coefficients of the polynomial $\prod_{1 \leq i < j \leq n} (x_i + x_j)$ are odd?
16. [3–]*
 a. Let p be a prime, and let A be the matrix $A = \left[\binom{j+k}{k} \right]_{j,k=0}^{p-1}$, taken over the field \mathbb{F}_p . Show that $A^3 = I$, the identity matrix. (Note that A vanishes below the main antidiagonal, i.e., $A_{jk} = 0$ if $j+k \geq p$.)
 b. How many eigenvalues of A are equal to 1?
17. a. [1+]* Let $m, n \in \mathbb{N}$. Prove the identity $\left(\binom{n}{m} \right) = \left(\binom{m+1}{n-1} \right)$.
 b. [2–] Give a combinatorial proof.
18. [2+]* Find a *simple* description of all $n \in \mathbb{P}$ with the following property: There exists $k \in [n]$ such that $\binom{n}{k-1}, \binom{n}{k}, \binom{n}{k+1}$ are in arithmetic progression.
19. a. [2+] Let $a_1, \dots, a_n \in \mathbb{N}$. Show that when we expand the product

$$\prod_{\substack{i,j=1 \\ i \neq j}}^n \left(1 - \frac{x_i}{x_j} \right)^{a_i}$$

as a Laurent polynomial in x_1, \dots, x_n (i.e., negative exponents allowed), then the constant term is the multinomial coefficient $\binom{a_1 + \dots + a_n}{a_1, \dots, a_n}$.

Hint: First prove the identity

$$1 = \sum_{i=1}^n \prod_{j \neq i} \left(1 - \frac{x_i}{x_j} \right)^{-1}. \quad (1.120)$$

- b. [2–] Put $n = 3$ to deduce the identity

$$\sum_{k=-a}^a (-1)^k \binom{a+b}{a+k} \binom{b+c}{b+k} \binom{c+a}{c+k} = \binom{a+b+c}{a, b, c}.$$

(Set $\binom{m}{i} = 0$ if $i < 0$.) Note that if we specialize $a = b = c$, then we obtain

$$\sum_{k=0}^{2a} (-1)^k \binom{2a}{k}^3 = \binom{3a}{a, a, a}.$$

- c. [3+] Let q be an additional indeterminate. Show that when we expand the product

$$\prod_{1 \leq i < j \leq k} \left(1 - q \frac{x_i}{x_j} \right) \left(1 - q^2 \frac{x_i}{x_j} \right) \cdots \left(1 - q^{a_i} \frac{x_i}{x_j} \right) \cdot \left(1 - \frac{x_j}{x_i} \right) \left(1 - q \frac{x_j}{x_i} \right) \cdots \left(1 - q^{a_j-1} \frac{x_j}{x_i} \right) \quad (1.121)$$

as a Laurent polynomial in x_1, \dots, x_n (whose coefficients are now polynomials in q), then the constant term is the q -multinomial coefficient $\binom{a_1 + \dots + a_n}{a_1, \dots, a_n}$.

- d. [3+] Let $k \in \mathbb{P}$. When the product

$$\prod_{1 \leq i < j \leq n} \left[\left(1 - \frac{x_i}{x_j} \right) \left(1 - \frac{x_j}{x_i} \right) (1 - x_i x_j) \left(1 - \frac{1}{x_i x_j} \right) \right]^k$$

is expanded as earlier, show that the constant term is

$$\binom{k}{k} \binom{3k}{k} \binom{5k}{k} \cdots \binom{(2n-3)k}{k} \cdot \binom{(n-1)k}{k}.$$

- e. [3–] Let $f(a_1, a_2, \dots, a_n)$ denote the constant term of the Laurent polynomial

$$\prod_{i=1}^n (q^{-a_i} + q^{-a_i+1} + \cdots + q^{a_i}),$$

where each $a_i \in \mathbb{N}$. Show that

$$\begin{aligned} & \sum_{a_1, \dots, a_n \geq 0} f(a_1, \dots, a_n) x_1^{a_1} \cdots x_n^{a_n} \\ &= (1+x_1) \cdots (1+x_n) \sum_{i=1}^n \frac{x_i^{n-1}}{(1-x_i^2) \prod_{j \neq i} (x_i - x_j)(1-x_i x_j)}. \end{aligned}$$

20. [2]* How many $m \times n$ matrices of 0's and 1's are there, such that every row and column contains an even number of 1's? An odd number of 1's?
21. [2]* Fix $n \in \mathbb{P}$. In how many ways (as a function of n) can one choose a composition α of n , and then choose a composition of each part of α ? (Give an elegant combinatorial proof.)
22. a. [2] Find the number of compositions of $n > 1$ with an even number of even parts. Naturally a combinatorial proof is preferred.
b. [2+] Let $e(n)$, $o(n)$, and $k(n)$ denote, respectively, the number of partitions of n with an even number of even parts, with an odd number of even parts, and that are self-conjugate. Show that $e(n) - o(n) = k(n)$. Is there a simple combinatorial proof?
23. [2] Give a simple “balls into boxes” proof that the total number of parts of all compositions of n is equal to $(n+1)2^{n-2}$. (The simplest argument expresses the answer as a sum of two terms.)
24. [2+] Let $1 \leq k < n$. Give a combinatorial proof that among all 2^{n-1} compositions of n , the part k occurs a total of $(n-k+3)2^{n-k-2}$ times. For instance, if $n=4$ and $k=2$, then the part 2 appears once in $2+1+1$, $1+2+1$, $1+1+2$, and twice in $2+2$, for a total of five times.
25. [2+] Let $n-r=2k$. Show that the number $f(n, r, s)$ of compositions of n with r odd parts and s even parts is given by $\binom{r+s}{r} \binom{r+k-1}{r+s-1}$. Give a generating function proof and a bijective proof.
26. [2]* Let $\bar{c}(m, n)$ denote the number of compositions of n with largest part at most m . Show that

$$\sum_{n \geq 0} \bar{c}(m, n) x^n = \frac{1-x}{1-2x+x^{m+1}}.$$

27. [2+] Find a simple explicit formula for the number of compositions of $2n$ with largest part exactly n .
28. [2]* Let $\kappa(n, j, k)$ be the number of weak compositions of n into k parts, each part less than j . Give a generating function proof that

$$\kappa(n, j, k) = \sum_{r+s=j=n} (-1)^s \binom{k+r-1}{r} \binom{k}{s},$$

where the sum is over all pairs $(r, s) \in \mathbb{N}^2$ satisfying $r+s=j=n$.

29. [2]* Fix $k, n \in \mathbb{P}$. Show that

$$\sum a_1 \cdots a_k = \binom{n+k-1}{2k-1},$$

where the sum ranges over all compositions (a_1, \dots, a_k) of n into k parts.

30. [2] Fix $1 \leq k \leq n$. How many integer sequences $1 \leq a_1 < a_2 < \cdots < a_k \leq n$ satisfy $a_i \equiv i \pmod{2}$ for all i ?
31. [2+]
- Let $\#N = n$, $\#X = x$. Find a simple explicit expression for the number of ways of choosing a function $f: N \rightarrow X$ and then linearly ordering each block of the coimage of f . (The elements of N and X are assumed to be distinguishable.)
 - How many ways as in (a) are there if f must be surjective? (Give a simple explicit answer.)
 - How many ways as in (a) are there if the elements of X are indistinguishable? (Express your answer as a finite sum.)
32. [2] Fix positive integers n and k . Let $\#S = n$. Find the number of k -tuples (T_1, T_2, \dots, T_k) of subsets T_i of S subject to each of the following conditions *separately*, that is, the three parts are independent problems (all with the same general method of solution).
- $T_1 \subseteq T_2 \subseteq \cdots \subseteq T_k$.
 - The T_i 's are pairwise disjoint.
 - $T_1 \cup T_2 \cup \cdots \cup T_k = S$.
33. a. [2-]* Let $k, n \geq 1$. Find the number of sequences $\emptyset = S_0, S_1, \dots, S_k$ of subsets of $[n]$ if for all $1 \leq i \leq k$ we have either (i) $S_{i-1} \subset S_i$ and $|S_i - S_{i-1}| = 1$, or (ii) $S_i \subset S_{i-1}$ and $|S_{i-1} - S_i| = 1$.
- b. [2+]* Suppose that we add the additional condition that $S_k = \emptyset$. Show that now the number $f_k(n)$ of sequences is given by

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

Note that $f_k(n) = 0$ if k is odd.

34. [2] Fix $n, j, k \in \mathbb{P}$. How many integer sequences are there of the form $1 \leq a_1 < a_2 < \cdots < a_k \leq n$, where $a_{i+1} - a_i \geq j$ for all $1 \leq i \leq k-1$?
35. The *Fibonacci numbers* are defined by $F_1 = 1$, $F_2 = 1$, $F_n = F_{n-1} + F_{n-2}$ if $n \geq 3$. Express the following numbers in terms of the Fibonacci numbers.
- [2-] The number of subsets S of the set $[n] = \{1, 2, \dots, n\}$ such that S contains no two consecutive integers.
 - [2] The number of compositions of n into parts greater than 1.
 - [2-] The number of compositions of n into parts equal to 1 or 2.
 - [2] The number of compositions of n into odd parts.
 - [2] The number of sequences $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ of 0's and 1's such that $\varepsilon_1 \leq \varepsilon_2 \geq \varepsilon_3 \leq \varepsilon_4 \geq \varepsilon_5 \leq \cdots$.
 - [2+] $\sum a_1 a_2 \cdots a_k$, where the sum is over all 2^{n-1} compositions $a_1 + a_2 + \cdots + a_k = n$.
 - [2+] $\sum (2^{a_1-1} - 1) \cdots (2^{a_k-1} - 1)$, summed over the same set as in (f).
 - [2+] $\sum 2^{\#\{i: a_i=1\}}$, summed over the same set as (f).
 - [2+] $\sum (-1)^{k-1} (5^{a_1-1} + 1) \cdots (5^{a_k-1} + 1)$, summed over the same set as (f).
 - [2+]* The number of sequences $(\delta_1, \delta_2, \dots, \delta_n)$ of 0's, 1's, and 2's such that 0 is never immediately followed by 1.

- k. [2+] The number of distinct terms of the polynomial

$$P_n = \prod_{j=1}^n (1 + x_j + x_{j+1}).$$

For instance, setting $x_1 = a$, $x_2 = b$, $x_3 = c$, we have $P_2 = 1 + a + 2b + c + ab + b^2 + ac + bc$, which has eight distinct terms.

36. [2] Fix $k, n \in \mathbb{P}$. Find a simple expression involving Fibonacci numbers for the number of sequences (T_1, T_2, \dots, T_k) of subsets T_i of $[n]$ such that

$$T_1 \subseteq T_2 \supseteq T_3 \subseteq T_4 \supseteq \dots.$$

37. [2] Show that

$$F_{n+1} = \sum_{k=0}^n \binom{n-k}{k}. \quad (1.122)$$

38. [2]* Show that the number of permutations $w \in \mathfrak{S}_n$ fixed by the fundamental transformation $\mathfrak{S}_n \xrightarrow{\wedge} \mathfrak{S}_n$ of Proposition 1.3.1 (i.e., $w = \widehat{w}$) is the Fibonacci number F_{n+1} .

39. [2+] Show that the number of ordered pairs (S, T) of subsets of $[n]$ satisfying $s > \#T$ for all $s \in S$ and $t > \#S$ for all $t \in T$ is equal to the Fibonacci number F_{2n+2} .

40. [2]* Suppose that n points are arranged on a circle. Show that the number of subsets of these points containing no two that are consecutive is the Lucas number L_n . This result shows that the Lucas number L_n may be regarded as a “circular analogue” of the Fibonacci number F_{n+2} (via Exercise 1.35(a)). For further explication, see Example 4.7.16.

41. a. [2] Let $f(n)$ be the number of ways to choose a subset $S \subseteq [n]$ and a permutation $w \in \mathfrak{S}_n$ such that $w(i) \notin S$ whenever $i \in S$. Show that $f(n) = F_{n+1}n!$.

- b. [2+] Suppose that in (a) we require w to be an n -cycle. Show that the number of ways is now $g(n) = L_n(n-1)!$, where L_n is a Lucas number.

42. [3] Let

$$\begin{aligned} F(x) &= \prod_{n \geq 2} (1 - x^{F_n}) = (1-x)(1-x^2)(1-x^3)(1-x^5)(1-x^8) \cdots \\ &= 1 - x - x^2 + x^4 + x^7 - x^8 + x^{11} - x^{12} - x^{13} + x^{14} + x^{18} + \cdots. \end{aligned}$$

Show that every coefficient of $F(x)$ is equal to $-1, 0$, or 1 .

43. [2–] Using only the combinatorial definitions of the Stirling numbers $S(n, k)$ and $c(n, k)$, give formulas for $S(n, 1)$, $S(n, 2)$, $S(n, n)$, $S(n, n-1)$, $S(n, n-2)$ and $c(n, 1)$, $c(n, 2)$, $c(n, n)$, $c(n, n-1)$, $c(n, n-2)$. For the case $c(n, 2)$, express your answer in terms of the harmonic number $H_m = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{m}$ for suitable m .

44. a. [2]* Show that the total number of cycles of all even permutations of $[n]$ and the total number of cycles of all odd permutations of $[n]$ differ by $(-1)^n (n-2)!$. Use generating functions.

- b. [3–]* Give a bijective proof.

45. [2+] Let $S(n, k)$ denote a Stirling number of the second kind. The generating function $\sum_n S(n, k)x^n = x^k / (1-x)(1-2x) \cdots (1-kx)$ implies the identity

$$S(n, k) = \sum 1^{a_1-1} 2^{a_2-1} \cdots k^{a_k-1}, \quad (1.123)$$

the sum being over all compositions $a_1 + \cdots + a_k = n$. Give a *combinatorial* proof of (1.123) analogous to the second proof of Proposition 1.3.7. That is, we want to associate with each partition π of $[n]$ into k blocks a composition $a_1 + \cdots + a_k = n$ such that exactly $1^{a_1-1}2^{a_2-1}\cdots k^{a_k-1}$ partitions π are associated with this composition.

46. a. [2] Let $n, k \in \mathbb{P}$, and let $j = \lfloor k/2 \rfloor$. Let $S(n, k)$ denote a Stirling number of the second kind. Give a generating function proof that

$$S(n, k) \equiv \binom{n-j-1}{n-k} \pmod{2}.$$

b. [3–] Give a combinatorial proof.

c. [2] State and prove an analogous result for Stirling numbers of the first kind.

47. Let D be the operator $\frac{d}{dx}$.

a. [2]* Show that $(xD)^n = \sum_{k=0}^n S(n, k)x^k D^k$.

b. [2]* Show that

$$x^n D^n = xD(xD-1)(xD-2)\cdots(xD-n+1) = \sum_{k=0}^n S(n, k)(xD)^k.$$

c. [2+]* Find the coefficients $a_{n,i,j}$ in the expansion

$$(x + D)^n = \sum_{i,j} a_{n,i,j} x^i D^j.$$

48. a. [3] Let $P(x) = a_0 + a_1x + \cdots + a_nx^n$, $a_i \geq 0$, be a polynomial all of whose zeros are negative real numbers. Regard $a_k/P(1)$ as the probability of choosing k , so we have a probability distribution on $[0, n]$. Let $\mu = \frac{1}{P(1)} \sum_k k a_k = P'(1)/P(1)$, the *mean* of the distribution; and let m be the *mode* (i.e., $a_m = \max_k a_k$). Show that

$$|\mu - m| < 1.$$

More precisely, show that

$$\begin{aligned} m &= k, & \text{if } k \leq \mu < k + \frac{1}{k+2}, \\ m &= k, \text{ or } k+1, \text{ or both,} & \text{if } k + \frac{1}{k+2} \leq \mu \leq k+1 - \frac{1}{n-k+1}, \\ m &= k+1, & \text{if } k+1 - \frac{1}{n-k+1} < \mu \leq k+1. \end{aligned}$$

b. [2] Fix n . Show that the signless Stirling number $c(n, k)$ is maximized at $k = \lceil 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \rceil$ or $k = \lfloor 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \rfloor$. In particular, $k \sim \log(n)$.

c. [3] Let $S(n, k)$ denote a Stirling number of the second kind, and define K_n by $S(n, K_n) \geq S(n, k)$ for all k . Let t be the solution of the equation $te^t = n$. Show that for sufficiently large n (and probably all n), either $K_n + 1 = \lfloor e^t \rfloor$ or $K_n + 1 = \lceil e^t \rceil$.

49. a. [2+] Deduce from equation (1.38) that all the (complex) zeros of $A_d(x)$ are real and simple. (Use Rolle's theorem.)

b. [2–]* Deduce from Exercise 1.133(b) that the polynomial $\sum_{k=1}^n k! S(n, k)x^k$ has only real zeros.

50. A sequence $\alpha = (a_0, a_1, \dots, a_n)$ of real numbers is *unimodal* if for some $0 \leq j \leq n$ we have $a_0 \leq a_1 \leq \cdots \leq a_j \geq a_{j+1} \geq a_{j+2} \geq \cdots \geq a_n$, and is *log-concave* if $a_i^2 \geq a_{i-1}a_{i+1}$ for $1 \leq i \leq n-1$. We also say that α has *no internal zeros* if there does not exist $i < j < k$ with $a_i \neq 0$, $a_j = 0$, $a_k \neq 0$, and that α is *symmetric* if $a_i = a_{n-i}$ for all i .

Define a polynomial $P(x) = \sum a_i x^i$ to be unimodal, log-concave, and so on, if the sequence (a_0, a_1, \dots, a_n) of coefficients has that property.

- a. [2-]* Show that a log-concave sequence of nonnegative real numbers with no internal zeros is unimodal.
- b. [2+] Let $P(x) = \sum_{i=0}^n a_i x^i = \sum_{i=0}^n \binom{n}{i} b_i x^i \in \mathbb{R}[x]$. Show that if all the zeros of $P(x)$ are real, then the sequence (b_0, b_1, \dots, b_n) is log-concave. (When all $a_i \geq 0$, this statement is stronger than the assertion that (a_0, a_1, \dots, a_n) is log-concave.)
- c. [2+] Let $P(x) = \sum_{i=0}^m a_i x^i$ and $Q(x) = \sum_{i=0}^n b_i x^i$ be symmetric, unimodal, and have nonnegative coefficients. Show that the same is true for $P(x)Q(x)$.
- d. [2+] Let $P(x)$ and $Q(x)$ be log-concave with no internal zeros and nonnegative coefficients. Show that the same is true for $P(x)Q(x)$.
- e. [2] Show that the polynomials $\sum_{w \in \mathfrak{S}_n} x^{\text{des}(w)}$ and $\sum_{w \in \mathfrak{S}_n} x^{\text{inv}(w)}$ are symmetric and unimodal.
- f. [4-] Let $1 \leq p \leq n-1$. Given $w = a_1 \cdots a_n \in \mathfrak{S}_n$, define

$$\text{des}_p(w) = \#\{(i, j) : i < j \leq i + p, a_i > a_j\}.$$

Thus $\text{des}_1 = \text{des}$ and $\text{des}_{n-1} = \text{inv}$. Show that the polynomial $\sum_{w \in \mathfrak{S}_n} x^{\text{des}_p(w)}$ is symmetric and unimodal.

- g. [2+] Let S be a subset of $\{(i, j) : 1 \leq i < j \leq n\}$. An S -inversion of $w = a_1 \cdots a_n \in \mathfrak{S}_n$ is a pair $(i, j) \in S$ for which $a_i > a_j$. Let $\text{inv}_S(w)$ denote the number of S -inversions of w . Find a set S (for a suitable value of n) for which the polynomial $P_S(x) := \sum_{w \in \mathfrak{S}_n} x^{\text{inv}_S(w)}$ is not unimodal.

51. [3-] Let $k, n \in \mathbb{P}$ with $k \leq n$. Let $V(n, k)$ denote the volume of the region \mathcal{R}_{nk} in \mathbb{R}^n defined by

$$0 \leq x_i \leq 1, \text{ for } 1 \leq i \leq n$$

$$k-1 \leq x_1 + x_2 + \cdots + x_n \leq k.$$

Show that $V(n, k) = A(n, k)/n!$, where $A(n, k)$ is an Eulerian number.

52. [3-] Fix $b \geq 2$. Choose n random N -digit integers in base b (allowing initial digits equal to 0). Add these integers using the usual addition algorithm. For $0 \leq j \leq n-1$, let $f(j)$ be the number of times that we carry j in the addition process. For instance, if we add 71801, 80914, and 62688 in base 10, then $f(0) = 1$ and $f(1) = f(2) = 2$. Show that as $N \rightarrow \infty$, the expected value of $f(j)/N$ (i.e., the expected proportion of the time we carry a j) approaches $A(n, j+1)/n!$, where $A(n, k)$ is an Eulerian number.
53. a. [2]* The *Eulerian Catalan number* is defined by $EC_n = A(2n+1, n+1)/(n+1)$. The first few Eulerian Catalan numbers, beginning with $EC_0 = 1$, are 1, 2, 22, 604, 31238. Show that $EC_n = 2A(2n, n+1)$, whence $EC_n \in \mathbb{Z}$.
b. [3-]* Show that EC_n is the number of permutations $w = a_1 a_2 \cdots a_{2n+1}$ with n descents, such that every left factor $a_1 a_2 \cdots a_i$ has at least as many ascents as descents. For $n = 1$ we are counting the two permutations 132 and 231.
54. [2]* How many n -element multisets on $[2m]$ are there satisfying: (i) $1, 2, \dots, m$ appear at most once each, and (ii) $m+1, m+2, \dots, 2m$ appear an even number of times each?
55. [2-]* If $w = a_1 a_2 \cdots a_n \in \mathfrak{S}_n$ then let $w^r = a_n \cdots a_2 a_1$, the reverse of w . Express $\text{inv}(w^r)$, $\text{maj}(w^r)$, and $\text{des}(w^r)$ in terms of $\text{inv}(w)$, $\text{maj}(w)$, and $\text{des}(w)$, respectively.
56. [2+] Let M be a finite multiset on \mathbb{P} . Generalize equation (1.41) by showing that

$$\sum_{w \in \mathfrak{S}_M} q^{\text{inv}(w)} = \sum_{w \in \mathfrak{S}_M} q^{\text{maj}(w)},$$

where $\text{inv}(w)$ and $\text{maj}(w)$ are defined in Section 1.7. Try to give a proof based on results in Section 1.4 rather than generalizing the proof of (1.41).

57. [2+] Let $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$. Show that the following conditions are equivalent.
- Let $C(i)$ be the set of indices j of the columns C_j that intersect the i th row of the diagram $D(w)$ of w . For instance, if $w = 314652$ as in Figure 1.5, then $C(1) = \{1, 2\}$, $C(3) = \{2\}$, $C(4) = \{2, 5\}$, $C(5) = \{2\}$, and all other $C(i) = \emptyset$. Then for every i, j , either $C(i) \subseteq C(j)$ or $C(j) \subseteq C(i)$.
 - Let $\lambda(w)$ be the entries of the inversion table $I(w)$ of w written in decreasing order. For instance, $I(52413) = (3, 1, 2, 1, 0)$ and $\lambda(52413) = (3, 2, 1, 1, 0)$. Regard λ as a partition of $\text{inv}(w)$. Then $\lambda(w^{-1}) = \lambda(w)'$, the conjugate partition to $\lambda(w)$.
 - The permutation w is 2143-avoiding (i.e., there do not exist $a < b < c < d$ for which $w_b < w_a < w_d < w_c$).
58. For $u \in \mathfrak{S}_k$, let $s_u(n) = \#\mathcal{S}_u(n)$, the number of permutations $w \in \mathfrak{S}_n$ avoiding u . If also $v \in \mathfrak{S}_k$, then write $u \sim v$ if $s_u(n) = s_v(n)$ for all $n \geq 0$ (an obvious equivalence relation). Thus by the discussion preceding Proposition 1.5.1, $u \sim v$ for all $u, v \in \mathfrak{S}_3$.
- [2]* Let $u, v \in \mathfrak{S}_k$. Suppose that the permutation matrix P_v can be obtained from P_u by one of the eight dihedral symmetries of the square. For instance, $P_{u^{-1}}$ and be obtained from P_u by reflection in the main diagonal. Show that $u \sim v$. We then say that u and v are *equivalent by symmetry*, denoted $u \approx v$. Thus \approx is a finer equivalence relation than \sim . What are the \approx equivalence classes for \mathfrak{S}_3 ?
 - [3] Show that there are exactly three \sim equivalence classes for \mathfrak{S}_4 . The equivalence classes are given by $\{1234, 1243, 2143, \dots\}$, $\{3142, 1342, \dots\}$, and $\{1342, \dots\}$, where the omitted permutations are obtained by \approx equivalence.
59. [3] Let $s_u(n)$ have the meaning of the previous exercise. Show that $c_u := \lim_{n \rightarrow \infty} s_u(n)^{1/n}$ exists and satisfies $1 < c_u < \infty$.
60. [2+] Define two permutations in \mathfrak{S}_n to be *equivalent* if one can be obtained from the other by interchanging adjacent letters that differ by at least two, an obvious equivalence relation. For instance, when $n = 3$ we have the four equivalence classes $\{123\}$, $\{132, 312\}$, $\{213, 231\}$, $\{321\}$. Describe the equivalence classes in terms of more familiar objects. How many equivalence classes are there?
61. a. [3-] Let $w = w_1 \cdots w_n$. Let

$$F(x; a, b, c, d) = \sum_{n \geq 1} \sum_{w \in \mathfrak{S}_n} a^{v(w)} b^{p(w)-1} c^{r(w)} d^{f(w)} \frac{x^n}{n!},$$

where $v(w)$ denotes the number of valleys w_i of w for $1 \leq i \leq n$ (where $w_0 = w_{n+1} = 0$ as preceding Proposition 1.5.3), $p(w)$ the number of peaks, $r(w)$ the number of double rises, and $f(w)$ the number of double falls. For instance, if $w = 32451$, then 3 is a peak, 2 is a valley, 4 is a double rise, 5 is a peak, and 1 is a double fall. Thus,

$$F(x; a, b, c, d) = x + (c + d) \frac{x^2}{2!} + (c^2 + d^2 + 2ab + 2cd) \frac{x^3}{3!} + (c^3 + d^3 + 3cd^2 + 3c^2d + 8abc + 8abd) \frac{x^4}{4!} + \cdots$$

Show that

$$F(x; a, b, c, d) = \frac{e^{vx} - e^{ux}}{ve^{ux} - ue^{vx}}, \quad (1.124)$$

where $uv = ab$ and $u + v = c + d$. In other words, u and v are zeros of the polynomial $z^2 - (c + d)z + ab$; it makes no difference which zero we call u and which v .

- b. [2–] Let $r(n, k)$ be the number of permutations $w \in \mathfrak{S}_n$ with k peaks. Show that

$$\sum_{n \geq 0} \sum_{k \geq 0} r(n, k) t^k \frac{x^n}{n!} = \frac{1 + u \tan(xu)}{1 - \frac{\tan(xu)}{u}}, \quad (1.125)$$

where $u = \sqrt{t-1}$.

- c. [2+] A *proper double fall* or *proper double descent* of a permutation $w = a_1 a_2 \cdots a_n$ is an index $1 < i < n$ for which $a_{i-1} > a_i > a_{i+1}$. (Compare with the definition of a double fall or double descent, where we also allow $i = 1$ and $i = n$ with the convention $a_0 = a_{n+1} = 0$.) Let $f(n)$ be the number of permutations $w \in \mathfrak{S}_n$ with no proper double descents. Show that

$$\begin{aligned} \sum_{n \geq 0} f(n) \frac{x^n}{n!} &= \frac{1}{\sum_{j \geq 0} \left(\frac{x^{3j}}{(3j)!} - \frac{x^{3j+1}}{(3j+1)!} \right)} \\ &= 1 + x + 2 \frac{x^2}{2!} + 5 \frac{x^3}{3!} + 17 \frac{x^4}{4!} + 70 \frac{x^5}{5!} + 349 \frac{x^6}{6!} \\ &\quad + 2017 \frac{x^7}{7!} + 13358 \frac{x^8}{8!} + \cdots \end{aligned} \quad (1.126)$$

62. In this exercise we consider one method for generalizing the disjoint cycle decomposition of permutations of sets to multisets. A *multiset cycle* of \mathbb{P} is a sequence $C = (i_1, i_2, \dots, i_k)$ of positive integers with repetitions allowed, where we regard (i_1, i_2, \dots, i_k) as equivalent to $(i_j, i_{j+1}, \dots, i_k, i_1, \dots, i_{j-1})$ for $1 \leq j \leq k$. Introduce indeterminates x_1, x_2, \dots , and define the *weight* of C by $w(C) = x_{i_1} \cdots x_{i_k}$. A *multiset permutation* or *multipermutation* of a multiset M is a multiset of multiset cycles, such that M is the multiset of all elements of the cycles. For instance, the multiset $\{1, 1, 2\}$ has the following four multipermutations: $(1)(1)(2)$, $(11)(2)$, $(12)(1)$, (112) . The *weight* $w(\pi)$ of a multipermutation $\pi = C_1 C_2 \cdots C_j$ is given by $w(\pi) = w(C_1) \cdots w(C_j)$.

- a. [2–]* Show that

$$\prod_C (1 - w(C))^{-1} = \sum_{\pi} w(\pi),$$

where C ranges over all multiset cycles on \mathbb{P} and π over all (finite) multiset permutations on \mathbb{P} .

- b. [2+] Let $p_k = x_1^k + x_2^k + \cdots$. Show that

$$\prod_C (1 - w(C))^{-1} = \prod_{k \geq 1} (1 - p_k)^{-1}.$$

- c. [1+] Let $f_k(n)$ denote the number of multiset permutations on $[k]$ of total size n . For instance, $f_2(3) = 14$, given by $(1)(1)(1)$, $(1)(1)(2)$, $(1)(2)(2)$, $(2)(2)(2)$, $(11)(1)$, $(11)(2)$, $(12)(1)$, $(12)(2)$, $(22)(1)$, $(22)(2)$, (111) , (112) , (122) , (222) . Deduce from (b) that

$$\sum_{n \geq 0} f_k(n) x^n = \prod_{i \geq 1} (1 - kx^i)^{-1}.$$

- d. [3–] Find a direct combinatorial proof of (b) or (c).

63. a. [2–] We are given n square envelopes of different sizes. In how many different ways can they be arranged by inclusion? For instance, if $n = 3$ there are six ways;

namely, label the envelopes A, B, C with A the largest and C the smallest, and let $I \in J$ mean that envelope I is contained in envelope J . Then the six ways are: (1) \emptyset , (2) $B \in A$, (3) $C \in A$, (4) $C \in B$, (5) $B \in A, C \in A$, (6) $C \in B \in A$.

- b. [2] How many arrangements have exactly k envelopes that are not contained in another envelope? That don't contain another envelope?
64. a. [2] Let $f(n)$ be the number of sequences a_1, \dots, a_n of positive integers such that for each $k > 1$, k only occurs if $k-1$ occurs before the last occurrence of k . Show that $f(n) = n!$. (For $n = 3$ the sequences are 111, 112, 121, 122, 212, 123.)
- b. [2] Show that $A(n, k)$ of these sequences satisfy $\max\{a_1, \dots, a_n\} = k$.
65. [3] Let $y = \prod_{n \geq 1} (1 - x^n)^{-1}$. Show that

$$4y^3 y'' + 5xy^3 y''' + x^2 y^3 y^{(iv)} - 16y^2 y'^2 - 15xy^2 y' y'' + 20x^2 y^2 y' y''' - 19x^2 y^2 y''^2 + 10xy y'^3 + 12x^2 y y'^2 y'' + 6x^2 y'^4 = 0. \quad (1.127)$$

66. [2-]* Let $p_k(n)$ denote the number of partitions of n into k parts. Give a bijective proof that

$$p_0(n) + p_1(n) + \dots + p_k(n) = p_k(n+k).$$

67. [2-]* Express the number of partitions of n with no part equal to 1 in terms of values $p(k)$ of the partition function.
68. [2]* Let $n \geq 1$, and let $f(n)$ be the number of partitions of n such that for all k , the part k occurs at most k times. Let $g(n)$ be the number of partitions of n such that no part has the form $i(i+1)$ (i.e., no parts equal to 2, 6, 12, 20, ...). Show that $f(n) = g(n)$.
69. [2]* Let $f(n)$ denote the number of self-conjugate partitions of n all of whose parts are even. Express the generating function $\sum_{n \geq 0} f(n)x^n$ as a simple product.
70. a. [2] Find a bijection between partitions $\lambda \vdash n$ of rank r and integer arrays

$$A_\lambda = \begin{pmatrix} a_1 & a_2 & \dots & a_r \\ b_1 & b_2 & \dots & b_r \end{pmatrix}$$

such that $a_1 > a_2 > \dots > a_r \geq 0$, $b_1 > b_2 > \dots > b_r \geq 0$, and $r + \sum (a_i + b_i) = n$.

- b. [2+] A *concatenated spiral self-avoiding walk* (CSSAW) on the square lattice is a lattice path in the plane starting at $(0,0)$, with steps $(\pm 1, 0)$ and $(0, \pm 1)$ and first step $(1, 0)$, with the following three properties: (i) the path is *self-avoiding* (i.e., it never returns to a previously visited lattice point), (ii) every step after the first must continue in the direction of the previous step or turn right, and (iii) at the end of the walk it must be possible to turn right and walk infinitely many steps in the direction faced without intersecting an earlier part of the path. For instance, writing $N = (0, 1)$, etc., the five CSSAW's of length four are $NNNN$, $NNNE$, $NNEE$, $NEEE$, and $NESS$. Note for instance that $NEES$ is not a CSSAW since continuing with steps $WWW \dots$ will intersect $(0,0)$. Show that the number of CSSAW's of length n is equal to $p(n)$, the number of partitions of n .
71. [2+] How many pairs (λ, μ) of partitions of integers are there such that $\lambda \vdash n$, and the Young diagram of μ is obtained from the Young diagram of λ by adding a single square? Express your answer in terms of the partition function values $p(k)$ for $k \leq n$. Give a simple combinatorial proof.

- 72. a.** [3–] Let $\lambda = (\lambda_1, \lambda_2, \dots)$ and $\mu = (\mu_1, \mu_2, \dots)$ be partitions. Define $\mu \leq \lambda$ if $\mu_i \leq \lambda_i$ for all i . Show that

$$\sum_{\mu \leq \lambda} q^{|\mu|+|\lambda|} = \frac{1}{(1-q)(1-q^2)^2(1-q^3)^2(1-q^4)^2 \dots}. \quad (1.128)$$

- b.** [3–] Show that the number of pairs (λ, μ) such that λ and μ have *distinct* parts, $\mu \leq \lambda$ as in (a), and $|\lambda| + |\mu| = n$, is equal to $p(n)$, the number of partitions of n . For instance, when $n = 5$ we have the seven pairs $(\emptyset, 5)$, $(\emptyset, 41)$, $(\emptyset, 32)$, $(1, 4)$, $(2, 3)$, $(1, 31)$, and $(2, 21)$.

- 73.** [2] Let λ be a partition. Show that

$$\sum_i \left\lceil \frac{\lambda_{2i-1}}{2} \right\rceil = \sum_i \left\lceil \frac{\lambda'_{2i-1}}{2} \right\rceil,$$

$$\sum_i \left\lfloor \frac{\lambda_{2i-1}}{2} \right\rfloor = \sum_i \left\lfloor \frac{\lambda'_{2i}}{2} \right\rfloor,$$

$$\sum_i \left\lfloor \frac{\lambda_{2i}}{2} \right\rfloor = \sum_i \left\lfloor \frac{\lambda'_{2i}}{2} \right\rfloor.$$

- 74.** [2] Let $p_k(n)$ denote the number of partitions of n into k parts. Fix $t \geq 0$. Show that as $n \rightarrow \infty$, $p_{n-t}(n)$ becomes eventually constant. What is this constant $f(t)$? What is the least value of n for which $p_{n-t}(n) = f(t)$? Your arguments should be combinatorial.
- 75.** [2–] Let $p_k(n)$ be as in Exercise 1.74, and let $q_k(n)$ be the number of partitions of n into k distinct parts. For example, $q_3(8) = 2$, corresponding to $(5, 2, 1)$ and $(4, 3, 1)$. Give a simple combinatorial proof that $q_k\left(n + \binom{k}{2}\right) = p_k(n)$.

- 76.** [2] Prove the partition identity

$$\prod_{i \geq 1} (1 + qx^{2i-1}) = \sum_{k \geq 0} \frac{x^{k^2} q^k}{(1-x^2)(1-x^4) \dots (1-x^{2k})}. \quad (1.129)$$

- 77.** [3–] Give a “subtraction-free” bijective proof of the pentagonal number formula by proving directly the identity

$$1 + \frac{\sum_{n \text{ odd}} (x^{n(3n-1)/2} + x^{n(3n+1)/2})}{\prod_{j \geq 1} (1 - x^j)} = \frac{1 + \sum_{n \text{ even}} (x^{n(3n-1)/2} + x^{n(3n+1)/2})}{\prod_{j \geq 1} (1 - x^j)}.$$

- 78. a.** [2] The *logarithmic derivative* of a power series $F(x)$ is $\frac{d}{dx} \log F(x) = F'(x)/F(x)$. By logarithmically differentiating the power series $\sum_{n \geq 0} p(n)x^n = \prod_{i \geq 1} (1 - x^i)^{-1}$, derive the recurrence

$$n \cdot p(n) = \sum_{i=1}^n \sigma(i) p(n-i),$$

where $\sigma(i)$ is the sum of the divisors of i .

- b.** [2+] Give a combinatorial proof.

79. a. [2+] Given a set $S \subseteq \mathbb{P}$, let $p_S(n)$ (resp. $q_S(n)$) denote the number of partitions of n (resp. number of partitions of n into distinct parts) whose parts belong to S . (These are special cases of the function $p(S, n)$ of Corollary 1.8.2.) Call a pair (S, T) , where $S, T \subseteq \mathbb{P}$, an *Euler pair* if $p_S(n) = q_T(n)$ for all $n \in \mathbb{N}$. Show that (S, T) is an Euler pair if and only if $2T \subseteq T$ (where $2T = \{2i : i \in T\}$) and $S = T - 2T$.
- b. [1+] What is the significance of the case $S = \{1\}$, $T = \{1, 2, 4, 8, \dots\}$?
80. [2+] If λ is a partition of an integer n , let $f_k(\lambda)$ be the number of times k appears as a part of λ , and let $g_k(\lambda)$ be the number of distinct parts of λ that occur at least k times. For example, $f_2(4, 2, 2, 2, 1, 1) = 3$ and $g_2(4, 2, 2, 2, 1, 1) = 2$. Show that $\sum f_k(\lambda) = \sum g_k(\lambda)$, where $k \in \mathbb{P}$ is fixed and both sums range over all partitions λ of a fixed integer $n \in \mathbb{P}$.
81. [2+] A *perfect partition* of $n \geq 1$ is a partition $\lambda \vdash n$ which “contains” precisely one partition of each positive integer $m \leq n$. In other words, regarding λ as the multiset of its parts, for each $m \leq n$ there is a unique submultiset of λ whose parts sum to m . Show that the number of perfect partitions of n is equal to the number of *ordered* factorizations (with any number of factors) of $n + 1$ into integers ≥ 2 .
Example. The perfect partitions of 5 are $(1, 1, 1, 1, 1)$, $(3, 1, 1)$, and $(2, 2, 1)$. The ordered factorizations of 6 are $6 = 2 \cdot 3 = 3 \cdot 2$.
82. [3] Show that the number of partitions of $5n + 4$ is divisible by 5.
83. [3–] Let $\lambda = (\lambda_1, \lambda_2, \dots) \vdash n$. Define

$$\alpha(\lambda) = \sum_i \lceil \lambda_{2i-1} / 2 \rceil,$$

$$\beta(\lambda) = \sum_i \lfloor \lambda_{2i-1} / 2 \rfloor,$$

$$\gamma(\lambda) = \sum_i \lceil \lambda_{2i} / 2 \rceil,$$

$$\delta(\lambda) = \sum_i \lfloor \lambda_{2i} / 2 \rfloor.$$

Let a, b, c, d be (commuting) indeterminates, and define

$$w(\lambda) = a^{\alpha(\lambda)} b^{\beta(\lambda)} c^{\gamma(\lambda)} d^{\delta(\lambda)}.$$

For instance, if $\lambda = (5, 4, 4, 3, 2)$, then $w(\lambda)$ is the product of the entries of the diagram

$$\begin{array}{ccccc} a & b & a & b & a \\ c & d & c & d & \\ a & b & a & b & \\ c & d & c & & \\ a & b & & & \end{array}$$

Show that

$$\sum_{\lambda \in \text{Par}} w(\lambda) = \prod_{j \geq 1} \frac{(1 + a^j b^{j-1} c^{j-1} d^{j-1})(1 + a^j b^j c^j d^{j-1})}{(1 - a^j b^j c^j d^j)(1 - a^j b^j c^{j-1} d^{j-1})(1 - a^j b^{j-1} c^j d^{j-1})}, \quad (1.130)$$

where Par denotes the set of all partitions λ of all integers $n \geq 0$.

84. [2]* Show that the number of partitions of n in which each part appears exactly 2, 3, or 5 times is equal to the number of partitions of n into parts congruent to $\pm 2, \pm 3, 6 \pmod{12}$.

85. [2+]* Prove that the number of partitions of n in which no part appears exactly once equals the number of partitions of n into parts not congruent to $\pm 1 \pmod{6}$.
86. [3] Prove that the number of partitions of n into parts congruent to 1 or 5 (mod 6) equals the number of partitions of n in which the difference between all parts is at least 3 and between multiples of 3 is at least 6.
87. [3-]* Let $A_k(n)$ be the number of partitions of n into odd parts (repetition allowed) such that exactly k distinct parts occur. For instance, when $n = 35$ and $k = 3$, one of the partitions being enumerated is $(9, 9, 5, 3, 3, 3, 3)$. Let $B_k(n)$ be the number of partitions $\lambda = (\lambda_1, \dots, \lambda_r)$ of n such that the sequence $\lambda_1, \dots, \lambda_r$ is composed of exactly k noncontiguous sequences of one or more consecutive integers. For instance, when $n = 44$ and $k = 3$, one of the partitions being enumerated is $(10, 9, 8, 7, 5, 3, 2)$, which is composed of 10, 9, 8, 7 and 5 and 3, 2. Show that $A_k(n) = B_k(n)$ for all k and n . Note that summing over all k gives Proposition 1.8.5 (i.e., $p_{\text{odd}}(n) = q(n)$).
88. a. [3] Prove the identities

$$\sum_{n \geq 0} \frac{x^{n^2}}{(1-x)(1-x^2) \cdots (1-x^n)} = \frac{1}{\prod_{k \geq 0} (1-x^{5k+1})(1-x^{5k+4})},$$

$$\sum_{n \geq 0} \frac{x^{n(n+1)}}{(1-x)(1-x^2) \cdots (1-x^n)} = \frac{1}{\prod_{k \geq 0} (1-x^{5k+2})(1-x^{5k+3})}.$$

- b. [2] Show that the identities in (a) are equivalent to the following combinatorial statements:
- The number of partitions of n into parts $\equiv \pm 1 \pmod{5}$ is equal to the number of partitions of n whose parts differ by at least 2.
 - The number of partitions of n into parts $\equiv \pm 2 \pmod{5}$ is equal to the number of partitions of n whose parts differ by at least 2 and for which 1 is not a part.
- c. [2]* Let $f(n)$ be the number of partitions $\lambda \vdash n$ satisfying $\ell(\lambda) = \text{rank}(\lambda)$. Show that $f(n)$ is equal to the number of partitions of n whose parts differ by at least 2.
89. [3] A *lecture hall partition* of length k is a partition $\lambda = (\lambda_1, \dots, \lambda_k)$ (some of whose parts may be 0) satisfying

$$0 \leq \frac{\lambda_k}{1} \leq \frac{\lambda_{k-1}}{2} \leq \cdots \leq \frac{\lambda_1}{k}.$$

Show that the number of lecture hall partitions of n of length k is equal to the number of partitions of n whose parts come from the set $1, 3, 5, \dots, 2k-1$ (with repetitions allowed).

90. [3] Let $f(n)$ be the number of partitions of n all of whose parts are Lucas numbers L_{2n+1} of odd index. For instance, $f(12) = 5$, corresponding to

$$\begin{aligned} &1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 \\ &4 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 \\ &4 + 4 + 1 + 1 + 1 + 1 \\ &4 + 4 + 4 \\ &11 + 1 \end{aligned}$$

Let $g(n)$ be the number of partitions $\lambda = (\lambda_1, \lambda_2, \dots)$ such that $\lambda_i / \lambda_{i+1} > \frac{1}{2}(3 + \sqrt{5})$ whenever $\lambda_{i+1} > 0$. For instance, $g(12) = 5$, corresponding to

$$12, \quad 11 + 1, \quad 10 + 2, \quad 9 + 3, \quad 8 + 3 + 1.$$

Show that $f(n) = g(n)$ for all $n \geq 1$.

91. a. [3–] Show that

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

b. [2] Deduce from (a) the Pentagonal Number Formula (Proposition 1.8.7).

c. [2] Deduce from (a) the two identities

$$\prod_{k \geq 1} \frac{1 - q^k}{1 + q^k} = \sum_{n \in \mathbb{Z}} (-1)^n q^{n^2}, \quad (1.131)$$

$$\prod_{k \geq 1} \frac{1 - q^{2k}}{1 - q^{2k-1}} = \sum_{n \geq 0} q^{\binom{n+1}{2}}. \quad (1.132)$$

d. [2+] Deduce from (a) the identity

$$\prod_{k \geq 1} (1 - q^k)^3 = \sum_{n \geq 0} (-1)^n (2n + 1) q^{n(n+1)/2}.$$

Hint. First substitute $-xq^{-1/2}$ for x and $q^{1/2}$ for q .

92. [3] Let $\mathcal{S} \subseteq \mathbb{P}$ and let $p(\mathcal{S}, n)$ denote the number of partitions of n whose parts belong to \mathcal{S} . Let

$$\mathcal{S} = \{n : n \text{ odd or } n \equiv \pm 4, \pm 6, \pm 8, \pm 10 \pmod{32}\},$$

$$\mathcal{T} = \{n : n \text{ odd or } n \equiv \pm 2, \pm 8, \pm 12, \pm 14 \pmod{32}\}.$$

Show that $p(\mathcal{S}, n) = p(\mathcal{T}, n - 1)$ for all $n \geq 1$. Equivalently, we have the remarkable identity

$$\prod_{n \in \mathcal{S}} \frac{1}{1 - x^n} = 1 + x \prod_{n \in \mathcal{T}} \frac{1}{1 - x^n}. \quad (1.133)$$

93. [3] Let

$$\mathcal{S} = \pm\{1, 4, 5, 6, 7, 9, 11, 13, 16, 21, 23, 28 \pmod{66}\},$$

$$\mathcal{T} = \pm\{1, 4, 5, 6, 7, 9, 11, 14, 16, 17, 27, 29 \pmod{66}\},$$

where

$$\pm\{a, b, \dots \pmod{m}\} := \{n \in \mathbb{P} : n \equiv \pm a, \pm b, \dots \pmod{m}\}.$$

Show that $p(\mathcal{S}, n) = p(\mathcal{T}, n)$ for all $n \geq 1$ except $n = 13$. Equivalently, we have another remarkable identity similar to equation (1.133):

$$\prod_{n \in \mathcal{S}} \frac{1}{1 - x^n} = x^{13} + \prod_{n \in \mathcal{T}} \frac{1}{1 - x^n}.$$

94. a. [3–] Let $n \geq 0$. Show that the following numbers are equal.
- The number of solutions to $n = \sum_{i \geq 0} a_i 2^i$, where $a_i = 0, 1$, or 2 .
 - Then number of odd integers k for which the Stirling number $S(n+1, k)$ is odd.
 - The number of odd binomial coefficients of the form $\binom{n-k}{k}$, $0 \leq k \leq n$.
 - The number of ways to write b_n as a sum of distinct Fibonacci numbers F_n , where

$$\prod_{i \geq 0} (1 + x^{F_{2i}}) = \sum_{n \geq 0} x^{b_n}, \quad b_0 < b_1 < \dots$$

- b. [2–] Denote by a_{n+1} the number being counted by (a), so $(a_1, a_2, \dots, a_{10}) = (1, 1, 2, 1, 3, 2, 3, 1, 4, 3)$. Deduce from (a) that

$$\sum_{n \geq 0} a_{n+1} x^n = \prod_{i \geq 0} (1 + x^{2^i} + x^{2^{i+1}}).$$

- c. [2] Deduce from (a) that $a_{2n} = a_n$ and $a_{2n+1} = a_n + a_{n+1}$.
- d. [3–] Show that every positive rational number can be written in exactly one way as a fraction a_n/a_{n+1} .
95. [3] At time $n = 1$ place a line segment (toothpick) of length one on the xy -plane, centered at $(0,0)$ and parallel to the y -axis. At time $n > 1$, place additional line segments that are centered at the end and perpendicular to an exposed toothpick end, where an *exposed end* is the end of a toothpick that is neither the end nor the midpoint of another toothpick. Figure 1.28 shows the configurations obtained for times $n \leq 6$. Let $f(n)$ be the total number of toothpicks that have been placed up to time n , and let

$$F(x) = \sum_{n \geq 1} f(n) x^n.$$

Figure 1.28 shows that

$$F(x) = x + 3x^2 + 7x^3 + 11x^4 + 15x^5 + 23x^6 + \dots$$

Show that

$$F(x) = \frac{x}{(1-x)(1-2x)} \left(1 + 2x \prod_{k \geq 0} (1 + x^{2^k-1} + 2x^{2^k}) \right).$$

96. Define

$$\begin{aligned} x \prod_{n \geq 1} (1 - x^n)^{24} &= \sum_{n \geq 1} \tau(n) x^n \\ &= x - 24x^2 + 252x^3 - 1472x^4 + 4830x^5 - 6048x^6 - 16744x^7 + \dots \end{aligned}$$

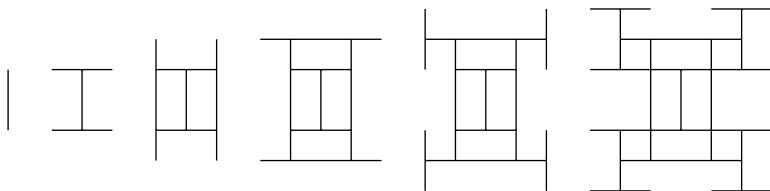


Figure 1.28 The growth of toothpicks.

- a. [3+] Show that $\tau(mn) = \tau(m)\tau(n)$ if m and n are relatively prime.
 b. [3+] Show that if p is prime and $n \geq 1$ then

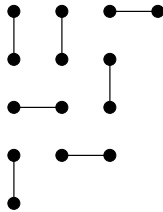
$$\tau(p^{n+1}) = \tau(p)\tau(p^n) - p^{11}\tau(p^{n-1}).$$

- c. [4] Show that if p is prime, then $|\tau(p)| < 2p^{11/2}$. Equivalently, write

$$\sum_{n \geq 0} \tau(p^n)x^n = \frac{P_p(x)}{1 - \tau(p)x + p^{11}x^2},$$

so by (b) and Theorem 4.4.1.1 the numerator $P_p(x)$ is a polynomial. Then the zeros of the denominator are not real.

- d. [5] Show that $\tau(n) \neq 0$ for all $n \geq 1$.
 97. [3–] Let $f(n)$ be the number of partitions of $2n$ whose Ferrers diagram can be covered by n edges, each connecting two adjacent dots. For instance, $(4, 3, 3, 3, 1)$ can be covered as follows:



Show that $\sum_{n \geq 0} f(n)x^n = \prod_{i \geq 1} (1 - x^i)^{-2}$.

98. [2+] Let $n, a, k \in \mathbb{N}$ and $\zeta = e^{2\pi i/n}$. Show that

$$\binom{na}{k}_{q=\zeta} = \begin{cases} \binom{a}{b}, & k = nb \\ 0, & \text{otherwise.} \end{cases}$$

99. [2] Let $0 \leq k \leq n$ and $f(q) = \binom{n}{k}$. Compute $f'(1)$. Try to avoid a lot of computation.
 100. [2+] State and prove a q -analogue of the Chu–Vandermonde identity

$$\sum_{i=0}^n \binom{a}{i} \binom{b}{n-i} = \binom{a+b}{n}$$

(Example 1.1.17).

101. [2]* Explain why we cannot set $q = 1$ on both sides of equation (1.85) to obtain the identity

$$1 = \sum_{k \geq 0} \frac{x^k}{k!}.$$

102. a. [2]* Let x and y be variables satisfying the commutation relation $yx = qxy$, where q commutes with x and y . Show that

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

- b. [2]* Generalize to $(x_1 + x_2 + \cdots + x_m)^n$, where $x_i x_j = q x_j x_i$ for $i > j$.
 c. [2+]* Generalize further to $(x_1 + x_2 + \cdots + x_m)^n$, where $x_i x_j = q_j x_j x_i$ for $i > j$, and where the q_j 's are variables commuting with all the x_i 's and with each other.

103. a. [3+] Given a partition λ (identified with its Young diagram) and $u \in \lambda$, let $a(u)$ (called the *arm length* of u) denote the number of squares directly to the right of u , counting u itself exactly once. Similarly, let $l(u)$ (called the *leg length* of u) denote the number of squares directly below u , counting u itself once. Thus, if $u = (i, j)$, then $a(u) = \lambda_i - j + 1$ and $l(u) = \lambda'_j - i + 1$. Define

$$\gamma(\lambda) = \#\{u \in \lambda : a(u) - l(u) = 0 \text{ or } 1\}.$$

Show that

$$\sum_{\lambda \vdash n} q^{\gamma(\lambda)} = \sum_{\lambda \vdash n} q^{\ell(\lambda)}, \quad (1.134)$$

where $\ell(\lambda)$ denotes the length (number of parts) of λ .

- b. [2]* Clearly the coefficient of x^n in the right-hand side of equation (1.134) is 1. Show directly (without using (a)) that the same is true for the left-hand side.

104. [2+] Let $n \geq 1$. Find the number $f(n)$ of integer sequences (a_1, a_2, \dots, a_n) such that $0 \leq a_i \leq 9$ and $a_1 + a_2 + \cdots + a_n \equiv 0 \pmod{4}$. Give a simple explicit formula (no sums) that depends on the congruence class of n modulo 4.

105. a. [3-] Let $n \in \mathbb{P}$, and let $f(n)$ denote the number of subsets of $\mathbb{Z}/n\mathbb{Z}$ (the integers modulo n) whose elements sum to 0 in $\mathbb{Z}/n\mathbb{Z}$. For instance, $f(4) = 4$, corresponding to $\emptyset, \{0\}, \{1, 3\}, \{0, 1, 3\}$. Show that

$$f(n) = \frac{1}{n} \sum_{\substack{d|n \\ d \text{ odd}}} \phi(d) 2^{n/d},$$

where ϕ denotes Euler's totient function.

- b. [5-] When n is odd, it can be shown using (a) (see Exercise 7.112) that $f(n)$ is equal to the number of necklaces (up to cyclic rotation) with n beads, each bead colored black or white. Give a combinatorial proof. (This is easy if n is prime.)
 c. [5-] Generalize. For instance, investigate the number of subsets S of $\mathbb{Z}/n\mathbb{Z}$ satisfying $\sum_{i \in S} p(i) \equiv \alpha \pmod{n}$, where p is a fixed polynomial and $\alpha \in \mathbb{Z}/n\mathbb{Z}$ is fixed.

106. [2] Let $f(n, k)$ be the number of sequences $a_1 a_2 \cdots a_n$ of positive integers such that the largest number occurring is k and such that the first occurrence of i appears before the first occurrence of $i + 1$ ($1 \leq i \leq k - 1$). Express $f(n, k)$ in terms of familiar numbers. Give a combinatorial proof. (It is assumed that every number $1, 2, \dots, k$ occurs at least once.)

107. [1+]* Give a direct combinatorial proof of equation (1.94e), namely,

$$B(n+1) = \sum_{i=0}^n \binom{n}{i} B(i), \quad n \geq 0.$$

108. a. [2+] Give a combinatorial proof that the number of partitions of $[n]$ such that no two consecutive integers appear in the same block is the Bell number $B(n-1)$.
 b. [2+]* Give a combinatorial proof that the number of partitions of $[n]$ such that no two *cyclically consecutive* integers (i.e., two integers i, j for which $j \equiv i + 1 \pmod{n}$) is equal to the number of partitions of $[n]$ with no singleton blocks.

109. [2+]

- Show that the number of permutations $a_1 \cdots a_n \in \mathfrak{S}_n$ for which there is no $1 \leq i < j \leq n-1$ satisfying $a_i < a_j < a_{j+1}$ is equal to the Bell number $B(n)$.
- Show that the same conclusion holds if the condition $a_i < a_j < a_{j+1}$ is replaced with $a_i < a_{j+1} < a_j$.
- Show that the number of permutations $w \in \mathfrak{S}_n$ satisfying the conditions of *both* (a) and (b) is equal to the number of involutions in \mathfrak{S}_n .

110. [3–] Let $f(n)$ be the number of partitions π of $[n]$ such that the union of no proper subset of the blocks of π is an interval $[a, b]$. For instance, $f(4) = 2$, corresponding to the partitions 13-24 and 1234, while $f(5) = 6$. Set $f(0) = 1$. Let

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

Find the coefficients of $(x/F(x))^{(-1)}$.

111. [3–] Let $f(n)$ be the number of partitions π of $[n]$ such that no block of π is an interval $[a, b]$ (allowing $a = b$). Thus, $f(1) = f(2) = f(3) = 0$ and $f(4) = 1$, corresponding to the partition 13-24. Let

$$F(x) = \sum_{n \geq 0} f(n)x^n = 1 + x^4 + 5x^5 + 21x^6 + \cdots.$$

Express $F(x)$ in terms of the *ordinary* generating function $G(x) = \sum_{n \geq 0} B(n)x^n = 1 + x + 2x^2 + 5x^3 + 15x^4 + \cdots$.

112. [2]* How many permutations $w \in \mathfrak{S}_n$ have the same number of cycles as weak excedances?

113. [2–]* Fix $k, n \in \mathbb{P}$. How many sequences (T_1, \dots, T_k) of subsets T_i of $[n]$ are there such that the *nonempty* T_i form a partition of $[n]$?

114. a. [2–]* How many permutations $w = a_1 a_2 \cdots a_n \in \mathfrak{S}_n$ have the property that for all $1 \leq i < n$, the numbers appearing in w between i and $i+1$ (whether i is to the left or right of $i+1$) are all less than i ? An example of such a permutation is 976412358.

b. [2–]* How many permutations $a_1 a_2 \cdots a_n \in \mathfrak{S}_n$ satisfy the following property: If $2 \leq j \leq n$, then $|a_i - a_j| = 1$ for some $1 \leq i < j$? Equivalently, for all $1 \leq i \leq n$, the set $\{a_1, a_2, \dots, a_i\}$ consists of consecutive integers (in some order). For example, for $n = 3$, there are the four permutations 123, 213, 231, 321. More generally, find the number of such permutations with descent set $S \subseteq [n-1]$.

115. [3–] Let $n = 2^{17} + 2$ and define $Q_n(t) = \sum_{S \subseteq [n-1]} t^{\beta_n(S)}$. Show that $e^{2\pi i/n}$ is (at least) a double root of $Q_n(t)$.

116. a. [2]* Show that the expected number of cycles of a random permutation $w \in \mathfrak{S}_n$ (chosen from the uniform distribution) is given by the harmonic number $H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \sim \log n$.

b. [3] Let $f(n)$ be the expected length of the longest cycle of a random permutation $w \in \mathfrak{S}_n$ (again from the uniform distribution). Show that

$$\lim_{n \rightarrow \infty} \frac{f(n)}{n} = \int_0^\infty \exp\left(-x - \int_x^\infty \frac{e^{-y}}{y} dy\right) dx = 0.62432965 \cdots.$$

117. [2+] Let w be a random permutation of $1, 2, \dots, n$ (chosen from the uniform distribution). Fix a positive integer $1 \leq k \leq n$. What is the probability p_{nk} that in the disjoint cycle decomposition of w , the length of the cycle containing 1 is k ? In other words,

what is the probability that k is the least positive integer for which $w^k(1) = 1$? Give a simple proof avoiding generating functions, induction, and so on.

- 118. a.** [2]* Let w be a random permutation of $1, 2, \dots, n$ (chosen from the uniform distribution), $n \geq 2$. Show that the probability that 1 and 2 are in the same cycle of w is $1/2$.
- b.** [2+] Generalize (a) as follows. Let $2 \leq k \leq n$, and let $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_\ell) \vdash k$, where $\lambda_\ell > 0$. Choose a random permutation $w \in \mathfrak{S}_n$. Let P_λ be the probability that $1, 2, \dots, \lambda_1$ are in the same cycle C_1 of w , and $\lambda_1 + 1, \dots, \lambda_1 + \lambda_2$ are in the same cycle C_2 of w different from C_1 , and so on. Show that

$$P_\lambda = \frac{(\lambda_1 - 1)! \cdots (\lambda_\ell - 1)!}{k!}.$$

- c.** [3−] Same as (b), except now we take w uniformly from the alternating group \mathfrak{A}_n . Let the resulting probability be Q_λ . Show that

$$Q_\lambda = \frac{(\lambda_1 - 1)! \cdots (\lambda_\ell - 1)!}{(k - 2)!} \left(\frac{1}{k(k - 1)} + (-1)^{n - \ell} \frac{1}{n(n - 1)} \right).$$

- 119.** [2+] Let P_n denote the probability that a random permutation (chosen from the uniform distribution) in \mathfrak{S}_{2n} has all cycle lengths at most n . Show that $\lim_{n \rightarrow \infty} P_n = 1 - \log 2 = 0.306852819 \dots$.

- 120.** [2+] Let $E_k(n)$ denote the expected number of k -cycles of a permutation $w \in \mathfrak{S}_n$, as discussed in Example 1.3.5. Give a simple combinatorial explanation of the formula $E_k(n) = 1/k$, $n \geq k$.

- 121. a.** [2]* Let $f(n)$ denote the number of fixed-point free involutions $w \in \mathfrak{S}_{2n}$ (i.e., $w^2 = 1$, and $w(i) \neq i$ for all $i \in [2n]$). Find a simple expression for $\sum_{n \geq 0} f(n)x^n/n!$. (Set $f(0) = 1$.)

- b.** [2−]* If $X \subseteq \mathbb{P}$, then write $-X = \{-i : i \in X\}$. Let $g(n)$ be the number of ways to choose a subset X of $[n]$, and then choose fixed point free involutions w on $X \cup (-X)$ and \bar{w} on $\bar{X} \cup (-\bar{X})$, where $\bar{X} = \{i \in [n] : i \notin X\}$. Use (a) to find a simple expression for $g(n)$.

- c.** [2+]* Find a combinatorial proof for the formula obtained for $g(n)$ in (b).

- 122.** [2−]* Find $\sum_w x^{\text{exc}(w)}$, where w ranges over all fixed-point free involutions in \mathfrak{S}_{2n} and $\text{exc}(w)$ denotes the number of excedances of w .

- 123.** [2]* Let \mathfrak{A}_n denote the alternating group on $[n]$ (i.e., the group of all permutations with an even number of cycles of even length). Define the *augmented cycle indicator* $\tilde{Z}_{\mathfrak{A}_n}$ of \mathfrak{A}_n by

$$\tilde{Z}_{\mathfrak{A}_n} = \sum_{w \in \mathfrak{A}_n} t^{\text{type}(w)},$$

as in equation (1.25). Show that

$$\sum_{n \geq 0} \tilde{Z}_{\mathfrak{A}_n} \frac{x^n}{n!} = \exp \left(t_1 x + t_3 \frac{x^3}{3} + t_5 \frac{x^5}{5} + \cdots \right) \cdot \cosh \left(t_2 \frac{x^2}{2} + t_4 \frac{x^4}{4} + t_6 \frac{x^6}{6} + \cdots \right).$$

- 124. a.** [2] Let $f_k(n)$ denote the number of permutations $w \in \mathfrak{S}_n$ with k inversions. Show combinatorially that for $n \geq k$,

$$f_k(n + 1) = f_k(n) + f_{k-1}(n + 1).$$

- b. [1+] Deduce from (a) that for $n \geq k$, $f_k(n)$ is a polynomial in n of degree k and leading coefficient $1/k!$. For instance, $f_2(n) = \frac{1}{2}(n+1)(n-2)$ for $n \geq 2$.
- c. [2+] Let $g_k(n)$ be the polynomial that agrees with $f_k(n)$ for $n \geq k$. Find $\Delta^j g_k(-n)$; that is, find the coefficients a_j in the expansion

$$g_k(-n) = \sum_{j=0}^k a_j \binom{n}{j}.$$

125. [2+]* Find the number $f(n)$ of binary sequences $w = a_1 a_2 \cdots a_k$ (where k is arbitrary) such that $a_1 = 1$, $a_k = 0$, and $\text{inv}(w) = n$. For instance, $f(4) = 5$, corresponding to the sequences 10000, 11110, 10110, 10010, 1100. How many of these sequences have exactly j 1's?

126. [2+]* Show that

$$\sum_w q^{\text{inv}(w)} = q^n \prod_{j=0}^{n-1} (1 + q^2 + q^4 + \cdots + q^{4j}),$$

where w ranges over all fixed-point free involutions in \mathfrak{S}_{2n} , and where $\text{inv}(w)$ denotes the number of inversions of w . Give a simple combinatorial proof analogous to the proof of Corollary 1.3.13.

127. [2]

- a. Let $w \in \mathfrak{S}_n$, and let $R(w)$ be the set of positions of the records (or left-to-right maxima) of w . For instance, $R(3265174) = \{1, 3, 6\}$. For any finite set S of positive integers, set $x^S = \prod_{i \in S} x_i$. Show that

$$\sum_{w \in \mathfrak{S}_n} q^{\text{inv}(w)} x^{R(w)} = x_1(x_2 + q)(x_3 + q + 1) \cdots (x_n + q + q^2 + \cdots + q^{n-1}). \quad (1.135)$$

- b. Let $V(w)$ be the set of the records themselves (e.g., $V(3265174) = \{3, 6, 7\}$). Show that

$$\begin{aligned} \sum_{w \in \mathfrak{S}_n} q^{\text{inv}(w)} x^{V(w)} &= (x_1 + q + q^2 + \cdots + q^{n-1})(x_2 + q + q^2 + \cdots + q^{n-2}) \cdots \\ &\quad \times (x_{n-1} + q)x_n. \end{aligned} \quad (1.136)$$

128. a. [2] A permutation $a_1 \cdots a_n$ of $[n]$ is called *indecomposable* or *connected* if n is the least positive integer j for which $\{a_1, a_2, \dots, a_j\} = \{1, 2, \dots, j\}$. Let $f(n)$ be the number of indecomposable permutations of $[n]$, and set $F(x) = \sum_{n \geq 0} n! x^n$. Show that

$$\sum_{n \geq 1} f(n) x^n = 1 - \frac{1}{F(x)}. \quad (1.137)$$

- b. [2+] If $a_1 \cdots a_n$ is a permutation of $[n]$, then a_i is called a *strong fixed point* if (1) $j < i \Rightarrow a_j < a_i$, and (2) $j > i \Rightarrow a_j > a_i$ (so in particular $a_i = i$). Let $g(n)$ be the number of permutations of $[n]$ with no strong fixed points. Show that

$$\sum_{n \geq 0} g(n) x^n = \frac{F(x)}{1 + x F(x)}.$$

- c. [2+] A permutation $w \in \mathfrak{S}_n$ is *stabilized-interval-free* (SIF) if there does not exist $1 \leq i < j \leq n$ for which $w \cdot [i, j] = [i, j]$ (as sets). For instance, 615342 fails to be SIF since $w \cdot [3, 5] = [3, 5]$. Let $h(n)$ be the number of SIF permutations $w \in \mathfrak{S}_n$, and set

$$H(x) = \sum_{n \geq 0} h(n)x^n = 1 + x + x^2 + 2x^3 + 7x^4 + 34x^5 + 206x^6 + \dots$$

Show that

$$H(x) = \frac{x}{\left(\sum_{n \geq 0} n!x^{n+1}\right)^{(-1)}},$$

where (-1) denotes compositional inverse (§5.4 of Vol. II). Equivalently, by the Lagrange inversion formula (Theorem 5.4.2 of Vol. II), $H(x)$ is uniquely defined by the condition

$$[x^{n-1}]H(x)^n = n!, \quad n \geq 1.$$

- d. [2+] A permutation $w \in \mathfrak{S}_n$ is called *simple* if it maps no interval $[i, j]$ of size $1 < j - i + 1 < n$ into another such interval. For instance, 3157462 is not simple, since it maps $[3, 6]$ into $[4, 7]$ (as sets). Let $k(n)$ be the number of simple permutations $w \in \mathfrak{S}_n$, and set

$$K(x) = \sum_{n \geq 1} k(n)x^n = x + 2x^2 + 2x^4 + 6x^5 + 46x^7 + 338x^8 + \dots$$

Show that

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

129. a. [2]* Let $f_k(n)$ be the number of indecomposable permutations $w \in \mathfrak{S}_n$ with k inversions. Generalizing equation (1.137), show that

$$\sum_{n \geq 1} f_k(n)q^k x^n = 1 - \frac{1}{F(q, x)},$$

where $F(q, x) = \sum_{n \geq 0} (n)!x^n$. As usual, $(n)! = (1+q)(1+q+q^2)\cdots(1+q+\dots+q^{n-1})$.

- b. [2] Write $1/F(q, x) = \sum_{n \geq 0} g_n(q)x^n$, where $g_n(q) \in \mathbb{Z}[q]$. Show that $\sum_{n \geq 0} g_n(q)$ is a well-defined formal power series, even though it makes no sense to substitute directly $x = 1$ in $1/F(q, x)$.
- c. [3] Write $1/F(q, x)$ in a form where it does make sense to substitute $x = 1$.

130. [2+] Let $u(n)$ be the number of permutations $w = a_1 \cdots a_n \in \mathfrak{S}_n$ such that $a_{i+1} \neq a_i \pm 1$ for $1 \leq i \leq n-1$. Equivalently, $f(n)$ is the number of ways to place n nonattacking kings on an $n \times n$ chessboard, no two on the same file or rank. Set

$$U(x) = \sum_{n \geq 0} u(n)x^n = 1 + x + 2x^4 + 14x^5 + 90x^6 + 646x^7 + 5242x^8 + \dots$$

Show that

$$U(x) = F\left(\frac{x(1-x)}{1+x}\right), \quad (1.138)$$

where $F(x) = \sum_{n \geq 0} n!x^n$ as in Exercise 1.128.

- 131.** [2+]* An n -dimensional cube K_n has $2n$ facets (or $(n-1)$ -dimensional faces), which come in n antipodal pairs. A *shelling* of K_n is equivalent to a linear ordering F_1, F_2, \dots, F_{2n} of its facets such that for all $1 \leq i \leq n-1$, the set $\{F_1, \dots, F_{2i}\}$ does not consist of i antipodal pairs. Let $f(n)$ be the number of shellings of K_n . Show that

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

- 132.** [1+]* Let $w \in \mathfrak{S}_n$. Which of the following items doesn't belong?

- $\text{inv}(w) = 0$
- $\text{maj}(w) = 0$
- $\text{des}(w) = 0$
- $\text{maj}(w) = \text{des}(w) = \text{inv}(w)$
- $D(w) = \emptyset$
- $c(w) = n$ (where $c(w)$ denotes the number of cycles of w)
- $w^5 = w^{12} = 1$

- 133. a.** [2+] Let $A_n(x)$ be the Eulerian polynomial. Give a combinatorial proof that $\frac{1}{2}A_n(2)$ is equal to the number of *ordered* set partitions (i.e., partitions whose blocks are linearly ordered) of an n -element set.
- b.** [2+]* More generally, show that

$$\frac{A_n(x)}{x} = \sum_{k=0}^{n-1} (n-k)! S(n, n-k) (x-1)^k.$$

Note that $(n-k)! S(n, n-k)$ is the number of ordered partitions of an n -set into $n-k$ blocks.

- 134.** [3-] Show that

$$A_n(x) = \sum_w x^{1+\text{des}(w)} (1+x)^{n-1-2\text{des}(w)},$$

where w ranges over all permutations in \mathfrak{S}_n with no proper double descents (as defined in Exercise 1.61) and with no descent at the end. For instance, when $n=4$, the permutations are 1234, 1324, 1423, 2134, 2314, 2413, 3124, 3412, 4123.

- 135. a.** [2] Let $A_n(x)$ be the Eulerian polynomial. Show that

$$A_n(-1) = \begin{cases} (-1)^{(n+1)/2} E_n, & n \text{ odd,} \\ 0, & n \text{ even.} \end{cases}$$

- b.** [3-] Give a combinatorial proof of (a) when n is odd.

- 136.** [2+] What sequence $\mathbf{c} = (c_1, \dots, c_n) \in \mathbb{N}^n$ with $\sum i c_i = n$ maximizes the number of $w \in \mathfrak{S}_n$ of type \mathbf{c} ? For instance, when $n=4$ the maximizing sequence is $(1, 0, 1, 0)$.

- 137.** [3-] Let ℓ be a prime number and write $n = a_0 + a_1 \ell + a_2 \ell^2 + \dots$, with $0 \leq a_i < \ell$ for all $i \geq 0$. Let $\kappa_\ell(n)$ denote the number of sequences $\mathbf{c} = (c_1, c_2, \dots, c_n) \in \mathbb{N}^n$ with $\sum i c_i = n$, such that the number of permutations $w \in \mathfrak{S}_n$ of type \mathbf{c} is relatively prime to ℓ . Show that

$$\kappa_\ell(n) = p(a_0) \prod_{i \geq 1} (a_i + 1),$$

where $p(a_0)$ is the number of partitions of a_0 . In particular, the number of \mathbf{c} such that an odd number of $w \in \mathfrak{S}_n$ have type \mathbf{c} is 2^b , where $[n/2]$ has b 1's in its binary expansion.

138. [2+]* Find a simple formula for the number of alternating permutations $a_1 a_2 \cdots a_{2n} \in \mathfrak{S}_{2n}$ satisfying $a_2 < a_4 < a_6 < \cdots < a_{2n}$.
139. [2+] An *even tree* is a (rooted) tree such that every vertex has an even number of children. (Such a tree must have an odd number of vertices.) Note that these are *not* plane trees (i.e., we don't linearly order the subtrees of a vertex). Express the number $g(n)$ of increasing even trees with $2n + 1$ vertices in terms of Euler numbers. Use generating functions.
140. [3–] Define a *simsun permutation* to be a permutation $w \in \mathfrak{S}_n$ such that w has no proper double descents (as defined in Exercise 1.61(c)) and such that for all $0 \leq k \leq n - 1$, if we remove $n, n - 1, \dots, n - k$ from w (written as a word) then the resulting permutation also has no proper double descents. For instance, $w = 3241$ is not *simsun* since if we remove 4 from w we obtain 321, which has a proper double descent. Show that the number of *simsun* permutations in \mathfrak{S}_n is equal to the Euler number E_{n+1} .
141. a. [2+] Let $E_{n,k}$ denote the number of alternating permutations of $[n + 1]$ with first term $k + 1$. For instance, $E_{n,n} = E_n$. Show that

$$E_{0,0} = 1, \quad E_{n,0} = 0 \quad (n \geq 1), \quad E_{n+1,k+1} = E_{n+1,k} + E_{n,n-k} \quad (n \geq k \geq 0). \quad (1.139)$$

Note that if we place the $E_{n,k}$'s in the triangular array

$$\begin{array}{ccccccccccc} & & & & & & & E_{00} & & & & \\ & & & & & & E_{10} & \rightarrow & E_{11} & & & \\ & & & E_{22} & \leftarrow & E_{21} & \leftarrow & E_{20} & & & & \\ & E_{30} & \rightarrow & E_{31} & \rightarrow & E_{32} & \rightarrow & E_{33} & & & & \\ E_{44} & \leftarrow & E_{43} & \leftarrow & E_{42} & \leftarrow & E_{41} & \leftarrow & E_{40} & & & \end{array} \quad (1.140)$$

...

and read the entries in the direction of the arrows from top-to-bottom (the so-called *boustrophedon* or *ox-plowing* order), then the first number read in each row is 0, and each subsequent entry is the sum of the previous entry and the entry above in the previous row. The first seven rows of the array are as follows:

$$\begin{array}{ccccccccccccccc} & & & & & & & 1 & & & & & & & \\ & & & & & & 0 & \rightarrow & 1 & & & & & & \\ & & & & 1 & \leftarrow & 1 & \leftarrow & 0 & & & & & & \\ & & 0 & \rightarrow & 1 & \rightarrow & 2 & \rightarrow & 2 & & & & & & \\ & 5 & \leftarrow & 5 & \leftarrow & 4 & \leftarrow & 2 & \leftarrow & 0 & & & & & \\ 0 & \rightarrow & 5 & \rightarrow & 10 & \rightarrow & 14 & \rightarrow & 16 & \rightarrow & 16 & & & & \\ 61 & \leftarrow & 61 & \leftarrow & 56 & \leftarrow & 46 & \leftarrow & 32 & \leftarrow & 16 & \leftarrow & 0 & & \\ & & & & & & \dots & & & & & & & \end{array}$$

- b. [3–] Define

$$[m, n] = \begin{cases} m, & m + n \text{ odd,} \\ n, & m + n \text{ even.} \end{cases}$$

Show that

$$\sum_{m \geq 0} \sum_{n \geq 0} E_{m+n, [m, n]} \frac{x^m}{m!} \frac{y^n}{n!} = \frac{\cos x + \sin x}{\cos(x + y)}. \quad (1.141)$$

- 142.** [3–] Define polynomials $f_n(a)$ for $n \geq 0$ by $f_0(a) = 1$, $f_n(0) = 0$ if $n \geq 1$, and $f'_n(a) = f_{n-1}(1-a)$. Thus,

$$\begin{aligned} f_1(a) &= a, \\ f_2(a) &= \frac{1}{2}(-a^2 + 2a), \\ f_3(a) &= \frac{1}{3!}(-a^3 + 3a), \\ f_4(a) &= \frac{1}{4!}(a^4 - 4a^3 + 8a), \\ f_5(a) &= \frac{1}{5!}(a^5 - 10a^3 + 25a), \\ f_6(a) &= \frac{1}{6!}(-a^6 + 6a^5 - 40a^3 + 96a). \end{aligned}$$

Show that $\sum_{n \geq 0} f_n(1)x^n = \sec x + \tan x$.

- 143. a.** [2–] Let $\text{fix}(w)$ denote the number of fixed points (cycles of length 1) of the permutation $w \in \mathfrak{S}_n$. Show that

$$\sum_{w \in \mathfrak{S}_n} \text{fix}(w) = n!.$$

Try to give a combinatorial proof, a generating function proof, and an algebraic proof.

- b.** [3+] Let Alt_n (respectively, Ralt_n) denote the set of alternating (respectively, reverse alternating) permutations $w \in \mathfrak{S}_n$. Define

$$f(n) = \sum_{w \in \text{Alt}_n} \text{fix}(w)$$

$$g(n) = \sum_{w \in \text{Ralt}_n} \text{fix}(w).$$

Show that

$$\begin{aligned} f(n) &= \begin{cases} E_n - E_{n-2} + E_{n-4} - \cdots + (-1)^{(n-1)/2} E_1, & n \text{ odd}, \\ E_n - 2E_{n-2} + 2E_{n-4} - \cdots + (-1)^{(n-2)/2} 2E_2 + (-1)^{n/2}, & n \text{ even}. \end{cases} \\ g(n) &= \begin{cases} E_n - E_{n-2} + E_{n-4} - \cdots + (-1)^{(n-1)/2} E_1, & n \text{ odd}, \\ E_n - (-1)^{n/2}, & n \text{ even}. \end{cases} \end{aligned}$$

- 144. a.** [2] Let

$$F(x) = 2 \sum_{n \geq 0} q^n \frac{\prod_{j=1}^n (1 - q^{2j-1})}{\prod_{j=1}^{2n+1} (1 + q^j)},$$

where $q = \left(\frac{1-x}{1+x}\right)^{2/3}$. Show that $F(x)$ is well-defined as a formal power series. Note that $q(0) = 1 \neq 0$, so some special argument is needed.

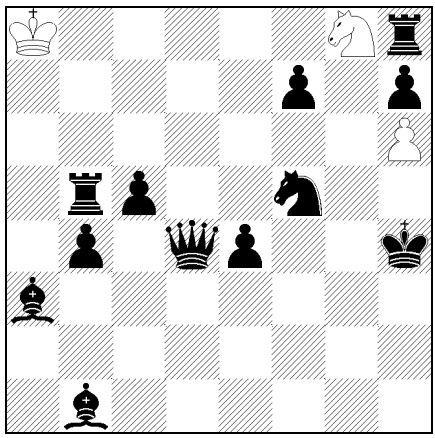
- b.** [3+] Let $F(x)$ be defined by (a), and write

$$\begin{aligned} F(x) = \sum_{n \geq 0} f(n)x^n &= 1 + x + x^2 + 2x^3 + 5x^4 + 17x^5 + 72x^6 \\ &\quad + 367x^7 + 2179x^8 + \cdots. \end{aligned}$$

Show that $f(n)$ is equal to the number of alternating fixed-point free involutions in \mathfrak{S}_{2n} (i.e., the number of permutations $w \in \mathfrak{S}_{2n}$ that are alternating permutations and have n cycles of length two). For instance, when $n = 3$ we have the two permutations 214365 and 645321, and when $n = 4$ we have the five permutations 21436587, 21867453, 64523187, 64827153, and 84627351.

145. [3–] Solve the following chess problem, where the condition “serieshelpmate” is defined in Exercise 1.11(c).

A. Karttunen, 2006



Serieshelpmate in 9: how many solutions?

146. [2+] Let $f_k(n)$ denote the number of permutations $w \in \mathfrak{S}_n$ such that

$$D(w) = \{k, 2k, 3k, \dots\} \cap [n-1],$$

as in equation (1.58). Let $1 \leq i \leq k$. Show that

$$\sum_{m \geq 0} f_k(mk+i) \frac{x^{mk+i}}{(mk+i)!} = \frac{\sum_{m \geq 0} (-1)^m \frac{x^{mk+i}}{(mk+i)!}}{\sum_{m \geq 0} (-1)^m \frac{x^{mk}}{(mk)!}}.$$

Note that when $i = k$ we can add 1 to both sides and obtain equation (1.59).

147. [2+] Call two permutations $u, v \in \mathfrak{S}_n$ *equivalent* if their min-max trees $M(u)$ and $M(v)$ are isomorphic as *unlabeled* binary trees. This notion of equivalence is clearly an equivalence relation. Show that the number of equivalence classes is the Motzkin number M_{n-1} defined in Exercise 6.37 and further explicated in Exercise 6.38.
148. [2+] Let $\Phi_n = \Phi_n(c, d)$ denote the cd -index of \mathfrak{S}_n , as defined in Theorem 1.6.3. Thus, $c = a + b$ and $d = ab + ba$. Let $S \subseteq [n-1]$, and let u_S be the variation of S as defined by equation (1.60). Show that

$$\Phi_n(a + 2b, ab + ba + 2b^2) = \sum_{S \subseteq [n-1]} \alpha(S) u_S,$$

where $\alpha(S)$ is given by equation (1.31).

149. [3–] If $F(x)$ is any power series with noncommutative coefficients such that $F(0) = 0$, then define $(1 - F(x))^{-1}$ to be the unique series $G(x)$ satisfying

$$(1 - F(x))G(x) = G(x)(1 - F(x)) = 1.$$

Equivalently, $G(x) = 1 + F(x) + F(x)^2 + \cdots$. Show that

$$\sum_{n \geq 1} \Phi_n(c, d) \frac{x^n}{n!} = \frac{\sinh(a-b)x}{a-b} \left[1 - \frac{1}{2} \left(\frac{c \cdot \sinh(a-b)x}{a-b} - \cosh(a-b)x + 1 \right) \right]^{-1}. \quad (1.142)$$

Note that the series on the right involves only *even* powers of $a-b$. Since $(a-b)^2 = c^2 - 2d$, it follows that the coefficients of this series are indeed polynomials in c and d .

- 150. a.** [3-]* Let $f(n)$ (respectively, $g(n)$) be the total number of c 's (respectively, d 's) that appear when we write the cd -index $\Phi_n(c, d)$ as a sum of monomials. For instance, $\Phi_4(c, d) = c^3 + 2cd + 2dc$, so $f(4) = 7$ and $g(4) = 4$. Show using generating functions that $f(n) = 2E_{n+1} - (n+1)E_n$ and $g(n) = nE_n - E_{n+1}$.
- b.** [5-] Is there a combinatorial proof?
- 151.** [3-] Let μ be a monomial of degree $n-1$ in the noncommuting variables c, d , where $\deg(c) = 1$ and $\deg(d) = 2$. Show that $[\mu]\Phi_n(c, d)$ is the number of sequences $\mu = v_0, v_1, \dots, v_{n-1} = 1$, where v_i is obtained from v_{i-1} by removing a c or changing a d to c . For instance, if $\mu = dcc$, there are three sequences: $(dcc, dcc, cc, c, 1)$, $(dcc, dc, cc, c, 1)$, $(dcc, dc, d, c, 1)$.
- 152.** [3-] Continue the notation from the previous exercise. Replace each c in μ with 0, each d with 10, and remove the final 0. We get the characteristic vector of a set $S_\mu \subseteq [n-2]$. For instance, if $\mu = cd^2c^2d$ then we get the characteristic vector 01010001 of the set $S_\mu = \{2, 4, 8\}$. Show that $[\mu]\Phi_n(c, d)$ is equal to the number of *simsum* permutations (defined in Exercise 1.140) in \mathfrak{S}_{n-1} with descent set S_μ .
- 153. a.** [2] Let $f(n)$ denote the coefficient of d^n in the cd -index Φ_{2n+1} . Show that $f(n) = 2^{-n} E_{2n+1}$.
- b.** [3] Show that $f(n)$ is the number of permutations w of the multiset $\{1^2, 2^2, \dots, (n+1)^2\}$ beginning with 1 such that between the two occurrences of i ($1 \leq i \leq n$) there is exactly one occurrence of $i+1$. For instance, $f(2) = 4$, corresponding to 123123, 121323, 132312, 132132.
- 154. a.** [1+] Let $F(x) = \sum_{n \geq 0} f(n)x^n/n!$. Show that

$$e^{-x} F(x) = \sum_{n \geq 0} [\Delta^n f(0)] x^n / n!.$$

- b.** [2] Find the unique function $f: \mathbb{P} \rightarrow \mathbb{C}$ satisfying $f(1) = 1$ and $\Delta^n f(1) = f(n)$ for all $n \in \mathbb{P}$.
- c.** [2] Generalize (a) by showing that

$$e^{-x} F(x+t) = \sum_{n \geq 0} \sum_{k \geq 0} \Delta^n f(k) \frac{x^n t^k}{n! k!}.$$

- 155. a.** [1+] Let $F(x) = \sum_{n \geq 0} f(n)x^n$. Show that

$$\frac{1}{1+x} F\left(\frac{x}{1+x}\right) = \sum_{n \geq 0} [\Delta^n f(0)] x^n.$$

- b.** [2+] Find the unique functions $f, g: \mathbb{N} \rightarrow \mathbb{C}$ satisfying $\Delta^n f(0) = g(n)$, $\Delta^{2n} g(0) = f(n)$, $\Delta^{2n+1} g(0) = 0$, $f(0) = 1$.
- c.** [2+] Find the unique functions $f, g: \mathbb{N} \rightarrow \mathbb{C}$ satisfying $\Delta^n f(1) = g(n)$, $\Delta^{2n} g(0) = f(n)$, $\Delta^{2n+1} g(0) = 0$, $f(0) = 1$.

- 156.** [2+] Let A be the abelian group of all polynomials $p: \mathbb{Z} \rightarrow \mathbb{C}$ such that $D^k p: \mathbb{Z} \rightarrow \mathbb{Z}$ for all $k \in \mathbb{N}$. (D^k denotes the k th derivative, and $D^0 p = p$.) Then A has a basis of the form $p_n(x) = c_n \binom{x}{n}$, $n \in \mathbb{N}$, where c_n is a constant depending only on n . Find c_n explicitly.

- 157.** [2] Let λ be a complex number (or indeterminate), and let

$$y = 1 + \sum_{n \geq 1} f(n)x^n, \quad y^\lambda = \sum_{n \geq 0} g(n)x^n.$$

Show that

$$g(n) = \frac{1}{n} \sum_{k=1}^n [k(\lambda + 1) - n] f(k) g(n-k), \quad n \geq 1.$$

This formula affords a method of computing the coefficients of y^λ much more efficiently than using (1.5) directly.

- 158.** [2+] Let f_1, f_2, \dots be a sequence of complex numbers. Show that there exist unique complex numbers a_1, a_2, \dots such that

$$F(x) := 1 + \sum_{n \geq 1} f_n x^n = \prod_{i \geq 1} (1 - x^i)^{-a_i}.$$

Set $\log F(x) = \sum_{n \geq 1} g_n x^n$. Find a formula for a_i in terms of the g_n 's. What are the a_i 's when $F(x) = 1 + x$ and $F(x) = e^{x/(1-x)}$?

- 159.** [2] Let $F(x) = 1 + a_1 x + \dots \in K[[x]]$, where K is a field satisfying $\text{char}(K) \neq 2$. Show that there exist unique series $A(x), B(x)$ satisfying $A(0) = B(0) = 1$, $A(x) = A(-x)$, $B(x)B(-x) = 1$, and $F(x) = A(x)B(x)$. Find simple formulas for $A(x)$ and $B(x)$ in terms of $F(x)$.

- 160. a.** [2] Let $0 \leq j < k$. The (k, j) -multisection of the power series $F(x) = \sum_{n \geq 0} a_n x^n$ is defined by

$$\Psi_{k,j} F(x) = \sum_{m \geq 0} a_{km+j} x^{km+j}.$$

Let $\zeta = e^{2\pi i/k}$ (where $i^2 = -1$). Show that

$$\Psi_{k,j} F(x) = \frac{1}{k} \sum_{r=0}^{k-1} \zeta^{-jr} F(\zeta^r x).$$

- b.** [2] As a simple application of (a), let $0 \leq j < k$, and let $f(n, k, j)$ be the number of permutations $w \in \mathfrak{S}_n$ satisfying $\text{maj}(w) \equiv j \pmod{k}$. Show that $f(n, k, j) = n!/k$ if $n \geq k$.

- c.** [2+] Show that

$$f(k-1, k, 0) = \frac{(k-1)!}{k} + \sum_{\xi} \frac{1}{(1-\xi)^{k-1}},$$

where ξ ranges over all primitive k th roots of unity. Can this expression be simplified?

- 161. a.** [2]* Let $F(x) = a_0 + a_1 x + \dots \in K[[x]]$, with $a_0 = 1$. For $k \geq 2$ define $F_k(x) = \Phi_{k,0}(x) = \sum_{m \geq 0} a_{km} x^{km}$. Show that for $n \geq 1$,

$$[x^{km}] \frac{F(x)}{\Phi_{k,0} F(x)} = 0.$$

- b. [2+] Let $\text{char } K \neq 2$. Given $G(x) = 1 + H(x)$ where $H(-x) = -H(x)$ (i.e., $H(x)$ has only odd exponents), find the general solution $F(x) = 1 + a_1x + \cdots$ to $F(x)/F_2(x) = G(x)$. Express your answer in the form $F(x) = \Phi(G(x))E(x)$, where $\Phi(x)$ is a function independent from $G(x)$, and where $E(x)$ ranges over some class \mathcal{E} of power series, also independent from $G(x)$.
162. [3–] Let $g(x) \in \mathbb{C}[[x]]$, $g(0) = 0$, $g(x) = g(-x)$. Find all power series $f(x)$ such that $f(0) = 0$ and

$$\frac{f(x) + f(-x)}{1 - f(x)f(-x)} = g(x).$$

Express your answer as an explicit algebraic function of $g(x)$ and a power series $h(x)$ (independent from $g(x)$) taken from some class of power series.

163. Let $f(x) \in \mathbb{C}[[x]]$, $f(x) = x + \text{higher order terms}$. We say that $F(x, y) \in \mathbb{C}[[x, y]]$ is a *formal group law* or *addition law* for $f(x)$ if $f(x + y) = F(f(x), f(y))$.
- a. [2–] Show that for every $f(x) \in \mathbb{C}[[x]]$ with $f(x) = x + \cdots$, there is a unique $F(x, y) \in \mathbb{C}[[x, y]]$ which is a formal group law for $f(x)$.
- b. [3] Show that $F(x, y)$ is a formal group law if and only if $F(x, y) = x + y + \text{higher order terms}$, and

$$F(F(x, y), z) = F(x, F(y, z)).$$

- c. [2] Find $f(x)$ so that $F(x, y)$ is a formal group law for $f(x)$ in the following cases:
- $F(x, y) = x + y$.
 - $F(x, y) = x + y + xy$.
 - $F(x, y) = (x + y)/(1 - xy)$.
 - $F(x, y) = x\sqrt{1 - y^2} + y\sqrt{1 - x^2}$.
- d. [2+] Using equation (5.128), show that the formal group law for $f(x) = xe^{-x}$ is given by

$$F(x, y) = x + y - \sum_{n \geq 1} (n-1)^{n-1} \frac{x^n y + xy^n}{n!},$$

where we interpret $0^0 = 1$ in the summand indexed by $n = 1$.

- e. [3] Find the formal group law for the function

$$f(x) = \int_0^x \frac{dt}{\sqrt{1 - t^4}}.$$

164. [3–] Solve the following equation for the power series $F(x, y) \in \mathbb{C}[[x, y]]$:

$$(xy^2 + x - y)F(x, y) = xF(x, 0) - y.$$

The point is to make sure that your solution has a power series expansion at $(0, 0)$.

165. [2+] Find a simple description of the coefficients of the power series $F(x) = x + \cdots \in \mathbb{C}[[x]]$ satisfying the functional equation

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

166. [2] Let $n \in \mathbb{P}$. Find a power series $F(x) \in \mathbb{C}[[x]]$ satisfying $F(F(x))^n = 1 + F(x)^n$, $F(0) = 1$.

- 167.** [2] Let $F(x) \in \mathbb{C}[[x]]$. Find a simple expression for the exponential generating function of the derivatives of $F(x)$, that is,

$$\sum_{n \geq 0} D^n F(x) \frac{t^n}{n!}, \quad (1.143)$$

where $D = d/dx$.

- 168.** Let K be a field satisfying $\text{char}(K) \neq 2$. If $A(x) = x + \sum_{n \geq 2} a_n x^n \in K[[x]]$, then let $A^{(-1)}(x)$ denote the compositional inverse of A ; that is, $A^{(-1)}(A(x)) = A(A^{(-1)}(x)) = x$.

- a.** [3–] Show that we can specify a_2, a_4, \dots arbitrarily, and they then determine uniquely a_3, a_5, \dots so that $A(-A(-x)) = x$. For instance,

$$\begin{aligned} a_3 &= a_2^2, \\ a_5 &= 3a_4a_2 - 2a_2^4, \\ a_7 &= 13a_2^6 - 18a_4a_2^3 + 2a_4^2 + 4a_2a_6. \end{aligned}$$

NOTE. Let $E(x) = A(-x)$. Then the conditions $A(x) = x + \dots$ and $A(-A(-x)) = x$ are equivalent to $E(x) = -x + \dots$ and $E(E(x)) = x$.

- b.** [5–] What are the coefficients when a_{2n+1} is written as a polynomial in a_2, a_4, \dots as in (a)?
- c.** [2+]* Show that $A(-A(-x)) = x$ if and only if there is a $B(x) = x + \sum_{n \geq 2} b_n x^n$ such that $A(x) = B^{(-1)}(-B(-x))$.
- d.** [2+] Show that if $A(-A(-x)) = x$, then there is a unique $B(x)$ as in (c) of the form $B(x) = x + \sum_{n \geq 1} b_{2n} x^{2n}$. For instance,

$$\begin{aligned} b_2 &= -\frac{1}{2}a_2, \\ b_4 &= \frac{1}{8}(5a_2^3 - 4a_4), \\ b_6 &= -\frac{1}{16}(49a_2^5 - 56a_2^2a_4 + 8a_6). \end{aligned}$$

- e.** [5–] What are the coefficients when b_{2n} is written as a polynomial in a_2, a_4, \dots as in (d)?
- f.** [2+] For any $C(x) = x + c_2x^2 + c_3x^3 + \dots$, show that there are unique power series

$$\begin{aligned} A(x) &= x + a_2x^2 + a_3x^3 + \dots, \\ D(x) &= x + d_3x^3 + d_5x^5 + \dots, \end{aligned}$$

such that $A(-A(-x)) = x$ and $C(x) = D(A(x))$. For instance,

$$\begin{aligned} a_2 &= c_2, \\ d_3 &= c_3 - c_2^2, \\ a_4 &= c_4 - 3c_3c_2 + 3c_2^3, \\ d_5 &= c_5 + 3c_2^2c_3 - 3c_2c_4 - c_2^4. \end{aligned}$$

- g.** [2+] Find $A(x)$ and $D(x)$ as in (f) when $C(x) = -\log(1-x)$.
- h.** [5–] What are the coefficients when a_{2n} and d_{2n+1} are written as a polynomial in c_2, c_3, \dots as in (f)?

- i. [2+] Note that if $A(x) = x/(1+2x)$, then $A(-A(-x)) = x$. Show that

$$B^{(-1)}(-B(-x)) = x/(1+2x)$$

if and only if $e^{-x} \sum_{n \geq 0} b_{n+1} x^n / n!$ is an even function of x (i.e., has only even exponents).

- j. [2+] Identify the coefficients b_{2n} of the unique $B(x) = x + \sum_{n \geq 1} b_{2n} x^{2n}$ satisfying $B^{(-1)}(-B(-x)) = x/(1+2x)$.

169. [2] Find a closed-form expression for the following generating functions:

a. $\sum_{n \geq 0} (n+2)^2 x^n$.

b. $\sum_{n \geq 0} (n+2)^2 \frac{x^n}{n!}$.

c. $\sum_{n \geq 0} (n+2)^2 \binom{2n}{n} x^n$.

170. a. [2-] Given $a_0 = \alpha$, $a_1 = \beta$, $a_{n+1} = a_n + a_{n-1}$ for $n \geq 1$, compute $y = \sum_{n \geq 0} a_n x^n$.
 b. [2+] Given $a_0 = 1$ and $a_{n+1} = (n+1)a_n - \binom{n}{2}a_{n-2}$ for $n \geq 0$, compute $y = \sum_{n \geq 0} a_n x^n / n!$.
 c. [2] Given $a_0 = 1$ and $2a_{n+1} = \sum_{i=0}^n \binom{n}{i} a_i a_{n-i}$ for $n \geq 0$, compute $\sum_{n \geq 0} a_n x^n / n!$ and find a_n explicitly. Compare equation (1.55), where (in the notation of the present exercise), $a_1 = 1$ and the recurrence holds for $n \geq 1$.
 d. [3] Let $a_k(0) = \delta_{0k}$, and for $1 \leq k \leq n+1$ let

$$a_k(n+1) = \sum_{j=0}^n \binom{n}{j} \sum_{\substack{2r+s=k-1 \\ r,s \geq 0}} (a_{2r}(j) + a_{2r+1}(j)) a_s(n-j).$$

Compute $A(x, t) := \sum_{k,n \geq 0} a_k(n) t^k x^n / n!$.

171. Given a sequence a_0, a_1, \dots of complex numbers, let $b_n = a_0 + a_1 + \dots + a_n$.

- a. [1+]* Let $A(x) = \sum_{n \geq 0} a_n x^n$ and $B(x) = \sum_{n \geq 0} b_n x^n$. Show that

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

- b. [2+] Let $A(x) = \sum_{n \geq 0} a_n \frac{x^n}{n!}$ and $B(x) = \sum_{n \geq 0} b_n \frac{x^n}{n!}$. Show that

$$B(x) = (I(e^{-x} A'(x)) + a_0) e^x, \quad (1.144)$$

where I denotes the *formal integral*, that is,

$$I\left(\sum_{n \geq 0} c_n x^n\right) = \sum_{n \geq 0} c_n \frac{x^{n+1}}{n+1} = \sum_{n \geq 1} c_{n-1} \frac{x^n}{n}.$$

172. [3-] The *Legendre polynomial* $P_n(x)$ is defined by

$$\frac{1}{\sqrt{1-2xt+t^2}} = \sum_{n \geq 0} P_n(x) t^n.$$

Show that $(1-x)^n P_n((1+x)/(1-x)) = \sum_{k=0}^n \binom{n}{k}^2 x^k$.

- 173.** [2+] Find simple closed expressions for the coefficients of the power series (expanded about $x = 0$):
- $\sqrt{\frac{1+x}{1-x}}.$
 - $2 \left(\sin^{-1} \frac{x}{2} \right)^2.$
 - $\sin(t \sin^{-1} x).$
 - $\cos(t \sin^{-1} x).$
 - $\sin(x) \sinh(x).$
 - $\sin(x) \sin(\omega x) \sin(\omega^2 x)$, where $\omega = e^{2\pi i/3}.$
 - $\cos(\log(1+x))$ (express the answer as the real part of a complex number)
- 174.** [1–] Find the order (number of elements) of the finite field \mathbb{F}_2 .
- 175.** [2+]* For $i, j \geq 0$ and $n \geq 1$, let $f_n(i, j)$ denote the number of pairs (V, W) of subspaces of \mathbb{F}_q^n such that $\dim V = i$, $\dim W = j$, and $V \cap W = \{0\}$. Find a formula for $f_n(i, j)$ which is a power of q times a q -multinomial coefficient.
- 176.** [2+] A sequence of vectors v_1, v_2, \dots is chosen uniformly and independently from \mathbb{F}_q^n . Let $E(n)$ be the expected value of k for which v_1, \dots, v_k span \mathbb{F}_q^n but v_1, \dots, v_{k-1} don't span \mathbb{F}_q^n . For instance

$$E(1) = \frac{q}{q-1},$$

$$E(2) = \frac{q(2q+1)}{(q-1)(q+1)},$$

$$E(3) = \frac{q(3q^3 + 4q^2 + 3q + 1)}{(q-1)(q+1)(q^2 + q + 1)}.$$

Show that

$$E(n) = \sum_{i=1}^n \frac{q^i}{q^i - 1}.$$

- 177. a.** [2+]* Let $f(n, q)$ denote the number of matrices $A \in \text{Mat}(n, q)$ satisfying $A^2 = 0$. Show that

$$f(n, q) = \sum_{2i+j=n} \frac{\gamma_n(q)}{q^{i(i+2j)} \gamma_i(q) \gamma_j(q)},$$

where $\gamma_m(q) = \#\text{GL}(m, q)$. (The sum ranges over all pairs $(i, j) \in \mathbb{N} \times \mathbb{N}$ satisfying $2i + j = n$.)

- b.** [2]* Write $f(n, q) = g(n, q)(q-1)^k$ so that $g(n, 1) \neq 0, \infty$. Thus, $f(n, q)$ may be regarded as a q -analogue of $g(n, 1)$. Show that

$$\sum_{n \geq 0} g(n, 1) \frac{x^n}{n!} = e^{x^2+x}.$$

- c.** [5–] Is there an intuitive explanation of why $f(n, q)$ is a “good” q -analogue of $g(n, 1)$?

- 178.** [2+]* Let $f(n)$ be the number of pairs (A, B) of matrices in $\text{Mat}(n, q)$ satisfying $AB = 0$. Show that

$$f(n) = \sum_{k=0}^n q^{n(n-k)} \binom{n}{k} (q^n - 1)(q^n - q) \cdots (q^n - q^{n-1}).$$

- 179.** [2-]* How many pairs (A, B) of matrices in $\text{Mat}(n, q)$ satisfy $A + B = AB$?
- 180.** [5-] How many matrices $A \in \text{Mat}(n, q)$ have square roots, i.e., $A = B^2$ for some $B \in \text{Mat}(n, q)$? The $q = 1$ situation is Exercise 5.11(a).
- 181.** [2]* Find a simple formula for the number $f(n)$ of matrices $A = (A_{ij}) \in \text{GL}(n, q)$ such that $A_{11} = A_{1n} = A_{n1} = A_{nn} = 0$.
- 182.** [2+] Let $f(n, q)$ denote the number of matrices $A = (A_{ij}) \in \text{GL}(n, q)$ such that $A_{ij} \neq 0$ for all i, j . Let $g(n, q)$ denote the number of matrices $B = (B_{ij}) \in \text{GL}(n-1, q)$ such that $B_{ij} \neq 1$ for all i, j . Show that

$$f(n, q) = (q-1)^{2n-1} g(n, q).$$

- 183.** [2] Prove the identity

$$\frac{1}{1-qx} = \prod_{d \geq 1} (1-x^d)^{-\beta(d)}, \quad (1.145)$$

where $\beta(d)$ is given by equation (1.103).

- 184. a.** [2]* Let $f_q(n)$ denote the number of monic polynomials $f(x)$ of degree n over \mathbb{F}_q that do not have a zero in \mathbb{F}_q (i.e., for all $\alpha \in \mathbb{F}_q$ we have $f(\alpha) \neq 0$). Find a simple formula for $F(x) = \sum_{n \geq 0} f_q(n)x^n$. Your answer should not involve any infinite sums or products.
NOTE. The constant polynomials $f(x) = \beta$ for $0 \neq \beta \in \mathbb{F}_q$ are included in the enumeration, but not the polynomial $f(x) = 0$.
- b.** [2]* Use (a) to find a simple explicit formula for $f(n, q)$ when n is sufficiently large (depending on q).
- 185. a.** [1]* Show that the number of monic polynomials of degree n over \mathbb{F}_q is q^n .
- b.** [2+] Recall that the *discriminant* of a polynomial $f(x) = (x - \theta_1) \cdots (x - \theta_n)$ is defined by

$$\text{disc}(f) = \prod_{1 \leq i < j \leq n} (\theta_i - \theta_j)^2.$$

Show that the number $D(n, 0)$ of monic polynomials $f(x)$ over \mathbb{F}_q with discriminant 0 (equivalently, $f(x)$ has an irreducible factor of multiplicity greater than 1) is q^{n-1} , $n \geq 2$.

- c.** [2+] Generalize (a) and (b) as follows. Fix $k \geq 1$, and let X be any subset of \mathbb{N}^k containing $(0, 0, \dots, 0)$. If f_1, \dots, f_k is a sequence of monic polynomials over \mathbb{F}_q , then set $f = (f_1, \dots, f_k)$ and $\deg(f) = (\deg(f_1), \dots, \deg(f_k))$. Given an irreducible polynomial $p \in \mathbb{F}_q[x]$, let $\text{mult}(p, f) = (\mu_1, \dots, \mu_k)$, where μ_i is the multiplicity of p in f_i . Given $\beta \in \mathbb{N}^k$, let $N(\beta)$ be the number of k -tuples $f = (f_1, \dots, f_k)$ of monic polynomials over \mathbb{F}_q such that $\deg(f) = \beta$ and such that for any irreducible polynomial p over \mathbb{F}_q we have $\text{mult}(p, f) \in X$. By a straightforward generalization of Exercise 1.158 to the multivariate case, there are unique $a_\alpha \in \mathbb{Z}$ such that

$$F_X(x) := \sum_{\alpha \in X} x^\alpha = \prod_{\substack{\alpha \in \mathbb{N}^k \\ \alpha \neq (0, 0, \dots, 0)}} (1 - x^\alpha)^{a_\alpha}, \quad (1.146)$$

where if $\alpha = (\alpha_1, \dots, \alpha_k)$, then $x^\alpha = x_1^{\alpha_1} \cdots x_k^{\alpha_k}$. Show that

$$\sum_{\beta \in \mathbb{N}^k} N(\beta) x^\beta = \prod_{\substack{\alpha \in \mathbb{N}^k \\ \alpha \neq (0, 0, \dots, 0)}} (1 - qx^\alpha)^{a_\alpha}.$$

Note that if $k = 1$ and $X = \mathbb{N}$, then $N(\beta)$ is the total number of monic polynomials of degree β . We have $f_{\mathbb{N}}(x) = 1/(1-x)$ and $\sum_{\beta \in \mathbb{N}} N(\beta)x^\beta = 1/(1-qx) = \sum_{n \geq 0} q^n x^n$, agreeing with (a).

186. Deduce from Exercise 1.185(c) the following results.

- a.** [2] The number $N_r(n)$ of monic polynomials $f \in \mathbb{F}_q[x]$ of degree n with no factor of multiplicity at least r is given by

$$N_r(n) = q^n - q^{n-r+1}, \quad n \geq r. \quad (1.147)$$

Note that the case $r = 2$ is equivalent to (b).

- b.** [2] Let $N(m, n)$ be the number of pairs (f, g) of monic relatively prime polynomials over \mathbb{F}_q of degrees m and n . In other words, f and g have nonzero resultant. Then

$$N(m, n) = q^{m+n-1}, \quad m, n \geq 1. \quad (1.148)$$

- c.** [2+] A polynomial f over a field K is *powerful* if every irreducible factor of f occurs with multiplicity at least two. Let $P(n)$ be the number of powerful monic polynomials of degree n over \mathbb{F}_q . Show that

$$P(n) = q^{\lfloor n/2 \rfloor} + q^{\lfloor n/2 \rfloor - 1} - q^{\lfloor (n-1)/3 \rfloor}, \quad n \geq 2. \quad (1.149)$$

187. a. [3–] Let q be an odd prime power. Show that as f ranges over all monic polynomials of degree $n > 1$ over \mathbb{F}_q , $\text{disc}(f)$ is just as often a nonzero square in \mathbb{F}_q as a nonsquare.

- b.** [2+] For $n > 1$ and $a \in \mathbb{F}_q$, let $D(n, a)$ denote the number of monic polynomials of degree n over \mathbb{F}_q with discriminant a . Thus by Exercise 1.185(b), we have $D(n, 0) = q^{n-1}$. Show that if $(n(n-1), q-1) = 1$ (so $q = 2^m$) or $(n(n-1), q-1) = 2$ (so q is odd) then $D(n, a) = q^{n-1}$ for all $a \in \mathbb{F}_q$. (Here (r, s) denotes the greatest common divisor of r and s .)

- c.** [5–] Investigate further the function $D(n, a)$ for general n and a .

188. [3] Give a direct proof of Corollary 1.10.11 (i.e., the number of nilpotent matrices in $\text{Mat}(n, q)$ is $q^{n(n-1)}$).

189. [3–] Let V be an $(m+n)$ -dimensional vector space over \mathbb{F}_q , and let $V = V_1 \oplus V_2$, where $\dim V_1 = m$ and $\dim V_2 = n$. Let $f(m, n)$ be the number of nilpotent linear transformations $A: V \rightarrow V$ satisfying $A(V_1) \subseteq V_2$ and $A(V_2) \subseteq V_1$. Show that

$$f(m, n) = q^{m(n-1)+n(m-1)}(q^m + q^n - 1),$$

190. a. [2] Let $\omega^*(n, q)$ denote the number of conjugacy classes in the group $\text{GL}(n, q)$. Show that $\omega^*(n, q)$ is a polynomial in q satisfying $\omega^*(n, 1) = 0$. For instance,

$$\begin{aligned} \omega^*(1, q) &= q - 1, \\ \omega^*(2, q) &= q^2 - 1, \\ \omega^*(3, q) &= q^3 - q, \\ \omega^*(4, q) &= q^4 - q, \\ \omega^*(5, q) &= q^5 - q^2 - q + 1, \\ \omega^*(6, q) &= q^6 - q^2, \\ \omega^*(7, q) &= q^7 - q^3 - q^2 + 1, \\ \omega^*(8, q) &= q^8 - q^3 - q^2 + q. \end{aligned}$$

b. [2+] Show that

$$\omega^*(n, q) = q^n - q^{\lfloor (n-1)/2 \rfloor} + O(q^{\lfloor (n-1)/2 \rfloor - 1}).$$

c. [3–] Evaluate the polynomial values $\omega^*(n, 0)$ and $\omega^*(n, -1)$. When is $\omega^*(n, q)$ divisible by q^2 ?

191. [3–] Give a more conceptual proof of Proposition 1.10.2, that is, the number $\omega(n, q)$ of orbits of $\text{GL}(n, q)$ acting adjointly on $\text{Mat}(n, q)$ is given by

$$\omega(n, q) = \sum_j p_j(n) q^j.$$

192. a. [2]* Find a simple formula for the number of surjective linear transformations $A: \mathbb{F}_q^n \rightarrow \mathbb{F}_q^k$.

b. [2]* Show that the number of $m \times n$ matrices of rank k over \mathbb{F}_q is given by

$$\binom{m}{k} (q^n - 1)(q^n - q) \cdots (q^n - q^{k-1}).$$

193. [2] Let p_n denote the number of projections $P \in \text{Mat}(n, q)$ (i.e., $P^2 = P$). Show that

$$\sum_{n \geq 0} p_n \frac{x^n}{\gamma_n} = \left(\sum_{k \geq 0} \frac{x^k}{\gamma(k)} \right)^2,$$

where as usual $\gamma(k) = \gamma(k, q) = \#\text{GL}(k, q)$.

194. [2+] Let r_n denote the number of regular (or cyclic) $M \in \text{Mat}(n, q)$ (i.e., the characteristic and minimal polynomials of A are the same). Equivalently, there is a column vector $v \in \mathbb{F}_q^n$ such that the set $\{A^i v : i \geq 0\}$ spans \mathbb{F}_q^n (where we set $A^0 = I$). Show that

$$\begin{aligned} \sum_{n \geq 0} r_n \frac{x^n}{\gamma(n)} &= \prod_{d \geq 1} \left(1 + \frac{x^d}{(q^d - 1)(1 - (x/q)^d)} \right)^{\beta(d)} \\ &= \frac{1}{1-x} \prod_{d \geq 1} \left(1 + \frac{x^d}{q^d(q^d - 1)} \right)^{\beta(d)}. \end{aligned}$$

195. [2] A matrix A is *semisimple* if it can be diagonalized over the algebraic closure of the base field. Let s_n denote the number of semisimple matrices $A \in \text{Mat}(n, q)$. Show that

$$\sum_{n \geq 0} s_n \frac{x^n}{\gamma(n, q)} = \prod_{d \geq 1} \left(\sum_{j \geq 0} \frac{x^{jd}}{\gamma(j, q^d)} \right)^{\beta(d)}.$$

196. a. [2+] Generalize Proposition 1.10.15 as follows. Let $0 \leq k \leq n$, and let $f_k(n)$ be the number of matrices $A = (a_{ij}) \in \text{GL}(n, q)$ satisfying $a_{11} + a_{22} + \cdots + a_{kk} = 0$. Then

$$f_k(n) = \frac{1}{q} \left(\gamma(n, q) + (-1)^k (q-1) q^{\frac{1}{2}k(2n-k-1)} \gamma(n-k, q) \right). \quad (1.150)$$

- b.** [2+] Let H be any linear hyperplane in the vector space $\text{Mat}(n, q)$. Find (in terms of certain data about H) a formula for $\#(\text{GL}(n, q) \cap H)$.
- 197.** [3] Let $f(n)$ be the number of matrices $A \in \text{GL}(n, q)$ with zero diagonal (i.e., all diagonal entries are equal to 0). Show that

$$f(n) = q^{\binom{n-1}{2}-1} (q-1)^n \sum_{i=0}^n (-1)^i \binom{n}{i} (n-i)!$$

For instance,

$$\begin{aligned} f(1) &= 0, \\ f(2) &= (q-1)^2, \\ f(3) &= q(q-1)(q^4 - 4q^2 + 4q - 1), \\ f(4) &= q^3(q-1)(q^8 - q^6 - 5q^5 + 3q^4 + 11q^3 - 14q^2 + 6q - 1). \end{aligned}$$

- 198. a.** [2+] Let $h(n, r)$ denote the number of $n \times n$ symmetric matrices of rank r over \mathbb{F}_q . Show that

$$h(n+1, r) = q^r h(n, r) + (q-1)q^{r-1} h(n, r-1) + (q^{n+1} - q^{r-1})h(n, r-2), \quad (1.151)$$

with the initial conditions $h(n, 0) = 1$ and $h(n, r) = 0$ for $r > n$.

- b.** [2] Deduce that

$$h(n, r) = \begin{cases} \prod_{i=1}^s \frac{q^{2i}}{q^{2i}-1} \cdot \prod_{i=0}^{2s-1} (q^{n-i} - 1), & 0 \leq r = 2s \leq n, \\ \prod_{i=1}^s \frac{q^{2i}}{q^{2i}-1} \cdot \prod_{i=0}^{2s} (q^{n-i} - 1), & 0 \leq r = 2s+1 \leq n. \end{cases}$$

In particular, the number $h(n, n)$ of $n \times n$ invertible symmetric matrices over \mathbb{F}_q is given by

$$h(n, n) = \begin{cases} q^{m(m-1)}(q-1)(q^3-1)\cdots(q^{2m-1}), & n = 2m-1, \\ q^{m(m+1)}(q-1)(q^3-1)\cdots(q^{2m-1}), & n = 2m. \end{cases}$$

- 199. a.** [3] Show that the following three numbers are equal:
- The number of symmetric matrices in $\text{GL}(2n, q)$ with zero diagonal,
 - The number of symmetric matrices in $\text{GL}(2n-1, q)$,
 - The number of skew-symmetric matrices ($A = -A^t$) in $\text{GL}(2n, q)$.
- b.** [5] Give a combinatorial proof of (a). (No combinatorial proof is known that two of these items are equal.)
- 200.** [3] Let $C_n(q)$ denote the number of $n \times n$ upper-triangular matrices X over \mathbb{F}_q satisfying $X^2 = 0$. Show that

$$\begin{aligned} C_{2n}(q) &= \sum_j \left[\binom{2n}{n-3j} - \binom{2n}{n-3j-1} \right] \cdot q^{n^2-3j^2-j} \\ C_{2n+1}(q) &= \sum_j \left[\binom{2n+1}{n-3j} - \binom{2n+1}{n-3j-1} \right] \cdot q^{n^2+n-3j^2-2j}. \end{aligned}$$

- 201.** This exercise and the next show that simply stated counting problems over \mathbb{F}_q can have complicated solutions beyond the realm of combinatorics. (See also Exercise 4.39(a).)
- a.** [3] Let

$$f(q) = \#\{(x, y, z) \in \mathbb{F}_q^3 : x + y + z = 0, xyz = 1\}.$$

Show that $f(q) = q + a - 2$, where

- if $q \equiv 2 \pmod{3}$, then $a = 0$,
- if $q \equiv 1 \pmod{3}$, then a is the unique integer such that $a \equiv 1 \pmod{3}$ and $a^2 + 27b^2 = 4q$ for some integer b .

- b.** [2+] Let

$$g(q) = \#\{A \in \text{GL}(3, q) : \text{tr}(A) = 0, \det(A) = 1\}$$

Express $g(q)$ in terms of the function $f(q)$ of part (a).

- 202.** [4–] Let p be a prime, and let N_p denote the number of solutions modulo p to the equation $y^2 + y = x^3 - x$. Let $a_p = p - N_p$. For instance, $a_2 = -2$, $a_3 = 1$, $a_5 = 1$, $a_7 = -2$, and so on. Show that if $p \neq 11$, then

$$\begin{aligned} a_p &= [x^p]x \prod_{n \geq 1} (1 - x^n)^2 (1 - x^{11n})^2 \\ &= [x^p](x - 2x^2 - x^3 + 2x^4 + x^5 + 2x^6 - 2x^7 - 2x^9 + \cdots). \end{aligned}$$

- 203.** [3] The following quotation is from Plutarch's *Table-Talk* VIII. 9, 732: "Chrysippus says that the number of compound propositions that can be made from only ten simple propositions exceeds a million. (Hipparchus, to be sure, refuted this by showing that on the affirmative side there are 103,049 compound statements, and on the negative side 310,952.)"

According to T. L. Heath, *A History of Greek Mathematics*, vol. 2, p. 245, "it seems impossible to make anything of these figures."

Can in fact any sense be made of Plutarch's statement?

Solutions to Exercises

- 1.** *Answer:* 2. There is strong evidence that human babies, chimpanzees, and even rats have an understanding of this problem. See S. Dehaene, *The Number Sense: How the Mind Creates Mathematics*, Oxford, New York, 1997, pp. 23–27, 52–56.
- 2.** Here is one possible way to arrive at the answers. There may be other equally simple (or even simpler) ways to solve these problems.
 - a.** $2^{10} - 2^5 = 992$
 - b.** $\frac{1}{2}(7-1)! = 360$
 - c.** $5 \cdot 5!$ (or $6! - 5!$) = 600
 - d.** $\binom{6}{1}4! + \binom{6}{2}3! + \frac{1}{2}\binom{6}{3}2!^2 = 274$
 - e.** $\binom{6}{4} + \binom{6}{1}\binom{5}{2} + \frac{1}{3!}\binom{6}{2}\binom{4}{2} = 90$

- f. $(6)_4 = 360$
- g. $1 \cdot 3 \cdot 5 \cdot 7 \cdot 9 = 945$
- h. $\binom{7}{2} + \binom{8}{3} + \binom{9}{1} = 86$
- i. $\binom{11}{1,2,4,4} - \binom{8}{1,1,2,4} = 33810$
- j. $\binom{8+1}{4} = 126$
- k. $2 \binom{8}{1,3,4} + 3 \binom{8}{2,3,3} + \binom{8}{2,2,4} = 2660$
- l. $5! + \binom{5}{2} (5)_4 + \frac{1}{2} \binom{5}{1} \binom{4}{2} (5)_3 = 2220$
3. a. Given any n -subset S of $[x+n+1]$, there is a largest k for which $\#(S \cap [x+k]) = k$. Given k , we can choose S to consist of any k -element subset in $\binom{x+k}{k}$ ways, together with $\{x+k+2, x+k+3, \dots, x+n+1\}$.
- b. *First proof.* Choose a subset of $[n]$ and circle one of its elements in $\sum k \binom{n}{k}$ ways. Alternatively, circle an element of $[n]$ in n ways, and choose a subset of what remains in 2^{n-1} ways.
Second proof (not quite so combinatorial, but nonetheless instructive). Divide the identity by 2^n . It then asserts that the average size of a subset of $[n]$ is $n/2$. This follows since each subset can be paired with its complement.
- c. To give a noncombinatorial proof, simply square both sides of the identity (Exercise 1.8(a))

$$\sum_{n \geq 0} \binom{2n}{n} x^n = \frac{1}{\sqrt{1-4x}}$$

and equate coefficients. The problem of giving a combinatorial proof was raised by P. Veress and solved by G. Hajos in the 1930s. For some published proofs, see D. J. Kleitman, *Studies in Applied Math.* **54** (1975), 289–292; M. Sved, *Math. Intelligencer* **6**(4) (1984), 44–45; and V. De Angelis, *Amer. Math. Monthly* **113** (2006), 642–644.

- d. G. E. Andrews, *Discrete Math.* **11** (1975), 97–106.
- e. Given an n -element subset S of $[2n-1]$, associate with it the two n -element subsets S and $[2n] - S$ of $[2n]$.
- f. What does it mean to give a combinatorial proof of an identity with minus signs? The simplest (but not the only) possibility is to rearrange the terms so that all signs are positive. Thus, we want to prove that

$$\sum_{k \text{ even}} \binom{n}{k} = \sum_{k \text{ odd}} \binom{n}{k}, \quad n \geq 1. \quad (1.152)$$

Let \mathcal{E}_n (respectively \mathcal{O}_n) denote the sets of all subsets of $[n]$ of even (respectively, odd) cardinality. The left-hand side of equation (1.152) is equal to $\#\mathcal{E}_n$, while the right-hand side is $\#\mathcal{O}_n$. Hence, we want to give a bijection $\varphi: \mathcal{E}_n \rightarrow \mathcal{O}_n$. The

definition of φ is very simple:

$$\varphi(S) = \begin{cases} S \cup \{n\}, & n \notin S, \\ S - \{n\}, & n \in S. \end{cases}$$

Another way to look at this proof is to consider φ as an involution on all of $2^{[n]}$. Every orbit of φ has two elements, and their contributions to the sum $\sum_{S \subseteq [n]} (-1)^{\#S}$ cancel out, that is, $(-1)^{\#S} + (-1)^{\#\varphi(S)} = 0$. Hence φ is a *sign-reversing involution* as in the proof of Proposition 1.8.7.

- g. The left-hand side counts the number of triples (S, T, f) , where $S \subseteq [n]$, $T \subseteq [n+1, 2n]$, $\#S = \#T$, and $f: S \rightarrow [x]$. The right-hand side counts the number of triples (A, B, g) , where $A \subseteq [n]$, $B \in \binom{[2n]-A}{n}$, and $g: A \rightarrow [x-1]$. Given (S, T, f) , define (A, B, g) as follows: $A = f^{-1}([x-1])$, $B = ([n] - S) \cup T$, and $g(i) = f(i)$ for $i \in [x-1]$.
- h. We have that $\binom{i+j}{i} \binom{j+k}{j} \binom{k+i}{i}$ is the number of triples (α, β, γ) , where (i) α is a sequence of $i+j+2$ letters a and b beginning with a and ending with b , with $i+1$ a 's (and hence $j+1$ b 's), (ii) $\beta = (\beta_1, \dots, \beta_{j+1})$ is a sequence of $j+1$ positive integers with sum $j+k+1$, and (iii) $\gamma = (\gamma_1, \dots, \gamma_{i+1})$ is a sequence of $i+1$ positive integers with sum $k+i+1$. Replace the r th a in α by the word $c^{\gamma_r} d$, and replace the r th b in α by the word $d^{\beta_r} c$. In this way, we obtain a word δ in c, d of length $2n+4$ with $n+2$ c 's and $n+2$ d 's. This word begins with c and ends with $d(dc)^m$ for some $m \geq 1$. Remove the prefix c and suffix $d(dc)^m$ from δ to obtain a word ϵ of length $2(n-m+1)$ with $n-m+1$ c 's and $n-m+1$ d 's. The map $(\alpha, \beta, \gamma) \mapsto \epsilon$ is easily seen to yield a bijective proof of (h). This argument is due to Roman Travin (private communication, October 2007).

Example. Let $n = 8$, $i = 2$, $j = k = 3$, $\alpha = abbaabb$, $\beta = (2, 3, 1, 1)$, $\gamma = (2, 3, 1)$. Then

$$\delta = (c^2 d)(d^2 c)(d^3 c)(c^3 d)(cd)(dc)(dc),$$

$$\text{so } \epsilon = cd^3 cd^3 c^4 dc.$$

NOTE. Almost any binomial coefficient identity can be proved nowadays automatically by computer. For an introduction to this subject, see M. Petkovšek, H. S. Wilf, and D. Zeilberger, *A = B*, A K Peters, Wellesley, Mass., 1996. Of course, it is still of interest to find elegant bijective proofs of such identities.

8. a. We have $1/\sqrt{1-4x} = \sum_{n \geq 0} \binom{-1/2}{n} (-4)^n x^n$. Now

$$\begin{aligned} \binom{-1/2}{n} (-4)^n &= \frac{(-\frac{1}{2})(-\frac{3}{2}) \cdots (-\frac{2n-1}{2}) (-4)^n}{n!} \\ &= \frac{2^n \cdot 1 \cdot 3 \cdots (2n-1)}{n!} = \frac{(2n)!}{n!^2}. \end{aligned}$$

- b. Note that $\binom{2n-1}{n} = \frac{1}{2} \binom{2n}{n}$, $n > 0$ (see Exercise 1.3(e)).
9. b. Powerful methods exist for solving this type of problem (see Example 6.3.8); however, we give here a “naïve” solution. Suppose the path has k steps of the form $(0, 1)$, and therefore k $(1, 0)$'s and $n-k$ $(1, 1)$'s. These $n+k$ steps may be chosen

in any order, so

$$\begin{aligned}
 f(n, n) &= \sum_k \binom{n+k}{n-k, k, k} = \sum_k \binom{n+k}{2k} \binom{2k}{k} \\
 \Rightarrow \sum_{n \geq 0} f(n, n) x^n &= \sum_k \binom{2k}{k} \sum_{n \geq 0} \binom{n+k}{2k} x^n \\
 &= \sum_k \binom{2k}{k} \frac{x^k}{(1-x)^{2k+1}} \\
 &= \frac{1}{1-x} \left(1 - \frac{4x}{(1-x)^2} \right)^{-1/2}, \text{ by Exercise 1.8(a)} \\
 &= \frac{1}{\sqrt{1-6x+x^2}}.
 \end{aligned}$$

- 10.** Let the elements of S be $a_1 < a_2 < \dots < a_{r+s}$. Then the multiset $\{a_1, a_2 - 2, a_3 - 4, \dots, a_{r+s} - 2(r+s-1)\}$ consists of r odd numbers and s even numbers in $[2(n-r-s+1)]$. Conversely, we can recover S from any r odd numbers and s even numbers (allowing repetition) in $[2(n-r-s+1)]$. Hence,

$$f(n, r, s) = \left(\binom{n-r-s+1}{r} \right) \left(\binom{n-r-s+1}{s} \right) = \binom{n-r}{s} \binom{n-s}{r}.$$

This result is due to Jim Propp, private communication dated 29 July 2006. Propp has generalized the result to any modulus $m \geq 2$ and has also given a q -analogue.

- 11. a.** Choose $m+n+1$ points uniformly and independently from the interval $[0, 1]$. The integral is then the probability that the last chosen point u is greater than the first m of the other points and less than the next n points. There are $(m+n+1)!$ orderings of the points, of which exactly $m!n!$ of them have the first m chosen points preceding u and the next n following u . Hence,

$$B(m+1, n+1) = \frac{m!n!}{(m+n+1)!}.$$

The function $B(x, y)$ for $\operatorname{Re}(x), \operatorname{Re}(y) > 0$ is the *beta function*.

There are many more interesting examples of the combinatorial evaluation of integrals. Two of the more sophisticated ones are P. Valtr, *Discrete Comput. Geom.* **13** (1995), 637–643; and *Combinatorica* **16** (1996), 567–573.

- b.** Choose $(1+r+s)n+2t \binom{n}{2}$ points uniformly and independently from $[0, 1]$. Label the first n chosen points x , the next r chosen points y_1 , and so on, so that the points are labeled by the elements of M . Let P be the probability that the order of the points in $[0, 1]$ is a permutation of M that we are counting. Then

$$\begin{aligned}
 P &= \frac{n!r!s!(2t)!\binom{n}{2}}{((r+s+1)n+tn(n-1))!} f(n, r, s, t) \\
 &= \int_0^1 \cdots \int_0^1 (x_1 \cdots x_n)^r ((1-x_1) \cdots (1-x_n))^s \prod_{1 \leq i < j \leq n} (x_i - x_j)^{2t} dx_1 \cdots dx_n.
 \end{aligned}$$

This integral is the famous *Selberg integral*; see for example G. E. Andrews, R. Askey, and R. Roy, *Special Functions*, Cambridge University Press, Cambridge/New York, 1999 (Chapter 8), and P. J. Forrester and S. O. Warnaar, *Bull.*

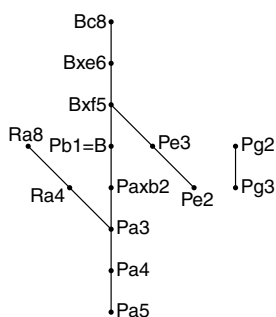


Figure 1.29 The solution poset for Exercise 1.11(c).

Amer. Math. Soc. **45** (2008), 489–534. The evaluation of this integral immediately gives equation (1.119). No combinatorial proof of (1.119) is known. Such a proof would be quite interesting since it would give a combinatorial evaluation of Selberg’s integral.

- c. One solution is 1.Pa5 2.Pa4 3.Pa3 4.Ra4 5.Ra8 6.Paxb2 7.Pb1=B 8.Pe2 9.Pe3 10.Bxf5 11.Bxe6 12.Bc8 13.Pg3 14.Pg2, after which White plays Bh2 mate. We attach indeterminates to each of the Black moves as follows: 1. a_{12} 2. a_{12} 3. x 4. a_{24} 5. a_{24} 6. a_{23} 7. a_{23} 8. a_{13} 9. a_{13} 10. x 11. a_{34} 12. a_{34} 13. a_{14} 14. a_{14} . We also place an indeterminate x before Black’s first move and after Black’s last move. All solutions are then obtained by permutations of Black’s 14 moves, together with x at the beginning and end, with the property that moves labeled by the same indeterminate must be played in the same order, and moves labeled a_{ij} must occur between the i th x and j th x . In the terminology of Chapter 3, the solutions correspond to the linear extensions of the poset shown in Figure 1.29. Hence the number of solutions is

$$f(4, 0, 0, 1) = 54054.$$

For similar serieshelpmates (called *queue problems*) whose number of solutions has some mathematical significance, see Exercises 1.145, 6.23, and 7.18. Some references are given in the solution to Exercise 6.23. The present problem comes from the article R. Stanley, *Suomen Tehtäväniekat* **59**, no. 4 (2005), 193–203.

13. Let S consist of all p -tuples (n_1, n_2, \dots, n_p) of integers $n_i \in [a]$ such that not all the n_i ’s are equal. Hence, $\#S = a^p - a$. Define two sequences in S to be equivalent if one is a cyclic shift of the other (clearly an equivalence relation). Since p is prime each equivalence class contains exactly p elements, and the proof follows. For additional results of this nature, see I. M. Gessel, in *Enumeration and Design* (Waterloo, Ont., 1982), Academic Press, Toronto, 1984, pp. 157–197, and G.-C. Rota and B. E. Sagan, *European J. Combin.* **1** (1980), 67–76.
14. a. We use the well-known and easily proved fact that $(x + 1)^p \equiv x^p + 1 \pmod{p}$, meaning that each coefficient of the polynomial $(x + 1)^p - (x^p + 1)$ is divisible by p . Thus,

$$\begin{aligned} (x + 1)^n &= (x + 1)^{\sum a_i p^i} \\ &\equiv \prod_i \left(x^{p^i} + 1 \right)^{a_i} \pmod{p} \\ &\equiv \prod_i \sum_{j=0}^{a_i} \binom{a_i}{j} x^{j p^i} \pmod{p}. \end{aligned}$$

- The coefficient of x^m on the left is $\binom{n}{m}$ and on the right is $\binom{a_0}{b_0}\binom{a_1}{b_1}\cdots$. This congruence is due to F. E. A. Lucas, *Bull. Soc. Math. France* **6** (1878), 49–54.
- b. The binomial coefficient $\binom{n}{m}$ is odd if and only if the binary expansion of m is “contained” in that of n ; that is, if m has a 1 in its i th binary digit, then so does n . Hence, $\binom{n}{m}$ is odd for all $0 \leq m \leq n$ if and only if $n = 2^k - 1$. More generally, the number of odd coefficients of $(1+x)^n$ is equal to $2^{b(n)}$, where $b(n)$ is the number of 1’s in the binary expansion of n . See Exercise 1.15 for some variations.
- c. Consider an $a \times p$ rectangular grid of squares. Choose pb of these squares in $\binom{pa}{pb}$ ways. We can choose the pb squares to consist of b entire rows in $\binom{a}{b}$ ways. Otherwise, in at least two rows, we will have picked between 1 and $p-1$ squares. For any choice of pb squares, cyclically shift the squares in each row independently. This partitions our choices into equivalence classes. Exactly $\binom{a}{b}$ of these classes contain one element; the rest contain a number of elements divisible by p^2 .
- d. Continue the reasoning of (c). If a choice of pb squares contains fewer than $b-2$ entire rows, then its equivalence class has cardinality divisible by p^3 . From this we reduce the problem to the case $a=2, b=1$. Now

$$\begin{aligned} \binom{2p}{p} &= \sum_{k=0}^p \binom{p}{k}^2 \\ &= 2 + p^2 \sum_{k=1}^{p-1} \frac{(p-1)^2(p-2)^2 \cdots (p-k+1)^2}{k!^2} \\ &\equiv 2 + p^2 \sum_{k=1}^{p-1} k^{-2} \pmod{p^3}. \end{aligned}$$

But as k ranges from 1 to $p-1$, so does k^{-1} modulo p . Hence,

$$\sum_{k=1}^{p-1} k^{-2} \equiv \sum_{k=1}^{p-1} k^2 \pmod{p}.$$

Now use, for example, the identity

$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$$

to get

$$\sum_{k=1}^{p-1} k^2 \equiv 0 \pmod{p}, \quad p \geq 5.$$

- e. The exponent of the largest power of p dividing $\binom{n}{m}$ is the number of carries needed to add m and $n-m$ in base p . See E. Kummer, *J. Math.* **44** (1852), 115–116, and L. E. Dickson, *Quart. J. Math.* **33** (1902), 378–384.

15. a. We have

$$1 + x + x^2 = \frac{1 - x^3}{1 - x} \equiv (1 - x)^2 \pmod{3}.$$

Hence, $(1 + x + x^2)^n \equiv (1 - x)^{2n} \pmod{3}$. It follows easily from Exercise 1.14(a) that if $2n$ has the ternary expansion $2n = \sum a_i 3^i$, then the number of coefficients of $(1 + x + x^2)^n$ not divisible by 3 is equal to $\prod (1 + a_i)$. This result was obtained in collaboration with T. Amdeberhan.

b. Let $f(n)$ be the desired number. First consider the case $n = 2^j(2^k - 1)$. Since $(1 + x + x^2)^{2^j} \equiv 1 + x^{2^j} + x^{2^{j+1}} \pmod{2}$, we have $f(n) = f(2^k - 1)$. Now

$$(1 + x + x^2)^{2^k - 1} \equiv \frac{1 + x^{2^k} + x^{2^{k+1}}}{1 + x + x^2} \pmod{2}.$$

It is easy to check that modulo 2 we have for k odd that

$$\begin{aligned} \frac{1 + x^{2^k} + x^{2^{k+1}}}{1 + x + x^2} &= 1 + x + x^3 + x^4 + x^6 + x^7 + \cdots + x^{2^k - 2} + x^{2^k - 1} + x^{2^k} \\ &\quad + x^{2^k + 2} + x^{2^k + 3} + x^{2^k + 5} + x^{2^k + 6} + \cdots + x^{2^{k+1} - 3} + x^{2^{k+1} - 2}. \end{aligned}$$

It follows that $f(2^k - 1) = (2^{k+2} + 1)/3$. Similarly, when k is even we have

$$\begin{aligned} \frac{1 + x^{2^k} + x^{2^{k+1}}}{1 + x + x^2} &= 1 + x + x^3 + x^4 + x^6 + x^7 + \cdots + x^{2^k - 4} + x^{2^k - 3} + x^{2^k - 1} \\ &\quad + x^{2^k + 1} + x^{2^k + 2} + x^{2^k + 4} + x^{2^k + 5} + \cdots + x^{2^{k+1} - 3} + x^{2^{k+1} - 2}. \end{aligned}$$

Hence, in this case, $f(2^k - 1) = (2^{k+2} - 1)/3$. For a generalization, see Exercise 4.25.

Now any positive integer n can be written uniquely as $n = \sum_{i=1}^r 2^{j_i} (2^{k_i} - 1)$, where $k_i \geq 1$, $j_i \geq 0$, and $j_{i+1} > j_i + k_i$. We are simply breaking up the binary expansion of n into the maximal strings of consecutive 1's. The lengths of these strings are k_1, \dots, k_r . Thus,

$$(1 + x + x^2)^n \equiv \prod_{i=1}^r (1 + x^{2^{j_i}} + x^{2^{j_i+1}})^{2^{k_i} - 1} \pmod{2}.$$

There is no cancellation among the coefficients when we expand this product since $j_{i+1} > j_i + 1$. Hence,

$$f(n) = \prod_{i=1}^r f(2^{k_i} - 1),$$

where $f(2^{k_i} - 1)$ is given earlier.

Example. The binary expansion of 6039 is 1011110010111. The maximal strings of consecutive 1's have lengths 1, 4, 1, and 3. Hence,

$$f(6039) = f(1)f(15)f(1)f(7) = 3 \cdot 21 \cdot 3 \cdot 11 = 2079.$$

c. We have

$$\prod_{1 \leq i < j \leq n} (x_i + x_j) \equiv \prod_{1 \leq i < j \leq n} (x_i - x_j) \pmod{2},$$

where the notation means that the corresponding coefficients of each side are congruent modulo 2. The latter product is just the value of the Vandermonde determinant $\det[x_i^{j-1}]_{i,j=1}^n$, so the number of odd coefficients is $n!$. This result can also be proved by a cancellation argument; see Exercise 2.34. A more subtle result, equivalent to Exercise 4.64(a), is that the number of *nonzero* coefficients of the polynomial $\prod_{1 \leq i < j \leq n} (x_i + x_j)$ is equal to the number of forests on an n -element vertex set.

Some generalizations of the results of this exercise appear in T. Amdeberhan and R. Stanley, Polynomial coefficient enumeration, preprint dated 3 February 2008;

(<http://math.mit.edu/~rstan/papers/coef.pdf>).

See also Exercise 4.24.

16. a. This result was first given by N. Strauss as Problem 6527, *Amer. Math. Monthly* **93** (1986), 659, and later as the paper *Linear Algebra Appl.* **90** (1987), 65–72. An elegant solution to Strauss's problem was given by I. M. Gessel, *Amer. Math. Monthly* **95** (1988), 564–565, and by W. C. Waterhouse, *Linear Algebra Appl.* **105** (1988), 195–198. Namely, let V be the vector space of all functions $\mathbb{F}_p \rightarrow \mathbb{F}_p$. A basis for V consists of the functions $f_j(a) = a^j$, $0 \leq j \leq p-1$. Let $\Phi: V \rightarrow V$ be the linear transformation defined by $(\Phi f)(x) = (1-x)^{p-1} f(1/(1-x))$. Then it can be checked that A is just the matrix of Φ with respect to the basis f_j . It is now routine to verify that $A^3 = I$.
- b. Answer: $(p+2\epsilon)/3$, where $\epsilon = 1$ if $p \equiv 1 \pmod{3}$ and $\epsilon = -1$ if $p \equiv -1 \pmod{3}$. Both Strauss, op. cit., and Waterhouse, op. cit., in fact compute the Jordan normal form of A . Waterhouse uses the linear transformation Φ to give a proof similar to that given in (a).
17. b. Think of a choice of m objects from n with repetition allowed as a placement of $n-1$ vertical bars in the slots between m dots (including slots at the beginning and end). For example,

|..||...|..

corresponds to the multiset $\{1^0, 2^2, 3^0, 4^3, 5^2\}$. Now change the bars to dots and vice versa:

.||..|||.||

yielding $\{1^1, 2^0, 3^2, 4^0, 5^0, 6^1, 7^0, 8^0\}$. This procedure gives the desired bijection. (Of course a more formal description is possible but only seems to obscure the elegance and simplicity of the above bijection.)

19. a. One way to prove (1.120) is to recall the Lagrange interpolation formula. Namely, if $P(x)$ is a polynomial of degree less than n and x_1, \dots, x_n are distinct numbers (or indeterminates), then

$$P(x) = \sum_{i=1}^n P(x_i) \prod_{j \neq i} \frac{x - x_j}{x_i - x_j}.$$

Now set $P(x) = 1$ and $x = 0$.

Applying the hint, we see that the constant term $C(a_1, \dots, a_n)$ satisfies the recurrence

$$C(a_1, \dots, a_n) = \sum_{i=1}^k C(a_1, \dots, a_i - 1, \dots, a_n),$$

if $a_i > 0$. If, on the other hand, $a_i = 0$, we have

$$C(a_1, \dots, a_{i-1}, 0, a_{i+1}, \dots, a_n) = C(a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n).$$

This is also the recurrence satisfied by $\binom{a_1+\dots+a_n}{a_1,\dots,a_n}$, and the initial conditions $C(0,\dots,0)=1$ and $\binom{0}{0,\dots,0}=1$ agree.

This result was conjectured by F. J. Dyson in 1962 and proved that same year by J. Gunson and K. Wilson. The elegant proof given here is due to I. J. Good in 1970. For further information and references, see [1.3, pp. 377–387].

- b. This identity is due to A. C. Dixon, *Proc. London Math. Soc.* **35**(1) (1903), 285–289.
- c. This is the “ q -Dyson conjecture,” due to G. E. Andrews, in *Theory and Application of Special Functions* (R. Askey, ed.), Academic Press, New York, 1975, pp. 191–224 (see §5). It was first proved by D. M. Bressoud and D. Zeilberger, *Discrete Math.* **54** (1985), 201–224. A more recent paper with many additional references is I. M. Gessel, L. Lv, G. Xin, and Y. Zhou, *J. Combinatorial Theory, Ser. A* **115** (2008), 1417–1435.
- d. I. G. Macdonald conjectured a generalization of (a) corresponding to any root system R . The present problem corresponds to $R = D_n$, while (a) is the case $R = A_{n-1}$ (when all the a_i ’s are equal). After many partial results, the conjecture was proved for all root systems by E. Opdam, *Invent. Math.* **98** (1989), 1–18. Macdonald also gave a q -analogue of his conjecture, which was finally proved by I. Cherednik in 1993 and published in *Ann. Math.* **141** (1995), 191–216. For the original papers of Macdonald, see *Sem. d’Alg. Paul Dubriel et Marie-Paule Malliavin*, Lecture Notes in Math., no. 867, Springer, Berlin, pp. 90–97, and *SIAM J. Math. Anal.* **13** (1982), 988–1007.
- e. Write

$$\begin{aligned}
 F(x) = F(x_1, \dots, x_n) &= \sum_{a_1, \dots, a_n \geq 0} \left[\prod_{i=1}^n (q^{-a_i} + \dots + q^{a_i}) \right] x_1^{a_1} \dots x_n^{a_n} \\
 &= \prod_{i=1}^n \sum_{j \geq 0} (q^{-j} + \dots + q^j) x_i^j \\
 &= \prod_{i=1}^n \sum_{j \geq 0} \left(\frac{q^{-j} - q^{j+1}}{1 - q} \right) x_i^j \\
 &= \frac{1}{(1-q)^n} \prod_{i=1}^n \left[\frac{1}{1 - q^{-1}x_i} - \frac{q}{1 - qx_i} \right] \\
 &= \prod_{i=1}^n \frac{1 + x_i}{(1 - q^{-1}x_i)(1 - qx_i)}.
 \end{aligned}$$

We seek the term $F_0(x)$ independent from q . By the Cauchy integral formula (letting each x_i be small),

$$\begin{aligned}
 F_0(x) &= \frac{1}{2\pi i} \oint \frac{dq}{q} \prod_{i=1}^n \frac{1 + x_i}{(1 - q^{-1}x_i)(1 - qx_i)} \\
 &= \frac{(1 + x_1) \dots (1 + x_n)}{2\pi i} \oint dq \prod_{i=1}^n \frac{q^{n-1}}{(q - x_i)(1 - qx_i)},
 \end{aligned}$$

where the integral is around the circle $|q| = 1$. The integrand has a simple pole at $q = x_i$ with residue $x_i^{n-1}/(1-x_i^2) \prod_{j \neq i} (x_i - x_j)(1 - x_i x_j)$, and the proof follows from the Residue Theorem.

NOTE. The complex analysis in the above proof can be replaced with purely formal computations using the techniques of Section 6.3.

22. a. Let $a_1 + \cdots + a_k$ be any composition of $n > 1$. If $a_1 = 1$, then associate the composition $(a_1 + a_2) + a_3 + \cdots + a_k$. If $a_1 > 1$, then associate $1 + (a_1 - 1) + a_2 + \cdots + a_k$. This defines an involution on the set of compositions of n that changes the parity of the number of even parts. Hence, the number in question is 2^{n-2} , $n \geq 2$. (Note the analogy with permutations: There are $\frac{1}{2}n!$ permutations with an even number of even cycles—namely, the elements of the alternating group.)
- b. It is easily seen that

$$\sum_{n \geq 0} (e(n) - o(n))x^n = \prod_{i \geq 1} (1 + (-1)^i x^i)^{-1}.$$

In the first proof of Proposition 1.8.5, it was shown that

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

Hence (putting $-x$ for x and taking reciprocals),

$$\begin{aligned} \prod_{i \geq 1} (1 + (-1)^i x^i)^{-1} &= \prod_{i \geq 1} (1 + x^{2i+1}) \\ &= \sum_{n \geq 0} k(n)x^n, \end{aligned}$$

by Proposition 1.8.4. A simple combinatorial proof of this exercise was given by the Cambridge Combinatorics and Coffee Club in December 1999.

23. Form all 2^{n-1} compositions of n as in (1.19). Each bar occurs in half the compositions, so there are $(n-1)2^{n-2}$ bars in all. The total number of parts is equal to the total number of bars plus the total number of compositions, so $(n-1)2^{n-2} + 2^{n-1} = (n+1)2^{n-2}$ parts in all. This argument is due to D. E. Knuth (private communication, 21 August 2007).


Variant argument. Draw n dots in a row. Place a double bar before the first dot or in one of the $n-1$ spaces between the dots. Choose some subset of the remaining spaces between dots, and place a bar in each of these spaces. The double bar and the bars partition the dots into compartments that define a composition α of n as in equation (1.19). The compartment to the right of the double bar specifies one of the parts of α . Hence, the total number $f(n)$ of parts of all compositions of n is equal to the number of ways of choosing the double bar and bars as described previously. As an example, the figure


..|.||..|...

corresponds to the composition $(2, 1, 2, 3)$ of 8 with the third part selected.

If we place the double bar before the first dot, then there are 2^{n-1} choices for the remaining bars. Otherwise there are $n-1$ choices for the double bar and then 2^{n-2} choices for the remaining bars. Hence, $f(n) = 2^{n-1} + (n-1)2^{n-2} = (n+1)2^{n-2}$.

24. Draw a line of n dots and circle k consecutive dots. Put a vertical bar to the left and right of the circled dots. For example, $n = 9$, $k = 3$; see Figure 1.30.
- Case 1.* The circled dots don't include an endpoint. The procedure can then be done in $n - k - 1$ ways. Then there remain $n - k - 2$ spaces between uncircled dots. Insert at

• • • • • |  | • Figure 1.30 First step of the solution to Exercise 1.24.

• • • | • | • |  | • Figure 1.31 Continuation of the solution to Exercise 1.24.

most one vertical bar in each space in 2^{n-k-2} ways. This defines a composition with one part equal to k circled. For example, if we insert bars as in Figure 1.31, then we obtain $3 + 1 + 1 + \textcircled{3} + 1$.

Case 2. The circled dots include an endpoint. This happens in two ways, and now there are $n - k - 1$ spaces into which bars can be inserted in 2^{n-k-1} ways.

Hence, we get the answer

$$(n - k - 1)2^{n-k-2} + 2 \cdot 2^{n-k-1} = (n - k + 3)2^{n-k-2}.$$

25. It is clear that

$$\sum_{n,r,s} f(n,r,s) q^r t^s x^n = \sum_{j \geq 0} \left(\frac{qx}{1-x^2} + \frac{tx^2}{1-x^2} \right)^j.$$

The coefficient of $q^r t^s$ is given by

$$\binom{r+s}{r} \frac{x^{r+2s}}{(1-x^2)^{r+s}} = \binom{r+s}{r} \sum_{m \geq 0} \binom{m+r+s-1}{r+s-1} x^{2m+r+2s},$$

and the proof follows.

For a bijective proof, choose a composition of $r + k$ into $r + s$ parts in $\binom{r+k-1}{r+s-1}$ ways. Multiply r of these parts by 2 in $\binom{r+s}{r}$ ways. Multiply each of the other parts by 2 and subtract 1. We obtain each composition of n with r odd parts and s even parts exactly once, and the proof follows.

27. Answer: $(n + 3)2^{n-2} - 1$.

30. Let $b_i = a_i - i + 1$. Then $1 \leq b_1 \leq b_2 \leq \dots \leq b_k \leq n - k + 1$ and each b_i is odd. Conversely, given the b_i 's we can uniquely recover the a_i 's. Hence, setting $m = \lfloor (n - k + 2)/2 \rfloor$, the number of odd integers in the set $[n - k + 1]$, we obtain the answer $\binom{m}{k} = \binom{m+k-1}{k} = \binom{q}{k}$, where $q = \lfloor (n + k)/2 \rfloor$.

This exercise is called *Terquem's problem*. For some generalizations, see M. Abramson and W. O. J. Moser, *J. Combinatorial Theory* **7** (1969), 171–180; S. M. Tanny, *Canad. Math. Bull.* **18** (1975), 769–770; J. de Biasi, *C. R. Acad. Sci. Paris Sér. A-B* **285** (1977), A89–A92; and I. P. Goulden and D. M. Jackson, *Discrete Math.* **22** (1978), 99–104. A further generalization is given by Exercise 1.10.

31. a. $x(x+1)(x+2)\cdots(x+n-1) = n! \left(\binom{n+1}{x-1} \right) = n! \left(\binom{x}{n} \right)$

b. $(n)_x (n-1)_{n-x} = n! \binom{n-1}{x-1}$

c. $\sum_{k=1}^x \frac{n!}{k!} \binom{n-1}{k-1}$

32. The key feature of this problem is that each element of S can be treated *independently*, as in Example 1.1.16.
- a. For each $x \in S$, we may specify the least i (if any) for which $x \in T_i$. There are $k + 1$ choices for each x , so $(k + 1)^n$ ways in all.
 - b. Now each x can be in at most one T_i , so again there are $k + 1$ choices for x and $(k + 1)^n$ choices in all. (In fact, there is a very simple bijection between the sequences enumerated by (a) and (b).)
 - c. Now each x can be in any subset of the T_i 's except the subset \emptyset . Hence there are $2^k - 1$ choices for each x and $(2^k - 1)^n$ ways in all.
34. Let $b_i = a_i - (i - 1)j$ to get $1 \leq b_1 \leq \dots \leq b_k \leq n - (k - 1)j$, so the number of sequences is $\binom{n - (k - 1)j}{k}$.
35. a. Obtain a recurrence by considering those subsets S which do or do not contain n . *Answer:* F_{n+2} .
- b. Consider whether the first part is 2 or at least 3. *Answer:* F_{n-1} .
 - c. Consider whether the first part is 1 or 2. *Answer:* F_{n+1} .
 - d. Consider whether the first part is 1 or at least 3. *Answer:* F_n .
 - e. Consider whether $\varepsilon = 0$ or 1. *Answer:* F_{n+2} .
 - f. The following proof, as well as the proofs of (g) and (h), are due to Ira Gessel. Gessel (private communication, 2 May 2007) has developed a systematic approach to “Fibonacci composition formulas” based on factorization in free monoids as discussed in Section 4.7. The sum $\sum a_1 a_2 \dots a_k$ counts the number of ways of inserting at most one vertical bar in each of the $n - 1$ spaces separating a line of n dots, and then circling one dot in each compartment. An example is shown in Figure 1.32. Replace each bar by a 1, each uncircled dot by a 2, and each circled dot by a 1. For example, Figure 1.32 becomes

3 1 2 2 1 1 2 1 1 1 1 1 2 1 1 1 2 2 1 1 1 2 1 2.

We get a composition of $2n - 1$ into 1's and 2's, and this correspondence is invertible. Hence, by (c) the answer is F_{2n} . A simple generating function proof can also be given using the identity

$$\begin{aligned} \sum_{k \geq 1} (x + 2x^2 + 3x^3 + \dots)^k &= \frac{x/(1-x)^2}{1 - x/(1-x)^2} \\ &= \frac{x}{1 - 3x + x^2} \\ &= \sum_{n \geq 1} F_{2n} x^n. \end{aligned}$$

- g. Given a composition (a_1, \dots, a_k) of n , replace each part a_i with a composition α_i of $2a_i$ into parts 1 and 2, such that α_i begins with a 1, ends in a 2, and for all j the $2j$ th 1 in α is followed by a 1, unless this $2j$ th 1 is the last 1 in α . For instance, the part $a_i = 4$ can be replaced by any of the seven compositions 1111112, 111122, 111212, 11222, 121112, 12122, 12212. It can be checked that (i) every composition of $2n$ into parts 1 and 2, beginning with 1 and ending with 2, occurs exactly once by applying this procedure to all compositions of n , and (ii)



Figure 1.32 An illustration of the solution to Exercise 1.35(f).

the number of compositions that can replace a_i is $2^{a_i-1} - 1$. It follows from part (c) that the answer is F_{2n-2} . A generating function proof takes the form

$$\begin{aligned}\sum_{k \geq 1} (x^2 + 3x^3 + 7x^4 + \dots)^k &= \frac{x^2/(1-x)(1-2x)}{1-x^2/(1-x)(1-2x)} \\ &= \frac{x^2}{1-3x+x^2} \\ &= \sum_{n \geq 2} F_{2n-2} x^n.\end{aligned}$$

- h.** Given a composition (a_1, \dots, a_k) of n , replace each 1 with either 2 or 1, 1, and replace each $j > 1$ with $1, 2, \dots, 2, 1$, where there are $j-1$ 2's. Every composition of $2n$ with parts 1 and 2 is obtained in this way, so from part (c) we obtain the answer F_{2n+1} . A generating function proof takes the form

$$\begin{aligned}\frac{1}{1-2x-x^2-x^3-x^4-\dots} &= \frac{1}{1-x-\frac{x}{1-x}} \\ &= \frac{1-x}{1-3x+x^2} \\ &= \sum_{n \geq 0} F_{2n+1} x^n.\end{aligned}$$

- i.** Answer: $2F_{3n-4}$ (with F_n defined for all $n \in \mathbb{Z}$ using the recurrence $F_n = F_{n-1} + F_{n-2}$), a consequence of the expansion

$$\frac{1}{1 + \frac{x}{1-5x} + \frac{x}{1-x}} = 1 - 2 \sum_{n \geq 1} F_{3n-4} x^n.$$

A bijective proof is not known. This result is due to D. E. Knuth (private communication, August 21, 2007).

- k.** Answer: F_{2n+2} . Let $f(n)$ be the number in question. Now

$$P_n = P_{n-1} + P_{n-1}x_n + P_n x_{n+1}. \quad (1.153)$$

Each term of the above sum has $f(n-1)$ terms when expanded as a polynomial in the x_i 's. Since

$$P_{n-1} + P_{n-1}x_n = P_{n-2}(1+x_{n-1}+x_n) + P_{n-2}(1+x_{n-1}+x_n)x_n,$$

the only overlap between the three terms in equation (1.153) comes from $P_{n-2}x_n$, which has $f(n-2)$ terms. Hence $f(n) = 3f(n-1) - f(n-2)$, from which the proof follows easily. This problem was derived from a conjecture of T. Amdeberhan (November 2007). For a variant, see Exercise 4.20.

- 36.** Let $f_n(k)$ denote the answer. For each $i \in [n]$ we can decide which T_j contains i independently of the other $i' \in [n]$. Hence, $f_n(k) = f_k(1)^n$. But computing $f_k(1)$ is equivalent to Exercise 1.35(e). Hence, $f_n(k) = F_{k+2}^n$.
- 37.** While it is not difficult to show that the right-hand side of equation (1.122) satisfies the Fibonacci recurrence and initial conditions, we prefer a more combinatorial proof. For instance, Exercise 1.34 in the case $j = 2$ shows that $\binom{n-k}{k}$ is the number of k -subsets of $[n-1]$ containing no two consecutive integers. Now use Exercise 1.35(a).

- 39.** *First Solution* (sketch). Let $a_{m,n}$ be the number of ordered pairs (S, T) with $S \subseteq [m]$ and $T \subseteq [n]$ satisfying $s > \#T$ for all $s \in S$ and $t > \#S$ for all $t \in T$. An easy bijection gives

$$a_{m,n} = a_{m-1,n} + a_{m-1,n-1}.$$

Using $a_{ij} = a_{ji}$, we get

$$a_{n,n} = a_{n,n-1} + a_{n-1,n-1}$$

$$a_{n,n-1} = a_{n-1,n-1} + a_{n-1,n-2},$$

from which it follows (using the initial conditions $a_{0,0} = 1$ and $a_{1,0} = 2$) that $a_{n,n} = F_{2n+2}$ and $a_{n,n-1} = F_{2n+1}$.

Second Solution (sketch). It is easy to see that

$$a_{m,n} = \sum_{\substack{i,j \geq 0 \\ i+j \leq \min\{m,n\}}} \binom{m-j}{i} \binom{n-i}{j}.$$

It can then be proved bijectively that $\sum_{\substack{i,j \geq 0 \\ i+j \leq n}} \binom{n-j}{i} \binom{n-i}{j}$ is the number of compositions of $2n+1$ with parts 1 and 2. The proof follows from Exercise 1.35(c).

This problem (for the case $n = 10$) appeared as Problem A-6 on the Fifty-First William Lowell Putnam Mathematical Competition (1990). These two solutions appear in K. S. Kedlaya, B. Poonen, and R. Vakil, *The William Lowell Putnam Mathematical Competition*, Mathematical Association of America, Washington, D.C., 2002, pp. 123–124.

- 41.** **a.** Perhaps the most straightforward solution is to let $\#S = k$, giving

$$\begin{aligned} f(n) &= \sum_{k=0}^n (n-k)_k (n-k)! \binom{n}{k} \\ &= n! \sum_{k=0}^n \binom{n-k}{k}. \end{aligned}$$

Now use Exercise 1.37. It is considerably trickier to give a direct bijective proof.

- b.** We now have

$$\begin{aligned} g(n) &= \sum_{k=0}^{n-1} (n-k)_k (n-k-1)! \binom{n}{k} \\ &= (n-1)! \sum_{k=0}^{n-1} \frac{n}{n-k} \binom{n-k}{k}. \end{aligned}$$

There are a number ways to show that $L_n = \sum_{k=0}^{n-1} \frac{n}{n-k} \binom{n-k}{k}$, and the proof follows. This result was suggested by D. E. Knuth (private communication, 21 August 2007) upon seeing (a). A simple bijective proof was suggested by R. X. Du (private communication, 27 March 2011); namely, choose an n -cycle C in $(n-1)!$ ways, and regard the elements of C as n points on a circle. We can choose S to be any subset of the points, no two consecutive. By Exercise 1.40 this can be done in L_n ways, so the proof follows.

42. Let $\prod_{n \geq 2} (1 - x^{F_n}) = \sum_{k \geq 0} a_k x^k$. Split the interval $[F_n, F_{n+1} - 1]$ into the three subintervals $[F_n, F_n + F_{n-3} - 2]$, $[F_n + F_{n-3} - 1, F_n + F_{n-2} - 1]$, and $[F_n + F_{n-2}, F_{n+1} - 1]$. The following results can be shown by induction:
- The numbers $a_{F_n}, a_{F_{n+1}}, \dots, a_{F_n + F_{n-3} - 2}$ are equal to the numbers $(-1)^{n-1} a_{F_{n-3} - 2}, (-1)^{n-1} a_{F_{n-3} - 3}, \dots, (-1)^{n-1} a_0$ in that order.
 - The numbers $a_{F_n + F_{n-3} - 1}, a_{F_n + F_{n-3}}, \dots, a_{F_n + F_{n-2} - 1}$ are equal to 0.
 - The numbers $a_{F_n + F_{n-2}}, a_{F_n + F_{n-2} + 1}, \dots, a_{F_{n+1} - 1}$ are equal to the numbers $a_0, a_1, \dots, a_{F_{n-3} - 1}$ in that order.

From these results the proof follows by induction.

N. Robbins, *Fibonacci Quart.* **34** (1996), 306–313, was the first to prove that the coefficients are $0, \pm 1$. The above explicit recursive description of the coefficients is due to F. Ardila, *Fibonacci Quart.* **42** (2004), 202–204. Another elegant proof was later given by Y. Zhao, The coefficients of a truncated Fibonacci series, *Fibonacci Quart.*, to appear, and a significant generalization by H. Diao, arXiv:0802.1293.

43. Answer:

$$\begin{aligned} S(n, 1) &= 1 & c(n, 1) &= (n-1)! \\ S(n, 2) &= 2^{n-1} - 1 & c(n, 2) &= (n-1)! H_{n-1} \\ S(n, n) &= 1 & c(n, n) &= 1 \\ S(n, n-1) &= \binom{n}{2} & c(n, n-1) &= \binom{n}{2} \\ S(n, n-2) &= \binom{n}{3} + 3 \binom{n}{4} & c(n, n-2) &= 2 \binom{n}{3} + 3 \binom{n}{4}. \end{aligned}$$

An elegant method for computing $c(n, 2)$ is the following. Choose a permutation $a_1 a_2 \dots a_n \in \mathfrak{S}_n$ with $a_1 = 1$ in $(n-1)!$ ways. Choose $1 \leq j \leq n-1$ and let w be the permutation whose disjoint cycle form is $(a_1, a_2, \dots, a_j)(a_{j+1}, a_{j+2}, \dots, a_n)$. We obtain exactly j times every permutation with two cycles such that the cycle not containing 1 has length $n-j$. Hence, $c(n, 2) = (n-1)! H_{n-1}$.

As a further example, let us compute $S(n, n-2)$. The block sizes of a partition of $[n]$ with $n-2$ blocks are either 3 (once) and 1 ($n-3$ times), or 2 (twice) and 1 ($n-4$ times). In the first case, there are $\binom{n}{3}$ ways of choosing the 3-element block. In the second case, there are $\binom{n}{4}$ ways of choosing the union of the two 2-element blocks, and then three ways to choose the blocks themselves. Hence, $S(n, n-2) = \binom{n}{3} + 3 \binom{n}{4}$ as claimed.

45. Define $a_{i+1} + a_{i+2} + \dots + a_k$ to be the least r such that when $1, 2, \dots, r$ are removed from π , the resulting partition has i blocks.
46. a. We have by equation (1.94c) that

$$\begin{aligned} \sum_{n \geq 0} S(n, k) x^n &= \frac{x^k}{(1-x)(1-2x) \dots (1-kx)} \\ &= \frac{x^k}{(1-x)^{\lceil k/2 \rceil} (\text{mod } 2)}. \end{aligned}$$

- b. The first of several persons to find a combinatorial proof were K. L. Collins and M. Hovey, *Combinatorica* **31** (1991), 31–32. For further congruence properties of $S(n, k)$, see L. Carlitz, *Acta Arith.* **10** (1965), 409–422.
- c. Taking equation (1.28) modulo 2 gives

$$\sum_{k=0}^n c(n, k) t^k = t^{\lceil n/2 \rceil} (t+1)^{\lfloor n/2 \rfloor} \pmod{2}.$$

Hence,

$$c(n, k) \equiv \binom{\lfloor n/2 \rfloor}{k - \lceil n/2 \rceil} = \binom{\lfloor n/2 \rfloor}{n-k} \pmod{2}.$$

48. a. This remarkable result is due to J. N. Darroch, *Ann. Math. Stat.* **35** (1964), 1317–1321. For a nice exposition including much related work, see J. Pitman, *J. Combinatorial Theory Ser. A* **77** (1997), 279–303.
- b. Let $P(x) = \sum_{k=0}^n c(n, k) x^k$. It is routine to compute from Proposition 1.3.7 that

$$\frac{P'(1)}{P(1)} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n},$$

and the proof follows from (a). For further information on the distribution of the number of cycles of a permutations $w \in \mathfrak{S}_n$, see Pitman, *ibid.*, pp. 289–290.

- c. This result is due to E. R. Canfield and C. Pomerance, *Integers* **2** (2002), A1 (electronic); *Corrigendum* **5**(1) (2005), A9, improving earlier expressions for K_n due to Canfield and Menon (independently). Previously it was shown by L. H. Harper, *Ann. Math. Stat.* **38** (1966), 410–414 (Lemma 1), that the polynomial $\sum_k S(n, k) x^k$ has real zeros. As Pitman points out in his paper cited in (a) (page 291), the result (a) of Darroch reduces the problem of estimating K_n to estimating the expected number of blocks of a random partition of $[n]$. For further discussion, see D. E. Knuth, *The Art of Computer Programming*, vol. 4, Fascicle 3, Addison-Wesley, Upper Saddle River, N.J., 2005 (Exercises 7.2.1.5–62 and 7.2.1.5–63(e)).
49. a. Let $F_d(x) = A_d(x)/(1-x)^{d+1}$. Differentiate equation (1.37) and multiply by x , yielding

$$F_{d+1}(x) = x \frac{d}{dx} F_d(x),$$

and so on.

- b. The proof is by induction on d . Since $A_1(x) = x$, the assertion is true for $d = 1$. Assume the assertion for d . By Rolle's theorem, the function $f(x) = \frac{d}{dx}(1-x)^{-d-1} A_d(x)$ has $d-1$ simple negative real zeros that interlace the zeros of $A_d(x)$. Since $\lim_{x \rightarrow -\infty} f(x) = 0$, there is an additional zero of $f(x)$ less than the smallest zero of $A_d(x)$. Using equation (1.38), we have accounted for d strictly negative simple zeros of $A_{d+1}(x)$, and $x = 0$ is an additional zero. The proof follows by induction. This result can be extended to permutations of a multiset; see R. Simion, *J. Combinatorial Theory, Ser. A* **36** (1984), 15–22.
50. b. Let $D = d/dx$. By Rolle's theorem, $Q(x) = D^{i-1} P(x)$ has real zeros, and thus also $R(x) = x^{n-i+1} Q(1/x)$. Again by Rolle's theorem, $D^{n-i-1} R(x)$ has real zeros. But one computes easily that

$$D^{n-i-1} R(x) = \frac{n!}{2} (b_{i-1} x^2 + 2b_i x + b_{i+1}).$$

In order for this quadratic polynomial to have real zeros, we must have $b_i^2 \geq b_{i-1}b_{i+1}$. This result goes back to I. Newton (see e.g. G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, 2nd ed., Cambridge University Press, Cambridge, England, 1952, p. 52).

- c. Let us say that a polynomial $P(x) = \sum_{i=0}^m a_i x^i$ with coefficients satisfying $a_i = a_{m-i}$ has center $m/2$. (We don't assume that $\deg P(x) = m$, i.e., we may have $a_m = 0$.) Thus $P(x)$ has center $m/2$ if and only if $P(x) = x^m P(1/x)$. If also $Q(x) = x^n Q(1/x)$ (so $Q(x)$ has center $n/2$), then $P(x)Q(x) = x^{m+n} P(1/x)Q(1/x)$. Thus, $P(x)Q(x)$ has symmetric coefficients (with center $(m+n)/2$). It is also easy to show this simply by computing the coefficients of $P(x)Q(x)$ in terms of the coefficients of $P(x)$ and $Q(x)$.

Now assume that $P(x) = \sum_{i=0}^m a_i x^i$ has center $m/2$ and has unimodal coefficients, and similarly for $Q(x) = \sum_{i=0}^n b_i x^i$. Let $A_j(x) = x^j + x^{j+1} + \cdots + x^{m-j}$, a polynomial with center $m/2$, and similarly $B_j(x) = x^j + x^{j+1} + \cdots + x^{n-j}$. It is easy to see that

$$P(x) = \sum_{i=0}^{\lfloor m/2 \rfloor} (a_i - a_{i-1}) A_i(x),$$

$$Q(x) = \sum_{j=0}^{\lfloor n/2 \rfloor} (b_j - b_{j-1}) B_j(x).$$

Thus,

$$P(x)Q(x) = \sum_{i=0}^{\lfloor m/2 \rfloor} \sum_{j=0}^{\lfloor n/2 \rfloor} (a_i - a_{i-1})(b_j - b_{j-1}) A_i(x) B_j(x).$$

It is easy to check by explicit computation that $A_i(x)B_j(x)$ has unimodal coefficients and center $(m+n)/2$. Since $P(x)$ and $Q(x)$ have unimodal coefficients, we have

$$(a_i - a_{i-1})(b_j - b_{j-1}) \geq 0.$$

Hence, we have expressed $P(x)Q(x)$ as a nonnegative linear combination of unimodal polynomials, all with the same center $(m+n)/2$. It follows that $P(x)Q(x)$ is also unimodal (with center $(m+n)/2$).

- d. Perhaps the most elegant proof (and one suggesting some nice generalizations) uses linear algebra. Write $P(x) = \sum_{i=0}^m a_i x^i$ and $Q(x) = \sum_{i=0}^n b_i x^i$. Set $a_i = 0$ if $i \notin [0, m]$, and similarly for b_i . If X and Y are $r \times r$ real matrices all of whose $k \times k$ minors are nonnegative, then the Cauchy-Binet theorem shows that the same is true for the matrix XY . Moreover, it is easily seen that if c_0, c_1, \dots, c_n is nonnegative and log-concave with no internal zeros, then $c_i c_j \geq c_{i-s} c_{j+s}$ whenever $i \leq j$ and $s \geq 0$. Now take $k = 2$, $X = [a_{j-i}]_{i,j=0}^{m+n}$, and $Y = [b_{j-i}]_{i,j=0}^{m+n}$, and the proof follows.
- e. The symmetry of the two polynomials is easy to see in various ways. The polynomial $x \sum_{w \in \mathfrak{S}_n} x^{\text{des}(w)}$ is the Eulerian polynomial $A_n(x)$ by equation (1.36); now use (a), (b), and Exercise 1.49. The unimodality of the polynomial $\sum_{w \in \mathfrak{S}_n} x^{\text{inv}(w)}$ follows from (c) and the product formula (1.30). NOTE. A combinatorial proof of the unimodality of $\sum_{w \in \mathfrak{S}_n} x^{\text{inv}(w)}$ is implicit in the proof we have given, whereas a combinatorial proof of the log-concavity and unimodality of $A_n(x)$ is due to V. Gasharov, *J. Combinatorial Theory Ser. A* **82** (1998), 134–146 (§§4–5).
- f. This result was proved by F. De Mari and M. Shayman, *Acta Appl. Math.* **12** (1988), 213–235, using the hard Lefschetz theorem from algebraic geometry. It would be interesting to give a more elementary proof. A related result was proved by M. Bóna, Generalized descents and normality, arXiv:0709.4483.

- g. Let $n = 4$ and $S = \{(1, 2), (2, 3), (3, 4), (1, 4)\}$. Then

$$P_S(x) = x^4 + 8x^3 + 6x^2 + 8x + 1.$$

Note that part (f) asserts that $P_S(x)$ is unimodal for $S = \{(i, j) : 1 \leq i < j \leq n, j \leq i + p\}$. It seems likely (though this has not been checked) that the proof of De Mari and Shayman can be extended to the case $S = \{(i, j) : 1 \leq i < j \leq n, j \leq i + p_i\}$, where p_1, \dots, p_{n-1} are any nonnegative integers. Can anything further be said about those S for which $P_S(x)$ is unimodal?

For further information on the fascinating topic of unimodal and log-concave sequences, see R. Stanley, in *Graph Theory and Its Applications: East and West*, Ann. New York Acad. Sci., vol. 576, 1989, pp. 500–535, and the sequel by F. Brenti, in *Contemp. Math.* **178**, Amer. Math. Soc., Providence, R.I., 1994, pp. 71–89. For the unimodality of the q -binomial coefficient $\binom{n}{k}_q$ and related results, see Exercise 7.75.

51. This result goes back to P. S. de Laplace. The following proof is due to R. Stanley, in *Higher Combinatorics (Proc. NATO Advanced Study Inst., Berlin, 1976; M. Aigner, ed.)*, Reidel, Dordrecht/Boston, 1977, p. 49. Given $w \in \mathfrak{S}_n$, let S_w denote the region (a simplex) in \mathbb{R}^n defined by

$$0 \leq x_{w(1)} \leq x_{w(2)} \leq \dots \leq x_{w(n)} \leq 1.$$

Define $S_{nk} = \bigcup_w S_w$, where w ranges over all permutations in \mathfrak{S}_n with exactly $k - 1$ descents. It is easy to see that $\text{vol}(S_w) = 1/n!$, so $\text{vol}(S_{nk}) = A(n, k)/n!$. Define a map $\varphi : S_{nk} \rightarrow \mathcal{R}_{nk}$ by $\varphi(x_1, \dots, x_n) = (y_1, \dots, y_n)$, where

$$y_i = \begin{cases} x_{i+1} - x_i, & \text{if } x_i < x_{i+1}, \\ 1 + x_{i+1} - x_i, & \text{if } x_i > x_{i+1}. \end{cases}$$

Here we set $x_{n+1} = 1$, and we leave φ undefined on the set of measure zero consisting of points where some $x_{i-1} = x_i$. One can check that φ is measure-preserving and a bijection up to a set of measure zero. Hence, $\text{vol}(\mathcal{R}_{nk}) = \text{vol}(S_{nk}) = A(n, k)/n!$. For some additional proofs, see W. Meyer and R. von Randow, *Math. Annalen* **193** (1971), 315–321, S. M. Tanny, *Duke Math. J.* **40** (1973), 717–722; and J. W. Pitman, *J. Combinatorial Theory Ser. A* **77** (1997), 279–303 (pp. 295–296). For a refinement and further references, see R. Ehrenborg, M. A. Readdy, and E. Steingrímsson, *J. Combinatorial Theory Ser. A* **81** (1998), 121–126. For some related results, see Exercise 4.62.

52. This amusing result is due to J. Holte, *Amer. Math. Monthly* **104** (1997), 138–149. Holte derived this result in the setting of Markov chains and obtained many additional results about the combinatorics of carrying. Further work on this subject is due to P. Diaconis and J. Fulman, *Amer. Math. Monthly* **116** (2009), 788–803, and *Advances in Applied Math.* **43** (2009), 176–196, and A. Borodin, P. Diaconis, and J. Fulman, *Bull. Amer. Math. Soc.* **47** (2009), 639–670. There is a simple intuitive reason, which is not difficult to make rigorous, why we get the Eulerian numbers. The probability that we carry j in a certain column is roughly the probability that if i_1, \dots, i_n are random integers in the interval $[0, b - 1]$, then $bj \leq i_1 + \dots + i_n < b(j + 1)$. Now divide by b and use Exercise 1.51.
56. Let $\phi(w)$ denote the standardization (as defined in the second proof of Proposition 1.7.1) of $w \in \mathfrak{S}_M$. If $M = \{1^{m_1}, 2^{m_2}, \dots\}$ and $\#M = n$, then $\{\phi(w) : w \in \mathfrak{S}_M\}$ consists of all permutations $v \in \mathfrak{S}_n$ such that $D(v^{-1}) = \{m_1, m_1 + m_2, \dots\} \cap [n - 1]$. It is easy to see that $\text{inv}(w) = \text{inv}(v)$ (a special case of (1.71)) and $\text{maj}(w) = \text{maj}(v)$. The proof now follows from equation (1.43) and Theorem 1.4.8. This result is due to P. A. MacMahon, stated explicitly on page 317 of his paper [1.54]. Some other classes

of permutations that are equidistributed with respect to inv and maj are given by A. Björner and M. L. Wachs, *J. Combinatorial Theory Ser. A* **52** (1989), 165–187, and D. Foata and D. Zeilberger, *J. Comput. Applied Math.* **68** (1996), 79–101. See also the solution to Exercise 5.49(e).

57. Condition (i) does not hold if and only if there are indices $i < i'$ and $j < j'$ such that $(i, j) \in D(w)$, $(i', j') \in D(w)$, $(i, j') \notin D(w)$, $(i', j) \notin D(w)$. Let $w(i'') = j$ and $w(i''') = j'$. It is easy to check by drawing a diagram that $i < i'' < i' < i'''$ and $w(i'') < w(i) < w(i''') < w(i')$, so w is not 2143-avoiding. The steps are reversible, so (i) and (iii) are equivalent. The equivalence of (i) and (ii) follows from the fact that the j th term of $I(w)$ (respectively, $I(w^{-1})$) is the number of elements of $D(w)$ in column (respectively, row) j .

The permutations of this exercise are called *vexillary*. For further information on their history and properties, see Exercise 7.22(d,e) of Vol. II.

58. b. The final step in obtaining this result was achieved by Z. Stankova, *Europ. J. Combin.* **17** (1996), 501–517. For further information, see H. S. Wilf, *Discrete Math.* **257** (2002), 575–583, and M. Bóna [1.11, §4.4].
59. This result is known as the *Stanley-Wilf conjecture*. It was shown by R. Arratia, *Electron. J. Combin.* **6**(1) (1999), N1, that the conjecture follows from the statement that there is a real number $c > 1$ (depending on u) for which $s_u(n) < c^n$ for all $n \geq 1$. This statement was given a surprisingly simple and elegant proof by A. Marcus and G. Tardos, *J. Combinatorial Theory Ser. A* **107** (2004), 153–160. A nice exposition of this proof due to D. Zeilberger is available at

(www.math.rutgers.edu/~zeilberg/mamaring/mamaringhtml/paramath.html).

Another nice exposition is given by M. Bóna [1.11, §4.5].

60. *Answer:* The equivalence classes consist of permutations whose inverses have a fixed descent set. The number of equivalence classes is therefore 2^{n-1} , the number of subsets of $[n-1]$.
It is not difficult to prove this result directly but, it also can be understood in a nice way using the “Cartier-Foata theory” of Exercise 3.123.
61. a. By the properties of the bijection $w \mapsto T(w)$ discussed in Section 1.5, we have that

$$F(x; a, b, c, d) = \sum_{n \geq 1} \sum_T a^{\text{lr}(T)} b^{e(T)-1} c^{r(T)} d^{l(T)} \frac{x^n}{n!},$$

where T ranges over all increasing binary trees on the vertex set $[n]$, with $\text{lr}(T)$ vertices with two children, $e(T)$ vertices that are endpoints, $l(T)$ vertices with just a left child, and $r(T)$ vertices with just a right child. By removing the root from T , we obtain the equation

$$\frac{\partial}{\partial x}(F - bx) = abF^2 + (c + d)F. \quad (1.154)$$

Solving this equation (a Riccati equation, with a well-known method of solution) with the initial condition $F(0; a, b, c, d) = 0$ yields equation (1.124).

This result is due to L. Carlitz and R. Scoville, *J. reine angew. Math.* **265** (1974), 110–137 (§7). Our presentation follows Exercise 3.3.46 of I. P. Goulden and D. M. Jackson, *Combinatorial Enumeration*, John Wiley & Sons, New York, 1983; reprinted by Dover, Mineola, N.Y., 2004. This latter reference contains more details on solving the differential equation (1.154).

- b. The generating function is given by $1 + tF(x; 1, t, 1, 1)$, which can be simplified to the right-hand side of equation (1.125).

The enumeration of permutations by number of peaks was first considered by F. N. David and D. E. Barton, *Combinatorial Chance*, Hafner, New York, 1962 (pp. 162–164). They obtain a generating function for $r(n, k)$ written in a different form from equation (1.125).

- c. We have that $f(n)$ is the number of increasing binary trees on $[n]$ such that no vertex has only a left child except possibly the last vertex obtained by beginning with the root and taking right children. Let $g(n)$ be the number of increasing binary trees on $[n]$ such that no vertex has only a left child. Then

$$f(n+1) = \sum_{k=0}^n \binom{n}{k} f(k)g(n-k),$$

$$g(n+1) = \sum_{k=0}^{n-1} \binom{n}{k} g(k)g(n-k),$$

with $f(0) = g(0) = 1$. Setting $F(x) = \sum f(n)x^n/n!$ and $G(x) = \sum g(n)x^n/n!$, we obtain $F' = FG$ and $G + G' = G^2 + 1$. We can solve these differential equations to obtain equation (1.126). Goulden and Jackson, op. cit. (Exercise 5.2.17, attribution on page 306) attribute this result to P. Flajolet (private communication, 1982). The proof in Goulden and Jackson is based essentially on the Principle of Inclusion-Exclusion and is given here in Exercise 2.23.

62. a. First note that

$$p_k^n = \sum_{d|n} \sum_A d(w(A))^{nk/d}, \quad (1.155)$$

where A ranges over all aperiodic cycles of length d (i.e., cycles of length d that are unequal to a proper cyclic shift of themselves). Now substitute (1.155) into the expansion of $\log \prod (1 - p_k)^{-1}$ and simplify.

This result is implicit in the work of R. C. Lyndon (see Lothaire [4.31, Thm. 5.1.5]). See also N. G. de Bruijn and D. A. Klarner, *SIAM J. Alg. Disc. Meth.* **3** (1982), 359–368. The result was stated explicitly by I. M. Gessel (unpublished). A different theory of cycles of multiset permutations, due to D. Foata, has a nice exposition in §5.1.2 of D. E. Knuth [1.48]. In Foata's theory, a multiset permutation has the meaning of Section 1.7.

- b. Let $x_1 = \cdots = x_k = x$, and $x_j = 0$ if $j > k$.
- c. Let $\sigma = (a_1, a_2, \dots, a_{jk})$ be a multiset cycle of length jk , where k is the largest integer for which the word $u = a_1 a_2 \cdots a_{jk}$ has the form v^k for some word v of length j (where v^k denotes the concatenation of k copies of v). Let $\Gamma(\sigma) = p_k^j$. Given a multiset permutation $\pi = \sigma_1 \sigma_2 \cdots \sigma_m$ where each σ_i is a multiset cycle, define $\Gamma(\pi) = \Gamma(\sigma_1) \cdots \Gamma(\sigma_m)$. It can then be verified combinatorially that the number of multiset permutations π with fixed $w(\pi)$ and $\Gamma(\pi)$ is equal to the coefficient of $w(\pi)$ in $\Gamma(\pi)$, leading to the desired bijection.

63. Label the envelopes $1, 2, \dots, n$ in decreasing order of size. Partially order an arrangement of envelopes by inclusion, and adjoin a root labeled 0 at the top. We obtain an (unordered) increasing tree on $n+1$ vertices, and this correspondence is clearly invertible. Hence by Proposition 1.5.5, there are $n!$ arrangements in all, of which $c(n, k)$ have k envelopes not contained in another and $A(n, k)$ have k envelopes not containing another.
64. a. Let u be a sequence being counted, with m_i occurrences of i . Replace the 1's in u from right-to-left by $1, 2, \dots, m_1$. Then replace the 2's from right-to-left by

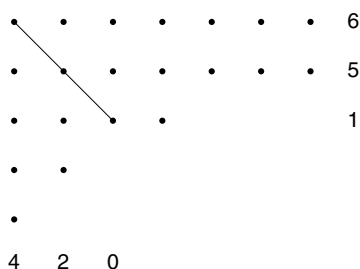


Figure 1.33 Frobenius notation.

$m_1 + 1, m_1 + 2, \dots, m_1 + m_2$, and so on. This procedure gives a bijection with \mathfrak{S}_n . For instance, 13213312 corresponds to 38527614. Note that this bijection could also be described as $u \mapsto \rho\psi\rho(u)$, where $\rho(v)$ is the reversal of v , and ψ denotes standardization (defined after the second proof of Proposition 1.7.1).

- b. The bijection in (a) has the property that $\max\{a_1, \dots, a_n\} = \text{des}(\rho(w)^{-1}) + 1$, and so on. This result was pointed out by D. E. Knuth (private communication, 21 August 2007) upon seeing (a).
65. It follows from a general theorem of Ramanujan (see D. Zagier, *The 1-2-3 of Modular Forms*, in (J. H. Bruinier, G. van der Geer, G. Harder and D. Zagier, eds.), Springer-Verlag, Berlin, 2008, Prop. 16, p. 49) that y satisfies a third order algebraic differential equation, but it is considerably more complicated than the fourth-degree equation (1.127). This equation was first computed by M. Rubey in 2010. See W. Heibisch and M. Rubey, *J. Symbolic Computation*, to appear.
70. a. Draw a line L along the main diagonal of the Ferrers diagram of λ . Then a_i is the number of dots in the i th row to the right of L , whereas b_i is the number of dots in the i th column below i . Figure 1.33 shows that $A_{77421} = \begin{pmatrix} 6 & 5 & 1 \\ 4 & 2 & 0 \end{pmatrix}$. This bijection is due to F. G. Frobenius, *Sitz. Preuss. Akad. Berlin* (1900), 516–534, and *Gesammelte Abh.* **3**, Springer, Berlin, 1969, pp. 148–166, and the array A_λ is called the *Frobenius notation* for λ .
- b. Suppose that the path P consists of c_1 steps N , followed by c_2 steps E , then c_3 steps S , and so on, ending in c_ℓ steps. If $\ell = 2r$, then associate with P the partition λ whose Frobenius notation is

$$A_\lambda = \begin{pmatrix} c_{\ell-1} & c_{\ell-3} & c_{\ell-5} & \cdots & c_1 \\ c_\ell - 1 & c_{\ell-2} - 1 & c_{\ell-4} - 1 & \cdots & c_2 - 1 \end{pmatrix}.$$

If $\ell = 2r - 1$ then associate with P the partition λ whose Frobenius notation is

$$A_\lambda = \begin{pmatrix} c_{\ell-1} & c_{\ell-3} & \cdots & c_2 & 0 \\ c_\ell - 1 & c_{\ell-2} - 1 & \cdots & c_3 - 1 & c_1 - 1 \end{pmatrix}.$$

This sets up the desired bijection. For instance, the CSSAW of Figure 1.34(a) corresponds to the partition $\lambda = (8, 6, 5, 2, 1)$ with $A_\lambda = \begin{pmatrix} 7 & 4 & 2 \\ 4 & 2 & 0 \end{pmatrix}$, while Figure 1.34(b) corresponds to $\lambda = (4, 3, 3, 3, 2, 1, 1)$ with $A_\lambda = \begin{pmatrix} 3 & 1 & 0 \\ 6 & 3 & 1 \end{pmatrix}$. This result is due to A. J. Guttmann and M. D. Hirschhorn, *J. Phys. A Math. Gen.* **17** (1984), 3613–3614. They give a combinatorial proof equivalent to the foregoing, though not stated in terms of Frobenius notation. The connection with Frobenius notation was given by G. E. Andrews, *Electronic J. Combinatorics* **18(2)** (2011), P6.

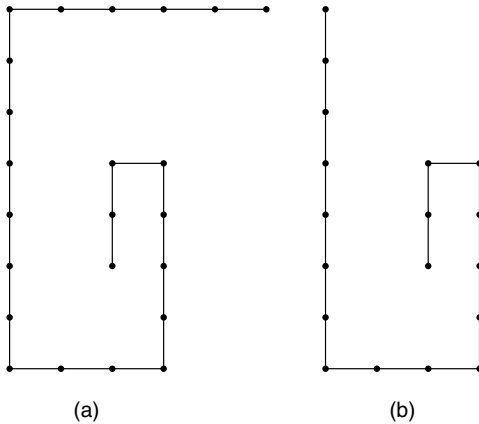
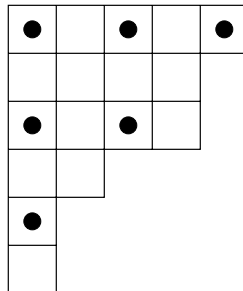


Figure 1.34 Two concatenated spiral self-avoiding walks.

71. *Answer.* $p(0) + p(1) + \cdots + p(n)$. Given $\nu \vdash k \leq n$, define λ to be ν with the part $n - k$ adjoined (in the correct position, so the parts remain weakly decreasing), and define μ to be ν with $n - k + 1$ adjoined. This yields the desired bijection. For some generalizations, see Theorem 3.21.11 and Exercise 3.150.
72. This exercise gives a glimpse of the fascinating subject of *plane partitions*, treated extensively in Sections 7.20–7.22 of Vol. II.
- a. Although equation (1.128) can be proved by ad hoc arguments, the “best” proof is a bijection using the RSK algorithm, the special case $q = 1$, $r = 2$, and $c \rightarrow \infty$ of Theorem 7.20.1 of Vol. II. A different generalization, but with a nonbijective proof, is given by Theorem 7.21.7 of Vol. II.
- b. This result is due to B. Gordon, *Proc. Amer. Math. Soc.* **13** (1962), 869–873. A bijective proof was given by C. Sudler, Jr., *Proc. Amer. Math. Soc.* **16** (1965), 161–168. This result can be generalized to a chain $\lambda^1 \subseteq \lambda^2 \subseteq \cdots \subseteq \lambda^k$ of any fixed number k of strict partitions, and with a fixed bound on the largest part of λ^k . See [7.146, Prop. 16.1] of Vol. II and G. E. Andrews, *Pacific J. Math.* **72** (1977), 283–291.
73. Consider for instance $\lambda = (5, 4, 4, 2, 1, 1)$, and put dots in the squares of the diagram of λ as follows:



Count the total number of dots by rows and by columns to obtain the first identity. The other formulas are analogous. There are many further variations.

74. Subtract one from each part of a partition of n into $n - t$ parts to deduce that $p_{n-t}(n) = p(t)$ if and only if $n \geq 2t$.

75. The partition $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$ corresponds to $\lambda_1 + k - 1 > \lambda_2 + k - 2 > \cdots > \lambda_k$.
76. By the bijection illustrated in Figure 1.16, the coefficient of $q^k x^n$ in the left-hand side of equation (1.129) is equal to the number of self-conjugate partitions λ of n whose rank is k . If we remove the Durfee square from the diagram of λ , then we obtain two partitions μ and μ' (the conjugate of μ) with largest part at most k . Hence we obtain the right-hand side of (1.129).
- One can also prove this identity by making the substitution $x \rightarrow x^2$ and $q \rightarrow qx^{-1}$ into equation (1.83).
77. Given $r \in \mathbb{Z}$, let λ be a partition satisfying $\lambda'_1 + r \geq \lambda_1 - 1$. Define $\psi_r(\lambda)$ to be the partition obtained by removing the first column of (the diagram of) λ and adding a new row at the top of length $\lambda'_1 + r$. We need to give a bijection

$$\gamma_n : \bigcup_{m \in 2\mathbb{Z}} \text{Par}(n - m(3m - 1)/2) \rightarrow \bigcup_{m \in 1 + 2\mathbb{Z}} \text{Par}(n - m(3m - 1)/2).$$

One can check that we can define γ_n as follows: for $\lambda \in \bigcup_{m \in 2\mathbb{Z}} \text{Par}(n - m(3m - 1)/2)$, let

$$\gamma_n(\lambda) = \begin{cases} \psi_{-3m-1}(\lambda), & \text{if } \lambda_1 - \lambda'_1 + 3m \leq 0, \\ \psi_{-3m+2}^{-1}(\lambda), & \text{if } \lambda_1 - \lambda'_1 + 3m \geq 0. \end{cases}$$

This proof appears in D. M. Bressoud and D. Zeilberger, *Amer. Math. Monthly* **92** (1985), 54–55. Our presentation follows Pak [1.62, §5.4.1].

78. a. Some related results are due to Euler and recounted in [1.55, §303].
- b. This problem was suggested by Dale Worley. For each $1 \leq i \leq n$, each partition λ of $n - i$, and each divisor d of i , we wish to associate a d -element multiset M of partitions of n so that every partition of n occurs exactly n times. Given i , λ , and d , simply associate d copies of the partition obtained by adjoining i/d d 's to λ .
79. a. See [1.2, Cor. 8.6].
- b. Clearly $p_S(n) = 1$ for all n , so the statement $q_S(n) = 1$ is just the uniqueness of the binary expansion of n .
80. For each partition λ of n and each part j of λ occurring at least k times, we need to associate a partition μ of n such that the total number of times a given μ occurs is the same as the number $f_k(\mu)$ of parts of μ that are equal to k . To do this, simply change k of the j 's in λ to j/k 's. For example, $n = 6$, $k = 2$:

λ	j	μ
1 1 1 1 1 1	1	2 1 1 1 1
2 1 1 1 1	1	2 2 1 1
3 1 1 1	1	3 2 1
4 1 1	1	4 2
2 2 1 1	2	2 2 1 1
2 2 1 1	1	2 2 2
2 2 2	2	2 2 2
3 3	3	2 2 2

This result was discovered by R. Stanley in 1972 and submitted to the Problems and Solutions section of *Amer. Math. Monthly*. It was rejected with the comment “A bit on the easy side, and using only a standard argument.” Daniel I. A. Cohen learned of this result and included the case $k = 1$ as Problem 75 of Chapter 3 in his book *Basic Techniques of Combinatorial Theory*, Wiley, New York, 1978. For this reason, the case $k = 1$ is sometimes called “Stanley’s theorem.” The generalization from $k = 1$ to arbitrary k was independently found by Paul Elder in 1984, as reported

by R. Honsberger, *Mathematical Gems III*, Mathematical Association of America, Washington, DC, 1985, p. 8. For this reason the general case is sometimes called “Elder’s theorem.” An independent proof of the general case was given by M. S. Kirdar and T. H. R. Skyrme, *Canad. J. Math.* **34** (1982), 194–195, based on generating functions. The bijection given here also appears in A. H. M. Hoare, *Amer. Math. Monthly* **93** (1986), 475–476. Another proof appears in L. Solomon, *Istituto Nazionale di Alta Matematica, Symposia Matematica*, vol. 13, Academic Press, London, 1974, pp. 453–466 (lemma on p. 461).

- 81.** Given an ordered factorization $n + 1 = a_1 a_2 \cdots a_k$, set $a_0 = 1$ and let λ be the partition for which the part $a_0 a_1 \cdots a_{j-1}$ occurs with multiplicity $a_j - 1$, $1 \leq j \leq k$. For instance, if $24 = 3 \cdot 2 \cdot 4$, then we obtain the partition 666311 of 23. This procedure sets up a bijection with perfect partitions of n , due to P. A. MacMahon, *Messenger Math.* **20** (1891), 103–119; reprinted in [1.3, pp. 771–787]. Note that if we have a perfect partition λ of n with largest part m , then there are exactly two ways to add a part p to λ to obtain another perfect partition, namely, $p = m$ or $p = n + 1$.
- 82.** This result is due to S. Ramanujan in 1919, who obtained the remarkable identity

$$\sum_{n \geq 0} p(5n + 4)x^n = 5 \frac{\prod_{k \geq 1} (1 - x^{5k})^5}{\prod_{k \geq 1} (1 - x^k)^6}.$$

F. J. Dyson conjectured in 1944 that for each $0 \leq i \leq 4$, exactly $p(5n + 4)/5$ partitions λ of $5n + 4$ satisfy $\lambda_1 - \lambda'_i \equiv i \pmod{5}$. This conjecture was proved by A. O. L. Atkin and H. P. F. Swinnerton-Dyer in 1953. Many generalizations of these results are known. For an introduction to the subject of partition congruences, see Andrews [1.2, Ch. 10]. For more recent work in this area, see K. Mahlburg, *Proc. Nat. Acad. Sci.* **102** (2005), 15373–15376.

- 83.** *Some Hints.* Let A be the set of all partitions λ such that $\lambda_{2i-1} - \lambda_{2i} \leq 1$ for all i , and let B be the set of all partitions λ such that λ' has only odd parts, each of which is repeated an even number of times. Verify the following statements.

- There is bijection $A \times B \rightarrow \text{Par}$ satisfying $w(\mu)w(v) = w(\lambda)$ if $(\mu, v) \mapsto \lambda$.
- We have

$$\sum_{\lambda \in B} w(\lambda) = \prod_{j \geq 1} \frac{1}{1 - a^j b^j c^{j-1} d^{j-1}}.$$

- Let $\lambda \in A$. Then the pairs $(\lambda_{2i-1}, \lambda_{2i})$ fall into two classes: (a, a) (which can occur any number of times), and $(a + 1, a)$ (which can occur at most once). Deduce that

$$\sum_{\lambda \in A} w(\lambda) = \prod_{j \geq 1} \frac{(1 + a^j b^{j-1} c^{j-1} d^{j-1})(1 + a^j b^j c^j d^{j-1})}{(1 - a^j b^j c^j d^j)(1 - a^j b^j c^{j-1} d^{j-1})}.$$

This elegant bijective proof is due to C. Boulet, *Ramanujan J.* **3** (2006), 315–320, simplifying and generalizing previous work of G. E. Andrews, A. V. Sills, R. P. Stanley, and A. J. Yee.

- 86.** This is *Schur’s partition theorem*. See G. E. Andrews, in *q-Series: Their Development and Application in Analysis, Number Theory, Combinatorics, Physics, and Computer Algebra*, American Mathematical Society, Providence, R.I., 1986, pp. 53–58. For a bijective proof, see D. M. Bressoud, *Proc. Amer. Math. Soc.* **79** (1980), 338–340. It is surprising that Schur’s partition theorem is easier to prove bijectively than the Rogers–Ramanujan identities (Exercise 1.88).
- 88.** **a.** These are the famous *Rogers–Ramanujan identities*, first proved by L. J. Rogers, *Proc. London Math. Soc.* **25** (1894), 318–343, and later rediscovered by I. Schur,

Sitzungsber. Preuss. Akad. Wiss. Phys.-Math. Klasse (1917), 302–321, S. Ramanujan (sometime before 1913, without proof), and others. For a noncombinatorial proof, see, for example, [1.2, §7.3]. For an exposition and discussion of bijective proofs, see Pak [1.62, §7 and pp. 62–63]. For an interesting recent bijective proof, see C. Boulet and I. Pak, *J. Combinatorial Theory Ser. A* **113** (2006), 1019–1030. None of the known bijective proofs of the Rogers–Ramanujan identities can be considered “simple,” comparable to the proof we have given of the pentagonal number formula (Proposition 1.8.7). An interesting reason for the impossibility of a nice proof was given by I. Pak, The nature of partition bijections II. Asymptotic stability, preprint.

- b. These combinatorial interpretations of the Rogers–Ramanujan identities are due to P. A. MacMahon, [1.55, §§276–280]. They can be proved similarly to the proof of Proposition 1.8.6, based on the observation that $(\lambda_1, \lambda_2, \dots, \lambda_k)$ is a partition of n with at most k parts if and only if $(\lambda_1 + 2k - 1, \lambda_2 + 2k - 3, \dots, \lambda_k + 1)$ is a partition of $n + k^2$ whose parts differ by at least two and with exactly k parts, and similarly for $(\lambda_1 + 2k, \lambda_2 + 2k - 2, \dots, \lambda_k + 2)$.

89. Let $\mu = (\mu_1, \dots, \mu_k)$ be a partition of n into k odd parts less than $2k$. We begin with the lecture hall partition $\lambda^0 = (0, \dots, 0)$ of length k and successively insert the parts $\mu_1, \mu_2, \dots, \mu_k$ to build up a sequence of lecture hall partitions $\lambda^1, \lambda^2, \dots, \lambda^k = \lambda$. The rule for inserting $\mu_i := 2v_i - 1$ into λ^{i-1} is the following. Add 1 to the parts of λ^{i-1} (allowing 0 as a part), beginning with the largest, until either (i) we have added 1 to μ_i parts of λ_{i-1} , or (ii) we encounter a value λ_{2c-1}^{i-1} for which

$$\frac{\lambda_{2c-1}^{i-1}}{n - 2c + 2} = \frac{\lambda_{2c}^{i-1}}{n - 2c + 1}.$$

In this case we add $v_i - c + 1$ to λ_{2c-1}^{i-1} and $v_i - c$ to λ_{2c}^{i-1} . It can then be checked that the map $\mu \mapsto \lambda$ gives the desired bijection.

Example. Let $k = 5$ and $\mu = (7, 5, 5, 3, 1)$. We have $\frac{\lambda_1^0}{5} = \frac{\lambda_2^0}{4} = 0$. Hence, $\lambda^1 = (4, 3, 0, 0, 0)$. We now have $\frac{\lambda_1^1}{5} \neq \frac{\lambda_2^1}{4}$, but $\frac{\lambda_3^1}{3} = \frac{\lambda_4^1}{2} = 0$. Hence, $\lambda^2 = (5, 4, 2, 1, 0)$. Continuing in this way, we get $\lambda^3 = (8, 6, 2, 1, 0)$, $\lambda^4 = (9, 7, 3, 1, 0)$, and $\lambda = \lambda^5 = (10, 7, 3, 1, 0)$.

Lecture hall partitions were introduced by M. Bousquet-Mélou and K. Eriksson, *Ramanujan J.* **1** (1997), 101–111, 165–185. They proved the result of this exercise as well as many generalizations and refinements. In our earlier sketch, we followed A. J. Yee, *Ramanujan J.* **5** (2001), 247–262. Her bijection is a simplified description of the bijection of Bousquet-Mélou and Eriksson. Much further work has been done in this area; see, for example, S. Corteel and C. D. Savage, *J. Combinatorial Theory Ser. A* **108** (2004), 217–245, for further information and references.

90. This curious result is connected with the theory of lecture hall partitions (Exercise 1.89). It was originally proved by M. Bousquet-Mélou and K. Eriksson, *Ramanujan J.* **1** (1997), 165–185 (end of Section 4). For a nice bijective proof of this result and related results, see C. D. Savage and A. J. Yee, *J. Combinatorial Theory Ser. A* **115** (2008), 967–996.
91. a. This famous result is the *Jacobi triple product identity*. It was first stated by C. F. Gauss (unpublished). The first published proof is due to C. G. J. Jacobi, *Fundamenta nova theoriae functionum ellipticarum*, Regiomonti, fratrum Bornträger, Paris, London, Amsterdam, St. Petersburg, 1829; reprinted in *Gesammelte Werke*, vol. 1, Reimer, Berlin, 1881, pp. 49–239. For a summary of its bijective proofs, see Pak [1.62, §6 and pp. 60–62].

- b. Substitute $q^{3/2}$ for q and $-q^{1/2}$ for x , and simplify.
- c. For the first, set $x = -1$ and use equation (1.81). For the second, substitute $q^{1/2}$ for both x and q . The right-hand side then has a factor equal to 2. Divide both sides by 2, and again use equation (1.81). These identities are due to Gauss, *Zur Theorie der neuen Transscendenten II, Werke*, Band III, Göttingen, 1866, pp. 436–445 (§4). For a cancellation proof, see Exercise 2.31. For another proof of equation (1.132) based on counting partitions of n with empty 2-core, see Exercise 7.59(g) of Vol. II.
- d. After making the suggested substitution, we obtain

$$\sum_{n \in \mathbb{Z}} (-1)^n x^n q^{\binom{n}{2}} = \prod_{k \geq 1} (1 - q^k)(1 - xq^{k-1})(1 - x^{-1}q^k).$$

Rewrite the left-hand side as

$$1 + \sum_{n \geq 1} (-1)^n (x^{-n} + x^n) q^{\binom{n}{2}}.$$

Now divide both sides by $1 - x$ and let $x \rightarrow 1$. The left-hand side becomes $\sum_{n \geq 0} (-1)^n (2n + 1) q^{\binom{n}{2}}$. The right-hand side has a factor equal to $1 - x$, so deleting this factor and then setting $x = 1$ gives

$$(1 - q)^2 \prod_{k \geq 2} (1 - q^{k-1})(1 - q^k)^2 = \prod_{k \geq 1} (1 - q^k)^3,$$

and the proof follows. This identity is due to C. G. J. Jacobi, *Fundamenta Nova Theoriae Functionum Ellipticarum*, Regiomonti, Sumtibus fratrum Borntraeger, Königsberg, Germany, 1829, p. 90.

92. This identity is due to G. E. Andrews, *Amer. Math. Monthly* **94** (1987), 437–439. A simple proof based on the Jacobi triple product identity (Exercise 1.91) is due to F. G. Garvan, in *Number Theory for the Millenium, II* (Urbana, IL, 2000), A K Peters, Natick, MA, 2002, pp. 75–92 (§1). This paper contains many further similar identities. For a continuation, see F. G. Garvan and H. Yesilyurt, *Int. J. Number Theory* **3** (2007), 1–42. No bijective proofs are known of any of these identities.
93. This identity is due to F. G. Garvan, op. cit. This paper and the continuation by Garvan and Yesilyurt, op. cit., contain many similar identities. No bijective proofs are known of any of them.
94. The sequence a_1, a_2, \dots (sometimes prepended with $a_0 = 0$) is called *Stern's diatomic sequence*, after the paper by M. A. Stern, *J. Reine angew. Math.* **55** (1858), 193–220. For a survey of its remarkable properties, see S. Northshield, *Amer. Math. Monthly* **117** (2010), 581–598.
95. This remarkable result is due to D. Applegate, O. E. Pol, and N. J. A. Sloane, *Congressus Numerantium* **206** (2010), 157–191.
96. a. The function $\tau(n)$ is *Ramanujan's tau function*. The function

$$\Delta(t) = (2\pi)^{12} \sum_{n \geq 1} \tau(n) e^{2\pi i t n}$$

plays an important role in the theory of modular forms; see, for example, T. Apostol, *Modular Forms and Dirichlet Series in Number Theory* 2nd ed., Springer-Verlag,

New York, 1997, p. 20, or J.-P. Serre, *A Course in Arithmetic*, Springer-Verlag, New York, 1973, §VII.4. The multiplicativity property of this exercise was conjectured by S. Ramanujan, *Trans. Cambridge Phil. Soc.* **22** (1916), 159–184, and proved by L. J. Mordell, *Proc. Cambridge Phil. Soc.* **19** (1917), 117–124.

- b. This result was also conjectured by Ramanujan, op. cit., and proved by Mordell, op. cit.
 - c. This inequality was conjectured by Ramanujan, op. cit., and proved by P. R. Deligne, *Inst. Hautes Études Sci. Publ. Math.* **43** (1974), 273–307; **52** (1980), 137–252. Deligne deduced Ramanujan's conjecture (in a nontrivial way) from his proof of the Riemann hypothesis for varieties over finite fields (the most difficult part of the “Weil conjectures”). Deligne in fact proved a conjecture of Petersson generalizing Ramanujan's conjecture.
 - d. This inequality was conjectured by D. H. Lehmer, *Duke Math. J.* **14** (1947), 429–492. It is known to be true for (at least) $n < 2.2 \times 10^{16}$.
97. This result follows from the case $p = 2$ and $\mu = \emptyset$ of Exercise 7.59(e). Greta Panova (October 2007) observed that it can also be deduced from Exercise 1.83. Namely, first prove by induction that the Ferrers diagram of λ can be covered by edges if and only if the Young diagram of λ has the same number of white squares as black squares in the usual chessboard coloring. Thus, $f(n)$ is the coefficient of q^n in the right-hand side of equation (1.130) after substituting $a = d = q/y$ and $b = c = y$. Apply the Jacobi triple product identity (Exercise 1.91) to the numerator and then set $y = 0$ to get $\sum_{n \geq 0} f(n)q^n = 1/\prod_{j \geq 1} (1 - q^j)^2$.
98. Substitute na for j , $-x$ for x , and ζ for q in the q -binomial theorem (equation (1.87)). The proof follows straightforwardly from the identity

$$\prod_{m=0}^{na-1} (1 - \zeta^m x) = (1 - x^n)^a.$$

For a host of generalizations, see V. Reiner, D. Stanton, and D. White, *J. Combinatorial Theory Ser. A* **108** (2004), 17–50.

99. It is an immediate consequence of the identity $f(q) = q^{k(n-k)} f(1/q)$ that

$$f'(1) = \frac{1}{2}k(n-k)f(1) = \frac{1}{2}k(n-k) \binom{n}{k}.$$

100. The Chu–Vandermonde identity follows from $(1+x)^{a+b} = (1+x)^a(1+x)^b$. Write $f_n(x) = (1+x)(1+qx) \cdots (1+q^{n-1}x)$. The q -analogue of $(1+x)^{a+b} = (1+x)^a(1+x)^b$ is $f_{a+b}(x) = f_b(x)f_a(q^b x)$. By the q -binomial theorem (equation (1.87)) we get

$$\sum_{n=0}^{a+b} q^{\binom{n}{2}} \binom{a+b}{n} x^n = \left(\sum_{k=0}^b q^{\binom{k}{2}} \binom{b}{k} x^k \right) \left(\sum_{k=0}^a q^{b k + \binom{k}{2}} \binom{a}{k} x^k \right).$$

Equating coefficients of x^n yields

$$\begin{aligned} q^{\binom{n}{2}} \binom{a+b}{n} &= \sum_{k=0}^n q^{\binom{n-k}{2} + b k + \binom{k}{2}} \binom{b}{n-k} \binom{a}{k} \\ &\Rightarrow \binom{a+b}{n} = \sum_{k=0}^n q^{k(k+b-n)} \binom{a}{k} \binom{b}{n-k}. \end{aligned}$$

103. See Lemma 3.1 of K. Liu, C. H. F. Yan, and J. Zhou, *Sci. China, Ser. A* **45** (2002), 420–431, for a proof based on the Hilbert scheme of n points in the plane. A combinatorial proof of a continuous family of results including this exercise appears in N. Loehr and G. S. Warrington, *J. Combinatorial Theory Ser. A* **116** (2009), 379–403.

104. Let $f(x) = 1 + x + \cdots + x^9$ and $i^2 = 1$. It is not hard to see that

$$\begin{aligned} f(n) &= \frac{1}{4} (f(1)^n + f(i)^n + f(-1)^n + f(-i)^n) \\ &= \frac{1}{4} (10^n + (1+i)^n + (1-i)^n) \\ &= \begin{cases} \frac{1}{4} (10^n + (-1)^k 2^{2k-1}), & n = 4k, \\ \frac{1}{4} (10^n + (-1)^k 2^{2k-1}), & n = 4k+1, \\ \frac{1}{4} 10^n, & n = 4k+2, \\ \frac{1}{4} (10^n + (-1)^{k+1} 2^{2k}), & n = 4k+3. \end{cases} \end{aligned}$$

105. a. Let $P(x) = (1+x)(1+x^2)\cdots(1+x^n) = \sum_{k \geq 0} a_k x^k$. Let $\zeta = e^{2\pi i/n}$ (or any primitive n th root of unity). Since for any integer k ,

$$\sum_{j=1}^n \zeta^{kj} = \begin{cases} n, & \text{if } n|k, \\ 0, & \text{otherwise,} \end{cases}$$

we have

$$\frac{1}{n} \sum_{j=1}^n P(\zeta^j) = \sum_j a_{jn} = f(n).$$

Now if ζ^j is a primitive d th root of unity (so $d = n/(j, n)$), then

$$x^d - 1 = (x - \zeta^j)(x - \zeta^{2j}) \cdots (x - \zeta^{dj}),$$

so putting $x = -1$ yields

$$(1 + \zeta^j)(1 + \zeta^{2j}) \cdots (1 + \zeta^{dj}) = \begin{cases} 2, & d \text{ odd,} \\ 0, & d \text{ even.} \end{cases}$$

Hence,

$$P(\zeta^j) = \begin{cases} 2^{n/d}, & d \text{ odd,} \\ 0, & d \text{ even.} \end{cases}$$

Since there are $\phi(d)$ values of $j \in [n]$ for which ζ^j is a primitive d th root of unity, we obtain

$$f(n) = \frac{1}{n} \sum_{j=1}^n P(\zeta^j) = \frac{1}{n} \sum_{\substack{d|n \\ d \text{ odd}}} \phi(d) 2^{n/d}.$$

This result appears in R. Stanley and M. F. Yoder, *JPL Technical Report 32-1526*, Deep Space Network **14** (1972), 117–123.

b. Suppose that n is an odd prime. Identify the beads of a necklace with $\mathbb{Z}/n\mathbb{Z}$ in an obvious way. Let $S \subseteq \mathbb{Z}/n\mathbb{Z}$ be the set of black beads. If $S \neq \emptyset$ and $S \neq \mathbb{Z}/n\mathbb{Z}$, then there is a unique $a \in \mathbb{Z}/n\mathbb{Z}$ for which

$$\sum_{x \in S} (x + a) = 0.$$

The set $\{x + a : x \in S\}$ represents the same necklace (up to cyclic symmetry), so we have associated with each nonmonochromatic necklace a subset of $\mathbb{Z}/n\mathbb{Z}$ whose elements sum to 0. Associate with the necklaces of all black beads and all white beads the subsets $S = \emptyset$ and $S = \mathbb{Z}/n\mathbb{Z}$, and we have the desired bijection.

A proof for any odd n avoiding roots of unity and generating functions was given by Anders Kaseorg (private communication) in 2004, though the proof is not a direct bijection.

c. See A. M. Odlyzko and R. Stanley, *J. Number Theory* **10** (1978), 263–272.

- 106.** We claim that $f(n, k)$ is just the Stirling number $S(n, k)$ of the second kind. We need to associate with a sequence $a_1 \cdots a_n$ being counted a partition of $[n]$ into k blocks. Simply put i and j in the same block when $a_i = a_j$. This yields the desired bijection. The sequences $a_1 \cdots a_n$ are called *restricted growth functions* or *restricted growth strings* (sometimes with 1 subtracted from each term). For further information, see S. Milne, *Advances in Math.* **26** (1977), 290–305.
- 108. a.** Given a partition π of $[n - 1]$, let $i, i + 1, \dots, j$ for $j > i$, be a maximal sequence of two or more consecutive integers contained in a block of π . Remove $j - 1, j - 3, j - 5, \dots$ from this sequence and put them in a block with n . Doing this for every such sequence $i, i + 1, \dots, j$ yields the desired bijection. See H. Prodinger, *Fibonacci Quart.* **19** (1981), 463–465, W. Y. C. Chen, E. Y. P. Deng, and R. R. X. Du, *Europ. J. Combin.* **26** (2005), 237–243, and W. Yang, *Discrete Math.* **156** (1996), 247–252.

Example. If $\pi = 1456-2378$, then the bijection gives 146-38-2579.

The preceding proof easily extends (as done in papers cited here) to show the following result: let $0 \leq k \leq n$, and let $B_k(n)$ be the number of partitions of $[n]$ so that if i and j are in a block then $|i - j| > k$. Then $B_k(n) = B(n - k)$.

- 109. a.** Given a partition $\pi \in \Pi_n$, list the blocks in decreasing order of their smallest element. Then list the elements of each block with the least element first, followed by the remaining elements in decreasing order, obtaining a permutation $w \in \mathfrak{S}_n$. The map $\pi \mapsto w$ is bijection from Π_n to the permutations being enumerated. For instance, if $\pi = 13569 - 248 - 7$, then $w = 728419653$. To obtain π from w , break w before each left-to-right minimum. This result, as well as those in (b) and (c), is due to A. Claesson, *Europ. J. Combinatorics* **22** (2001), 961–971.
- b.** Now write the blocks in decreasing order of their smallest element, with the elements of each block written in increasing order.
- c.** Let w be the permutation corresponding to π as defined in (a). Then w also satisfies the condition of (b) if and only if each block of π has size one or two.
- 110.** *Answer:* The coefficient of x^n is $B(n - 1)$, $n \geq 1$. See Proposition 2.6 of M. Klazar, *J. Combinatorial Theory Ser. A* **102** (2003), 63–87.
- 111.** The number of ways to partition a k -element subset of $[n]$ into j intervals is $\binom{k-1}{j-1} \binom{n-k+j}{j}$, since we can choose the interval sizes from left-to-right in $\binom{k-1}{j-1}$ ways (the number of compositions of k into j parts), and then choose the intervals themselves in $\binom{n-k+j}{j}$ ways. Hence by the Principle of Inclusion-Exclusion (Theorem 2.1.1),

$$f(n) = B(n) + \sum_{k=1}^n \sum_{j=1}^k B(n-k)(-1)^j \binom{k-1}{j-1} \binom{n-k+j}{j}.$$

Now

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

Hence

$$\begin{aligned} f(n) &= \sum_{k=0}^n B(n-k) (-1)^k \binom{n-k+1}{k} \\ &= \sum_{k=0}^n B(k) (-1)^{n-k} \binom{k+1}{n-k}. \end{aligned}$$

(Is there some way to see this directly from Inclusion-Exclusion?) Now multiply by x^n and sum on $n \geq 0$. Since by the binomial theorem

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

we get

$$\begin{aligned} F(x) &= \sum_{k \geq 0} B(k) x^k (1-x)^{k+1} \\ &= (1-x)G(x(1-x)). \end{aligned}$$

115. See D. Chebikin, R. Ehrenborg, P. Pylyavskyy, and M. A. Readdy, *J. Combinatorial Theory Ser. A* **116** (2009), 247–264. The polynomials $Q_n(t)$ are introduced in this paper and are shown to have many cyclotomic factors, but many additional such factors are not yet understood.

116. b. See L. A. Shepp and S. P. Lloyd, *Trans. Amer. Math. Soc.* **121** (1966), 340–357.

117. Answer: $p_{nk} = 1/n$ for $1 \leq k \leq n$. To see this, consider the permutations $v = b_1 \cdots b_{n+1}$ of $[n] \cup \{*\}$ beginning with 1. Put the elements to the left of $*$ in a cycle in the order they occur. Regard the elements to the right of $*$ as a word that defines a permutation of its elements (say with respect to the elements listed in increasing order). This defines a bijection between the permutations v and the permutations $w \in \mathfrak{S}_n$. The length of the cycle containing 1 is k if $b_{k+1} = *$. Since $*$ is equally likely to be any of b_2, \dots, b_{n+1} , the proof follows.

Example. Let $v = 1652 * 4873$. Then w has the cycle $(1, 6, 5, 2)$. The remaining elements are permuted as 4873 with respect to the increasing order 3478 (i.e., $w(3) = 4$, $w(4) = 8$, $w(7) = 7$, and $w(8) = 3$). In cycle form, we have $w = (1, 6, 5, 2)(3, 4, 8)(7)$.

118. b. We compute equivalently the probability that $n, n-1, \dots, n-\lambda_1+1$ are in the same cycle C_1 , and $n-\lambda_1, \dots, n-\lambda_1-\lambda_2+1$ are in the same cycle C_2 different from C_1 , and so on. Apply the fundamental bijection of Proposition 1.3.1 to w , obtaining a permutation $v = b_1 \cdots b_n$. It is easy to check that w has the desired properties if and only if the restriction u of v to $n-k+1, n-k+2, \dots, n$ has $n-k+\lambda_\ell$ appearing first, then the elements $n-k+1, n-k+2, \dots, n-k+\lambda_\ell-1$ in some order, then $n-k+\lambda_{\ell-1}+\lambda_\ell$, then the elements $n-k+\lambda_\ell+$

$1, \dots, n - k + \lambda_{\ell-1} + \lambda_{\ell} + 1$ in some order, then $n - k + \lambda_{\ell-2} + \lambda_{\ell-1} + \lambda_{\ell}$, and so on. Hence, of the $k!$ permutations of $n - k + 1, \dots, n$ there are $(\lambda_1 - 1)! \cdots (\lambda_{\ell} - 1)!$ choices for u , and the proof follows. For a variant of this problem when the distribution isn't uniform, see O. Bernardi, R. X. Du, A. Morales, and R. Stanley, in preparation.

- c. Let v be as in (b), and let $v' = b_2 b_1 b_3 b_4 \cdots b_n$. Exactly one of v and v' is even. Moreover, the condition in (b) on the restriction u is unaffected unless $b_1 = n - k + \lambda_{\ell}$ and $b_2 = n - k + i$ for some $1 \leq i \leq \lambda_{\ell} - 1$. In this case, v has exactly ℓ records, so w has exactly ℓ cycles. Hence, w is even if and only if $n - \ell$ is even. Moreover, the number of choices for u is

$$\frac{(n-2)!}{(k-2)!} (\lambda_1 - 1)! \cdots (\lambda_{\ell} - 1)!,$$

and the proof follows easily.

119. If a permutation $w \in \mathfrak{S}_{2n}$ has a cycle C of length $k > n$, then it has exactly one such cycle. There are $\binom{2n}{k}$ ways to choose the elements of C , then $(k-1)!$ ways to choose C , and finally $(2n-k)!$ ways to choose the remainder of w . Hence,

$$\begin{aligned} P_n &= 1 - \frac{1}{(2n)!} \sum_{k=n+1}^{2n} \binom{2n}{k} (k-1)! (2n-k)! \\ &= 1 - \sum_{k=n+1}^{2n} \frac{1}{k} \\ &= 1 - \sum_{k=1}^{2n} \frac{1}{k} + \sum_{k=1}^n \frac{1}{k} \\ &\sim 1 - \log(2n) + \log(n) \\ &= 1 - \log 2, \end{aligned}$$

and the proof follows. For an amusing application of this result, see P. M. Winkler, *Mathematical Mind-Benders*, A K Peters, Wellesley, MA, 2007 (pp. 12, 18–20).

120. *First Solution.* There are $\binom{n}{k} (k-1)!$ k -cycles, and each occurs in $(n-k)!$ permutations $w \in \mathfrak{S}_n$. Hence,

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

Second Solution. By Exercise 1.117 (for which we gave a simple bijective proof), the probability that some element $i \in [n]$ is in a k -cycle is $1/n$. Since there are n elements and each k -cycle contains k of them, the expected number of k -cycles is $(1/n)(n/k) = 1/k$.

124. a. Let $w = a_1 a_2 \cdots a_{n+1} \in \mathfrak{S}_{n+1}$ have k inversions, where $n \geq k$. There are $f_k(n)$ such w with $a_{n+1} = n+1$. If $a_i = n+1$ with $i < n+1$, then we can interchange a_i and a_{i+1} to form a permutation $w' \in \mathfrak{S}_{n+1}$ with $k-1$ inversions. Since $n \geq k$, every $w' = b_1 b_2 \cdots b_{n+1} \in \mathfrak{S}_{n+1}$ with $k-1$ inversions satisfies $b_1 \neq n+1$ and thus can be obtained from a $w \in \mathfrak{S}_{n+1}$ with k inversions as previously.
- b. Use induction on k .

c. By Corollary 1.3.13 we have

$$\begin{aligned}\sum_{k \geq 0} f_k(n) q^k &= (1+q)(1+q+q^2) \cdots (1+q+\cdots+q^{n-1}) \\ &= \frac{(1-q)(1-q^2) \cdots (1-q^n)}{(1-q)^n} \\ &= (1-q)(1-q^2) \cdots (1-q^n) \sum_{k \geq 0} \binom{-n}{k} (-1)^k q^k.\end{aligned}$$

Hence, if $\prod_{i \geq 1} (1 - q^i) = \sum_{j \geq 0} b_j q^j$, then

$$f_k(n) = \sum_{j=0}^k (-1)^j b_{k-j}, \quad n \geq k.$$

Moreover, it follows from the Pentagonal Number Formula (1.88) that

$$b_r = \begin{cases} (-1)^i, & \text{if } r = i(3i \pm 1)/2, \\ 0, & \text{otherwise.} \end{cases}$$

See pp. 15–16 of D. E. Knuth [1.48].

- 127. a.** We can reason analogously to the proofs of Proposition 1.3.12 and Corollary 1.3.13. Given $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$ and $1 \leq i \leq n$, define

$$r_i = \#\{j : j < i, w_j > w_i\}$$

and $\text{code}'(w) = (r_1, \dots, r_n)$. For instance, $\text{code}'(3265174) = (0, 1, 0, 1, 4, 0, 3)$. Note that $\text{code}'(w)$ is just a variant of $\text{code}(w)$ and gives a bijection from \mathfrak{S}_n to sequences (r_1, \dots, r_n) satisfying $0 \leq r_i \leq i - 1$. Moreover, $\text{inv}(w) = \sum r_i$, and w_i is a left-to-right maximum if and only if $r_i = 0$. From these observations, equation (1.135) is immediate.

- b.** Let $I(w) = (a_1, \dots, a_n)$, the inversion table of w . Then $\text{inv}(w) = \sum a_i$ (as noted in the proof of Corollary 1.3.13), and i is the value of a record if and only if $a_i = 0$. From these observations, equation (1.136) is immediate.

- 128. a.** First establish the recurrence

$$\sum_{j=1}^n f(j)(n-j)! = n!, \quad n \geq 1,$$

where we set $g(0) = 1$. Then multiply by x^n and sum on $n \geq 0$. This result appears in L. Comtet, *Comptes Rend. Acad. Sci. Paris* **A275** (1972), 569–572, and is also considered by Comtet in his book *Advanced Combinatorics*, Reidel, Dordrecht/Boston, 1974 (Exercise VII.16). For an extension of this exercise and further references, see Exercise 2.13.

- b.** (I. M. Gessel) Now we have

$$n! = g(n) + \sum_{j=1}^n g(j-1)(n-j)!, \quad n \geq 1,$$

where we set $g(0) = 1$.

- c.** See D. Callan, *J. Integer Sequences* **7** (2004), article 04.1.8.

- d. See M. H. Albert, M. D. Atkinson, and M. Klazar, *J. Integer Sequences* **6** (2003), article 02.4.4. For a survey of simple permutations, see R. Brignall, in *Permutation Patterns* (S. Linton, N. Ruškuc and V. Vatter, eds.), London Mathematical Society Lecture Note Series, vol. 376, Cambridge University Press, Cambridge, 2010, pp. 41–65. For some analogous results for set partitions, see M. Klazar, *J. Combinatorial Theory Ser. A* **102** (2003), 63–87.
129. a. It is easy to see that if w is an indecomposable permutation in \mathfrak{S}_n with k inversions, then $n \leq k + 1$. (Moreover, there are exactly 2^{k-1} indecomposable permutations in \mathfrak{S}_{k+1} with k inversions.) Hence, $g_n(q)$ has smallest term of degree $n - 1$, and the proof follows.
- b. *Answer.* We have the continued fraction

$$1 - \frac{1}{F(q, x)} = \frac{a_0}{1 - \frac{a_1}{1 - \frac{a_2}{1 - \dots}}},$$

where

$$a_n = (q^{\lfloor (n+1)/2 \rfloor} + q^{\lfloor (n+1)/2 \rfloor + 1} + \dots + q^n)x.$$

See A. de Medicis and X. G. Viennot, *Advances in Appl. Math.* **15** (1994), 262–304 (equations (1.24) and (1.25) and Theorem 5.3).

130. This result, stated in a less elegant form, is due to M. Abramson and W. O. J. Moser, *Ann. Math. Statist.* **38** (1967), 1245–1254. The solution in the form of equation (1.138) is due to L. W. Shapiro and A. B. Stephens, *SIAM J. Discrete Math.* **4** (1991), 275–280.
133. a. We have $\frac{1}{2}A_n(2) = \sum_{k=0}^{n-1} A(n, k+1)2^k$, where $A(n, k+1)$ permutations of $[n]$ have k descents. Thus we need to associate an ordered partition τ of $[n]$ with a pair (w, S) , where $w \in \mathfrak{S}_n$ and $S \subseteq D(w)$. Given $w = a_1 a_2 \dots a_n$, draw a vertical bar between a_i and a_{i+1} if $a_i < a_{i+1}$ or if $a_i > a_{i+1}$ and $i \in S$. The sets contained between bars (including the beginning and end) are read from left to right and define τ .
- Example.* Let $w = 724531968$ and $S = \{1, 5\}$. Write $7|2|4|53|1|96|8$, so $\tau = (7, 2, 4, 35, 1, 69, 8)$.
134. See D. Foata and M.-P. Schützenberger, [1.26, Thm. 5.6]. For a vast generalization of this kind of formula, see E. Nevo and T. K. Petersen, *Discrete Computational Geometry* **45** (2011), 503–521.
135. a. Put $x = -1$ in equation (1.40) and compare with (1.54).
- b. Let $n = 2m + 1$. Since $\text{des}(w) = m$ if w is alternating, it suffices to show combinatorially that $\sum_w (-1)^{\text{des}(w)} = 0$, where w ranges over all nonalternating permutations in \mathfrak{S}_n . For a nonalternating permutation $w \in \mathfrak{S}_n$, let $T = T(w)$ be the increasing binary tree corresponding to w , as defined in Section 1.5. Since w is not alternating, it follows from the table preceding Proposition 1.5.3 that T has a vertex j with only one successor. For definiteness, choose the least such vertex j , and let T' be the flip of T at j , as defined in Section 1.6.2. Define $w' \in \mathfrak{S}_n$ by $T(w') = T'$. Clearly $w'' = w$, so we have defined an involution $w \mapsto w'$ on all nonalternating permutations in \mathfrak{S}_n . Since n is odd, it again follows from the table preceding Proposition 1.5.3 that $\text{des}(w)$ is the number of vertices of $T(w)$ with a left successor. Hence, $(-1)^{\text{des}(w)} + (-1)^{\text{des}(w')} = 0$, and the proof follows. For further aspects of this line of reasoning, see D. Foata and M.-P. Schützenberger, [1.26, Thm. 5.6].
136. *Answer.* $c_1 = c_{n-1} = 1$, all other $c_i = 0$.

- 137.** The number of $w \in \mathfrak{S}_n$ of type \mathbf{c} is $\tau(\mathbf{c}) = n! / 1^{c_1} c_1! \cdots n^{c_n} c_n!$. Let $n = a_0 + a_1 \ell$. It is not hard to see that $\tau(\mathbf{c})$ is prime to ℓ if and only if, setting $k = c_\ell$, we have $c_1 \geq (n_1 - k)\ell$ where $\binom{n_1}{k}$ is prime to ℓ . It follows from Exercise 1.14 that the number of binomial coefficients $\binom{n_1}{k}$ prime to ℓ is $\prod_{i \geq 1} (a_i + 1)$. Since $(c_1 - (n_1 - k)\ell, c_2, \dots, c_{\ell-1})$ can be the type of an arbitrary partition of a_0 , the proof follows.

This result first appeared in I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, Oxford University Press, Oxford, 1979; 2nd ed., 1995 (Ex. 10 of Ch. I.2). The proof given here appears on pp. 260–261 of R. Stanley, *Bull. Amer. Math. Soc.* **4** (1981), 254–265.

- 139.** Let $z = \sum_{n \geq 1} g(n)x^n/n!$. Then $z' = 1 + \frac{1}{2}z^2 + \frac{1}{4!}z^4 + \cdots = \cosh(z)$. The solution to this differential equation satisfying $z(0) = 0$ is

$$z(x) = \log(\sec x + \tan x).$$

Since $z'(x) = \sec x$, it follows easily that $g(2n+1) = E_{2n}$. For further information and a bijective proof, see Section 3 of A. G. Kuznetsov, I. M. Pak, and A. E. Postnikov, [1.51].

- 140.** *Hint.* Let $f_k(n)$ be the number of simsun permutations in \mathfrak{S}_n with k descents. By inserting $n+1$ into a simsun permutation in \mathfrak{S}_n , establish the recurrence

$$f_k(n+1) = (n-2k+2)f_{k-1} + (k+1)f_k(n),$$

with the initial conditions $f_0(1) = 1$, $f_k(n) = 0$ for $k > \lfloor n/2 \rfloor$. Further details may be found in S. Sundaram, *Advances in Math.* **104** (1994), 225–296 (§3) in the context of symmetric functions. We can also give a bijective proof, as follows. Let \mathcal{E} be a flip equivalence class of binary trees on the vertex set $[n+1]$. There are E_{n+1} such flip equivalence classes (Proposition 1.6.2). There is a unique tree $T' \in \mathcal{E}$ such that (i) the path from the root 1 to $n+1$ moves to the right, (ii) for every vertex not on this path with two children, the largest child is on the left, and (iii) any vertex with just one child has this child on the right. Let $w' \in \mathfrak{S}_{n+1}$ satisfy $T' = T(w')$ (as in Section 1.5). Then w' ends in $n+1$; let $w \in \mathfrak{S}_n$ be w' with $n+1$ removed. It is not hard to check that the map $\mathcal{E} \mapsto w$ gives a bijection between flip equivalence classes and simsun permutations. This proof is due to Maria Monks (October 2007).

Simsun permutations are named after Rodica Simion and Sheila Sundaram. They first appear in the paper S. Sundaram, *ibid.* (p. 267). They are variants of the *André permutations* of Foata and Schützenberger [1.27]. The terminology “simsun permutation” is due to S. Sundaram (after they were originally called “Sundaram permutations” by R. Stanley) in *J. Algebraic Combin.* **4** (1995), 69–92 (p. 75). For some further work on simsun permutations, see G. Hetyei, *Discrete Comput. Geom.* **16** (1996), 259–275.

- 141. a.** *Hint.* Show that $E_{n+1,k}$ is the number of alternating permutations of $[n+2]$ with first term $k+1$ and second term unequal to k , and that $E_{n,n-k}$ is the number of alternating permutations of $[n+2]$ with first term $k+1$ and second term k .

The numbers $E_{n,k}$ are called *Entringer numbers*, after R. C. Entringer, *Nieuw. Arch. Wisk.* **14** (1966), 241–246. The triangular array (1.140) is due to L. Seidel, *Sitzungsber. Münch. Akad.* **4** (1877), 157–187 (who used the word “boustrophedon” to describe the triangle). It was rediscovered by A. Kempner, *Tôhoku Math. J.* **37** (1933), 347–362; R. C. Entringer, *op. cit.*; and V. I. Arnold, *Duke Math. J.* **63** (1991), 537–555. For further information and references, see J. Millar, N. J. A. Sloane, and N. E. Young, *J. Combinatorial Theory Ser. A* **76** (1996), 44–54. A more recent reference is R. Ehrenborg and S. Mahajan, *Ann. Comb.* **2** (1998),

111–129 (§2). The boustrophedon triangle was generalized to permutations with an arbitrary descent set by Viennot [1.75].

- b. Rotate the triangle and change the sign of E_{mn} when $m+n \equiv 1, 2 \pmod{4}$ to obtain the array

$$\begin{array}{cccccccc}
 1 & & 0 & & -1 & & 0 & & 5 & & 0 & \cdots \\
 & -1 & & -1 & & 1 & & 5 & & -5 & & \\
 & & 0 & & 2 & & 4 & & -10 & & & \\
 & & & 2 & & 2 & & -14 & & & & \\
 & & & & 0 & & -16 & & & & & \\
 & & & & & -16 & & \cdots & & & & \\
 & & & & & & \ddots & & & & &
 \end{array}$$

This array is just a difference table, as defined in Section 1.9. By (a) the exponential generating function for the first row is $\sec(ix) = \operatorname{sech}(x)$. By Exercise 1.154(c) we get

$$\sum_{m \geq 0} \sum_{n \geq 0} (-1)^{\lfloor (2m+2n+3)/4 \rfloor} E_{m+n, [m, n]} \frac{x^m}{m!} \frac{y^n}{n!} = e^{-x} \operatorname{sech}(x+y).$$

If we convert all the negative coefficients to positive, it's not hard to see that the generating function becomes the right-hand side of equation (1.141), as claimed.

The transformation into a difference table that we have used here appears in Seidel, op. cit., and is treated systematically by D. Dumont, *Sém. Lotharingien de Combinatoire* 5 (1981), B05c (electronic). Equation (1.141) appears explicitly in R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*, 2nd ed., Addison-Wesley, Reading, Mass., 1994 (Exercise 6.75).

142. It is easy to verify that

$$\sum_{n \geq 0} f_n(a) x^n = (\sec x)(\cos(a-1)x + \sin ax),$$

and the proof follows. The motivation for this problem comes from the fact that for $0 \leq a \leq 1$, $f_n(a)$ is the volume of the convex polytope in \mathbb{R}^n given by

$$x_i \geq 0 \ (1 \leq i \leq n), \ x_1 \leq a, \ x_i + x_{i+1} \leq 1 \ (1 \leq i \leq n-1).$$

For further information on the case $a = 1$, see Exercise 4.56(c).

143. a. *Combinatorial Proof.* Let $1 \leq i \leq n$. The number of permutations $w \in \mathfrak{S}_n$ fixing i is $(n-1)!$. Hence the total number of fixed points of all $w \in \mathfrak{S}_n$ is $n \cdot (n-1)! = n!$. *Generating Function Proof.* We have

$$f(n) := \sum_{w \in \mathfrak{S}_n} \operatorname{fix}(w) = n! \frac{d}{dt_1} Z_n|_{t_i=1},$$

where Z_n is defined by (1.25). Hence by Theorem 1.3.3 we get

$$\begin{aligned}
 \sum_{n \geq 0} f(n) \frac{x^n}{n!} &= \frac{d}{dt_1} \exp \left(t_1 x + t_2 \frac{x^2}{2} + t_3 \frac{x^3}{3} + \cdots \right) \Big|_{t_i=1} \\
 &= x \exp \left(x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots \right) \\
 &= \frac{x}{1-x},
 \end{aligned}$$

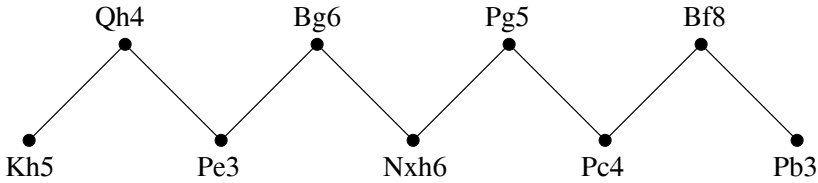


Figure 1.35 The solution poset for Exercise 1.145.

whence $f(n) = n!$.

Algebraic Proof. Let G be a finite group acting a set Y . By Burnside's lemma (Lemma 7.24.5, Vol. II), also called the Cauchy–Frobenius lemma, the average number of fixed points of $w \in G$ is the number of orbits of the action. Since the “defining representation” of \mathfrak{S}_n on $[n]$ has one orbit, the proof follows.

- b.** This result is a straightforward consequence of Proposition 6.1 of R. Stanley, *J. Combinatorial Theory Ser. A* **114** (2007), 436–460. Is there a combinatorial proof?

- 144. a.** It is in fact not hard to see that

$$2q^n \frac{\prod_{j=1}^n (1 - q^{2j-1})}{\prod_{j=1}^{2n+1} (1 + q^j)} = \frac{2(2n-1)!!}{3^n} x^n + O(x^{n+1}),$$

where $(2n-1)!! = 1 \cdot 3 \cdot 5 \cdots (2n-1)$.

- b.** See page 450 of R. Stanley, *J. Combinatorial Theory Ser. A* **114** (2007), 436–460.

- 145.** One solution is 1.Kh5 2.Pe3 3.Nxh6 4.Pc4 5.Pb3 6.Qh4 7.Bg6 8.Rg5 9.Bf8, followed by Nf6 mate. Label these nine Black moves as 1,3,5,7,9,2,4,6,8 in the order given. All solutions are a permutation of these nine moves. If a_1, a_2, \dots, a_9 is a permutation w of the labels of the moves, then they correspond to a solution if and only if w^{-1} is reverse alternating. (In other words, Qh4 must occur after both Kh5 and Pe3, Bg6 must occur after both Pe3 and Nxh6, etc.). In the terminology of Chapter 3, the solutions correspond to the linear extensions of the “zigzag poset” shown in Figure 1.35. Hence, the number of solutions is $E_9 = 7936$. For some properties of zigzag posets, see Exercise 3.66.

- 146.** The proof is a straightforward generalization of the proof we indicated of equation (1.59). For a q -analogue, see Proposition 3.3.19.4 and the discussion following it.

- 147.** A binary tree is an unlabeled min-max tree if and only if every nonendpoint vertex has a nonempty left subtree. Let f_n be the number of such trees on n vertices. Then

$$f_{n+1} = \sum_{k=1}^n f_k f_{n-k}, \quad n \geq 1.$$

Setting $y = \sum_{n \geq 0} f_n x^n$, we obtain

$$\frac{y - 1 - x}{x} = y^2 - y.$$

It follows that

$$y = \frac{1 + x - \sqrt{1 - 2x - 3x^2}}{2x}.$$

Comparing with the definition of M_n in Exercise 6.27 shows that $f_n = M_{n-1}$, $n \geq 1$.

148. It is easy to see from equations (1.33) and (1.63) that

$$\Psi_n(a+b, b) = \sum_{S \subseteq [n-1]} \alpha(S) u_S.$$

The proof follows from the formula $\Psi(a, b) = \Phi(a+b, ab+ba)$ (Theorem 1.6.3).

149. *Hint.* First establish the recurrence

$$2\Phi_n = \sum_{\substack{0 \leq i \leq n \\ n-i=2j-1}} \binom{n}{i} \Phi_i c(c^2 - 2d)^{j-1} - \sum_{\substack{0 \leq i \leq n \\ n-i=2j}} \binom{n}{i} \Phi_i (c^2 - 2d)^j \\ + \begin{cases} 2(c^2 - 2d)^{k-1}, & n = 2k - 1, \\ 0, & n = 2k. \end{cases}$$

The generating function follows easily from multiplying this recurrence by $x^n/n!$ and summing on $n \geq 1$.

This result is due to R. Stanley, *Math. Z.* **216** (1994), 483–499 (Corollary 1.4).

151. This elegant result is due to R. Ehrenborg, private communication (2007), based on the Pyr operator of R. Ehrenborg and M. Readdy, *J. Algebraic Combin.* **8** (1998), 273–299. Using concepts from Chapter 3, the present exercise has the following interpretation. Let P be the poset whose elements are all cd -monomials. Define α to cover β in P if β is obtained from α by removing a c or changing a d to c . Then $[\mu]\Phi_n(c, d)$ is equal to the number of maximal chains of the interval $[1, \mu]$. The problem of counting such chains was considered by F. Bergeron, M. Bousquet-Mélou, and S. Dulucq, *Ann. Sci. Math. Québec* **19** (1995), 139–151. They showed that the total number of saturated chains from 1 to rank n is E_{n+1} (the sum of the coefficients of Φ_{n+1}), though they did not interpret the number of maximal chains in each interval. Further properties of the poset P (and some generalizations) were given by B. Sagan and V. Vatter, *J. Algebraic Combinatorics* **24** (2006), 117–136.
152. An analogous result where simsum permutations are replaced by “André permutations” was earlier proved by M. Purtill [1.65]. The result for simsun permutations was stated without proof by R. Stanley, *Math. Zeitschrift*, **216** (1994), 483–499 (p. 498), saying that it can be proved by “similar reasoning” to Purtill’s. This assertion was further explicated by G. Hetyei, *Discrete Comput. Geom.* **16** (1996), 259–275 (Remark on p. 270).
153. a. *First Solution.* Put $c = 0$ and $d = 1$ in equation (1.142) (so $a - b = \sqrt{-2}$) and simplify. We obtain

$$\sum_{n \geq 0} f(n) \frac{x^{2n+1}}{(2n+1)!} = \sqrt{2} \tan(x/\sqrt{2}).$$

The proof follows from Proposition 1.6.1.

Second Solution. By equations (1.62) and (1.64) we have that $2^n f(n)$ is the number of complete (i.e., every internal vertex has two children) min-max trees with n internal vertices. A complete min-max tree with $n+1$ internal vertices is obtained by placing either 1 or $2n+3$ at the root, forming a left complete min-max subtree whose vertices are $2k+1$ elements from $\{2, 3, \dots, 2n+2\}$ ($0 \leq k \leq n$), and forming a right complete min-max subtree with the remaining elements. Hence setting $g(n) = 2^n f(n)$, we obtain the recurrence

$$g(n+1) = 2 \sum_{k=0}^n \binom{2n+1}{2k+1} g(k) g(n-k).$$

It is then straightforward to show that $g(n) = E_{2n+1}$. The result of this exercise was first proved by Foata and Schützenberger [1.27, Propriété 2.6] in the context of André polynomials.

R. Ehrenborg (private communication, 2007) points out that there is a similar formula for the coefficient of any monomial in Φ_n not containing two consecutive c 's.

- b. See R. L. Graham and N. Zang, Enumerating split-pair arrangements, preprint dated 10 January 2007. For some further combinatorial interpretations of F_n , see C. Poupard, *Europ. J. Combinatorics* **10** (1989), 369–374; A. G. Kuznetsov, I. M. Pak and A. E. Postnikov, *Uspekhi Mat. Nauk* **49** (1994), 79–110; and M. P. Develin and S. P. Sullivant, *Ann. Combinatorics* **7** (2003), 441–466 (Corollary 5.7).

154. a. Use equation (1.98).

- b. By (a), $e^{-x}F'(x) = F(x)$, from which $F(x) = e^{e^x-1}$, so $f(n)$ is the Bell number $B(n)$. The difference table in question looks like

1	2	5	15	52	203	...
	1	3	10	37	151	...
		2	7	27	114	...
			5	20	87	...
				15	67	...
					52	...
						...

Note that the first row is identical to the leftmost diagonal below the first row. This “Bell number triangle” is due to C. S. Peirce, *Amer. J. Math.* **3** (1880), 15–57 (p. 48). It gained some popularity by appearing in the “Mathematical Games” column of M. Gardner [1.33, Fig. 13]. D. E. Knuth uses it to develop properties of Bell numbers in *The Art of Computer Programming*, vol. 4, Fascicle 3, Addison-Wesley, Upper Saddle River, N.J., 2005, Section 7.2.1.5, and gives some further properties in Exercises 7.2.1.5–26 to 7.2.1.5–31.

- c. By Taylor’s theorem and (a), we have

$$\sum_{n \geq 0} \sum_{k \geq 0} \Delta^n f(k) \frac{x^n}{n!} \frac{t^k}{k!} = e^{-x} \left(F(x) + F'(x)t + F''(x)\frac{t^2}{2!} + \cdots \right) \\ = e^{-x} F(x+t).$$

This result appears in D. Dumont and X. G. Viennot, *Ann. Discrete Math.* **6** (1980), 77–87, but is undoubtedly much older.

155. a. For further information related to this problem and Exercise 1.154(a), see D. Dumont, in *Séminaire Lotharingien de Combinatoire*, 5ème Session, Institut de Recherche Mathématique Avancée, Strasbourg, 1982, pp. 59–78.

- b. One computes $f(0) = 1$, $f(1) = 2$, $f(2) = 6$, $f(3) = 20, \dots$. Hence, guess $f(n) = \binom{2n}{n}$ and $F(x) := \sum f(n)x^n = (1-4x)^{-1/2}$. By (a) we then have $G(x) := \sum g(n)x^n = \frac{1}{1+x} F\left(\frac{x}{1+x}\right) = (1-2x-3x^2)^{-1/2}$. To verify the guess, one must check that $\frac{1}{1+x} G\left(\frac{x}{1+x}\right) = F(x^2)$, which is routine.
- c. (suggested by L. W. Shapiro) One computes $f(0) = 1$, $f(1) = 1$, $f(2) = 2$, $f(3) = 5$, $f(4) = 14, \dots$. Hence, guess $f(n) = \frac{1}{n+1} \binom{2n}{n}$ (the Catalan number C_n) and

$F(x) := \sum f(n)x^n = \frac{1}{2x}(1 - (1 - 4x)^{1/2})$. Then

$$F_1(x) := \sum f(n+1)x^n = \frac{1}{x}(F(x) - 1) = \frac{1}{2x^2}(1 - 2x - (1 - 4x)^{1/2}),$$

so by (a),

$$G(x) := \sum g(n)x^n = \frac{1}{1+x}F_1\left(\frac{1}{1+x}\right) = \frac{1}{2x^2}(1 - x - (1 - 2x - 3x^2)^{-1/2}).$$

To verify this guess, one must check that $\frac{1}{1+x}G\left(\frac{1}{1+x}\right) = F(x^2)$, which is routine.

156. Answer: $c_n = \prod_p p^{\lfloor n/p \rfloor}$, where p ranges over all primes. Thus, $c_0 = 1$, $c_1 = 1$, $c_2 = 2$, $c_3 = 6$, $c_4 = 12$, $c_5 = 60$, $c_6 = 360$, and so on. See E. G. Straus, *Proc. Amer. Math. Soc.* **2** (1951), 24–27. The sequence c_n can also be defined by the recurrence $c_0 = 1$ and $c_{n+1} = s_{n+1}c_n$, where s_{n+1} is the largest squarefree divisor of $n+1$.

157. Let $z = y^\lambda$, and equate coefficients of x^{n-1} on both sides of $(\lambda+1)y'z = (yz)'$. This result goes back to Euler and is discussed (with many similar methods for manipulating power series) in D. E. Knuth, *The Art of Computer Programming*, vol. 2, 3rd ed., Addison-Wesley, Upper Saddle River, N.J., 1997, Section 4.7. It was rediscovered by H. W. Gould, *Amer. Math. Monthly* **81** (1974), 3–14.

158. Let $\log F(x) = \sum_{n \geq 1} g_n x^n$. Then

$$\sum_{n \geq 1} g_n x^n = \sum_{i \geq 1} \sum_{j \geq 1} \frac{a_i x^{ij}}{j} = \sum_{n \geq 1} \frac{x^n}{n} \sum_{d|n} da_d.$$

Hence,

$$ng_n = \sum_{d|n} da_d,$$

so by the Möbius inversion formula of elementary number theory,

$$a_n = \frac{1}{n} \sum_{d|n} dg_d \mu(n/d). \quad (1.156)$$

We have $1+x = (1-x)^{-1}(1-x^2)$ (no need to use (1.156)).

If $F(x) = e^{x/(1-x)}$ then $g_n = 1$ for all n , so by (1.156) we have $a_n = \phi(n)/n$, where $\phi(n)$ is Euler's totient function.

159. Answer: $A(x) = \sqrt{F(x)F(-x)}$, $B(x) = \sqrt{F(x)/F(-x)}$. This result is due to Marcelo Aguiar (private communication, 2006) as part of his theory of combinatorial Hopf algebras and noncommutative diagonalization.

160. a. This formula is a standard result of hoary provenance which follows readily from

$$\sum_{r=0}^{k-1} \zeta^{rj} = \begin{cases} 0, & 0 < j < k, \\ k, & j = 0. \end{cases}$$

b. Let $\zeta = e^{2\pi i/k}$. According to (a) and Proposition 1.4.6, we have

$$f(n, k, j) = \frac{1}{k} \sum_{r=0}^{k-1} \zeta^{-jr} (n)!|_{q=\zeta^r}. \quad (1.157)$$

If $n \geq k$ then at least one factor $1+q+\cdots+q^m$ of $(n)!$ will vanish at $q = \zeta^r$ for $1 \leq r \leq k-1$. Thus, the only surviving term of the sum is $(n)!|_{q=1} = n!$, and the proof follows.

- c. When $n = k - 1$, we have $(n)!|_{q=\zeta^r} = 0$ unless $r = 0$ or ζ^r is a primitive k th root of unity. In the former case, we get the term $(k - 1)!/k$. In the latter case, write $\xi = \zeta^r$. Then

$$(k - 1)!|_{q=\xi} = \frac{(1 - \xi)(1 - \xi^2) \cdots (1 - \xi^{k-1})}{(1 - \xi)^{k-1}}. \quad (1.158)$$

Now

$$\prod_{j=1}^{k-1} (q - \xi^j) = \frac{q^k - 1}{q - 1},$$

Letting $q \rightarrow 1$ gives $\prod_{j=1}^{k-1} (1 - \xi^j) = k$. Hence from equation (1.158), we have

$$(k - 1)!|_{q=\xi} = \frac{k}{(1 - \xi)^{k-1}},$$

and the proof follows from setting $n = k - 1$ and $j = 0$ in equation (1.157).

NOTE. Let $\Phi_n(x)$ denote the (monic) n th cyclotomic polynomial, i.e., its zeros are the primitive n th roots of unity. It can be shown that if $n \geq 2$ then

$$f(n - 1, n, 0) = \frac{(n - 1)!}{n} + (-1)^n (n - 1) [x^{n-1}] \log \frac{\Phi_n(1 + x)}{\Phi_n(1)}.$$

Let us also note that $\Phi_n(1) = p$ if n is the power of a prime p ; otherwise $\Phi_n(1) = 1$.

- 161. b.** We have

$$\frac{H(x)}{H(x) + H(-x)} = \frac{G(x)}{2}.$$

Hence,

$$\begin{aligned} \frac{H(-x)}{H(x)} &= \frac{2}{G(x)} - 1 \\ \Rightarrow \log H(-x) - \log H(x) &= \log \left(\frac{2}{G(x)} - 1 \right). \end{aligned}$$

If we divide the left-hand side by -2 , then we obtain the odd part of $\log H(x)$. Hence,

$$\log H(x) = -\frac{1}{2} \log \left(\frac{2}{G(x)} - 1 \right) + E_1(x),$$

where $E_1(x)$ is any even power series in x with $E_1(0) = 0$. Thus, $E(x) := e^{E_1(x)}$ is an arbitrary even power series with $E(0) = 1$. Therefore, we get the general solution

$$H(x) = \left(\frac{2}{G(x)} - 1 \right)^{-1/2} E(x).$$

- 162.** Using the formulas

$$\tan(x + y) = \frac{\tan x + \tan y}{1 - (\tan x)(\tan y)},$$

$$\tan x/2 = \frac{\pm \sqrt{1 + \tan^2 x} - 1}{\tan x},$$

we have $\tan(\tan^{-1} f(x) + \tan^{-1} f(-x)) = g(x)$

$$\begin{aligned}\Rightarrow \tan^{-1} f(x) &= \frac{1}{2} \tan^{-1} g(x) + k(x), \quad k(x) = -k(-x) \\ \Rightarrow f(x) &= \tan\left(\frac{1}{2} \tan^{-1} g(x) + k(x)\right) \\ &= \frac{\tan \frac{1}{2} \tan^{-1} g(x) + \tan k(x)}{1 - (\tan \frac{1}{2} \tan^{-1} g(x)) \tan k(x)} \\ &= \frac{\frac{\sqrt{1+g(x)^2}-1}{g(x)} + h(x)}{1 - \frac{\sqrt{1+g(x)^2}-1}{g(x)} h(x)}, \quad h(x) = -h(-x).\end{aligned}$$

Choosing the correct sign gives

$$f(x) = \frac{-\sqrt{1+g(x)^2}-1+g(x)h(x)}{g(x)-(\sqrt{1+g(x)^2}-1)h(x)},$$

where $h(x)$ is any even power series.

- 163. a.** We have $F(x, y) = f(f^{(-1)}(x) + f^{(-1)}(y))$. The concept of a formal group law goes back to S. Bocher, *Ann. Math.* **47** (1946), 192–201.
- b.** See for instance A. Fröhlich, *Lecture Notes in Math.*, no. 74, Springer-Verlag, Berlin/New York, 1968. For a combinatorial approach to formal groups via Hopf algebras, see C. Lenart, Ph.D. thesis, University of Manchester, 1996, and C. Lenart and N. Ray, Some applications of incidence Hopf algebras to formal group theory and algebraic topology, preprint, University of Manchester, 1995.
- c.** $f(x) = x, e^x - 1, \tan x, \sin x$, respectively.
- d.** Let $R(x) = (xe^{-x})^{(-1)}$. Thus,

$$\begin{aligned}F(x, y) &= (R(x) + R(y))e^{-R(x)-R(y)} \\ &= xe^{-R(y)} + ye^{-R(x)}.\end{aligned}$$

The proof follows from equation (5.128), which asserts that

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

- e.** Euler, *Institutiones Calculi integralis*, *Ac. Sc. Petropoli*, 1761, showed that

$$F(x, y) = \frac{x\sqrt{1-y^4} + y\sqrt{1-x^4}}{1+x^2y^2}.$$

- 164.** Note that setting $x = 0$ is useless. Instead, write

$$F(x, y) = \frac{x F(x, 0) - y}{xy^2 + x - y}.$$

The denominator factors as $x(y - \theta_1(x))(y - \theta_2(x))$, where

$$\theta_1(x), \theta_2(x) = \frac{1 \mp \sqrt{1-4x^2}}{2x}.$$

Now $y - \theta_1(x) \sim y - x$ as $x, y \rightarrow 0$, so the factor $1/(y - \theta_1(x))$ has no power series expansion about $(0, 0)$. Since $F(x, y)$ has such an expansion, the factor $y - \theta_1(x)$ must appear in the numerator. Hence $x F(x, 0) = \theta_1(x)$, yielding

$$F(x, 0) = \frac{1 - \sqrt{1 - 4x^2}}{2x^2} = \sum_{n \geq 0} C_n x^{2n},$$

$$F(x, y) = \frac{2}{1 - 2xy + \sqrt{1 - 4x^2}}.$$

The solution to this exercise is a simple example of a technique known as the *kernel method*. This method originated in Exercise 2.2.1–4 of Knuth's book *The Art of Computer Programming*, vol. 1, Addison-Wesley, Reading, Massachusetts, 1973, 3rd ed., 1997. The present exercise is the same as Knuth's (after omitting some preliminary steps). See Section 1 of H. Prodinger, *Sém. Lotharingien de Combinatoire* **50** (2004), article B50f, for further information and examples. An interesting variant of the kernel method applied to queuing theory appears in Chapter 14 of L. Flatto, *Poncelet's Theorem*, American Math. Society, Providence, R.I., 2009.

165. Answer. The coefficient $f(n)$ of $F(x)$ is the number of 1's in the binary expansion of n .

166. Answer. $F(x) = (1 + x^n)^{1/n} = \sum_{k \geq 0} \binom{1/n}{k} x^{kn}$.

167. Equation (1.143) is just the Taylor series expansion of $F(x + t)$ at $t = 0$.

168. a. It is not hard to check that for general $A(x) = x + a_2 x^2 + a_3 x^3 + \dots$, we have

$$A(-A(-x)) = x + p_2 x^2 + p_3 x^3 + \dots,$$

where p_{2n-1} and p_{2n} are polynomials in $a_2, a_3, \dots, a_{2n-1}$. (It's easy to see that, in fact, $p_2 = 0$.) Moreover, the only term of p_{2n-1} involving a_{2n-1} is $2a_{2n-1}$. Hence, if $A(-A(-x)) = x$ then once $a_2, a_3, \dots, a_{2n-2}$ are specified, we have that a_{2n-1} is uniquely determined. Thus, we need to show that if $a_2, a_4, \dots, a_{2n-2}$ are specified, thereby determining $a_3, a_5, \dots, a_{2n-1}$, then $p_{2n} = 0$. For instance, equating coefficients of x^3 in $A(-A(-x)) = x$ gives $a_3 = a_2^2$. Then

$$p_4 = a_2^3 - a_2 a_3 = a_2^3 - a_2(a_2^2) = 0.$$

We can reformulate the result we need to prove more algebraically. Given $A(x) = x + a_2 x^2 + \dots$, let $B(x) = A(-A(-x)) = x + p_2 x^2 + \dots$. Let $A^{(-1)}(x) = x + \alpha_2 x^2 + \alpha_3 x^3 + \dots$. Then

$$A(-x) = B(-A^{(-1)}(x)) = A^{(-1)}(x) + p_2 A^{(-1)}(x)^2 - \dots.$$

Taking the coefficient of x^{2n} gives

$$a_{2n} \equiv -\alpha_{2n} + p_{2n} \pmod{I}. \quad (1.159)$$

But also

$$\begin{aligned} -A(-x) &= A^{(-1)}(B(x)) \\ &= B(x) + \alpha_2 B(x)^2 + \dots \end{aligned}$$

Taking coefficients of x^{2n} yields

$$\begin{aligned} -a_{2n} &\equiv p_{2n} + [x^{2n}] \sum_{i=2}^{2n} \alpha_i (x + p_{2n} x^{2n})^i \pmod{I} \\ &\equiv p_{2n} + \alpha_{2n} \pmod{I}. \end{aligned} \quad (1.160)$$

Equations (1.159) and (1.160) imply $p_{2n} \in I$, as desired. This proof was obtained in collaboration with Whan Ghang. NOTE. It was shown by Ghang that a_{2n+1} is a polynomial in a_2, a_4, \dots, a_{2n} with integer coefficients.

NOTE. An equivalent reformulation of the result of this item is the following. For any $A(x) = x + a_2x^2 + \dots \in K[[x]]$, either $A(-A(-x)) = x$ or $A(-A(-x)) - x$ has odd degree. This result can be considerably generalized. For instance, if $C(x) = -x + c_2x^2 + \dots$ and $C(C(x)) = x$, then (writing composition of functions as juxtaposition) either $ACAC(x) = x$ or $ACAC(x) - x$ has odd degree. More generally, if ζ is a primitive k th root of unity and $C(x) = \zeta x + c_2x^2 + \dots$, where $C^k = x$, then either $(AC)^k(x) = x$ or $(AC)^k(x) - x$ has degree $d \equiv 1 \pmod{k}$. The possibility of such a generalization was suggested by F. Bergeron (private communication, 2007).

- b. Use induction on n .
 c. Marcelo Aguiar (private communication, 2006) first obtained this result as part of his theory of combinatorial Hopf algebras and noncommutative diagonalization.
 d. Answer: $A(x) = 2x/(2-x)$ and $D(x) = \log \frac{2+x}{2-x}$. This example is due to Aguiar.
 e. First show the following.

$$\bullet \sum_{n \geq 1} a_n \left(\frac{x}{1-x} \right)^n = \sum_{n \geq b_n} x^n \iff e^x \sum_{j \geq 0} a_{j+1} \frac{x^j}{j!} = \sum_{j \geq 0} b_{j+1} \frac{x^j}{j!}.$$

(See Exercises 154(a) and 155(a).)

- For any $F(x) = x + \sum_{n \geq 2} a_n x^n$ and $H(x) = x + \sum_{n \geq 2} b_n x^n$, we have

$$F^{(-1)}(-F(-x)) = H^{(-1)}(-H(-x)),$$

if and only if $F(x)/H(x)$ is odd.

- f. Answer: $b_{2n} = (-1)^{n-1} E_{2n-1}$, where E_{2n-1} is an Euler number.
 169. There are many possible methods. A uniform way to do all three parts is to note that for any power series $F(x) = \sum_{n \geq 0} a_n x^n$, we have

$$x D F(x) = \sum_{n \geq 0} n a_n x^n,$$

where $D = \frac{d}{dx}$. Hence,

$$(xD + 2)^2 F(x) = \sum_{n \geq 0} (n+2)^2 a_n x^n.$$

Letting $F(x) = 1/(1-x)$, e^x , and $1/\sqrt{1-4x}$ yields, after some routine computation, the three answers:

$$\begin{aligned} \sum_{n \geq 0} (n+2)^2 x^n &= \frac{4-3x+x^2}{(1-x)^3}, \\ \sum_{n \geq 0} (n+2)^2 \frac{x^n}{n!} &= (x^2+5x+4)e^x, \\ \sum_{n \geq 0} (n+2)^2 \binom{2n}{n} x^n &= \frac{4-22x+36x^2}{(1-4x)^{3/2}}. \end{aligned}$$

170. a. Answer: $y = (\alpha + (\beta - \alpha)x)/(1 - x - x^2)$. The general theory of linear recurrence relations with constant coefficients is developed in Sections 4.1–4.4.

- b. The recurrence yields $y' = (xy)' - \frac{1}{2}xy^2$, $y(0) = 1$, from which we obtain

$$y = \frac{\exp\left(\frac{x}{2} + \frac{x^2}{4}\right)}{\sqrt{1-x}}.$$

For the significance of this generating function, see Example 5.2.9.

- c. We obtain $2y' = y^2$, $y(0) = 1$, whence $y = 1/(1 - \frac{1}{2}x)$. Thus, $a_n = 2^{-n}n!$.
- d. (sketch) Let $F_k(x) = \sum_{n \geq 0} a_k(n)x^n/n!$, so $A(x, t) = \sum_{k \geq 0} F_k(x)t^k$. The recurrence for $a_k(n)$ gives

$$F'_k(x) = \sum_{2r+s=k-1} (F_{2r}(x) + F_{2r+1}(x))F_s(x). \quad (1.161)$$

Let $A_e(x) = \frac{1}{2}(A(x, t) + A(x, -t))$ and $A_o(x, t) = \frac{1}{2}(A(x, t) - A(x, -t))$. From equation (1.161) and some manipulations, we obtain the system of differential equations

$$\begin{aligned} \frac{\partial A_e}{\partial x} &= tA_eA_o + A_o^2, \\ \frac{\partial A_o}{\partial x} &= tA_e^2 + A_eA_o. \end{aligned} \quad (1.162)$$

To solve this system, note that

$$\frac{\partial A_e/\partial x}{\partial A_o/\partial x} = \frac{A_o}{A_e}.$$

Hence, $\frac{\partial}{\partial x}(A_e^2 - A_o^2) = 0$, so $A_e^2 - A_o^2$ is independent of x . Some experimentation suggests that $A_e^2 - A_o^2 = 1$, which together with (1.162) yields

$$\frac{\partial A_e}{\partial x} = tA_e\sqrt{A_e^2 - 1} + A_e^2 - 1.$$

This equation can be routinely solved by separation of variables (though some care must be taken to choose the correct branch of the resulting integral, including the correct sign of $\sqrt{A_e^2 - 1}$). A similar argument yields A_o , and we finally obtain the following expression for $A = A_e + A_o$:

$$A(x, t) = \sqrt{\frac{1-t}{1+t}} \left(\frac{2}{1 - \frac{1-\rho}{t}e^{\rho x}} - 1 \right),$$

where $\rho = \sqrt{1-t^2}$. It can then be checked that this formula does indeed give the correct solution to the original differential equations, justifying the assumption that $A_e^2 - A_o^2 = 1$. For further details and motivation, see Section 2 of R. Stanley, *Michigan Math. J.*, **57** (2008), 675–687.

- 171.** While this problem can be solved by the “brute force” method of computing the coefficients on the right-hand side of equation (1.144), it is better to note that $B'(x) - B(x) = A'(x)$ and then solve this differential equation for $B(x)$ with the initial condition $B(0) = a_0$. Alternatively, one could start with $A(x)$:

$$B(x) = (1 + I + I^2 + \cdots)A(x) = (1 - I)^{-1}A(x).$$

Multiplying by $1 - I$ and differentiating both sides results in the same differential equation $B'(x) - B(x) = A'(x)$. (It isn't difficult to justify these formal manipulations of the operator I .)

172. One method of proof is to first establish the three term recurrence

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x)$$

and then use induction.

173. a.

$$\begin{aligned}\sqrt{\frac{1+x}{1-x}} &= (1+x)(1-x^2)^{-1/2} \\ &= \sum_{n \geq 0} 4^{-n} \binom{2n}{n} (x^{2n} + x^{2n+1}).\end{aligned}$$

b. $\sum_{n \geq 1} \frac{x^{2n}}{n^2 \binom{2n}{n}}.$

c. $\sum_{n \geq 0} t(t^2 - 1^2)(t^2 - 3^2) \cdots (t^2 - (2n-1)^2) \frac{x^{2n+1}}{(2n+1)!}.$

d. $\sum_{n \geq 0} t^2(t^2 - 2^2)(t^2 - 4^2) \cdots (t^2 - (2n-2)^2) \frac{x^{2n}}{(2n)!}.$

e. $2 \sum_{n \geq 0} (-1)^n 2^{2n} \frac{x^{4n+2}}{(4n+2)!}.$ Similar results hold for $\cos(x) \cosh(x)$, $\cos(x) \sinh(x)$, and $\sin(x) \cosh(x)$.

f. $6 \sum_{n \geq 0} (-1)^n 2^{6n} \frac{x^{6n+3}}{(6n+3)!}.$ Similar results hold when any subset of the three \sin 's is replaced by \cos . There seems to be no analogous result for *four* factors.

To do (c), for instance, first observe that the coefficient of $x^{2n+1}/(2n+1)!$ in $\sin(t \sin^{-1} x)$ is a polynomial $P_n(t)$ of degree $2n+1$ and leading coefficient $(-1)^n$. If $k \in \mathbb{Z}$, then $\sin(2k+1)\theta$ is an odd polynomial in $\sin \theta$ of degree $2k+1$. Hence, $P_n(\pm(2k+1)) = 0$ for $n > k$. Moreover, $\sin 0 = 0$ so $P_n(0) = 0$. We now have sufficient information to determine $P_n(t)$ uniquely. To get (b), consider the coefficient of t^2 in (d). For (g), note that

$$\cos(\log(1+x)) = \Re(1+x)^i.$$

174. Hint. What is the number of elements of the set $\{0, 1\}$?

176. Induction on n . We have $E(0) = 0$. For $n \geq 1$, choose the first vector v_1 at random. If $v_1 = 0$, we expect $E(n)$ further steps, and this occurs with probability $1/q^n$. Otherwise, v_1 is not the zero vector. Consider the projection of our space to a subspace complementary to v_1 . The uniform distribution over \mathbb{F}_q^n projects to the uniform distribution over this copy of \mathbb{F}_q^{n-1} , and our sequence of vectors will span \mathbb{F}_q^n precisely when the set of their projections spans \mathbb{F}_q^{n-1} . It follows that we expect $E(n-1)$ further steps, and so

$$E(n) = 1 + \frac{E(n) + (q^n - 1)E(n-1)}{q^n}.$$

Solving this equation gives $E(n) = E(n-1) + q^n/(q^n - 1)$, and so

$$E(n) = \sum_{i=1}^n q^i / (q^i - 1).$$

This argument was suggested by J. Lewis (October 2009).

- 182.** Suppose that $A \in \text{GL}(n, q)$ has no 0 entries. There are exactly $(q-1)^{2n-1}$ matrices of the form DAD' , where D, D' are diagonal matrices in $\text{GL}(n, q)$. Exactly one of the matrices $C = DAD'$ has the first entry in every row and column equal to -1 . Subtract the first column of C from every other column, obtaining a matrix D . Let B be obtained from D by removing the first row and column. Then B is a matrix in $\text{GL}(n-1, q)$ with no entry equal to 1, and every such matrix is obtained exactly once by this procedure.
- 183.** *First Solution* (sketch). The identity asserts that each of the q^n monic polynomials of degree n can be written uniquely as a product of monic irreducible polynomials.
Second Solution (sketch). Take logarithms of both sides and simplify the right-hand side.
- 185. a.** Note that $q^n - D(n, 0)$ is the number of monic polynomials of degree n over \mathbb{F}_q with nonzero discriminant. In the same way that we obtained the first solution to Exercise 1.183, we get

$$\sum_{n \geq 0} (q^n - D(n, 0))x^n = \prod_{d \geq 1} (1 + x^d)^{\beta(d)}.$$

Hence,

$$\begin{aligned} \sum_{n \geq 0} (q^n - D(n, 0))x^n &= \left(\frac{1-x^2}{1-x} \right)^{\beta(d)} \\ &= \frac{1-qx^2}{1-qx}, \end{aligned}$$

the last step by Exercise 1.183. The proof follows easily. This result appears in D. E. Knuth, *The Art of Computer Programming*, vol. 2, 3rd ed., Addison-Wesley, Reading, Mass., 1997 (Exercise 4.6.2-2(b)) and is attributed to E. R. Berlekamp. Greta Panova (November 2007) showed that this problem can also be solved by establishing the recurrence

$$D(n, 0) = \sum_{k \geq 1} q^k (q^{2n-k} - D(n-2k, 0)).$$

b. We have

$$\begin{aligned} \sum_{\beta \in \mathbb{N}^k} N(\beta) x^\beta &= \prod_{d \geq 1} \left(\sum_{\alpha \in X} x^{\alpha d} \right)^{\beta(d)} \\ &= \prod_{d \geq 1} \prod_{\substack{\alpha \in \mathbb{N}^k \\ \alpha \neq (0, 0, \dots, 0)}} (1 - x^{\alpha d})^{a_\alpha \beta(d)}. \end{aligned}$$

The proof follows from Exercise 1.183.

- 186. a.** Let $k = 1$ and $X = \{0, 1, \dots, r-1\}$ in Exercise 1.185(c). We get

$$\sum_{n \in X} x^n = 1 + x + \dots + x^{r-1} = \frac{1-x^r}{1-x}.$$

Hence,

$$\sum_{n \geq 0} N_r(n) x^n = \frac{1-qx^r}{1-qx},$$

yielding equation (1.147). This result is stated in D. E. Knuth, *The Art of Computer Programming*, vol. 2, 3rd ed., Addison-Wesley, Reading, Mass., 1997 (solution to Exercise 4.6.2-2(b)).

- b. Set $k = 2$ and $X = \{(m, 0) : m \in \mathbb{N}\} \cup \{(0, n) : n \in \mathbb{P}\}$ to get

$$\sum_{m, n \geq 0} N(m, n) x^m y^n = \frac{1 - qxy}{(1 - qx)(1 - qy)},$$

from which equation (1.148) follows.

- c. Take $k = 1$ and $X = \mathbb{N} - \{1\}$. We get

$$\sum_{n \geq 0} P(n) x^n = \frac{1 - qx^6}{(1 - qx^2)(1 - qx^3)},$$

via the identity

$$1 + \frac{t^2}{1 - t} = \frac{1 - t^6}{(1 - t^2)(1 - t^3)}.$$

Using the partial fraction expansion

$$\frac{1 - qx^6}{(1 - qx^2)(1 - qx^3)} = -\frac{x}{q} + \frac{(1 + q)(1 + x)}{q(1 - qx^2)} - \frac{1 + qx + qx^2}{q(1 - qx^3)},$$

it is routine to obtain equation (1.149). This result can also be obtained by noting that every monic powerful polynomial can be written uniquely in the form $f^2 g^3$, where f and g are monic and g is squarefree. Hence, $P(n) = \sum_{2i+3j=n} q^i (q^j - D(j, 0))$, where $D(j, 0)$ is defined in Exercise 1.185(b), and so on. This result is due to R. Stanley (proposer), Problem 11348, *Amer. Math. Monthly* **115** (2008), 262; R. Stong (solution), **117** (2010), 87–88.

NOTE. The term “powerful polynomial” is borrowed from the corresponding notion for integers. See for instance the Wikipedia entry “Powerful number” at

(http://en.wikipedia.org/wiki/Powerful_number).

- 187. a.** The resultant $\text{res}(f, g)$ of two polynomials $f(x) = \prod (x - \theta_i)$ and $g(x) = \prod (x - \tau_j)$ over a field K is defined by

$$\text{res}(f, g) = \prod_{i, j} (\theta_i - \tau_j).$$

It is a standard fact (a consequence of the fact that $\text{res}(f, g)$ is invariant under any permutation of the θ_i 's and of the τ_j 's) that $\text{res}(f, g) \in K$. Suppose that $f(x) = f_1(x) \cdots f_k(x)$ where each $f_i(x)$ is irreducible. Clearly,

$$\text{disc}(f) = \prod_{i=1}^k \text{disc}(f_i) \cdot \prod_{1 \leq i < j \leq k} \text{res}(f_i, f_j)^2. \quad (1.163)$$

A standard result from Galois theory states that the discriminant of an irreducible polynomial $g(x)$ of degree n over a field K is a square in K if and only if the Galois group of $g(x)$ (regarded as a group of permutations of the zeros of $g(x)$) is contained in the alternating group \mathfrak{A}_n . Now the Galois group of an irreducible polynomial of degree n over \mathbb{F}_q is generated by an n -cycle and hence is contained in \mathfrak{A}_n if and only if n is odd. It follows from equation (1.163) that if $\text{disc}(f) \neq 0$, then $\text{disc}(f)$

is a square in \mathbb{F}_q if and only if $n - k$ is even. This result goes back to L. Stickelberger, *Verh. Ersten Internationaler Mathematiker-Kongresses (Zürich, 1897)*, reprinted by Kraus Reprint Limited, Nendeln/Liechtenstein, 1967, pp. 182–193. A simplification of Stickelberger's argument was given by K. Dalen, *Math. Scand.* **3** (1955), 124–126. See also L. E. Dickson, *Bull. Amer. Math. Soc.* **13** (1906/07), 1–8, and R. G. Swan, *Pacific J. Math.* **12** (1962), 1099–1106 (Corollary 1). The foregoing proof is possibly new. NOTE: Swan, *ibid.* (§3), uses this result to give a simple proof of the law of quadratic reciprocity.

Now let $N_e(n)$ (respectively, $N_o(n)$) denote the number of monic polynomials of degree n which are a product of an even number (respectively, odd number) of distinct irreducible factors. It is easy to see (analogous to the solution to Exercise 1.183) that

$$\sum_{n \geq 0} (N_e(n) - N_o(n))x^n = \prod_{d \geq 1} (1 - x^d)^{\beta(d)}.$$

But

$$\prod_{d \geq 1} (1 - x^d)^{\beta(d)} = 1 - qx$$

by Exercise 1.183. Hence, $N_e(n) = N_o(n)$ for $n > 1$, and the proof follows.

- b.** Let $f(x) = \prod_{i=1}^n (x - \theta_i)$ be a monic polynomial of degree n over \mathbb{F}_q . For $a \in \mathbb{F}_q^* = \mathbb{F}_q - \{0\}$, write $f_a(x) = a^n f(x/a)$, so $f_a(x) = \prod_{i=1}^n (x - a\theta_i)$. It follows that

$$\text{disc}(f_a(x)) = a^{n(n-1)} \text{disc}(f(x)).$$

If $(n(n-1), q-1) = 1$ then the map $a \mapsto a^{n(n-1)}$ is a bijection on \mathbb{F}_q^* . Hence, if $\text{disc}(f) \neq 0$, then we have $\{\text{disc}(f_a) : a \in \mathbb{F}_q^*\} = \mathbb{F}_q^*$. It follows that $D(n, a) = D(n, b)$ for all $a, b \in \mathbb{F}_q^*$. Since $D(n, 0) = q^{n-1}$, we have $D(n, a) = q^{n-1}$ for all $a \in \mathbb{F}_q$.

Now assume that $(n(n-1), q-1) = 2$. Thus, as a ranges over \mathbb{F}_q^* , $a^{n(n-1)}$ ranges over all squares in \mathbb{F}_q^* twice each. Some care must be taken since we can have $f_a(x) = f_b(x)$ for $a \neq b$. (This issue did not arise in the case $(n(n-1), q-1) = 1$ since the $f_a(x)$'s had distinct discriminants.) Thus for each f , let P_f be the multiset of all f_a , $a \in \mathbb{F}_q^*$. The multiset union $\bigcup_f P_f$ contains each monic polynomial of degree n over \mathbb{F}_q exactly $q-1$ times. For each $a, b \in \mathbb{F}_q^*$ such that either both a, b or neither a, b are squares, the same number of polynomials (counting multiplicity) $g \in \bigcup_f P_f$ satisfy $\text{disc}(g) = a$ as satisfy $\text{disc}(g) = b$. Finally, by (a) it follows that the number of $g \in \bigcup_f P_f$ with square discriminants is the same as the number with nonsquare discriminants. Hence, $D(n, a) = D(n, b)$ for all $a, b \in \mathbb{F}_q^*$, and thus as previously for all $a, b \in \mathbb{F}_q$.

- 188. First Solution.** Let V be an n -dimensional vector space over a field K , and fix an ordered basis $\mathbf{v} = (v_1, \dots, v_n)$ of V . Let \mathcal{N}_n denote the set of all nilpotent linear transformations $A : V \rightarrow V$. We will construct a bijection $\varphi : \mathcal{N}_n \rightarrow V^{n-1}$. Letting $V = \mathbb{F}_q$, it follows that $\#\mathcal{N}_n = \#(\mathbb{F}_q)^{n-1} = q^{n(n-1)}$.

The bijection is based on a standard construction in linear algebra known as *adapting* an ordered basis $\mathbf{w} = (w_1, \dots, w_n)$ of a vector space V to an m -dimensional subspace U of V . It constructs from \mathbf{w} in a canonical way a new ordered basis $w_{i_1}, \dots, w_{i_{n-m}}, u_1, \dots, u_m$ of V such that the first $n-m$ elements form a subsequence of \mathbf{w} and the last m form an ordered basis of u . See, for example, M. C. Crabb, *Finite Fields and Their Applications* **12** (2006), 151–154 (p. 153) for further details.

Now let $A \in \mathcal{N}_n$, and write $V_i = A^i(V)$, $i \geq 0$. Let r be the least integer for which $V_r = 0$, so we have a strictly decreasing sequence

$$V = V_0 \supset V_1 \supset \dots \supset V_r = 0.$$

Set $n_i = \dim V_i$ and $m_i = n_{i-1} - n_i$. Adapt the ordered basis \mathbf{v} of V to V_1 . Then adapt this new ordered basis to V_2 , etc. After $r - 1$ steps we have constructed in a canonical way an ordered basis $\mathbf{y} = (y_1, \dots, y_n)$ such that y_{n-n_i+1}, \dots, y_n is a basis for V_i , $1 \leq i \leq r - 1$. We associate with A the $(n - 1)$ -tuple $\varphi(A) = (A(y_1), \dots, A(y_{n-1})) \in V^{n-1}$. (Note that $A(y_n) = 0$.) It is straightforward to check that this construction gives a bijection $\varphi : \mathcal{N}_n \rightarrow V^{n-1}$ as desired.

This argument is due to M. C. Crabb, *ibid.*, and we have closely followed his presentation (though with fewer details). As Crabb points out, this bijection can be regarded as a generalization of the Prüfer bijection (first proof of Proposition 5.3.2, specialized to rooted trees) for counting rooted trees on an n -element set. Further connections between the enumeration of trees and linear transformations were obtained by J.-B. Yin, Ph.D. thesis, M.I.T., 2009. For a further result of this nature, see Exercise 1.189. *Second Solution* (sketch), due to Hansheng Diao, November 2007. Induction on n , the base case $n = 1$ being trivial. The statement is true for $k < n$. Let Q be the matrix in $\text{Mat}(n, q)$ with 1's on the diagonal above the main diagonal and 0's elsewhere (i.e., a Jordan block of size n with eigenvalue 0). Let

$$\mathcal{A} = \{(M, N) \in \text{Mat}(n, q) \times \text{Mat}(n, q) : N \text{ is nilpotent, } QM = MN\}.$$

We compute $\#\mathcal{A}$ in two ways. Let $f(n)$ be the number of nilpotent matrices in $\text{Mat}(n, q)$. We can choose N in $f(n)$ ways. Choose $v \in \mathbb{F}_q^n$ in q^n ways. Then there is a unique matrix $M \in \text{Mat}(n, q)$ with first row v such that $QM = MN$. Hence,

$$\#\mathcal{A} = q^n f(n). \quad (1.164)$$

On the other hand, one can show that if M has rank r , then the number of choices for N so that $QM = MN$ is $f(n - r)q^{r(n-r)}$. Using Exercise 1.192(b) and induction, we get

$$\begin{aligned} \#\mathcal{A} &= f(n) + \sum_{r=1}^n (q^n - 1) \cdots (q^n - q^{r-1}) q^{(n-r)(n-r-1)} \cdot q^{r(n-r)} \\ &= f(n) + q^{n(n-1)} (q^n - 1). \end{aligned}$$

Comparing with equation (1.164) completes the proof.

Third Solution (sketch), due to Greta Panova and Yi Sun (independently), November 2007. Count in two ways the number of $(n + 1)$ -tuples $(N, v_1, v_2, \dots, v_n)$ with N nilpotent in $\text{Mat}(n, q)$, and $v_i \in \mathbb{F}_q^n$ such that $N(v_i) = v_{i+1}$ ($1 \leq i \leq n - 1$) and $v_1 \neq 0$. On the one hand, there are $f(n)(q^n - 1)$ such $(n + 1)$ -tuples since they are determined by N and v_1 . On the other hand, one can show that the number of such $(n + 1)$ -tuples such that $v_k \neq 0$ and $v_{k+1} = 0$ (with $v_{n+1} = 0$ always) is $f(n - k)q^{k(n-k)}(q^n - 1) \cdots (q^n - q^{k-1})$, yielding the recurrence

$$f(n)(q^n - 1) = \sum_{k=1}^n f(n - k)q^{k(n-k)}(q^n - 1) \cdots (q^n - q^{k-1}).$$

The proof follows straightforwardly by induction on n .

For some additional work on counting nilpotent matrices, see G. Lusztig, *Bull. London Math. Soc.* **8** (1976), 77–80.

- 189.** See Proposition 4.27 of J. Yin, *A q -Analogue of Spanning Trees: Nilpotent Transformations over Finite Fields*, Ph.D. thesis, M.I.T., 2009. This result may be regarded as a q -analogue of the fact that the number of spanning trees of the complete bipartite graph $K_{m,n}$ is $m^{n-1}n^{m-1}$ (see Exercise 5.30, Vol. II).

- 190. a.** We can imitate the proof of Proposition 1.10.2, using $\mathcal{I}^* := \mathcal{I} - \{x\}$ instead of \mathcal{I} and β^* (defined by equation (1.112)) instead of β . We therefore get

$$\begin{aligned} \sum_{n \geq 0} \omega^*(n, q) x^n &= \prod_{n \geq 1} \prod_{j \geq 1} (1 - x^{jn})^{-\beta^*(n)} \\ &= \prod_{j \geq 1} \frac{1 - x^j}{1 - qx^j}, \end{aligned} \quad (1.165)$$

from which the proof is immediate.

- b.** This result follows easily from the Pentagonal Number Formula (1.88) and Exercise 1.74. A more careful analysis shows that if $m = \lfloor (n-1)/2 \rfloor$, then

$$\omega^*(n, q) = q^n - q^m - q^{m-1} - q^{m-2} - \dots - q^{2\lfloor (n+5)/6 \rfloor} + O(q^{\lfloor (n+5)/6 \rfloor - 1}).$$

- c.** It follows from the Pentagonal Number Formula and equation (1.165) that

$$\omega^*(n, 0) = \begin{cases} (-1)^k, & \text{if } n = k(3k \pm 1)/2, \\ 0, & \text{otherwise,} \end{cases}$$

We also have

$$\omega^*(n, -1) = \begin{cases} 2(-1)^k, & \text{if } n = k^2, \\ 0, & \text{otherwise,} \end{cases}$$

a consequence of the identity (1.131) due to Gauss.

By differentiating (1.165) with respect to q and setting $q = 0$, it is not hard to see that $\omega^*(n, q)$ is divisible by q^2 if and only if

$$\frac{k(3k-1)}{2} < n < \frac{k(3k+1)}{2}$$

for some $k \geq 1$.

- 191. First Solution** (in collaboration with G. Lusztig). Let F be an algebraic closure of \mathbb{F}_q . We claim that the set Ω of orbits of the adjoint representation of $\mathrm{GL}(n, F)$ has the structure

$$\Omega \cong \bigoplus_{\lambda \vdash n} F^{\ell(\lambda)}. \quad (1.166)$$

Let $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k) \vdash n$, where $\lambda_k > 0$. Given $\alpha = (\alpha_1, \dots, \alpha_k) \in F^k$, let $M = M(\lambda, \alpha) \in \mathrm{Mat}(n, F)$ be defined as follows: M is a direct sum of k Jordan blocks J_1, \dots, J_k , with J_i containing λ_i main diagonal elements equal to α_i . We do yet have a set of orbit representatives, since if we have j blocks of the same size, then they can appear in any order. Hence, the different conjugacy classes formed by j blocks of size m has the structure F^j / \mathfrak{S}_j , where \mathfrak{S}_j acts on F^j by permuting coordinates. But it is well known that $F^j / \mathfrak{S}_j \cong F^j$, namely, the elements of F^j / \mathfrak{S}_j correspond to k -element multisets $\{\alpha_1, \dots, \alpha_k\}$ of elements of F which we associate with $(\beta_1, \dots, \beta_k) \in F^k$ by

$$\prod_{i=1}^k (x - \alpha_i) = x^k + \sum_{j=1}^k \beta_j x^{k-j}.$$

Hence, (1.166) follows. It is now a consequence of standard properties of the Frobenius map $\alpha \mapsto \alpha^q$ that the space Ω_q of orbits of the adjoint representation of $\mathrm{GL}(n, q)$ has an analogous decomposition

$$\Omega_q \cong \bigoplus_{\lambda \vdash n} \mathbb{F}_q^{\ell(\lambda)},$$

and the proof follows.

Second Solution. Let $f(z) \in \mathbb{F}_q[z]$ be a monic polynomial of degree k . Let $f(z) = \prod f_i(z)^{r_i}$ be its factorization into irreducible factors (over \mathbb{F}_q). Let $M_f \in \mathrm{Mat}(n, q)$ be a matrix whose adjoint orbit is indexed by $\Phi: \mathcal{I}(q) \rightarrow \mathrm{Par}$ satisfying $\Phi(f_i) = (r_i)$ (the partition with one part equal to r_i). A specific example of such a matrix is the *companion matrix*

$$M_f = \begin{bmatrix} 0 & 0 & \cdots & 0 & -\beta_0 \\ 1 & 0 & \cdots & 0 & -\beta_1 \\ 0 & 1 & \cdots & 0 & -\beta_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\beta_{k-1} \end{bmatrix},$$

where $f(z) = \beta_0 + \beta_1 z + \cdots + \beta_{k-1} z^{k-1} + z^k$. For fixed k , the space of all such M_f is just an affine space \mathbb{F}_q^k (since it is isomorphic to the space of all monic polynomials of degree k). Now given a partition $\lambda \vdash n$ with conjugate $\lambda' = (\lambda'_1, \lambda'_2, \dots)$, choose polynomials $f_i(z) \in \mathbb{F}_q[z]$ such that $\deg f_i = \lambda'_i$ and $f_{i+1} | f_i$ for all $i \geq 1$. Let $M = M_{f_1} \oplus M_{f_2} \oplus \cdots \in \mathrm{Mat}(n, q)$. For fixed λ , the space of all such M has the structure $\mathbb{F}_q^{\lambda'_1} = \mathbb{F}_q^{\ell(\lambda)}$ (since once f_{i+1} is chosen, there are $q^{\lambda'_{i+1} - \lambda'_i}$ choices for f_i). It is easy to check that the M 's form a cross-section of the orbits as λ ranges over all partitions of n , so the number of orbits is $\sum_{\lambda \vdash n} q^{\ell(\lambda)}$. This argument appears in J. Hua, *J. Combinatorial Theory Ser. A* **79** (1997), 105–117 (Theorem 11).

Third Solution, due to Gabriel Tavares Bujokas and Yufei Zhao (independently), November 2007. We want the number of functions $\Phi: \mathcal{I}(q) \rightarrow \mathrm{Par}$ satisfying $\sum_{f \in \mathcal{I}(q)} |\Phi_M(f)| \cdot \deg(f) = n$. For each $i \geq 1$, let

$$p_i = \prod_{f \in \mathcal{I}(q)} f^{m_i(\Phi(f))},$$

where $m_i(\Phi(f))$ denotes the number of parts of $\Phi(f)$ equal to i . Thus, the p_i 's are arbitrary monic polynomials satisfying $\sum i \deg(p_i) = n$. First, choose $\lambda \vdash \langle 1^{d_1}, 2^{d_2}, \dots \rangle \vdash n$ and then each p_i so that $\deg p_i = d_i$. There are thus $q^{\sum d_i} = q^{\ell(\lambda)}$ choices for the p_i 's, so

$$\omega(n, q) = \sum_{\lambda \vdash n} q^{\ell(\lambda)} = \sum_j p_j(n) q^j.$$

193. A matrix P is a projection if and only if $\Phi_P(z) = \langle 1^k \rangle$ for some k , $\Phi_P(z-1) = \langle 1^{n-k} \rangle$, and otherwise $\Phi_P(f) = \emptyset$. The proof now follows from Theorem 1.10.4 and Lemma 1.10.5 exactly as does Corollary 1.10.6.

194. A matrix M is regular if and only if for all $f \in \mathcal{I}(q)$ there is an integer $k \geq 0$ such that $\Phi_M(f) = (k)$. Write $c_f(k)$ for $c_f(\lambda)$ when $\lambda = (k)$. From Theorem 1.10.7, we have

$$c_f(k) = q^{kd} - q^{(k-1)d}, \quad k \geq 1,$$

where $d = \deg(f)$. Substitute $t_{f,\lambda} = 1$ if $\lambda = (k)$ and $t_{f,\lambda} = 0$ otherwise in Theorem 1.10.4 to get

$$\begin{aligned} \sum_{n \geq 0} r_n \frac{x^n}{\gamma_n} &= \prod_{f \in \mathcal{I}} \left(1 + \sum_{k \geq 1} \frac{x^{k \cdot \deg(f)}}{q^{k \cdot \deg(f)} - q^{(k-1) \cdot \deg(f)}} \right) \\ &= \prod_{d \geq 1} \left(1 + \sum_{k \geq 1} \frac{x^{kd}}{q^{kd}(1 - q^{-d})} \right)^{\beta(d)} \\ &= \prod_{d \geq 1} \left(1 + \frac{(x/q)^d}{(1 - q^{-d})(1 - (x/q)^d)} \right)^{\beta(d)} \\ &= \prod_{d \geq 1} \left(1 + \frac{x^d}{(q^d - 1)(1 - (x/q)^d)} \right)^{\beta(d)}. \end{aligned}$$

We can write this identity in the alternative form

$$\sum_{n \geq 0} r_n \frac{x^n}{\gamma_n} = \frac{1}{1-x} \prod_{d \geq 1} \left(1 + \frac{x^d}{q^d(q^d - 1)} \right)^{\beta(d)}$$

by using equation (1.145) with x/q substituted for x .

195. A matrix M is semisimple if and only if for all $f \in \mathcal{I}(q)$ there is an integer $k \geq 0$ such that $\Phi_M(f) = \langle 1^k \rangle$. The proof now follows from Theorem 1.10.4 and Lemma 1.10.5 exactly as does Corollary 1.10.6.

196. a. The proof parallels that of Proposition 1.10.15. We partition \mathfrak{S}_n into two classes \mathcal{A} and \mathcal{B} , where

$$\mathcal{A} = \{w \in \mathfrak{S}_n : w \neq 12 \cdots ku \text{ for some } u \in \mathfrak{S}_{[k+1,n]}\},$$

$$\mathcal{B} = \{w \in \mathfrak{S}_n : w = 12 \cdots ku \text{ for some } u \in \mathfrak{S}_{[k+1,n]}\}.$$

Let

$$G(n, k, q) = \{A \in \text{GL}(n, q) : A_{11} + \cdots + A_{kk} = 0\}.$$

For $w \in \mathcal{A}$, we have

$$\#(\Gamma_w \cap G(n, k, q)) = \frac{1}{q} \# \Gamma_w.$$

For $w \in \mathcal{B}$, we have

$$\begin{aligned} \#(\Gamma_w \cap G(n, k, q)) &= q^{\binom{n}{2} + \text{inv}(w)} (q-1)^{n-k} a_k \\ &= q^{\binom{n}{2} + \text{inv}(w)} (q-1)^{n-k} ((q-1)^k + (-1)^k (q-1)). \end{aligned}$$

Hence,

$$\sum_{w \in \mathcal{B}} \#(\Gamma_w \cap G(n, k, q)) = \frac{1}{q} \left(\sum_{w \in \mathcal{B}} q^{\binom{n}{2} + \text{inv}(w)} (q-1)^n \right)$$

$$\begin{aligned}
& + (-1)^k (q-1) q^{\binom{n}{2} - \binom{n-k}{2}} \sum_{u \in \mathfrak{S}_{[k+1, n]}} q^{\binom{n-k}{2} + \text{inv}(u)} (q-1)^{n-k} \Big) \\
& = \frac{1}{q} \left(\sum_{w \in \mathcal{B}} (\# \Gamma_w) + (-1)^k (q-1) q^{\frac{1}{2}k(2n-k-1)} \gamma_{n-k}(q) \right),
\end{aligned}$$

and the proof follows.

- b.** The hyperplane H can be defined by $H = \{M \in \text{Mat}(n, q) : M \cdot N = 0\}$, where N is a fixed nonzero matrix in $\text{Mat}(n, q)$ and $M \cdot N = \text{tr}(MN^t)$, the standard dot product in the vector space $\text{Mat}(n, q)$. If $P, Q \in \text{GL}(n, q)$, then $M \cdot N = 0$ if and only if $((P')^{-1}M(Q')^{-1}) \cdot (PNQ) = 0$. Since two matrices $N, N' \in \text{Mat}(n, q)$ are related by $N' = PNQ$ for some $P, Q \in \text{GL}(n, q)$ if and only if they have the same rank, it follows that $\#(\text{GL}(n, q) \cap H)$ depends only on $\text{rank}(N)$. If $\text{rank}(N) = k$, then we may take

$$N_{ij} = \begin{cases} 1, & 1 \leq i = j \leq k, \\ 0, & \text{otherwise.} \end{cases}$$

Hence, $\#(\text{GL}(n, q) \cap H)$ is given by the right-hand side of equation (1.150).

- 197.** *Hint.* Let $f(n, k)$ be the number of $k \times n$ matrices over \mathbb{F}_q of rank k with zero diagonal, where $1 \leq k \leq n$. Show that

$$f(n, k+1) = q^{k-1} (q-1) (f(n, k) \cdot (n-k) - f(n-1, k)),$$

with the initial condition $f(n, 1) = q^{n-1} - 1$. The solution to this recurrence is

$$f(n, k) = q^{\binom{k-1}{2}} (q-1)^k \left(\sum_{i=0}^k (-1)^i \binom{k}{i} \frac{(n-i)!}{(n-k)!} \right).$$

Now set $k = n$.

This result is due to J. B. Lewis, R. I. Liu, A. H. Morales, G. Panova, S. V. Sam, and Y. Zhang, *Matrices with restricted entries and q -analogues of permutations*, arXiv:1011.4539 (Proposition 2.2). This paper contains a host of other results about counting matrices over \mathbb{F}_q . A further result in this paper is given by Exercise 1.199.

- 198. a.** An $(n+1) \times (n+1)$ symmetric matrix may be written as

$$N = \begin{bmatrix} \beta & y \\ y^t & M \end{bmatrix},$$

where M is an $n \times n$ symmetric matrix, $\beta \in \mathbb{F}_q$, and $y \in \mathbb{F}_q^n$. Elementary linear algebra arguments show that from a particular matrix M of rank r we obtain:

- $q^{n+1} - q^{r+1}$ matrices N of rank $r+2$,
- $(q-1)q^r$ matrices N of rank $r+1$,
- q^r matrices N of rank r ,
- no matrices of other ranks.

The recurrence (1.151) follows. This recurrence (with more details of the proof) was given by J. MacWilliams, *Amer. Math. Monthly* **76** (1969), 152–164, and was used to prove (b). A simpler recurrence for $h(n, n)$ alone was given by G. Lusztig, *Transformation Groups* **10** (2005), 449–487 (end of §3.14).

- b.** We can simply verify that the stated formula for $h(n, r)$ satisfies the recurrence (1.151), together with the initial conditions. For some generalizations and further

information, see R. Stanley, *Ann. Comb.* **2** (1998), 351–363; J. R. Stembridge, *Ann. Comb.* **2** (1998), 365–385; F. Chung and C. Yang, *Ann. Comb.* **4** (2000), 13–25; and P. Belkale and P. Brosnan, *Duke Math. J.* **116** (2003), 147–188.

NOTE. There is a less ad hoc way to compute the quantity $h(n, n)$. Namely, $\mathrm{GL}(n, q)$ acts on $n \times n$ invertible symmetric matrices M over \mathbb{F}_q by $A \cdot M = A^t M A$. This action has two orbits whose stabilizers are the two forms of the orthogonal group $O(n, q)$. The orbit sizes can be easily computed from standard facts about $O(n, q)$. For further details, see R. Stanley, op. cit. (§4).

- 199. a.** The equality of the first two items when q is even is due to J. MacWilliams, *Amer. Math. Monthly* **76** (1969), 152–164 (Theorems 2, 3). The equality of the second two items appears in O. Jones, *Pacific J. Math.* **180** (1997), 89–100. For the remainder of the exercise, see Section 3 of the paper of Lewis–Liu–Morales–Panova–Sam–Zhang cited in the solution to Exercise 1.197.
- 200.** This result was conjectured by A. A. Kirillov and A. Melnikov, in *Algèbre non commutative, groupes quantiques et invariants* (Reims, 1995), Sémin. Congr. **2**, Soc. Math. France, Paris, 1997, pp. 35–42, and proved by S. B. Ekhad and D. Zeilberger, *Electron. J. Combin.* **3**(1) (1996), R2. No conceptual reason is known for such a simple formula.
- 201. a.** The result follows from the theory of Gauss sums as developed, for example, in K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, 2nd ed., Springer-Verlag, New York, 1990, and may have been known to Gauss or Eisenstein. This information was provided by N. Elkies (private communication, 1 August 2006).
- b.** The argument is analogous to the proof of Proposition 1.10.15. Let

$$G = \{A \in \mathrm{GL}(3, q) : \mathrm{tr}(A) = 0, \det(A) = 1\}.$$

If $123 \neq w \in \mathfrak{S}_3$, then $\#(\Gamma_2 \cap G) = \frac{1}{q(q-1)} \# \Gamma_w$. On the other hand, $\#(\Gamma_{123} \cap G) = q^3 f(q)$. Hence, we get

$$\begin{aligned} \#G &= \frac{1}{q(q-1)} (\gamma(3, q) - \# \Gamma_{123}) + q^3 f(q) \\ &= q^3 (q-1)^2 (q^2 + 2q + 2) + q^3 f(q). \end{aligned}$$

- 202.** This result is an instance of the Shimura–Taniyama–Weil conjecture, namely, every elliptic curve is modular. An important special case of the conjecture (sufficient to imply Fermat’s Last Theorem) was proved by A. Wiles in 1993, with a gap fixed by Wiles and R. Taylor in 1994. The full conjecture was proved by Breuil, Conrad, Diamond, and Taylor in 1999. Our example follows H. Darmon, *Notices Amer. Math. Soc.* **46** (1999), 1397–1401, which has much additional information.
- 203.** The statement about 103,049 was resolved in January 1994 when David Hough, then a graduate student at George Washington University, noticed that 103,049 is the total number of bracketings of a string of 10 letters. The problem of finding the number of bracketing of a string of n letters is known as *Schröder’s second problem* and is discussed in Section 6.2, Vol. II. See also the Notes to Chapter 6, Vol. II, where also a possible interpretation of 310,952 is discussed. Hough’s discovery was first published by R. Stanley, *Amer. Math. Monthly* **104** (1997), 344–350. A more scholarly account was given by F. Acerbi, *Archive History Exact Sci.* **57** (2003), 465–502.